Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2020, Vol. 56, No. 1, 265–295
https://doi.org/10.1214/19-AIHP961

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

On laws of large numbers in L^2 for supercritical branching Markov processes beyond λ-positivity

Matthieu Jonckheere[a] and Santiago Saglietti[b]

[a]*Departamento de Matemática, Universidad de Buenos Aires, 1428 Buenos Aires, Argentina. E-mail: mjonckhe@dm.uba.ar*
[b]*Faculty of Industrial Engineering and Management, Technion, 3200003 Haifa, Israel. E-mail: saglietti.s@campus.technion.ac.il*

Received 16 November 2017; revised 27 January 2019; accepted 28 January 2019

Abstract. We give necessary and sufficient conditions for laws of large numbers to hold in L^2 for the empirical measure of a large class of branching Markov processes, including λ-positive systems but also some λ-transient ones, such as the branching Brownian motion with drift and absorption at 0. This is a significant improvement over previous results on this matter, which had only dealt so far with λ-positive systems. Our approach is purely probabilistic and is based on spinal decompositions and many-to-few lemmas. In addition, we characterize when the limit in question is always strictly positive on the event of survival, and use this characterization to derive a simple method for simulating (quasi-)stationary distributions.

Résumé. Nous obtenons des conditions nécessaires et suffisantes pour des lois des grands nombres dans L^2 concernant les mesures empiriques d'une large classe de processus de Markov branchants, comme le mouvement Brownien branchant avec dérive et absorption en 0. Cela constitue un pas significatif pour ce genre de résultats qui étaient jusqu'à présent limités aux processus λ-positifs. Notre approche est purement probabiliste et est basée sur des décompositions en épine (spine) et des lemmes associés. De plus, nous caractérisons la stricte positivité de la limite quand le processus de branchement survit et utilisons cette caractérisation pour donner une méthode simple de simulation de distributions (quasi-)stationaires.

MSC: 60F99; 60J80

Keywords: Law of large numbers; Branching Markov processes; Spine decomposition; h-transform; λ-positivity

1. Introduction

Since the late seventies there has been an intensive effort of research dedicated to proving laws of large numbers for general branching Markov processes (or BMPs, for short). For us, a BMP will be characterized by the following dynamics:

i. It starts with a single particle located at some point x of a measurable space $(J, \mathcal{B}_J)$.
ii. This particle moves within J according to some prescribed Markovian dynamics.
iii. Independently of its motion, the particle also branches at constant rate $r > 0$, dying on the spot and giving birth to a random number of new particles having some prescribed distribution υ on $\mathbb{N}_0$.
iv. Each of these new particles then independently mimics the stochastic behavior of its parent, relative to their own starting point.

Given a BMP $\xi = (\xi_t)_{t\geq 0}$ on some space $(J, \mathcal{B}_J)$, by a law of large numbers (LLN, for short) we shall understand the following: there exists a nonempty class $\mathcal{C} \subseteq \mathcal{B}_J$, a measure ν on $(J, \mathcal{B}_J)$ and a nonnegative random variable D_∞ such that for all pairs $B, B' \in \mathcal{C}$ with $\nu(B') \neq 0$

$$\frac{\xi_t(B)}{\mathbb{E}(\xi_t(B'))} \longrightarrow \frac{\nu(B)}{\nu(B')} \cdot D_\infty \tag{1}$$

where, for any $A \in \mathcal{B}_J$, we denote by $\xi_t(A)$ the number of particles of ξ inside the set A at time t. The earliest results in this regard can be found in [2,5,47,51]. Later, the renovated approach in [39] introducing the notion of spine for the branching process sprouted a multitude of new results, see for instance [21,31,32]. In the recent works [15,16], functional analytic methods were used to obtain results in the setting of branching symmetric Hunt processes. LLNs were also investigated in the related context of superprocesses, see [13,14,23,24,41]. For a more detailed overview of past results and recent developments on these matters, we refer the interested reader to [18,20]. See also [36, Section 2.5].

In the study of LLNs, a property which has proved crucial in most of the approaches developed so far is the λ-positivity of the BMP ξ. Essentially, λ-positivity means that the motion of a certain spine describing the genealogy of the branching process (which is sometimes referred to in the literature as the *immortal particle*) is positive recurrent. To the best of our knowledge, all previous works on the topic have required some form of this hypothesis, with [19, 51] being the only remarkable exceptions we could find. However, recent questions stemming from particle systems demand a better understanding of BMPs which are not λ-positive. Indeed, there exists a large body of literature studying empirical measures of population models with mutations and selection [4,6,10–12,17,25,27,44], for which hydrodynamic limits were obtained on finite time windows [3,6,17,50]. However, it is still an open problem in many of these systems how to obtain scaling limits of the empirical measure in its stationary regime, see [4,27,44]. Different couplings with branching Markov processes have been proposed to study this problem, but typically the resulting branching process is not λ-positive, see for example [4,44]. Thus, this highlights the need of a convergence theory for empirical measures of branching Markov processes going past this assumption of λ-positivity. A canonical example appearing in this context is the Branching Brownian Motion with drift and absorption at 0. Kesten introduced this model in his paper [38] from 1978 and stated there that a strong law of large numbers holds for this process whenever it is supercritical, although he did not provide a proof of this fact nor did he made any similar assertions regarding L^2-convergence. As it turns out, the method of Watanabe in [51] may be adapted to establish the strong LLN for this model, but only in the (strictly smaller) subregion of the supercritical regime in which the LLN also holds in L^2, see Section 3.6 for details. However, it was not until very recently that the full validity of Kesten's claim was proved in [43], almost forty years later, building upon the results from the present article. A weak analogue of this LLN for super-Brownian motion (without drift) was also obtained in [19], though it does not address the question of convergence in L^2.

Once a law of large numbers as in (1) is established, it is then of particular interest to determine whether $D_\infty > 0$ holds almost surely in the event the branching process survives forever (assuming it is supercritical, so that there is a positive probability that this occurs). Indeed, assuming that $J \in \mathcal{C}$ and $\nu(J) = 1$, it is usually simple to check that the former statement implies the following convergence for the empirical measure associated with ξ: conditionally on the event of survival, for all $B \in \mathcal{C}$ one has that

$$\frac{\xi_t(B)}{\xi_t(J)} \longrightarrow \nu(B). \tag{2}$$

(The sense in which the convergence in (2) holds will depend on that of (1), although for example convergence in probability in (1) already gives convergence in probability and in L^p for all $p \geq 1$ in (2).) In turn, we see that whenever $D_\infty > 0$ holds almost surely in the event of survival, we obtain a more complete description of the asymptotic behavior of the branching process as $t \to +\infty$:

i. For any $B \in \mathcal{C}$ with $\nu(B) \neq 0$, the number of particles $\xi_t(B)$ in B grows like its expectation (times D_∞, which acts as a random scale constant).
ii. If $J \in \mathcal{C}$ and $\nu(J) = 1$, the proportion of particles $\frac{\xi_t(B)}{\xi_t(J)}$ inside any given set $B \in \mathcal{C}$ behaves asymptotically as ν (in the sense of (2)).

Furthermore, we also see that (2) yields a simple and direct method to simulate the distribution ν which, in many cases, might not be known explicitly. However, showing that $D_\infty > 0$ upon survival is a subtle question (see [29,30] for some specific examples) and, in fact, it may not always be true, see Section 3.5. In general, the literature addressing this matter is very limited, specially if not under the λ-positivity assumption and whenever dealing with more involved situations with absorption and/or infinite state spaces.

Our contribution in this article is two-fold. First, we derive a necessary and sufficient condition for laws of large numbers to hold in L^2 for a wide class of supercritical branching Markov processes having *fixed* (i.e., not state-dependent) branching rates and offspring distributions with a finite second moment, which includes many λ-positive systems but also other highly relevant examples without this property.

More precisely, we show in Theorem 2.6 that, whenever the immortal particle process mentioned above (which is well-defined even for non-λ-positive systems) has a distribution which is of regular variation as $t \to \infty$ (see Section 2.2 for details), then (1) holds in L^2 if and only if a specific additive martingale associated with ξ is bounded in L^2. Furthermore, we show that, in the latter case, D_∞ is precisely the L^2-limit of this martingale. Then, in Proposition 2.5, we obtain an explicit formula for the asymptotic variance of this martingale, so that one can determine whether it is indeed bounded in L^2 by performing a direct computation. Our approach is purely probabilistic and is based solely on simple spinal decomposition techniques, namely the "many-to-one/two" lemmas, which allow us to effectively control the particle correlations as time tends to infinity.

We then focus on studying conditions which guarantee that $D_\infty > 0$ almost surely upon survival, i.e. $P(D_\infty > 0|\text{survival}) = 1$ (notice that this is in fact stronger than just the non-degeneracy of D_∞, which only amounts to having $P(D_\infty > 0|\text{survival}) > 0$). In Theorem 2.10 we show that, whenever (1) holds in L^2, the equality $P(D_\infty > 0|\text{survival}) = 1$ is equivalent to the process ξ being *strongly supercritical*: heuristically, this means that, in the event of survival, particles of ξ can never accumulate all together on the boundary of the state space (or, equivalently, that as $t \to \infty$ one can always see infinitely often at least one particle of ξ lying in the "bulk" of the state space, where by bulk we understand the region of the state space which is not adjacent to the boundary, see (13) below for a more precise statement). Strong supercriticality is related to the notion of *strong local survival* studied in [7] and references mentioned therein, although we are not aware of any previous connections made between this and the strict positivity of D_∞ on survival.

Finally, we illustrate our results through a series of examples. First, Proposition 8.1 shows that any λ-positive system whose associated immortal particle relaxes sufficiently fast to equilibrium (i.e. it admits a geometric Lyapunov functional growing sufficiently fast at infinity) verifies our hypotheses and thus satisfies a law of large numbers in L^2. Moreover, Theorem 2.16 shows that, whenever this Lyapunov functional does not grow too fast, the almost sure convergence also holds. Afterwards, we use this to obtain laws of large numbers for several classic λ-positive systems: branching ergodic motions, branching Galton–Watson processes, branching contact processes and branching inward/outward Ornstein–Uhlenbeck processes. Finally, we study the emblematic case of the Branching Brownian Motion with drift and absorption at 0 presented in [38], which is not λ-positive. For this system we completely characterize the region of parameters for which a LLN holds in L^2 and, in particular, show that it is strictly smaller than the region of parameters for which the process is supercritical. The method we use to show this is simpler and more general than the one developed by Watanabe in [51], which relies heavily on Fourier analysis.

As a final remark, we believe that all of our results should also extend to the more general case of non-constant (i.e. state-dependent) branching rates and/or offspring distributions without any major modifications to our hypotheses nor to the overall method. However, we chose to leave this outside the scope of the present article since it would somewhat complicate the overall notation as well as hide the intuition behind our current assumptions.

The rest of the article is structured as follows. In Section 2 we describe the basic setup and notation, and then state our main results. Section 3 discusses several examples and applications. In Section 4 we recall the many-to-few lemma, which constitutes one of the main tools of our analysis, while Sections 5, 6, 7 and 8 contain the proofs of our main results.

2. Preliminaries and main results

2.1. *Preliminaries*

Let $X = (X_t)_{t\geq 0}$ be a homogeneous Markov process with cadlag trajectories on some metric space $\overline{J}$. We will assume throughout that:

- X is allowed to have *absorbing states*, i.e. $x \in \overline{J}$ such that, whenever $X_t = x$ for some t, one has $X_s = x$ for all times $s > t$.
- The set $\partial_* \overline{J}$ of all absorbing states of X belongs to $\mathcal{B}$, the Borel σ-algebra of $\overline{J}$.
- $J := \overline{J} - \partial_* \overline{J}$ is locally compact and separable.

Let us now recall how the dynamics of a BMP goes:

i. The dynamics starts with a single particle, located initially at some starting point $x \in \overline{J}$, whose position evolves according to a copy of the Markov process X.
ii. This initial particle branches at a fixed rate $r > 0$ (independently of the motion it describes) and, whenever it does so, it dies and gets replaced at its current position by an independent random number of particles m having some fixed distribution υ on $\mathbb{N}_0$.
iii. Starting from their birth position, now each of these new particles independently mimics the same stochastic behavior of its parent.
iv. If a particle has 0 children, then it dies and moves to a graveyard state Δ forever.

A rigorous construction of such dynamics can be given in terms of *marked Galton–Watson trees*, where particles of our branching process are identified with the individuals on the GW-tree, each individual carrying a mark which represents the trajectory of the associated particle, see e.g. [32]. However, since we will not make any explicit use of this construction, we shall not give the details of it here and instead refer the interested reader to [32, Section 5].

Given any time $t \geq 0$, for each particle u in the dynamics at time t we shall write u_t to indicate its position at time t. Also, we let $\overline{\chi}_t$ denote the collection of particles in the branching dynamics which are alive at time t, i.e. $u_t \notin \Delta$. We identify $\overline{\chi}_t$ with a finite measure χ_t on $(\overline{J}, \mathcal{B})$ by setting

$$\chi_t := \sum_{u \in \overline{\chi}_t} \delta_{u_t}.$$

Furthermore, let $\overline{\xi}_t$ denote the collection of particles u in $\overline{\chi}_t$ which have not been absorbed yet, i.e. such that $u_t \in J$, and define its induced measure ξ_t on $(J, \mathcal{B}_J)$ as

$$\xi_t = \sum_{u \in \overline{\xi}_t} \delta_{u_t}.$$

Finally, we write $|\xi_t| := \xi_t(J)$ for the total mass of ξ_t, i.e. the number of living particles at time t which have not been absorbed yet, and define the *empirical measure* ν_t as

$$\nu_t := \frac{1}{|\xi_t|} \cdot \xi_t$$

with the convention that $\infty \cdot 0 = 0$, used whenever $|\xi_t| = 0$.

Since throughout most of the article the starting position x of our branching dynamics is fixed, we shall typically omit the dependence on x of the processes under consideration from the notation, the only exception being throughout Section 7 where we will be forced to work simultaneously with different initial conditions and so writing out this dependence will become necessary. However, since many of the quantities computed in the sequel will indeed depend on the precise value of x, we shall use the subscript x in the notation, e.g. by writing P_x or $\mathbb{E}_x$, to remind the reader that the process involved in the corresponding probability or expectation starts at x.

2.2. *A necessary and sufficient condition for laws of large numbers in L^2*

We now begin to present and discuss our main results. Before we can do so, however, we must introduce some assumptions on our branching dynamics. Our initial assumptions on the underlying motion X are the following.

Assumptions 2.1.

A1. There exists $\lambda \geq 0$ and a nonnegative $\mathcal{B}$-measurable function $h : \overline{J} \to \mathbb{R}_{\geq 0}$ such that:
  i. $h(x) = 0$ if and only if $x \in \partial_* \overline{J}$.
  ii. The process $M = (M_t)_{t \geq 0}$ given by the formula

$$M_t := \frac{h(X_t)}{h(X_0)} e^{\lambda t},$$

  is a (mean-one) square-integrable martingale, i.e. $-\lambda$ is a (right) eigenvalue of $\mathcal{L}$ with associated eigenfunction h satisfying $\mathbb{E}_x(h^2(X_t)) < +\infty$ for all $t \geq 0$ and $x \in J$.

A2. There exists a nonempty class of subsets $\mathcal{C}_X \subseteq \mathcal{B}_J$ such that for each $x \in \overline{J}$ and $B \in \mathcal{C}_X$ one has the asymptotic formula

$$P_x(X_t \in B) = h(x)p(t)e^{-\lambda t}\big(\nu(B) + s_B(x,t)\big), \tag{3}$$

for all $t > 0$, where λ and h are those from (A1) and:
  i. ν is a (non-necessarily finite) measure on $(J, \mathcal{B}_J)$ satisfying:
  - $\nu(B) \in [0, +\infty)$ for all $B \in \mathcal{C}_X$,
  - there exists at least one $B' \in \mathcal{C}_X$ such that $\nu(B') > 0$.

  ii. $p(t)$ is a regularly varying function at infinity, i.e. a function $p : (0, +\infty) \to (0, +\infty)$ such that the limit

$$\ell(a) := \lim_{t \to +\infty} \frac{p(at)}{p(t)}$$

  exists and is finite for all $a > 0$.

iii. $s_B(\cdot,t)$ converges to zero as $t \to +\infty$ uniformly over $J_n := \{x \in J : \frac{1}{n} \le h(x) \le n\}$ for each $n \in \mathbb{N}$.
iv. There exist $t_0, \overline{s}_B > 0$ such that $\sup_{x\in \overline{J}} s_B(x,t) \le \overline{s}_B$ for all $t > t_0$.

Remark 2.2. Note that, if we consider the change of measure $\tilde{P}$ (also known as h-transform) given by

$$\frac{d\tilde{P}}{dP}\bigg|_{\mathcal{F}_t} = M_t, \tag{4}$$

where $(\mathcal{F}_t)_{t\ge 0}$ denotes the filtration generated by X, and also define the measure μ on $(J, \mathcal{B}_J)$ via the formula

$$\frac{d\mu}{d\nu} = h, \tag{5}$$

then

$$P_x(X_t \in B) = \mathbb{E}_x\big(\mathbb{1}_B(X_t)\big) = h(x)e^{-\lambda t}\tilde{\mathbb{E}}_x\left(\frac{\mathbb{1}_B}{h}(X_t)\right) = h(x)p(t)e^{-\lambda t}\big(\nu(B) + s_B(x,t)\big) \tag{6}$$

where $\tilde{\mathbb{E}}$ denotes expectation with respect to the measure $\tilde{P}$ and s_B is given by

$$s_B(x,t) := \frac{1}{p(t)}\tilde{\mathbb{E}}_x\left(\frac{\mathbb{1}_B}{h}(X_t)\right) - \nu(B) = \frac{1}{p(t)}\tilde{\mathbb{E}}_x\left(\frac{\mathbb{1}_B}{h}(X_t)\right) - \mu\left(\frac{\mathbb{1}_B}{h}\right). \tag{7}$$

Thus, (A2) can be reformulated as the assumption that there exists a regularly varying function p at infinity such that

$$\lim_{t\to+\infty} \frac{1}{p(t)}\tilde{\mathbb{E}}_x\big(f(X_t)\big) = \mu(f) \tag{8}$$

for all f of the form $f = \frac{\mathbb{1}_B}{h}$ with $B \in \mathcal{C}_X$ and, moreover, that this limit is uniform over each J_n. In particular, it follows that, for any $B \in \mathcal{C}_X$ such that $\nu(B) > 0$, the function

$$\ell_B^{(x)}(t) := \tilde{\mathbb{E}}_x\left(\frac{\mathbb{1}_B}{h}(X_t)\right)$$

is of regular variation at infinity. In conclusion, we may regard (A2) as requiring that, under $\tilde{P}$, the distribution of X be (in some sense) of regular variation at infinity uniformly over each J_n.

Assumptions 2.1 are satisfied in many different situations. Indeed, as we will see in Section 2.4, the underlying motion of any λ-positive branching process (see Section 2.4 for a precise definition) is a natural example of system verifying these conditions, and this already covers a wide range of possibilities: not only are ergodic motions in this category, but also certain transient systems as well as many examples of almost-surely absorbed motions having ν as their Yaglom limit. Moreover, as mentioned in the Introduction, Assumptions 2.1 are also satisfied by processes which are not λ-positive, see [38,51] and Section 3.6 below for Brownian motion with nonpositive drift and absorption at 0 and also [40] for other examples of Lévy processes satisfying these assumptions.

On the other hand, we point out that the class $\mathcal{C}_X$ was introduced in Assumptions 2.1 because, in general, one cannot expect the asymptotics in (3) to be valid for *every* $B \in \mathcal{B}_J$, see Section 3.5. Nevertheless, even if this is not the case one can still produce convergence results which hold for all B in this smaller class $\mathcal{C}_X$. Finally, we observe that in all the examples of Section 3 the measure ν actually corresponds to the left eigenmeasure of the generator $\mathcal{L}$ associated to $-\lambda$. However, this fact will not be used throughout our analysis.

Now, since we are interested in understanding the evolution in L^2 of the branching dynamics ξ in the supercritical case in which $|\xi_t|$ remains positive for all times $t > 0$ with positive probability, we must also make the following assumptions on m and r.

Assumptions 2.3. We shall assume throughout that the pair (m, r) satisfies:

I1. $m_2 := \mathbb{E}(m^2) < +\infty$ and $m_1 := \mathbb{E}(m) > 1$.
I2. $r(m_1 - 1) > \lambda$, where λ is the parameter from Assumptions 2.1.

It follows from Assumptions 2.3 and Lemmas 4.1–4.3 below that all second moments $\mathbb{E}_x(|\xi_t|^2)$ are well-defined for any $x \in J$ and $t \geq 0$, and also $\mathbb{E}_x(|\xi_t|) \to +\infty$ as $t \to +\infty$ for all x as long as there exists at least one $B \in \mathcal{C}_X$ with $\nu(B) > 0$, which is precisely the case that interests us.

Now, our first result is concerned with the L^2-convergence of the so-called *Malthusian martingale* associated to our branching dynamics, which we define below.

Definition 2.4. We define the *Malthusian martingale* $D = (D_t)_{t\geq 0}$ as

$$D_t = D_t^{(x)} := \frac{1}{h(x)} \sum_{u \in \bar{\xi}_t} h(u_t) e^{-(r(m_1-1)-\lambda)t}.$$

It follows from the many-to-one lemma in Section 4 and (A1) that D is indeed a martingale. Furthermore, (A1) implies in fact that D is square-integrable. Being also nonnegative, we know that there exists an almost sure limit D_∞. Our first result is then the following, variants of which can also be found in [33–35,51].

Proposition 2.5. *For every $x \in J$ we have that*

$$\lim_{t\to+\infty} \mathbb{E}_x\big(D_t^2\big) = (m_2 - m_1) \int_0^\infty \mathbb{E}_x\big(M_s^2\big) r e^{-r(m_1-1)s}\, ds =: \Phi_x,$$

so that D converges in L^2 to D_∞ if and only if $\Phi_x < +\infty$. In this case, for each $x \in J$ we have

$$\mathbb{E}_x(D_\infty) = 1 \quad \textit{and} \quad \mathbb{E}_x\big(D_\infty^2\big) = \Phi_x.$$

We should point out that the $\Phi_x < +\infty$ condition is not trivial under Assumptions 2.1, so that L^2-convergence may not always hold, see e.g. Section 3.6. However, the L^2-convergence of D is crucial as it dictates the validity of a law of large numbers in L^2 for ξ, as our next result shows.

Theorem 2.6. *Let $\mathcal{C}_X \subseteq \mathcal{B}_J$ be any nonempty class of subsets satisfying the conditions in* (A2). *If for $x \in J$ and $B, B' \in \mathcal{C}_X$ with $\nu(B') > 0$ we define $W(B, B') = (W_t(B, B'))_{t\geq 0}$ by the formula*

$$W_t\big(B, B'\big) = W_t^{(x)}\big(B, B'\big) := \frac{\xi_t(B)}{\mathbb{E}_x(\xi_t(B'))}$$

(*which is well-defined for t large enough since* $\liminf_{t\to+\infty} \mathbb{E}_x(\xi_t(B')) > 0$ *by Lemma* 4.1 *and* (3)), *then the following holds*:

i. *The sequence $W(B, B')$ satisfies*

$$\lim_{t\to+\infty} \mathbb{E}_x\big(W_t^2\big(B, B'\big)\big) = \left[\frac{\nu(B)}{\nu(B')}\right]^2 \Phi_x.$$

In particular, it is bounded in L^2 if and only if $\Phi_x < +\infty$.

ii. *If $W(B, B')$ is bounded in L^2 then we have that as $t \to +\infty$*

$$W_t\big(B, B'\big) \xrightarrow{L^2} \frac{\nu(B)}{\nu(B')} \cdot D_\infty. \tag{9}$$

In particular, conditionally on the event $\{D_\infty > 0\}$, we have that as $t \to +\infty$

$$\nu_t\big(B, B'\big) := \frac{\xi_t(B)}{\xi_t(B')} \xrightarrow{P} \frac{\nu(B)}{\nu(B')}. \tag{10}$$

We note that $W(B, B')$ will not be a martingale in general, not even when $B = B' = J$ due to the presence of absorption, so that the existence of a limit in L^2 whenever it is bounded is by no means a trivial statement. Still, the main idea behind Theorem 2.6 is that, whenever $W(B, B')$ is L^2-bounded, it behaves asymptotically as the martingale $\frac{\nu(B)}{\nu(B')} \cdot D$ and thus it must also converge.

2.3. *Strict positivity of D_∞ on the event of non-extinction*

Let us observe that the event $\Lambda := \{D_\infty > 0\}$ is contained in the event of non-extinction $\Theta := \{|\xi_t| > 0 \text{ for all } t\}$. Ideally, we would like both events to be almost surely equal, i.e.

$$P_x(\Lambda|\Theta) = 1, \tag{11}$$

since, in that case, Theorem 2.6 would tell us that, almost surely on the event of non-extinction, for any $B \in \mathcal{C}_X$ with $\nu(B) > 0$ the number of particles $\xi_t(B)$ grows like its expectation which, using the results from Section 4, can be explicitly computed and, furthermore, that these particles distribute themselves according to ν. The following result intends to give conditions under which the equality in (11) is guaranteed. It is based on the study of the moment generating operator associated to the branching dynamics, which we define now.

Definition 2.7. Let $\mathfrak{B}$ denote the class of measurable functions $g : (J, \mathcal{B}_J) \to ([0,1], \mathcal{B}_{[0,1]})$, where $\mathcal{B}_{[0,1]}$ denotes the Borel σ-algebra on $[0,1]$. We define the *moment generating operator* $G : \mathfrak{B} \to \mathfrak{B}$ by the formula

$$G(g)(x) := \mathbb{E}_x\left(\prod_{u \in \bar{\xi}_1} g(u_1)\right)$$

with the convention that $\prod_{u\in\varnothing} = 1$, used whenever $|\xi_1| = 0$.

Notice that the fact that G is well-defined, i.e. that $G(f)$ is measurable for all $f \in \mathcal{B}$, follows from the measurability of the map $x \mapsto \mathbb{E}_x(f(X_t))$ for all $t \geq 0$ and $f \in \mathfrak{B}$, which in turn holds if the process X admits a Markov kernel.

Now, it is immediate to see that $\mathbf{1}$, the function constantly one on J, is a fixed point of G, i.e. $G(\mathbf{1}) = \mathbf{1}$. Furthermore, by the branching property of the dynamics one has that the functions

$$\eta(x) := P_x\big(\Theta^c\big) \quad \text{and} \quad \sigma(x) := P_x(D_\infty = 0) \tag{12}$$

are also fixed points of G, see Proposition 7.1 and Assumptions 2.8 below.[1] Since clearly $\eta \leq \sigma$ and we also have $\sigma \neq \mathbf{1}$ since $\mathbb{E}_x(D_\infty) = 1$ for all $x \in J$ by Theorem 2.6, if we show that G has at most two fixed points then this would imply that $\eta \equiv \sigma$ and so (11) would follow at once. Unfortunately, it is not always the case that G has only two fixed points, see e.g. Section 3.5. Hence, we must impose some additional conditions for this to occur. First, we make some further assumptions.

Assumptions 2.8. Throughout Section 2.3 we will make the following additional assumptions:

B0. The function σ is measurable.
B1. $\Phi_x < +\infty$ for all $x \in J$.
B2. For any $B \in \mathcal{B}_J$ with $\nu(B) > 0$ there exists $B^* \in \mathcal{C}_X$ such that $B^* \subseteq B$ and $\nu(B^*) > 0$.
B3. For any pair $x \neq x' \in J$ and $B \in \mathcal{B}_J$ there exists $n = n(x, x', B) \in \mathbb{N}$ such that

$$P_{x'}(X_1 \in B) > 0 \quad \Longrightarrow \quad P_x(X_{n+1} \in B) > 0.$$

Assumption (B0) will need to be verified on a case-by-case basis, but for example it will always hold if either J is countable or if the branching dynamics is monotone in its starting position, see [36, Proposition 7.3]. This will be always the case in all of our applications. Furthermore, (B1) is not really restrictive, as we will focus here only on the case in which there is convergence in L^2. On the other hand, we impose (B2) in order to obtain an appropriate control on the growth of $\xi_t(B)$, namely that for each $n \in \mathbb{N}$ and starting position $x \in J$

$$\lim_{t\to+\infty}\Big[\inf_{y\in J_n} \mathbb{E}_y\big(\xi_t(B)\big)\Big] = +\infty \quad \text{and} \quad \lim_{t\to+\infty} \xi_t(B) = +\infty \quad \text{on } \{D_\infty > 0\},$$

which follows from (B1)–(B2) by Theorem 2.6 (see Section 7 below). Furthermore, typically (B2) is very easy to check, see [36, Section 9] for details. Finally, (B3) amounts to a sort of irreducibility for the evolution of X when conditioned on remaining unabsorbed. This notion of irreducibility is different than that of ψ-irreducibility featured in [46] and weaker

[1] The fact that η is indeed measurable will follow from Proposition 7.1, whereas the measurability of σ is an additional assumption that we will have to make, see Assumptions 2.8.

than the standard definition of irreducibility when J is countable. Although not entirely standard, it is nevertheless the notion which appears naturally in our analysis and it is satisfied in all applications of interest, see [36].

Next, we introduce the notion of strong supercriticality which plays a key role in what follows.

Definition 2.9. We shall say that the branching dynamics $\xi = (\xi_t)_{t\geq 0}$ is *strongly supercritical* if:

i. ξ is supercritical, i.e. $P_x(\Theta) > 0$.
ii. $\eta(x) = P_x(\Gamma)$, where Γ is the event defined as

$$\Gamma := \left\{ \lim_{t\to+\infty} \left[\min_{u\in\overline{\xi}_t} \Phi_{u_t} \right] = +\infty \right\}$$

with the convention that $\min_{u\in\varnothing} \Phi_{u_t} := +\infty$, used whenever $|\xi_t| = 0$.

Note that, provided (i) holds, (ii) is equivalent to the condition $P_x(\Gamma|\Theta) = 0$.

One can check (see once again [36, Section 9]) that in all the examples of Section 3 the mapping $x \mapsto \Phi_x$ is bounded over subsets of J which are at a positive distance from $\partial_* \overline{J}$ and, on the other hand, that it tends to infinity as x approaches $\partial_* \overline{J}$. Thus, one can interpret strong supercriticality as the condition stating that on the event of non-extinction, we can always see at least one particle lying in the "bulk" of the state space J (away from $\partial_* J$) for infinitely many arbitrarily large times. On the other hand, we will see later in Section 7 that under Assumptions 2.8 strong supercriticality is equivalent to having

$$P_x\Big(\limsup_{t\to+\infty} \xi_t(B) > 0\Big) = P_x(\Theta) > 0 \tag{13}$$

for every $x \in J$ and all $B \in \mathcal{B}$ with $\nu(B) > 0$, which is the analogue in our context of the notion of strong local survival studied in [7] and other references therein. However, in general it will not be equivalent to the concept of (plain) local survival introduced in [22], which is said to take place whenever there exists a compact set $\mathcal{K} \subseteq J$ such that

$$P_x\Big(\limsup_{t\to+\infty} \xi_t(\mathcal{K}) > 0\Big) > 0. \tag{14}$$

See Section 3.5 for further details.

Our next result states that strong supercriticality is a necessary and sufficient condition for G to have exactly two fixed points whenever under Assumptions 2.8.

Theorem 2.10. *If Assumptions 2.1, 2.3 and 2.8 hold, the following statements are equivalent:*

i. *G has exactly two fixed points, η and $\mathbf{1}$.*
ii. *$\eta(x) = \sigma(x)$ for all $x \in J$.*
iii. *$\eta(x) = \sigma(x)$ for some $x \in J$.*
iv. *ξ is strongly supercritical for some starting position $x \in J$.*
v. *ξ is strongly supercritical for all starting positions $x \in J$.*

We note that strong supercriticality is not a trivial condition under our current assumptions, not even for the particular case of λ-positive systems to be considered in Section 2.4 below. Indeed, Section 3.5 shows an example of a λ-positive system which satisfies Assumptions 2.1, 2.3 and 2.8 but is not strongly supercritical. In particular, we have in this example that the random variable D_∞ is zero with positive probability on the event Θ of non-extinction. Nonetheless, whenever ξ is strongly supercritical this is not the case and so one obtains the following corollary combining both Theorems 2.5 and 2.10.

Corollary 2.11. *If Assumptions 2.1, 2.3 and 2.8 hold, and ξ is strongly supercritical then $D_\infty > 0$ almost surely on Θ. In particular, for every $B, B' \in \mathcal{C}_X$ with $\nu(B') > 0$ and starting position $x \in J$ we have, conditionally on Θ, that as $t \to +\infty$*

$$\nu_t(B, B') \xrightarrow{P} \frac{\nu(B)}{\nu(B')}.$$

Still, strong supercriticality appears to be a hard condition to check directly, at least in principle. In the extended version of this article [36], we introduce via examples some general methods to establish strong supercriticality which apply to a wide range of systems.

2.4. *The case of λ-positive systems*

Perhaps the simplest example of an underlying motion satisfying Assumptions 2.1 is that of a λ-positive process, which we formally introduce now.

Definition 2.12. Given $\lambda \in \mathbb{R}$ and a semigroup $S = (S_t)_{t\geq 0}$, we will say that S is λ-positive if there exist a nonnegative measurable function $h : \overline{J} \to [0,+\infty)$ satisfying $h|_J > 0$ and a (not necessarily finite) measure ν on $(\overline{J}, \mathcal{B})$, both unique up to constant multiples, such that:

- $S_t[h] = e^{-\lambda t} h$ for all t,
- For any nonnegative measurable $f : \overline{J} \to \mathbb{R}_{\geq 0}$ and all t,

$$\int_{\overline{J}} S_t[f](x)\, d\nu(x) = e^{-\lambda t} \int_{\overline{J}} f(x)\, d\nu(x),$$

- The function h and the measure ν are such that

$$\nu(h) := \int_{\overline{J}} h(x)\, d\nu(x) < +\infty.$$

In particular, we shall say that our branching dynamics ξ is λ-positive whenever its associated expectation semigroup $S^{(\xi)} = (S_t^{(\xi)})_{t\geq 0}$ is λ-positive according to the definition given above, where for each $t \geq 0$ and nonnegative measurable $f : \overline{J} \to \mathbb{R}_{\geq 0}$ we define

$$S_t^{(\xi)}[f](x) := \mathbb{E}_x\left(\sum_{u\in N_t} f(u_t)\right).$$

Remark 2.13. Using the many-to-one lemma (Lemma 4.1 below) it is straightforward to see that, in the current case of a constant branching rate $r > 0$ (and only in this case), ξ will be λ-positive if and only if the underlying motion X is $(r(m_1 - 1) - \lambda)$-positive. For this reason, in the following we shall focus only on λ-positivity of underlying motions rather than that of branching dynamics, as it makes no difference in our current setting.

Note that if S is a λ-positive Markov semigroup then its right-eigenfunction h satisfies $h \equiv 0$ on $\partial_* J$. With this, we may define a Markov semigroup $\tilde{S} = (\tilde{S}_t)_{t\geq 0}$ on the space J by the formula

$$\tilde{S}_t[f](x) = \frac{e^{\lambda t}}{h(x)} S_t[hf](x)$$

for any nonnegative $\mathcal{B}_J$-measurable $f : J \to \mathbb{R}_{\geq 0}$, $x \in J$ and $t \geq 0$, where we set $hf \equiv 0$ on $\partial_* J$. We will call $\tilde{S}$ the *h-transform* of S. Let us notice that, in the case S is the semigroup belonging to a λ-positive underlying motion X, $\tilde{S}$ is none other than the semigroup corresponding to X under the h-transform defined in (4). It is straightforward to see that, if one normalizes h and ν so that $\nu(h) = 1$, then the probability distribution μ on $(J, \mathcal{B}_J)$ defined through the formula

$$\frac{d\mu}{d\nu} = h$$

is invariant for $\tilde{S}$, i.e. for all nonnegative $\mathcal{B}_J$-measurable $f : J \to \mathbb{R}_{\geq 0}$ and all $t \geq 0$

$$\int_J \tilde{S}_t[f](x)\, d\mu(x) = \int_J f(x)\, d\mu(x).$$

Let us assume further that X is ergodic under the h-transform, with limiting distribution μ. Then, for any $B \in \mathcal{B}_J$ such that the function $\frac{\mathbb{1}_B}{h}$ is bounded and μ-almost surely continuous, we have

$$\lim_{t\to+\infty} \tilde{\mathbb{E}}_x\left(\frac{\mathbb{1}_B}{h}(X_t)\right) = \mu\left(\frac{\mathbb{1}_B}{h}\right) = \nu(B)$$

as $t \to +\infty$, so that by (6) the asymptotic formula (3) holds for any such B by taking $p(t) \equiv 1$. Hence, we see that such λ-positive motions fall naturally into the context of Assumptions 2.1. However, neither the uniform convergence of s_B over each J_n nor the square-integrability of M_t will follow immediately from the ergodicity of X under the h-transform, so that further conditions will need to be imposed on the process to guarantee them. This is not a disadvantage when compared to other approaches in the literature, as additional conditions are always imposed in order to obtain a law of large numbers, see [15,21]. Here, we propose an alternative condition to check the remainder of Assumptions 2.1 in the λ-positive setting and obtain Theorem 2.6, based on the existence of a Lypaunov functional for the process X under the measure $\tilde{P}$.

Definition 2.14. A $\mathcal{B}_J$-measurable $V : J \to \mathbb{R}_{\geq 0}$ is called a (*geometric*) *Lyapunov functional for* X (under the measure $\tilde{P}$) whenever it satisfies:

V1. There exists $t > 0$ such that for every $R > 0$ one can find $\alpha_R \in (0, 1)$ verifying

$$\left|\tilde{\mathbb{E}}_x\big(f(X_t)\big) - \tilde{\mathbb{E}}_y\big(f(X_t)\big)\right| \leq 2(1 - \alpha_R)\|f\|_\infty$$

for any bounded $\mathcal{B}_J$-measurable $f : J \to \mathbb{R}$ and all $x, y \in J$ such that $V(x) + V(y) \leq R$.

V2. There exist constants $\gamma, K > 0$ such that for all $t \geq 0$ and $x \in J$ one has

$$\tilde{\mathbb{E}}_x\big(V(X_t)\big) \leq e^{-\gamma t} V(x) + K.$$

Having a Lyapunov functional ensures that, under the h-transform $\tilde{P}$, the process X converges to equilibrium exponentially fast and, furthermore, that it does so uniformly over subsets of J where V is bounded, see Proposition 8.1 below. As a consequence, one has the following result relating the validity of Theorem 2.6 for λ-positive processes to the existence of a h-locally bounded Lyapunov functional for X with a large enough growth at infinity.

Proposition 2.15. *If X is λ-positive and admits a Lyapunov functional V such that*:

V3. *V is h-locally bounded, i.e.* $\sup_{x \in J_n} V(x) < +\infty$ *for each* $n \in \mathbb{N}$,

V4. $\|\frac{h}{1+V}\|_\infty < +\infty$,

then Assumptions 2.1 *are satisfied for all $B \in \mathcal{C}_X$, where $\mathcal{C}_X$ here is given by*

$$\mathcal{C}_X := \left\{ B \in \mathcal{B}_J : \left\| \frac{\mathbb{1}_B}{h} \right\|_\infty < +\infty \right\}. \tag{15}$$

Furthermore, there exists a constant $C = C(r(m_1 - 1), \lambda, V) > 0$ such that for all $x \in J$

$$\Phi_x \leq C \cdot \frac{1 + V(x)}{h(x)} < +\infty. \tag{16}$$

In particular, the convergences in (9) *and* (10) *hold for any $x \in J$ and $B, B' \in \mathcal{C}_X$ with $\nu(B') > 0$, and assumptions* (B1)–(B2) *are also satisfied.*

Finally, whenever X is λ-positive and admits such a Lyapunov functional, as a matter of fact one can show (9)–(10) in the almost sure sense provided that, on the other hand, V does not grow too fast at infinity. This is the content of our last result.

Theorem 2.16. *Suppose that ξ is a λ-positive process such that the eigenfunction h is continuous and X admits a Lyapunov functional V verifying* (V3)–(V4). *Then, for each starting position $x \in J$ such that*

$$\overline{\Phi}_x := \frac{(m_2 - m_1)r}{h(x)} \int_0^\infty \tilde{\mathbb{E}}_x\big(h(X_s)\big(1 + V(X_s)\big)\big)e^{-(r(m_1-1)-\lambda)s}\, ds < +\infty \tag{17}$$

there exists a full P-measure set $\Omega = \Omega^{(x)}$ satisfying that:

i. *For any $\omega \in \Omega$ one has*

$$\lim_{t \to +\infty} \frac{\xi_t(B)(\omega)}{\mathbb{E}_x(\xi_t(B'))} = \frac{\nu(B)}{\nu(B')} \cdot D_\infty(\omega) \tag{18}$$

for all pairs $B, B' \in \mathcal{C}_X$ satisfying $\nu(\partial B) = 0$ and $\nu(B') > 0$, where $\mathcal{C}_X$ is given by (15).

ii. *For any* $\omega \in \Omega \cap \Lambda$ *one has*

$$\lim_{t\to+\infty} \frac{\xi_t(B)(\omega)}{\xi_t(B')(\omega)} = \frac{\nu(B)}{\nu(B')} \tag{19}$$

for all pairs $B, B' \in \mathcal{C}_X$ *with* $\nu(\partial B) = \nu(\partial B') = 0$ *and* $\nu(B') > 0$.

Note that Theorem 2.16 can be combined with Theorem 2.10 in order to obtain (19) for any $\omega \in \Omega \cap \Theta$ whenever ξ is strongly supercritical.

3. Examples and applications

We now illustrate our results through a series of examples. We present first the case of generic ergodic motions, which fall in the category of 0-positive processes, and then proceed on to study four different models with λ-positive motions for $\lambda > 0$. Finally, we conclude in Section 3.6 with our results for the Branching Brownian Motion with a negative drift and absorption at the origin, which is a canonical example of a system with an underlying motion which is not λ-positive and constitutes our most significant and novel contribution. Further details as well as the verification that all of our required assumptions are met in each of the examples can be found in the extended version of the article, [36, Section 9].

3.1. *Ergodic motions*

Suppose that X is a motion without absorbing states, i.e. $\overline{J} = J$, which is ergodic and has stationary probability distribution ν on $(J, \mathcal{B}_J)$. In this case, it is easy to check that X is 0-positive and verifies (3) with $h \equiv 1$, $p \equiv 1$, $\lambda = 0$ for any $B \in \mathcal{B}_J$ such that

$$s_B(x,t) := P_x(X_t \in B) - \nu(B) \underset{t\to+\infty}{\longrightarrow} 0.$$

One can then show that Theorem 2.6 holds with $\mathcal{C}_X$ given by

$$\mathcal{C}_X := \left\{ B \in \mathcal{B}_J : \lim_{t\to+\infty} s_B(\cdot,t) = 0 \text{ uniformly over compact sets of } J \right\}. \tag{20}$$

In this case, for any starting position x, the random variable D_∞ satisfies

$$\mathbb{E}_x(D_\infty) = 1 \quad \text{and} \quad \mathbb{E}_x\left(D_\infty^2\right) =: \Phi_x = \frac{m_2 - m_1}{m_1 - 1}.$$

Since there is no absorption here and $\sup_{x\in J} \Phi_x < +\infty$, the corresponding branching dynamics is immediately strongly supercritical. Thus, if in addition X is irreducible in the sense of (B3) then Corollary 2.11 holds and $D_\infty > 0$ on Θ. Finally, if X admits a Lyapunov functional (as defined in Definition 2.14) which is bounded over compact subsets of J then the contents of Theorem 2.16 can also be shown to hold. See [36, Section 3.1] for details.

3.2. *Subcritical Galton–Watson process*

Let us now consider λ-positive motions with $\lambda > 0$. As a first example, let X be a continuous-time Galton–Watson process, i.e. a process on $\overline{J} := \mathbb{N}_0$ with transition rates q given for any $x \in \mathbb{N}_0$ and $y \in \mathbb{N}^* := \{-1\} \cup \mathbb{N}_0$ by

$$q(x, x+y) := x\rho(y),$$

where ρ is some probability vector in $\mathbb{N}^*$ representing the offspring distribution of each individual in the branching process (minus 1). We assume that X is subcritical, i.e. that $-\lambda := \sum_y y\rho(y) < 0$, so that X is almost surely absorbed at 0. If also $\sum_y y\log(y)\rho(y+1) < \infty$, then X is λ-positive (see [45,49]), with associated eigenfunction h given by $h(x) \propto x$ and that ν is a finite measure assigning positive mass to all $x \in \mathbb{N}$, although an explicit expression for ν is, in general, not known.

If ρ has a finite second moment then D is bounded in L^2 and thus the contents of Theorem 2.6 all hold in this case. Furthermore, using h as a Lyapunov functional for X, one can exploit (16) to verify that ξ is strongly supercritical. Thus, if we suppose that $\rho(-1) \in (0, 1)$, then X is irreducible (in the classical sense) and, as a consequence, we obtain Theorem 2.10 and Corollary 2.11. Finally, if ρ has a finite *third* moment then (17) is satisfied and thus the contents of Theorem 2.16 also hold.

3.3. *Subcritical contact process on* $\mathbb{Z}^d$ (*modulo translations*)

Let $\mathcal{P}_f(\mathbb{Z}^d)$ denote the class of all finite subsets of $\mathbb{Z}^d$ and $Y = (Y_t)_{t\geq 0}$ be the contact process on $\mathbb{Z}^d$, i.e. the Markov process on $\mathcal{P}_f(\mathbb{Z}^d)$ with transition rates q given, for any $\sigma \in \mathcal{P}_f(\mathbb{Z}^d)$ and $x \in \mathbb{Z}^d$, by

$$q\big(\sigma, \sigma \cup \{x\}\big) = \gamma \big|\big\{y \in \sigma : |y - x|_1 = 1\big\}\big| \quad \text{and} \quad q\big(\sigma, \sigma - \{x\}\big) = \mathbb{1}_\sigma(x),$$

where $\gamma > 0$ is a fixed constant called the *infection rate*. Observe that Y is translation invariant, so that there are no finite measures $\nu \neq 0$ verifying (3) for Y. This can be fixed if one considers the process modulo translations. Indeed, say that two non-empty sets $\sigma, \sigma' \in \mathcal{P}_f(\mathbb{Z}^d)$ are *equivalent* if they are translations of each other. Let J denote the quotient space obtained from this equivalence and, for any non-empty $\sigma \in \mathcal{P}_f(\mathbb{Z}^d)$, let $\langle\sigma\rangle$ denote its corresponding equivalence class in J. Also, set $\langle\varnothing\rangle := \varnothing$ and $\overline{J} := J \cup \{\varnothing\}$. Then, we define $X_t := \langle Y_t\rangle$. We call X the *contact process on* $\mathbb{Z}^d$ *modulo translations*.

It is well-known, see [8], that J is an irreducible class for the process X and that there exists $\gamma_c = \gamma_c(d) > 0$ such that the absorbing state $\varnothing$ is reached almost surely if and only if $\gamma \leq \gamma_c$. Moreover, it has been shown that for $\gamma < \gamma_c$ then λ, h and ν as in Assumptions 2.1 indeed exist and that the process is in fact λ-positive, see [1,26], although neither h nor ν are explicitly known. What is known, however, is that ν is finite and it assigns positive mass to every $x \in J$, see [26]. Using h again as a Lyapunov functional, one can show that for any subcritical infection rate $\gamma < \gamma_c$ the Malthusian martingale D is bounded in L^2 for all initial conditions ζ and that the branching dynamics ξ is strongly supercritical, so that the contents of all our main results always hold for this model.

3.4. *Recurrent Ornstein–Uhlenbeck process killed at* 0

Consider a 1-dimensional recurrent Ornstein–Ulhenbeck process which is killed at 0, i.e. the stopped process $X = (X_t)_{t\geq 0}$ on $\mathbb{R}_{\geq 0}$ defined as $X_t := Y_{t\wedge\tau_0}$, where $\tau_0 := \inf\{t \geq 0 : Y_t = 0\}$ and Y is given by the SDE

$$dY_t = -\lambda Y_t\, dt + dB_t,$$

with B a standard (1-dimensional) Brownian motion and $\lambda > 0$ a fixed parameter called the *drift*. It is well-known, see [42], that X is λ-positive with eigenfunction $h(x) := \sqrt{\frac{4\lambda}{\pi}}x$ (when h, ν are normalized so that ν is a probability measure and $\nu(h) = 1$) and eigenmeasure ν having density $f_X(x) := 2\lambda x e^{-\lambda x^2}\mathbb{1}_{(0,+\infty)}(x)$ with respect to the Lebesgue measure l on $\mathbb{R}$. As in the previous example, the Malthusian martingale D is bounded in L^2 and the branching dynamics ξ is strongly supercritical for all starting positions $x \in (0, +\infty)$, so that the contents of all our main results always hold for this model.

3.5. *Transient Ornstein–Uhlenbeck process*

This example was considered originally in [21]. Let X be the process given by the SDE

$$dX_t = \lambda X_t\, dt + dB_t,$$

with B a standard (1-dimensional) Brownian motion and $\lambda > 0$ a fixed parameter called the *drift*.

In this case, one can see that X is λ-positive with $h(x) \propto \exp\{-\frac{\lambda}{x}^2\}$ and ν equal to the Lebesgue measure on $\mathbb{R}$. Unlike previous examples, here ν is an infinite measure so that, in particular, the asymptotics in (3) does not hold for every $B \in \mathcal{B}_\mathbb{R}$ (for example, it does not hold if $B = \mathbb{R}$ because in this case (3) cannot decay exponentially) but, by the transience of X, it does hold for any B which is bounded. Still, this is enough to yield the following results for any starting position x:

i. D converges almost surely and in L^2 to some random variable $D_\infty \in L^2$.
ii. The contents of Theorem 2.6 hold with $\mathcal{C}_X$ the class of *bounded* Borel subsets of $\mathbb{R}$.
iii. ξ is *not* strongly supercritical. In fact, $0 < \sigma(x) < 1$ so that $\sigma \neq \eta, \mathbf{1}$.
iv. $\overline{\Phi}_x < +\infty$, so that the contents of Theorem 2.16 also hold in this case.

Parts (i) and (iv) above can also be found in [21] (together with the L^1 convergence of $W(B, B')$), whereas (ii) and (iii) are new results. On the other hand, we note that it follows from (iv) that for any compact set $\mathcal{K} \subseteq \mathbb{R}$ and $x \in \mathbb{R}$ we have

$$P_x\Big(\limsup_{t\to+\infty} \xi_t(\mathcal{K}) > 0\Big) \geq P_x(\Lambda) > 0$$

but, however, from (iii) that ξ is not strongly supercritical. This confirms our statement about the notion of local survival introduced in (14) being, in general, weaker than strong supercriticality.

3.6. *Brownian motion with drift killed at the origin*

Finally, we conclude with an example of an underlying motion which is not λ-positive. Consider a Brownian motion with negative drift $-c < 0$ killed at the origin, i.e. the stopped process $X = (X_t)_{t\geq 0}$ on $\mathbb{R}_{\geq 0}$ defined as $X_t := Y_{t\wedge\tau_0}$, where $\tau_0 := \inf\{t \geq 0 : Y_t = 0\}$ and Y is given by SDE

$$dY_t = -c\,dt + dB_t,$$

with B a standard (1-dimensional) Brownian motion. It was shown in [48] that the process X satisfies Assumptions 2.1 for $\lambda := \frac{c^2}{2}$ and $\mathcal{C}_X := \mathcal{B}_{(0,+\infty)}$, the class of Borel subsets of $(0, +\infty)$, and with $h(x) := \frac{1}{\sqrt{2\pi\lambda^2}} x e^{cx}$, $p(t) := t^{-\frac{3}{2}}$ and ν given by the density $f_X(x) := 2\lambda x e^{-cx}\mathbb{1}_{(0,+\infty)}(x)$ with respect to Lebesgue. However, X is not a λ-positive motion since one can easily verify that $\nu(h) = +\infty$.[2] Nonetheless, our approach still applies and we can thus obtain the following for any starting position $x > 0$:

i. The martingale D is bounded in L^2 if and only if $r(m_1 - 1) > c^2 = 2\lambda$ (note that this is strictly contained in the supercritical region $r(m_1 - 1) > \lambda$). In this case, the contents of Theorem 2.6 all hold.
ii. ξ is strongly supercritical. In particular, the contents of Theorem 2.10 and Corollary 2.11 all hold.

As underlined in the introduction, the almost sure convergence for $r(m_1 - 1) > 2\lambda$ can already be obtained by the methods developed by Watanabe in [51]. Moreover, the L^1 and almost sure convergences in the entire supercritical region $r(m_1 - 1) > \lambda$ were recently proved in [43], using some of our results.

4. The many-to-few lemmas

An element which will prove to be crucial in the proof of Theorem 2.6 is the ability to compute the first and second moments of the process $|\xi| = (|\xi_t|)_{t\geq 0}$ in exact form. We do this with the help of the many-to-few lemmas we state below. For simplicity, we will state only a reduced version of the many-to-one and many-to-two lemmas, which are all we need. For the many-to-few lemma in its full generality (and its proof) we refer to [32]. Note that equivalent formulas for the first and higher moments of additive functionals of the BMP have been known for a long time, e.g. [33–35].

First, we state the many-to-one lemma. It receives this name because it reduces expectations involving random sums over *many* particles, i.e. over all those in ξ_t, to expectations involving only one particle.

Lemma 4.1 (Many-to-one Lemma). *Given a nonnegative measurable function $f : (\overline{J}, \mathcal{B}) \to \mathbb{R}_{\geq 0}$, for every $t \geq 0$ and $x \in J$ we have*

$$\mathbb{E}_x\Bigg(\sum_{u\in\overline{X}_t} f(u_t)\Bigg) = e^{r(m_1-1)t}\mathbb{E}_x\big(f(X_t)\big).$$

Next, we state the many-to-two lemma, used to compute correlations between pairs of particles. Before we can do so, however, we must introduce the notion of 2-spine for our branching dynamics.

Definition 4.2. Consider the following coupled evolution on $\overline{J}$:

i. The dynamics starts with 2 particles, both located initially at some $x \in J$, whose positions evolve together randomly, i.e. describing the same random trajectory, according to $\mathcal{L}$.
ii. The particles wait for an independent random exponential time E of parameter $(m_2 - m_1)r$ and then split at their current position, each of them then evolving independently afterwards according to $\mathcal{L}$.

Now, for $i = 1, 2$, let $X^{(i)} = (X_t^{(i)})_{t\geq 0}$ be the process which indicates the position of the i-th particle. We call the pair $(X^{(1)}, X^{(2)})$ a 2-spine associated to the triple $(m, r, \mathcal{L})$ and E its splitting time.

The many-to-two lemma then goes as follows.

[2] Nor is it λ'-positive for any other $\lambda' \neq \frac{c^2}{2}$ because otherwise (3) would not hold for $\lambda = \frac{c^2}{2}$ and all $B \in \mathcal{B}_{(0,+\infty)}$.

Lemma 4.3 (Many-to-two Lemma). *Given any pair of measurable functions $f, g : (\overline{J}, \mathcal{B}) \to \mathbb{R}_{\geq 0}$, for every $t \geq 0$ and $x \in J$ we have*

$$\mathbb{E}_x\left(\sum_{u,v\in\overline{X}_t} f(u_t)g(v_t)\right) = e^{2r(m_1-1)t}\mathbb{E}_x\left(e^{[\mathrm{Var}(m)+(m_1-1)^2]r(E\wedge t)} f\left(X_t^{(1)}\right)g\left(X_t^{(2)}\right)\right),$$

where $(X^{(1)}, X^{(2)})$ is a 2-spine associated to $(m, r, \mathcal{L})$ and E denotes its splitting time.

5. Proof of Proposition 2.5

We first compute $\mathbb{E}_x(D_t^2)$ for every $t \geq 0$ and $x \in J$. Note that, by the many-to-two lemma and the definition of 2-spine, a straightforward computation (see the proof of Theorem 2.6 for details) yields that

$$\begin{aligned}\mathbb{E}_x\left(D_t^2\right) &= \frac{1}{h^2(x)}\mathbb{E}_x\left(\sum_{u,v\in\overline{\xi}_t} h(u_t)h(v_t)e^{-2(r(m_1-1)-\lambda)t}\right) \\ &= \frac{e^{2\lambda t}}{h^2(x)}\mathbb{E}_x\left(e^{[\mathrm{Var}(m)+(m_1-1)^2]r(E\wedge t)}h\left(X_t^{(1)}\right)h\left(X_t^{(2)}\right)\right) \\ &= [1]_t + [2]_t\end{aligned}$$

where

$$[1]_t := \frac{e^{2\lambda t}}{h^2(x)}\mathbb{E}_x\left(e^{[\mathrm{Var}(m)+(m_1-1)^2]r(E\wedge t)}h^2\left(X_t^{(1)}\right)\mathbb{1}_{\{E>t\}}\right) = \mathbb{E}_x\left(M_t^2\right)e^{-r(m_1-1)t}$$

and

$$[2]_t := (m_2 - m_1)r\int_0^t \mathbb{E}_x\left(M_s^2\mathbb{E}_{X_s}^2(M_{t-s})\right)e^{-r(m_1-1)s}\,ds.$$

Now, by (A1) we have that M is a mean-one martingale so that $\mathbb{E}_{X_s}(M_{t-s}) = 1$ for all $s \in [0, t]$. Thus, we obtain that

$$[2]_t = (m_2 - m_1)r\int_0^t \mathbb{E}_x\left(M_s^2\right)e^{-r(m_1-1)s}\,ds.$$

Recalling the definition of Φ_x, it is then clear that $[2]_t \to \Phi_x$ as $t \to +\infty$ and, on the other hand, that whenever $\Phi_x < +\infty$ we have that $\liminf_{t\to+\infty}[1]_t \to 0$, so that $\liminf_{t\to+\infty}\mathbb{E}_x(D_t^2) = \Phi_x$. But, since $\lim_{t\to+\infty}\mathbb{E}_x(D_t^2)$ always exists (although it can be $+\infty$, in principle) because $(D)^2$ is a submartingale, we conclude that

$$\lim_{t\to+\infty}\mathbb{E}_x\left(D_t^2\right) = \Phi_x.$$

Being D a martingale, this implies that it converges in L^2 if and only if $\Phi_x < +\infty$ and that, in this case, one has $\mathbb{E}_x(D_\infty^2) = \Phi_x$. Moreover, since $\mathbb{E}_x(D_t) = 1$ for all $t \geq 0$, it also follows that $\mathbb{E}_x(D_\infty) = 1$ and so this concludes the proof.

6. Proof of Theorem 2.6

This section contains the proof of Theorem 2.6. We will split the proof into two parts:

I. First, we will show that, given $B, B' \in \mathcal{C}_X$ with $\nu(B') > 0$, for any $x \in J$ one has

$$\lim_{t\to+\infty}\mathbb{E}_x\left(W_t^2\left(B, B'\right)\right) = \left[\frac{\nu(B)}{\nu(B')}\right]\Phi_x.$$

II. Then, we use (I) to conclude the convergence in (9) whenever $\Phi_x < +\infty$. In particular, the convergence $W_t(B', B') \xrightarrow{P} D_\infty$ together with (9) yields that for any $B \in \mathcal{C}_X$

$$\nu_t\left(B, B'\right) \xrightarrow{P} \frac{\nu(B)}{\nu(B')}$$

as $t \to +\infty$, conditionally on the event $\{D_\infty > 0\}$.

We dedicate a separate subsection to each parts, but begin first with a section devoted to proving two auxiliary lemmas to be used throughout the proof.

6.1. *Preliminary lemmas*

The first lemma we shall require is the following.

Lemma 6.1. *If assumption* (A1) *holds then for any $T > 0$ we have*

$$\lim_{n\to+\infty}\Big[\sup_{t\in[0,T]} \mathbb{E}_x\big(M_t^2 \mathbb{1}_{\{X_t\notin J_n\}}\big)\Big] = 0. \tag{21}$$

Proof. Notice that for any $t \in [0, T]$ we have the bound

$$\mathbb{E}_x\big(M_t^2 \mathbb{1}_{\{X_t\notin J_n\}}\big) \le \frac{e^{2\lambda T}}{h^2(x)} \cdot \frac{1}{n^2} + \mathbb{E}_x\Big(\Big(\sup_{s\in[0,T]} M_s^2\Big)\mathbb{1}_{\{\sup_{s\in[0,T]} h(X_s)>n\}}\Big)$$

so that it will suffice to show that

$$\lim_{n\to+\infty} \mathbb{E}_x\Big(\Big(\sup_{s\in[0,T]} M_s^2\Big)\mathbb{1}_{\{\sup_{s\in[0,T]} h(X_s)>n\}}\Big) = 0. \tag{22}$$

But since $\sup_{s\in[0,T]} M_s^2$ is P_x-integrable by Doob's inequality and (A1), and we also have that

$$\lim_{n\to+\infty} \mathbb{1}_{\{\sup_{s\in[0,T]} h(X_s)>n\}} = \mathbb{1}_{\{\sup_{s\in[0,T]} M_s^2=+\infty\}},$$

where the right-hand side is now P_x-almost surely null by the integrability of $\sup_{s\in[0,T]} M_s^2$, using the dominated convergence theorem we can conclude (22). □

The second lemma concerns the asymptotic behavior of the function p in (3).

Lemma 6.2. *The function p from assumption* (A2) *satisfies*:

i. *p has subexponential growth*, *i.e.* $\lim_{t\to+\infty} \frac{\log p(t)}{t} = 0$.
ii. *If we define the function $q(t_1,t_2) := \frac{p(t_2)}{p(t_1+t_2)}$ then for any $C > 0$ we have that*

$$\limsup_{t_2\to+\infty}\Big[\sup_{t_1\in[0,Ct_2]} q(t_1,t_2)\Big] =: q_C < +\infty \quad \textit{and} \quad \lim_{t_2\to+\infty}\Big[\sup_{t_1\in[0,C]} \big|q(t_1,t_2)-1\big|\Big] = 0.$$

This Lemma follows directly from: (a) Potter's bounds and (b) the uniform convergence theorem for regularly varying functions, see for instance [9].

6.2. *Part I*

Assume first that $\Phi_x < +\infty$ and let us show that then one has

$$\lim_{t\to+\infty} \mathbb{E}_x\big(W_t^2(B,B')\big) = \left[\frac{\nu(B)}{\nu(B')}\right]^2 \Phi_x. \tag{23}$$

To this end, take $t > 0$ and notice that

$$\mathbb{E}_x\big(W_t^2(B,B')\big) = \frac{\mathbb{E}_x(\xi_t^2(B))}{\mathbb{E}_x^2(\xi_t(B'))}. \tag{24}$$

Let us compute the expectations in the right-hand side of (24) by using the many-to-few lemmas. On the one hand, by the many-to-one lemma we have that

$$\mathbb{E}_x\big(\xi_t(B')\big) = \mathbb{E}_x\Big(\sum_{u\in\overline{\mathcal{X}}_t} \mathbb{1}_{\{u_t\in B'\}}\Big) = e^{r(m_1-1)t} P_x\big(X_t \in B'\big). \tag{25}$$

On the other hand, the many-to-two lemma yields

$$\mathbb{E}_x\big(\xi_t^2(B)\big) = \mathbb{E}_x\Bigg(\sum_{u,v\in\overline{\xi}_t} \mathbb{1}_{\{u_t\in B\}}\mathbb{1}_{\{v_t\in B\}}\Bigg)$$

$$= e^{2r(m_1-1)t}\mathbb{E}_x\big(\mathbb{1}_{\{X_t^{(1)}\in B\}}\mathbb{1}_{\{X_t^{(2)}\in B\}}e^{[\mathrm{Var}(m)+(m_1-1)^2]r(E\wedge t)}\big).$$

By separating in cases depending on whether $E > t$ or not, we obtain

$$\mathbb{E}_x\big(\xi_t^2(B)\big) = (1)_t + (2)_t,$$

where

$$(1)_t := e^{r(m_2-1)t}\mathbb{E}_x(\mathbb{1}_{\{X_t\in B\}}\mathbb{1}_{\{E>t\}})$$

and

$$(2)_t := e^{2r(m_1-1)t}\mathbb{E}_x\big(\mathbb{1}_{\{X_t^{(1)}\in B\}}\mathbb{1}_{\{X_t^{(2)}\in B\}}e^{[\mathrm{Var}(m)+(m_1-1)^2]rE}\mathbb{1}_{\{E\le t\}}\big).$$

Now, using the independence of E from the motion of the 2-spine, the Markov property yields

$$(1)_t = e^{r(m_1-1)t}P_x(X_t\in B)$$

and

$$(2)_t = (m_2-m_1)re^{2r(m_1-1)t}\int_0^t P_x\big(X_t^{(1),s}\in B, X_t^{(2),s}\in B\big)e^{-r(m_1-1)s}\,ds,$$

where $X^{(1),s}$ and $X^{(2),s}$ are two coupled copies of the Markov process X which coincide until time s and then evolve independently after s. If we condition on the position of these coupled processes at time s, then we obtain

$$(2)_t = (m_2-m_1)re^{2r(m_1-1)t}\int_0^t \mathbb{E}_x\big(P^2_{X_s}(X_{t-s}\in B)\big)e^{-r(m_1-1)s}\,ds. \tag{26}$$

Now, from (25) and (3) we conclude that

$$[1]_t := \frac{(1)_t}{\mathbb{E}_x^2(\xi_t(B'))} = \frac{P_x(X_t\in B)}{P_x(X_t\in B')}\cdot\frac{1}{\mathbb{E}_x(\xi_t(B'))} = \frac{\nu(B)+s_B(x,t)}{[\nu(B')+s_{B'}(x,t)]^2}\cdot\frac{1}{h(x)p(t)}e^{-(r(m_1-1)-\lambda)t}$$

which, by (i) in Lemma 6.2, shows that if t is taken sufficiently large then

$$\big|[1]_t\big| \le 2\frac{\nu(B)}{\nu(B')}\frac{e^{-\frac{1}{2}(r(m_1-1)-\lambda)t}}{h(x)}. \tag{27}$$

Similarly, one has that

$$[2]_t := \frac{(2)_t}{\mathbb{E}_x^2(\xi_t(B))} = \int_0^t \Psi_{x,t}(s)\,ds,$$

where

$$\Psi_{x,t}(s) := (m_2-m_1)r\frac{\mathbb{E}_x(P^2_{X_s}(X_{t-s}\in B))}{P_x^2(X_t\in B')}e^{-r(m_1-1)s}.$$

To treat the term $[2]_t$ we split the integral into three separate parts, i.e. for $\alpha\in(0,1)$ and $T>0$ to be specified later we write

$$[2]_t = [a]_t + [b]_t + [c]_t$$

where

$$[a]_t := \int_{\alpha t}^{t} \Psi_{x,t}(s)\,ds, \qquad [b]_t := \int_{T}^{\alpha t} \Psi_{x,t}(s)\,ds, \qquad [c]_t := \int_{0}^{T} \Psi_{x,t}(s)\,ds.$$

The first term $[a]_t$ deals with the case in which $s \to t$ and the asymptotics in (3) for $P_y(X_{t-s} \in B)$ may not hold. In this case, if α is taken close enough to 1 then $[a]_t$ tends to zero as $t \to +\infty$. Indeed, notice that

$$\mathbb{E}_x\big(P_{X_s}^2(X_{t-s} \in B)\big) \le \mathbb{E}_x\big(P_{X_s}(X_{t-s} \in B)\big) = P_x(X_t \in B)$$

by the Markov property, so that

$$[a]_t \le (m_2 - m_1) r \frac{P_x(X_t \in B)}{P_x^2(X_t \in B')} \int_{\alpha t}^{\infty} e^{-r(m_1-1)s}\,ds = \frac{m_2 - m_1}{m_1 - 1} e^{(1-\alpha)r(m_1-1)t}[1]_t.$$

By recalling (27), if α is chosen sufficiently close to 1 and t taken sufficiently large then

$$\big|[a]_t\big| \le e^{-\frac{1}{4}(r(m_1-1)-\lambda)t}. \tag{28}$$

Similarly to $[a]_t$, the term $[b]_t$ can also be made arbitrarily small as $t \to +\infty$ if T is large enough. Indeed, if $q_{\frac{\alpha}{1-\alpha}}$ is the quantity q_C from Lemma 6.2 for $C := \frac{\alpha}{1-\alpha}$, then by (3) we have that

$$\begin{aligned}\Psi_{x,t}(s) &= (m_2 - m_1) r \left[\frac{q(s,t-s)}{\nu(B') + s_{B'}(x,t)}\right]^2 \mathbb{E}_x\big(M_s^2\big(\nu(B) + s_B(X_s, t-s)\big)^2\big) e^{-r(m_1-1)s} \\ &\le (m_2 - m_1) 8 r q_{\frac{\alpha}{1-\alpha}}^2 \left(\frac{\nu(B) + \bar{s}_B}{\nu(B')}\right)^2 \mathbb{E}_x\big(M_s^2\big) e^{-r(m_1-1)s}\end{aligned} \tag{29}$$

if $s \le \alpha t$ and t is large enough so as to have

- $\sup_{y \in J, u \ge (1-\alpha)t} s_B(y,u) \le \bar{s}_B$,
- $q(s, t-s) \le 2 q_{\frac{\alpha}{1-\alpha}}$ (which can be done by (ii) in Lemma 6.2 since $\frac{s}{t-s} \le \frac{\alpha}{1-\alpha}$),
- $s_{B'}(x,t) \ge -\frac{\nu(B')}{2}$,

so that

$$\big|[b]_t\big| \le 8 q_{\frac{\alpha}{1-\alpha}}^2 \left(\frac{\nu(B) + \bar{s}_B}{\nu(B')}\right)^2 \cdot (m_2 - m_1) \int_T^{\infty} \mathbb{E}_x\big(M_s^2\big) r e^{-r(m_1-1)s}\,ds. \tag{30}$$

Since $\Phi_x < +\infty$, the right-hand side of (30) can be made arbitrarily small if T is chosen sufficiently large depending on α.

Finally, let us treat the last term $[c]_t$. By (A2) and (ii) in Lemma 6.2, for $s \le T$ we may write

$$\Psi_{x,t}(s) = \frac{(m_2 - m_1) r}{\nu^2(B')} \big(1 + o_t(1)\big) \mathbb{E}_x\big(M_s^2\big(\nu(B) + s_B(X_s, t-s)\big)^2\big) e^{-r(m_1-1)s}$$

where $o_t(1)$ (which depends on x, t, s and B') tends to zero uniformly in $s \le T$ as $t \to +\infty$. Thus, we may decompose

$$[c]_t = [c_1]_t + \big[c_1^*\big]_t$$

with

$$\big[c_1^*\big]_t = \frac{m_2 - m_1}{\nu^2(B')} \int_0^T \mathbb{E}_x\big(M_s^2\big(\nu(B) + s_B(X_s, t-s)\big)^2\big)) r e^{-r(m_1-1)s}\,ds$$

and

$$\big|[c_1]_t\big| \le \frac{(\nu(B) + \bar{s}_B)^2}{\nu^2(B')} \Phi_x \Big[\sup_{s \le T} o_t(1)\Big], \tag{31}$$

where the right-hand side of (31) tends to zero as $t \to +\infty$ since $\Phi_x < +\infty$. Finally, given $n \in \mathbb{N}$ we decompose $[c_1^*]_t$ by splitting the expectation inside into two depending on whether $X_s \in J_n$ or not. More precisely, we write

$$\left[c_1^*\right]_t = [c_2]_t + \left[c_2^*\right]_t$$

where

$$\left[c_2^*\right]_t = \frac{(m_2 - m_1)}{\nu^2(B')} \int_0^T \mathbb{E}_x\left(M_s^2\left(\nu(B) + s_B(X_s, t-s)\right)^2 \mathbb{1}_{\{X_s \in J_n\}}\right) r e^{-r(m_1-1)s}\, ds$$

and

$$\left|[c_2]_t\right| \le \frac{m_2 - m_1}{m_1 - 1} \frac{(\nu(B) + \overline{s}_B)^2}{\nu^2(B')} \sup_{s \in [0,T]} \mathbb{E}_x\left(M_s^2 \mathbb{1}_{\{X_s \notin J_n\}}\right). \tag{32}$$

Notice that the right-hand side of (32) is independent of t and tends to zero as n tends to infinity for any fixed $T > 0$ by Lemma 6.1. On the other hand, observe that by (A2) and Lemma 6.1

$$\mathbb{E}_x\left(M_s^2\left(\nu(B) + s_B(X_s, t-s)\right)^2\right)\mathbb{1}_{\{X_s \in J_n\}}) = \nu^2(B)\mathbb{E}_x\left(M_s^2\right)\left(1 + \overline{o}_t(1)\right)$$

where the term $\overline{o}_t$ (which depends on x, n, t, s and B) tends to zero uniformly in $s \le T$ as $t \to +\infty$ since $\sup_{s\in[0,T]} \mathbb{E}_x(M_t^2) \le 4\mathbb{E}_x(M_T^2)$ by Doob's inequality. By repeating the same argument that lead us to (31), we conclude that

$$\left[c_2^*\right]_{t,h} = [c_3]_t + [c_4]$$

where

$$[c_4] = \left[\frac{\nu(B)}{\nu(B')}\right]^2 \cdot (m_2 - m_1) \int_0^T \mathbb{E}_x\left(M_s^2\right) r e^{-r(m_1-1)s}\, ds$$

and $|[c_3]_t|$ tends to zero as $t \to +\infty$. Thus, we find that if we write $\Gamma := \{1, a, b, c_1, c_2, c_3\}$ then

$$\left| W_t\left(B, B'\right) - \left[\frac{\nu(B)}{\nu(B')}\right]^2 \Phi_x \right| \le \sum_{i \in \Gamma} \left|[i]_t\right| + \left[\frac{\nu(B)}{\nu(B')}\right]^2 \cdot (m_2 - m_1) \int_T^\infty \mathbb{E}_x\left(M_s^2\right) r e^{-r(m_1-1)s}\, ds. \tag{33}$$

By taking α adequately close to 1, T large enough (depending on α) and then n sufficiently large (depending on T), the right-hand side of (33) can be made arbitrarily small for all t large enough and so (23) follows.

Now, let us assume that $\Phi_x = +\infty$ and show that $\lim_{t\to+\infty} \mathbb{E}_x(W_t^2(B, B')) = +\infty$ in this case, proving (23). To see this, we notice that for any fixed $T > 0$ we have

$$\mathbb{E}_x\left(W_t^2\left(B, B'\right)\right) \ge [c]_{t,0} = \int_0^T \Psi_{x,t,0}(s)\, ds.$$

If n is chosen large enough so that $\sup_{s\in[0,T]} \mathbb{E}_x(M_s^2 \mathbb{1}_{\{X_s \notin J_n\}}) < 1$, then for all t large enough to guarantee that

- $\inf_{s\in[0,T]}\left(\frac{q(s,t-s)}{\nu(B')+s_{B'}(x,t)}\right)^2 \ge \frac{1}{2\nu^2(B')}$
- $\inf_{s\in[0,T], y\in J_n} (\nu(B) + s_B(y, t-s))^2 \ge \frac{\nu^2(B)}{2}$,

by (29) we obtain

$$\begin{aligned}
[c]_{t,0} &\ge \left[\frac{\nu(B)}{\nu(B')}\right]^2 \frac{(m_2 - m_1)}{4} \int_0^T \mathbb{E}_x\left(M_s^2 \mathbb{1}_{\{X_s \in J_n\}}\right) r e^{-r(m_1-1)s}\, ds \\
&\ge \left[\frac{\nu(B)}{\nu(B')}\right]^2 \frac{(m_2 - m_1)}{4} \int_0^T \mathbb{E}_x\left(M_s^2\right) r e^{-r(m_1-1)s}\, ds - \left[\frac{\nu(B)}{\nu(B')}\right]^2 \frac{(m_2 - m_1)}{4(m_1 - 1)}.
\end{aligned}$$

The right-hand side of this last inequality can be made arbitrarily large by taking T big enough, due to the fact that $\Phi_x = +\infty$. In particular, this implies that

$$\lim_{t\to+\infty} \mathbb{E}_x\left(W_t^2\left(B, B'\right)\right) = +\infty$$

and thus concludes the proof of Part I.

6.3. *Part II*

We now check that, whenever $\Phi_x < +\infty$, one has

$$W_t(B, B') \xrightarrow{L^2} \frac{\nu(B)}{\nu(B')} \cdot D_\infty,$$

for every $B, B' \in \mathcal{C}_X$ with $\nu(B') > 0$. Notice that by Proposition 2.5 it suffices to show that

$$\lim_{t \to +\infty} \left\| W_t(B, B') - \frac{\nu(B)}{\nu(B')} \cdot D_t \right\|_{L^2} = 0. \tag{34}$$

Now, observe that

$$\left\| W_t(B, B') - \frac{\nu(B)}{\nu(B')} \cdot D_t \right\|_{L^2}^2 = \mathbb{E}_x\left(W_t^2(B, B')\right) - 2\frac{\nu(B)}{\nu(B')}\mathbb{E}_x\left(W_t(B, B')D_t\right) + \left[\frac{\nu(B)}{\nu(B')}\right]^2 \mathbb{E}_x\left(D_t^2\right)$$

so that, by (23) and Proposition 2.5, (34) will follow if we show that

$$\lim_{t \to +\infty} \mathbb{E}_x\left(W_t(B, B')D_t\right) = \frac{\nu(B)}{\nu(B')}\Phi_x.$$

But this can be done by proceeding exactly as in Part I. We omit the details.

Finally, that

$$\nu_t(B, B') \xrightarrow{P} \frac{\nu(B)}{\nu(B')}$$

conditionally on the event $\{D_\infty > 0\}$ follows from the fact that

$$\frac{\mathbb{E}_x(\xi_t(B'))}{\xi_t(B')} = \frac{1}{W_t(B', B')} \xrightarrow{P} \frac{1}{D_\infty} \quad \text{and} \quad \frac{\xi_t(B)}{\mathbb{E}_x(\xi_t(B'))} = W_t(B, B') \xrightarrow{P} \frac{\nu(B)}{\nu(B')} \cdot D_\infty \tag{35}$$

conditionally on the event $\{D_\infty > 0\}$, which in turn follows from (9). This concludes Part II and thus the proof of Theorem 2.6.

7. Proof of Theorem 2.10

We also divide the proof of Theorem 2.10, now into four parts. First, we show that η and σ are indeed fixed points of G and, using (B3), that η and $\mathbf{1}$ cannot intersect other fixed points of G. Next, we show that (B1)–(B2) imply that our branching dynamics can be dominated from below by a supercritical Galton–Watson process. Using this domination, we then show that the notion of strong supercriticality can be reformulated in terms of certain fixed points of the operator G. Finally, we use this alternative formulation to show both implications of Theorem 2.10.

For improved clarity, throughout this section we shall use the superscript x, i.e. $\xi^{(x)}$ or $D^{(x)}$, to denote the initial condition of the underlying process whenever it becomes necessary.

7.1. *Part I*

We begin by checking first that both η and σ are fixed points of G.

Proposition 7.1. *The functions η and σ are fixed points of G.*

Proof. We begin by showing that η is a fixed point of G. First, observe that the measurability of η follows from the fact that $\eta = \lim_{n \to +\infty} G^{(n)}(\mathbf{0})$, where $G^{(n)}$ denotes the n-th composition of G with itself and $\mathbf{0}$ is the function constantly equal to 0. Now, to see that it is indeed a fixed point, observe that for any $t > 0$ and $x \in J$ we have the relation

$$\left|\xi_{1+t}^{(x)}\right| = \sum_{u \in \overline{\xi}_1^{(x)}} \left|\xi_t^{(u_1)}\right|,$$

which implies that, for any $t > 0$, $|\xi_{1+t}^{(x)}|$ equals zero if and only if $|\xi_t^{(u_1)}|$ is zero for every $u \in \overline{\xi}_1^{(x)}$. Thus, if we take $t \to +\infty$ then the former yields

$$\mathbb{1}_{(\Theta^{(x)})^c} = \prod_{u \in \overline{\xi}_1^{(x)}} \mathbb{1}_{(\Theta^{(u_1)})^c}. \tag{36}$$

By taking expectations $\mathbb{E}_x$ on the equality in (36), we obtain that $\eta(x) = G(\eta)(x)$. Furthermore, since this holds for any $x \in J$, we conclude that η is a fixed point of G.

Now, to see that σ is a fixed point of G (we already know that it is measurable due to (B0)), we observe the analogous relation

$$D_{1+t}^{(x)} = \frac{1}{h(x)} \sum_{u \in \overline{\xi}_1^{(x)}} h(u_1) e^{-(r(m_1-1)-\lambda)} D_t^{(u_1)}$$

which, upon taking the limit $t \to +\infty$, becomes

$$D_\infty^{(x)} = \frac{1}{h(x)} \sum_{u \in \overline{\xi}_1^{(x)}} h(u_1) e^{-(r(m_1-1)-\lambda)} D_\infty^{(u_1)}.$$

Since $h(y) = 0$ if and only if $y \in \partial_* \overline{J}$, that σ is a fixed point of G now follows as before. $\square$

Next, we use irreducibility to see that η and $\mathbf{1}$ cannot intersect other fixed points of G.

Proposition 7.2. *Assume that* (B3) *holds. Then, if g is a fixed point of G we have that*:

i. $\eta(x) \leq g(x) \leq 1$ *for all* $x \in J$.
ii. $g(x) = \eta(x)$ *for some* $x \in J \Longrightarrow g \equiv \eta$.
iii. $g(x) = 1$ *for some* $x \in J \Longrightarrow g \equiv \mathbf{1}$.

Proof. We show first that if g is a fixed point of G then $\eta \leq g \leq \mathbf{1}$. Indeed, the $g \leq \mathbf{1}$ inequality is immediate whereas the $\eta \leq g$ inequality follows from the fact that G is an increasing operator, i.e. $G(f_1) \leq G(f_2)$ if $f_1 \leq f_2$, together with the fact that $\eta = \lim_{n \to +\infty} G^{(n)}(\mathbf{0})$, where $G^{(n)}$ denotes the n-th composition of G with itself and $\mathbf{0}$ is the function constantly equal to 0.

Now, let us prove (ii). First, we observe that it is easy to check by induction that for any $n \in \mathbb{N}$

$$G^{(n)}(g)(x) = \mathbb{E}_x\left(\prod_{u \in \overline{\xi}_n} g(u_n)\right).$$

In particular, if $x \in J$ satisfies $P_x(X_n \in \{y \in J : \eta(y) < g(y)\}) > 0) > 0$ for some $n \in \mathbb{N}$ then, by considering only evolutions of ξ in which there is no branching until time n, it is clear that

$$P_x\left(\prod_{u \in \overline{\xi}_n} \eta(u_n) < \prod_{u \in \overline{\xi}_n} g(u_n)\right) > 0$$

so that

$$\eta(x) = G^{(n)}(\eta)(x) = \mathbb{E}_x\left(\prod_{u \in \overline{\xi}_n} \eta(u_n)\right) < \mathbb{E}_x\left(\prod_{u \in \overline{\xi}_n} g(u_n)\right) = G^{(n)}(h)(x) = g(x).$$

Therefore, if $\eta(x) = g(x)$ then we must have $P_x(X_n \in \{y \in J : \eta(y) < g(y)\}) = 0$ for every $n \in \mathbb{N}$. By irreducibility (assumption (B3)) we then obtain that $P_{x'}(X_1 \in \{y \in J : \eta(y) < g(y)\}) = 0$ and, as a consequence, that $P_{x'}(\xi_1(\{y \in J : \eta(y) < g(y)\}) > 0) = 0$ holds for every $x' \in J$ since the particle positions $(u_1)_{u \in \overline{\xi}_1^{(x')}}$ are all distributed as $X_1^{(x')}$. Since η and h are fixed points of G, this implies that $g(x') \leq \eta(x')$ for every $x' \in J$ which, together with (i) shown above, allows us to conclude that $\eta \equiv g$. The proof of (iii) is analogous. $\square$

7.2. *Part II*

The following step is to show that, under (B1)–(B2), one has a suitable lower bound on the growth of our dynamics. To this end, for each $n \in \mathbb{N}$ we define the set

$$\tilde{J}_n := \{x \in J : \Phi_x \leq n\} \tag{37}$$

and then write $\hat{J}_n := J_n \cap \tilde{J}_n$. Notice that the sequence $(\hat{J}_n)_{n\in\mathbb{N}}$ is increasing and, furthermore, that $\bigcup_{n\in\mathbb{N}} \hat{J}_n = J$ by (B1). Now, the precise meaning of lower bound on the growth of our dynamics is formulated in the following definition.

Definition 7.3. We say that the *lower bound condition* holds, and denote it in the sequel by (LB), if for any $n \in \mathbb{N}$ and $B \in \mathcal{B}_J$ with $\nu(B) > 0$ there exists a time $T_{n,B}$ and a random variable $L_{n,B}$ satisfying $\mathbb{E}(L_{n,B}) > 1$ such that for all $x \in \hat{J}_n$ and every $t > T_{n,B}$ one has

$$L_{n,B} \preceq \xi_t(B)$$

where $\preceq$ denotes stochastic domination, i.e. for any bounded measurable and increasing $f : \mathbb{R} \to \mathbb{R}$ one has that

$$\mathbb{E}\big(f(L_{n,B})\big) \leq \inf_{x\in\hat{J}_n} \mathbb{E}_x\big(f\big(\xi_t(B)\big)\big).$$

Remark 7.4. Note that, by (B2) and Lemma 4.1 below, for any $B \in \mathcal{B}_J$ with $\nu(B) > 0$ and $x \in J_n$ we have that

$$\begin{aligned}
\mathbb{E}_x\big(\xi_t(B)\big) &\geq \mathbb{E}_x\big(\xi_t\big(B^*\big)\big) = h(x)p(t)e^{(r(m_1-1)-\lambda)t}\big(\nu\big(B^*\big) + s_{B^*}(x,t)\big) \\
&\geq \frac{1}{n}p(t)e^{(r(m_1-1)-\lambda)t}\Big(\nu\big(B^*\big) + \inf_{y\in J_n} s_{B^*}(y,t)\Big)
\end{aligned} \tag{38}$$

so that, by Lemma 6.2 and (A2), for all t large enough depending on B and n we have that

$$\inf_{x\in\hat{J}_n} \mathbb{E}_x\big(\xi_t(B)\big) > 1. \tag{39}$$

Thus, condition (LB) is simply a stronger form of (39), one in which we ask the entire distributions of the random variables $(\xi_t(B))_{x\in\hat{J}_n}$ to be uniformly supercritical rather than just their means.

The lower bound (LB) will be the main tool in the proof of Lemma 7.8 in Part III below, which is crucial for proving Theorem 2.10. Our next result states that (LB) holds under (B2).

Proposition 7.5. *Assumption* (B2) *implies condition* (LB).

Proof. Let us fix $n \in \mathbb{N}$ and notice that, by the Cauchy–Schwarz inequality, we have for any $x \in J_n$, $B \in \mathcal{B}_J$ and $K, T \in \mathbb{N}$ that

$$\mathbb{E}_x^2\big(\xi_T(B)\mathbb{1}_{\xi_T(B)\geq K}\big) \leq \mathbb{E}_x\big(\xi_T^2(B)\big)P_x\big(\xi_T(B) \geq K\big). \tag{40}$$

On the other hand, if $\nu(B) > 0$ then it follows from (B2), (38) and Assumptions 2.1 that

$$\lim_{T\to+\infty}\Big[\inf_{y\in J_n} \mathbb{E}_y\big(\xi_T(B)\big)\Big] = +\infty.$$

Therefore, by (40) we conclude that if T is sufficiently large (depending only on K, n and B) then for all $x \in J_n$ we have

$$P_x\big(\xi_T(B) \geq K\big) \geq \frac{[\mathbb{E}_x(\xi_T(B)) - K]^2}{\mathbb{E}_x(\xi_T^2(B))} \geq \frac{1}{2} \cdot \frac{\mathbb{E}_x^2(\xi_T(B))}{\mathbb{E}_x(\xi_T^2(B))}.$$

Now, a careful inspection of the proof of Theorem 2.6 shows that there exists a constant $C_n > 0$ and a time $T_n > 0$ such that

$$\frac{\mathbb{E}_x(\xi_T^2(B))}{\mathbb{E}_x^2(\xi_T(B))} \leq C_n\Phi_x + 1$$

for all $x \in J_n$ and $T > T_n$. We stress that C_n and T_n do *not* depend on $x \in J_n$, only on n and B. Therefore, since $\sup_{y \in \tilde{J}_n} \Phi_y < +\infty$, we may take $K \in \mathbb{N}$ sufficiently large and $T \in \mathbb{N}$ accordingly so that

$$\inf_{x \in \hat{J}_n} P_x\big(\xi_T(B) \geq K\big) \geq \frac{1}{K-1}.$$

It follows that $L_{n,B} \preceq \xi_T(B)$ for any such T and all $x \in \hat{J}_n$, where $L_{n,B}$ has distribution given by

$$P(L_{n,B} = K) = \frac{1}{K-1} = 1 - P(L_{n,B} = 0).$$

Since in this case $\mathbb{E}(L_{n,B}) = \frac{K}{K-1} > 1$, this concludes the proof. □

7.3. *Part III*

We continue by using Proposition 7.5 to show that strong supercriticality can be reformulated in terms of certain fixed points of G. More precisely, we have the following result.

Proposition 7.6. *If Assumptions 2.8 are satisfied then $\xi^{(x)}$ is strongly supercritical if and only if the following two conditions hold*:

i. $\xi^{(x)}$ *is supercritical*, *i.e.* $P_x(\Theta) > 0$.
ii′. *There exists $n \in \mathbb{N}$ such that*

$$P_x(\Theta) = P_x\Big(\limsup_{k \to +\infty} \xi_k(\tilde{J}_n) > 0\Big)$$

where $\tilde{J}_n$ is given by (37).

Remark 7.7. Observe that for $B \in \mathcal{B}$ the function g_B defined as

$$g_B(x) := P_x\Big(\limsup_{n \to +\infty} \xi_n(B) = 0\Big)$$

is a fixed point of G. Indeed, the proof of this statement is analogous to that of Proposition 7.1. Thus, Proposition 7.6 states that $\xi^{(x_0)}$ is strongly supercritical if and only if for some $n \in \mathbb{N}$

$$g_{\tilde{J}_n}(x_0) = \eta(x_0) < 1.$$

But then by Proposition 7.2 we conclude that the same statement must hold for all $x \in J$, so that $\xi^{(x)}$ is strongly supercritical for some $x \in J$ if and only if it is strongly supercritical for all $x \in J$.

We now prove Proposition 7.6. Let us notice that it suffices to show that (ii) in Definition 2.9 is equivalent to (ii′) in the statement above. To see the (ii′) $\Longrightarrow$ (ii) implication, notice the inclusions

$$\Theta^c \subseteq \Gamma = \bigcap_{n \in \mathbb{N}} \Big\{\lim_{t \to +\infty} \xi_t(\tilde{J}_n) = 0\Big\} \subseteq \bigcap_{n \in \mathbb{N}} \Big\{\lim_{k \to +\infty} \xi_k(\tilde{J}_n) = 0\Big\}. \tag{41}$$

Now, for each $n \in \mathbb{N}$ let us write $A_n := \{\lim_{k \to +\infty} \xi_k(\tilde{J}_n) = 0\}$. Notice that, since the sequence $(\tilde{J}_n)_{n \in \mathbb{N}}$ is increasing, we have that $(A_n)_{n \in \mathbb{N}}$ is decreasing and therefore that

$$P_x\bigg(\bigcap_{n \in \mathbb{N}} \Big\{\lim_{k \to +\infty} \xi_k(\tilde{J}_n) = 0\Big\}\bigg) = \lim_{n \to +\infty} P_x(A_n). \tag{42}$$

Therefore, if (ii′) holds then it follows that $P_x(\Theta^c) = P_x(A_n)$ for some n and, by (41) and (42), we conclude that (ii) holds. On the other hand, if (ii) holds then by (41) we have

$$P_x\big(\Theta^c\big) = \lim_{n \to +\infty} P_x\Big(\lim_{t \to +\infty} \xi_t(\tilde{J}_n) = 0\Big). \tag{43}$$

Thus, if we show that for all $n \in \mathbb{N}$ sufficiently large so that $x \in \hat{J}_n$ and $\nu(\tilde{J}_n) > 0$ we have

$$\lim_{k\to+\infty} \xi_k^{(x)}(\tilde{J}_n) = 0 \quad \Longrightarrow \quad \lim_{t\to+\infty} \xi_t^{(x)}(\tilde{J}_{n+1}) = 0 \tag{44}$$

then, from (43) and the inclusion $(\Theta^{(x)})^c \subseteq \{\lim_{k\to+\infty} \xi_k^{(x)}(\tilde{J}_n) = 0\}$, by iterating (44) we conclude that for n sufficiently large

$$P_x\big(\Theta^c\big) \le P_x\Big(\lim_{k\to+\infty} \xi_k(\tilde{J}_n) = 0\Big) \le \lim_{m\to+\infty} P_x\Big(\lim_{t\to+\infty} \xi_t(\tilde{J}_m) = 0\Big) = P_x\big(\Theta^c\big),$$

which immediately gives (ii′). Now, (44) follows at once from the next lemma.

Lemma 7.8. *For any $m \in \mathbb{N}$ and $B \in \mathcal{B}_J$ such that $\hat{J}_m \neq \varnothing$ and $\nu(B) > 0$ there exists $T \in \mathbb{N}$ satisfying that for any $x \in \hat{J}_m$ one has*

$$\limsup_{t\to+\infty} \xi_t^{(x)}(\tilde{J}_m) > 0 \quad \Longrightarrow \quad \lim_{k\to+\infty} \xi_{kT}^{(x)}(B) = +\infty.$$

Proof. The idea is to couple the sequence $(\xi_{kT}^{(x)}(B))_{k\in\mathbb{N}}$ together with an i.i.d. sequence $(Z^{(n)})_{n\in\mathbb{N}}$ of supercritical single-type Galton–Watson branching processes such that if at least one $Z^{(n)}$ survives on the event $\{\limsup_{t\to+\infty} \xi_t^{(x)}(\tilde{J}_m) > 0\}$ then $\xi_{kT}(B)$ tends to infinity as $k \to +\infty$.

We proceed as follows. First we notice that, since condition (LB) holds by Proposition 7.5, there exists a random variable $L_{m,B}$ with $\mathbb{E}(L_{m,B}) > 1$ and a time $T \in \mathbb{N}$ such that for all $y \in \hat{J}_m$ and $t \ge T$ we have

$$L_{m,B} \preceq \xi_t^{(y)}(B). \tag{45}$$

Next, given a fixed $x \in \hat{J}_m$, we define the process $V^{(1)} := (V_j^{(1)})_{j\in\mathbb{N}}$ by the formula

$$V_j^{(1)} := \xi_{jT}^{(x)}(B)$$

and observe that for each $j \in \mathbb{N}$ we have

$$V_{j+1}^{(1)} \ge \sum_{u\in\overline{\xi}_{jT}^{(x)}(B)} \xi_T^{(u)}(B),$$

where $\overline{\xi}_{jT}^{(x)}(B)$ denotes the subcollection of particles of $\overline{\xi}_{jT}^{(x)}$ which are located inside the subset B and, for $u \in \overline{\xi}_t^{(x)}$, $\xi^{(u)}$ is the sub-dynamics of $\xi^{(x)}$ associated to the particle u starting at time t. Since for each $u \in \overline{\xi}_{jT}^{(x)}(B)$ we have that $L_{B,m} \preceq \xi_T^{(u)}(B)$ by (45), it follows that (by enlarging the current probability space if necessary) one can couple $V^{(1)}$ with a Galton–Watson process $Z^{(1)} := (Z_j^{(1)})_{j\in\mathbb{N}}$ with offspring distribution given by $L_{m,B}$, in such a way that $V_j^{(1)} \ge Z_j^{(1)}$ holds for all $j \in \mathbb{N}$. Therefore, if $Z^{(1)}$ survives then $Z_j^{(1)}$ must tend to infinity as $j \to +\infty$ and, consequently, so must $\xi_{jT}^{(x)}(B)$. In this case, we decouple $\xi^{(x)}$ from the remaining $Z^{(n)}$ (for $n \ge 2$) by taking these to be independent from $\xi^{(x)}$. If $Z^{(1)}$ dies out, however, we proceed as follows:

i. Define $\tau^{(1)} := \inf\{j \in \mathbb{N} : Z_j^{(1)} = 0\}$. Notice that $Z^{(1)}$ dies out if and only if $\tau^{(1)} < +\infty$.
ii. If $\xi_{sT}^{(x)}(\tilde{J}_m) = 0$ for all $s \ge \tau^{(1)}$, then decouple the $(Z^{(n)})_{n\ge2}$ from ξ as before.
iii. If $\xi_{sT}^{(x)}(\tilde{J}_m) > 0$ for some (random) $s \ge \tau^{(1)}$, then choose some $y \in \overline{\xi}_{sT}(\tilde{J}_m)$ at random and define the process $V^{(2)} = (V_j^{(2)})_{j\in\mathbb{N}}$ according to the formula

$$V_j^{(2)} := \xi_{(\lceil s\rceil - s + j)T}^{(y)}(B),$$

where $\lceil s\rceil$ here denotes the smallest integer greater than or equal to s. Let us observe that, by construction, there exists a (random) $k \in \mathbb{N}$ such that $V_j^{(2)} \le \xi_{(k+j)T}^{(x)}(B)$ for all $j \in \mathbb{N}$. By a similar argument than the one carried out for $V^{(1)}$, it is possible to couple $V^{(2)}$ with a Galton–Watson process $Z^{(2)}$ which is independent of $Z^{(1)}$ but has the same distribution, in such a way that $V_j^{(2)} \ge Z_j^{(2)}$ for all $j \in \mathbb{N}$. If $Z^{(2)}$ survives then, by the considerations above, $\xi_{jT}^{(x)}$ must tend to infinity as $j \to +\infty$. If not, then one can repeat this procedure to obtain a branching process $Z^{(3)}$ and so on.

Since every $Z^{(n)}$ has the same *positive* probability of survival, it follows that at least one of them will survive on the event $\{\limsup_{t\to+\infty} \xi_t^{(x)}(\tilde{J}_m) > 0\}$, and so the result now follows. □

7.4. *Part IV*

We now conclude by showing all implications in the statement of Theorem 2.10.

First, let us observe that the condition $\Phi_x < +\infty$ implies that $\sigma(x) < 1$. Indeed, if $\Phi_x < +\infty$ then by Proposition 2.5 we have $\mathbb{E}_x(D_\infty) = 1$ so that $\sigma(x) = P_x(D_\infty = 0) < 1$ necessarily holds. Thus, if (B1) holds then $\sigma \neq \mathbf{1}$ and therefore (i) must imply (ii).

That (ii) $\Longrightarrow$ (iii) is obvious, so we move on to (iii) $\Longrightarrow$ (iv). Take $x \in J$ such that $\eta(x) = \sigma(x)$. Note that, by the argument given above, $\sigma(x) < 1$ so that if $\eta(x) = \sigma(x)$ then $\xi^{(x)}$ is supercritical. It remains to verify (ii′) of Proposition 7.6. But (B2) together with Theorem 2.6 and (38) imply that for any $B \in \mathcal{B}_J$ with $\nu(B) > 0$

$$\limsup_{n\to+\infty} P_x\big(\xi_n(B) = 0|\Lambda\big) \le \limsup_{n\to+\infty} P_x\big(\xi_n\big(B^*\big) = 0|\Lambda\big) = 0,$$

from which a straightforward calculation yields that

$$P_x\Big(\Big\{\limsup_{n\to+\infty} \xi_n(B) > 0\Big\} \cap \Lambda\Big) = P_x(\Lambda).$$

Therefore, since we also have the inequalities

$$P_x\Big(\Big\{\limsup_{n\to+\infty} \xi_n(B) > 0\Big\} \cap \Lambda\Big) \le P_x\Big(\limsup_{n\to+\infty} \xi_n(B) > 0\Big) \le P_x(\Theta),$$

if $\eta(x) = \sigma(x)$ then we have $P_x(\Theta) = P_x(\Lambda)$ and so (ii′) follows. This shows that (iii) $\Longrightarrow$ (iv).

The implication (iv) $\Longrightarrow$ (v) is also obvious, so it remains to show (v) $\Longrightarrow$ (i). For this purpose, we note that, if $g \neq \mathbf{1}$ is a fixed point of G, Proposition 7.2 yields that $g(y) < 1$ for *all* $y \in J$. Thus, by (B2) one can find $B \in \mathcal{C}_X$ with $\nu(B) > 0$ and $\varepsilon > 0$ such that $\sup_{y\in B} g(y) < 1 - \varepsilon$. Now, since g is a fixed point of G, we have

$$g(x) = \lim_{n\to+\infty} G^{(n)}(g)(x) = \lim_{n\to+\infty} \mathbb{E}_x\Big(\prod_{u\in\overline{\xi}_n} g(u_n)\Big) \le \liminf_{n\to+\infty} \mathbb{E}_x\big((1-\varepsilon)^{\xi_n(B)}\big)$$

where for the last inequality we have used the fact that $g \le 1$. Moreover, let us observe that

$$\mathbb{E}_x\big((1-\varepsilon)^{\xi_n(B)}\big) \le P_x\big(|\xi_n| = 0\big) + P_x\big(\Theta^c \cap \big\{|\xi_n| > 0\big\}\big) + \mathbb{E}_x\big((1-\varepsilon)^{\xi_n(B)} \mathbb{1}_\Theta\big), \tag{46}$$

where, since $\{|\xi_n^{(x)}| > 0\} \searrow \Theta^{(x)}$ as $n \to +\infty$, we have that

$$\lim_{n\to+\infty} P_x\big(|\xi_n| = 0\big) = \eta(x) \quad \text{and} \quad \lim_{n\to+\infty} P_x\big(\Theta^c \cap \big\{|\xi_n| > 0\big\}\big) = P_x\big(\Theta^c \cap \Theta\big) = 0.$$

Hence, if we could show that

$$\liminf_{n\to+\infty} \mathbb{E}_x\big((1-\varepsilon)^{\xi_n(B)} \mathbb{1}_\Theta\big) = 0, \tag{47}$$

then we would immediately obtain that $g(x) \le \eta(x)$. Together with the obvious reverse inequality, this would yield $\eta(x) = g(x)$ and hence, by Proposition 7.2, that $\eta \equiv g$. Since we have that $\eta \neq \mathbf{1}$ by the strong supercriticality of $\xi^{(x)}$, (v) $\Longrightarrow$ (i) would follow at once. Thus, let us show (47). Observe that (47) immediately follows if we can show that for any $K > 0$

$$\liminf_{n\to+\infty} P_x\big(\xi_n(B) \le K|\Theta\big) = 0. \tag{48}$$

Since $\xi^{(x)}$ is strongly supercritical, by (ii′) of Proposition 7.6 we have that (48) is then equivalent to

$$\liminf_{n\to+\infty} P_x\Big(\xi_n(B) \le K \,\Big|\, \limsup_{j\to+\infty} \xi_j(\tilde{J}_k) > 0\Big) = 0 \tag{49}$$

if $k \in \mathbb{N}$ is taken sufficiently large. But (49) is now a straightforward consequence of Lemma 7.8, so that (47) follows.

8. Proofs of Proposition 2.15 and Theorem 2.16

We now prove our results for λ-positive systems, namely Proposition 2.15 and Theorem 2.16.

8.1. *Proof of Proposition* 2.15

Proposition 2.15 is essentially a consequence of the following.

Proposition 8.1. *If X is λ-positive and admits a Lyapunov functional V as in Definition 2.14, then $\mu(g) < +\infty$ for any $\mathcal{B}_J$-measurable $g : J \to \mathbb{R}$ such that $\|\frac{g}{1+V}\|_\infty < +\infty$ and, furthermore, there exist constants $C, \gamma > 0$ such that for all $x \in J$, $t \geq 0$ and any g as above one has*

$$\left|\tilde{\mathbb{E}}_x\big(g(X_t)\big) - \mu(g)\right| \leq C\big(1 + V(x)\big)e^{-\gamma t}\left\|\frac{g - \mu(g)}{1 + V}\right\|_\infty.$$

Proof. This result is a careful combination of [28, Theorem 3.6] and [46, Theorem 4.3, Theorem 6.1], using the fact that since

$$\mu(g) \leq \left\|\frac{g}{1 + V}\right\|_\infty \mu(1 + V) = \left\|\frac{g}{1 + V}\right\|_\infty \big(1 + \mu(V)\big)$$

it suffices to check that $\mu(V) < +\infty$ to see that $\mu(g) < +\infty$ for any function g as in the statement. We omit the details. □

Observe that if X admits a Lyapunov functional V satisfying (V4) then, by Proposition 8.1, if we set

$$\mathcal{C}_X := \left\{B \in \mathcal{B}_J : \left\|\frac{\mathbb{1}_B}{h}\right\|_\infty < +\infty\right\}$$

then for any $B \in \mathcal{C}_X$ and $x \in J$ we have that

$$\left|s_B(x,t)\right| = \left|\tilde{\mathbb{E}}_x\left(\frac{\mathbb{1}_B}{h}(X_t)\right) - \mu\left(\frac{\mathbb{1}_B}{h}\right)\right| \leq 2C\big(1 + V(x)\big)e^{-\gamma t}\left\|\frac{\mathbb{1}_B}{h}\right\|_\infty$$

so that if (V3) holds then (A2-iii) is automatically satisfied. Furthermore, we always have that

$$s_B(x,t) \leq \tilde{\mathbb{E}}_x\left(\frac{\mathbb{1}_B}{h}(X_t)\right) \leq \left\|\frac{\mathbb{1}_B}{h}\right\|_\infty < +\infty$$

for any such B, so that (A2-iv) is also satisfied. Since we have already seen that (3) and (A2-i–ii) hold whenever X is λ-positive, this shows that (A2) is satisfied. On the other hand, if (V4) holds then again by Proposition 8.1 we have that for any $x \in J$ and $t \geq 0$

$$\mathbb{E}_x\big(M_t^2\big) = \frac{e^{\lambda t}}{h(x)}\tilde{\mathbb{E}}_x\big(h(X_t)\big) \leq \frac{e^{\lambda t}}{h(x)}\left(\mu(h) + 2C\big(1 + V(x)\big)\left(\left\|\frac{h}{1 + V}\right\|_\infty + \mu(h)\right)\right) < +\infty \tag{50}$$

so that (A1) is also satisfied. Finally, since from (50) we can obtain in fact the bound

$$\mathbb{E}_x\big(M_t^2\big) \leq C_{h,\mu}\frac{e^{\lambda t}}{h(x)}\big(1 + V(x)\big)$$

for some constant $C_{h,\mu} > 0$, a straightforward computation shows that

$$\Phi_x \leq \frac{m_2 - m_1}{r(m_1 - 1) - \lambda}C_{h,\mu} \cdot \frac{1 + V(x)}{h(x)}.$$

In particular, this yields that (B1) immediately holds. Furthermore, since for any given $B \in \mathcal{B}_J$ we have that $B \cap J_n \in \mathcal{C}_X$ for all n, we obtain also (B2) by taking $B^* = B \cap J_n$ for n large enough so as to guarantee that $\nu(B \cap J_n) \geq \frac{\nu(B)}{2}$. This concludes the proof of Proposition 2.15.

8.2. *Proof of Theorem* 2.16

For the proof we shall essentially follow the approach in [2,16,21], introducing a few modifications to their methods to adapt to our different hypotheses.

First, given any bounded and $\mathcal{B}_J$-measurable function $f : J \to \mathbb{R}$ and any starting point $x \in J$, let us define the process $U^f = (U_t^f)_{t\geq 0}$ by the formula

$$U_t^f := \frac{1}{h(x)} \sum_{u \in \overline{\xi}_t} h(u_t) f(u_t) e^{-(r(m_1-1)-\lambda)t}. \tag{51}$$

Notice that if we take $f \equiv 1$ then we recover our Malthusian martingale, i.e $U_t^f = D_t$ for all t. Our goal is to show that, for any starting point $x \in J$, there exists a full P-measure set $\Omega = \Omega^{(x)}$ such that for any $\omega \in \Omega$

$$\lim_{t\to+\infty} U_t^f(\omega) = \nu(hf) \cdot D_\infty(\omega), \tag{52}$$

for all functions of the form $f = \frac{\mathbb{1}_B}{h}$ with $B \in \mathcal{C}_X$ such that $\nu(\partial B) = 0$. By Lemma 4.1, this will automatically imply (18) since $p(t) \equiv 1$ in the λ-positive case. The first step in the proof of (52) is to obtain the desired convergence along lattice times, i.e. along time sequences $(t_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}_{>0}$ of the form $t_n = n\delta$ for $\delta > 0$ arbitrarily small. This will be a consequence of the next lemma. For simplicity, in the sequel we suppress the dependence of U^f on f from the notation unless it becomes strictly necessary.

Lemma 8.2. *For any $\delta > 0$ and starting position $x \in J$, we have almost surely*

$$\lim_{n\to+\infty} U_{2n\delta} - \mathbb{E}_x(U_{2n\delta}|\mathcal{F}_{n\delta}) = 0,$$

where $(\mathcal{F}_t)_{t\geq 0}$ is the filtration generated by ξ.

Proof. It will suffice to show that

$$\sum_{n\in\mathbb{N}} \mathbb{E}_x\big(\big(U_{2n\delta} - \mathbb{E}_x(U_{2n\delta}|\mathcal{F}_{n\delta})\big)^2\big) < +\infty. \tag{53}$$

To this end, we observe that, by proceeding as in the proof of [21, Lemma 18] for the case $p = 2$, one has for any $t \geq 0$

$$\begin{aligned}
\mathbb{E}_x\big(\big(U_{2t} - \mathbb{E}_x(U_{2t}|\mathcal{F}_t)\big)^2\big|\mathcal{F}_t\big) &= \left[\frac{e^{-(r(m_1-1)-\lambda)t}}{h(x)}\right]^2 \sum_{u\in\overline{\xi}_t} h^2(u_t)\mathbb{E}_{u_t}\big(\big(U_t - \mathbb{E}_u(U_t)\big)^2\big) \\
&\leq \left[\frac{e^{-(r(m_1-1)-\lambda)t}}{h(x)}\right]^2 \sum_{u\in\overline{\xi}_t} h^2(u_t)\mathbb{E}_{u_t}\big(U_t^2\big) \\
&\leq \left[\frac{\|f\|_\infty e^{-(r(m_1-1)-\lambda)t}}{h(x)}\right]^2 \sum_{u\in\overline{\xi}_t} h^2(u_t)\mathbb{E}_{u_t}\big(D_t^2\big)
\end{aligned}$$

so that by Lemma 4.1 we obtain

$$\mathbb{E}_x\big(\big(U_{2t} - \mathbb{E}_x(U_{2t}|\mathcal{F}_t)\big)^2\big) \leq \left[\frac{\|f\|_\infty e^{-(r(m_1-1)-\lambda)t}}{h(x)}\right]^2 e^{r(m_1-1)t}\mathbb{E}_x\big(h^2(X_t)\mathbb{E}_{X_t}\big(D_t^2\big)\big).$$

Now, by the proof of Theorem 2.5 we now that

$$\mathbb{E}_{X_t}\big(D_t^2\big) = \mathbb{E}_{X_t}\big(M_t^2\big)e^{-r(m_1-1)t} + (m_2 - m_1)r\int_0^t \mathbb{E}_{X_t}\big(M_s^2\big)e^{-r(m_1-1)s}\,ds, \tag{54}$$

so that, using the Markov property for X, we conclude that

$$\mathbb{E}_x\big(h^2(X_t)\mathbb{E}_{X_t}\big(D_t^2\big)\big) = \mathbb{E}_x\big(h^2(X_{2t})\big)e^{-(r(m_1-1)-2\lambda)t} + (m_2 - m_1)r\int_0^t \mathbb{E}_x\big(h^2(X_{t+s})\big)e^{-(r(m_1-1)-2\lambda)s}\,ds. \tag{55}$$

Upon observing that the second term in the right-hand side of (55) can be rewritten as

$$e^{(r(m_1-1)-2\lambda)t}\left[(m_2-m_1)r\int_t^{2t}\mathbb{E}_x\big(h^2(X_s)\big)e^{-(r(m_1-1)-2\lambda)s}\,ds\right],$$

a straightforward calculation now yields that $\mathbb{E}_x((U_{2t}-\mathbb{E}_x(U_{2t}|\mathcal{F}_t))^2)$ can be bounded from above by

$$\frac{\|f\|_\infty^2}{h(x)}\left(\tilde{\mathbb{E}}_x\big(h(X_{2t})\big)e^{-(r(m_1-1)-\lambda)2t}+(m_2-m_1)r\int_t^{2t}\tilde{\mathbb{E}}_x\big(h(X_s)\big)e^{-(r(m_1-1)-\lambda)s}\,ds\right). \tag{56}$$

Now, since $\tilde{\mathbb{E}}_x(h(X_s))\le\|\frac{h}{1+V}\|_\infty\tilde{\mathbb{E}}_x(1+V(X_s))\le\|\frac{h}{1+V}\|_\infty(1+V(x)+K)$ for any s by (V2)–(V4), we see that

$$\mathbb{E}_x\big(\big(U_{2t}-\mathbb{E}_x(U_{2t}|\mathcal{F}_t)\big)^2\big)\le\frac{\|f\|_\infty^2}{h(x)}\cdot\left\|\frac{h}{1+V}\right\|_\infty\big(1+V(x)+K\big)\cdot\left(1+\frac{(m_2-m_1)r}{r(m_1-1)-\lambda}\right)\cdot e^{-(r(m_1-1)-\lambda)t}$$

which is enough to imply (53). □

Now, the convergence along lattice times will follow if we can show that, as n tends to infinity, for any $\delta>0$ and starting position $x\in J$ one has almost surely

$$\lim_{n\to+\infty}\big(\mathbb{E}_x(U_{2n\delta}|\mathcal{F}_{n\delta})-\nu(hf)\cdot D_\infty\big)=0$$

or, equivalently, that

$$\lim_{n\to+\infty}\big(\mathbb{E}_x(U_{2n\delta}|\mathcal{F}_{n\delta})-\nu(hf)\cdot D_{n\delta}\big)=0.$$

By Lemma 4.1, the former reduces to showing that P-almost surely

$$\Sigma_{n\delta}:=\frac{1}{h(x)}\sum_{u\in\bar{\xi}_{n\delta}}e^{-(r(m_1-1)-\lambda)n\delta}h(u_{n\delta})\big[\tilde{\mathbb{E}}_{u_{n\delta}}\big(f(X_{n\delta})\big)-\nu(hf)\big]\longrightarrow 0, \tag{57}$$

which in turn will follow if we show that

$$\sum_{n\in\mathbb{N}}\mathbb{E}_x\big(\Sigma_{n\delta}^2\big)<+\infty. \tag{58}$$

But notice that, by proceeding as in the computation leading to (54), we have

$$\mathbb{E}_x\big(\Sigma_{n\delta}^2\big)=\langle 1\rangle_{n\delta}+\langle 2\rangle_{n\delta}$$

where

$$\langle 1\rangle_{n\delta}:=\mathbb{E}_x\big(M_{n\delta}^2\cdot s_{hf}^2(X_{n\delta},n\delta)\big)e^{-r(m_1-1)n\delta}$$

and

$$\langle 2\rangle_{n\delta}:=(m_2-m_1)\int_0^{n\delta}\mathbb{E}_x\big(M_s^2\cdot\mathbb{E}_{X_s}^2\big(M_{n\delta-s}\cdot s_{hf}(X_{n\delta-s},n\delta)\big)\big)re^{-r(m_1-1)s}\,ds$$

where, for $y\in J$ and $s\ge 0$, we write

$$s_{hf}(y,s):=\tilde{\mathbb{E}}_y\big(f(X_s)\big)-\nu(hf).$$

Now, since $\|\frac{f}{1+V}\|_\infty\le\|f\|_\infty<+\infty$ by assumption on f, Proposition 8.1 yields the existence of some constants $C,\alpha>0$ such that for all $y\in J$ and $s\ge 0$

$$\big|s_{hf}(y,s)\big|\le C\big(1+V(y)\big)e^{-\alpha s}\|f\|_\infty.$$

On the other hand, since f is bounded we always have that $\|s_{hf}\|_\infty \le 2\|f\|_\infty$. Hence, we obtain

$$\begin{aligned}\langle 1\rangle_{n\delta} &\le \frac{4\|f\|_\infty^2}{h(x)}\tilde{\mathbb{E}}_x\big(h(X_{n\delta})\big)e^{-(r(m_1-1)-\lambda)n\delta}\\ &\le \frac{4\|f\|_\infty^2}{h(x)}\left\|\frac{h}{1+V}\right\|_\infty \tilde{\mathbb{E}}_x\big(1+V(X_{n\delta})\big)e^{-(r(m_1-1)-\lambda)n\delta}\\ &\le \frac{4\|f\|_\infty^2}{h(x)}\left\|\frac{h}{1+V}\right\|_\infty \big(1+V(x)+K\big)e^{-(r(m_1-1)-\lambda)n\delta}\end{aligned}$$

so that $\sum_{n\in\mathbb{N}}\langle 1\rangle_{n\delta} < +\infty$ and, similarly, using (V2) we obtain

$$\begin{aligned}\mathbb{E}_x\big(M_s^2\cdot\mathbb{E}_{X_s}^2\big(M_{n\delta-s}\cdot s_{hf}(X_{n\delta-s},n\delta)\big)\big) &\le \frac{2\|f\|_\infty}{h(x)}\tilde{\mathbb{E}}_x\big(h(X_s)\tilde{\mathbb{E}}_{X_s}\big(s_{hf}(X_{n\delta-s},n\delta)\big)e^{\lambda s}\\ &\le \frac{2C\|f\|_\infty^2}{h(x)}\tilde{\mathbb{E}}_x\big(h(X_s)\tilde{\mathbb{E}}_{X_s}\big(1+V(X_{n\delta-s})\big)\big)e^{-\alpha n\delta}e^{\lambda s}\\ &\le \frac{2C\|f\|_\infty^2}{h(x)}\tilde{\mathbb{E}}_x\big(h(X_s)\big(1+K+e^{-\gamma(n\delta-s)}V(X_s)\big)\big)e^{-\alpha n\delta}e^{\lambda s}\\ &\le \frac{2C(1+K)\|f\|_\infty^2}{h(x)}\tilde{\mathbb{E}}_x\big(h(X_s)\big(1+V(X_s)\big)\big)e^{-\alpha n\delta}e^{\lambda s},\end{aligned}$$

so that

$$\langle 2\rangle_{n\delta} \le 2C(1+K)\|f\|_\infty^2\overline{\Phi}_x e^{-\alpha n\delta}$$

and thus $\sum_{n\in\mathbb{N}}\langle 2\rangle_{n\delta} < +\infty$ since $\overline{\Phi}_x < +\infty$. Hence, we conclude that (58) holds and therefore that (57) as well, from which the convergence along lattice times now immediately follows.

To establish the full limit as $t \to +\infty$, suppose first that f in (51) above is the indicator function of some open set $B \in \mathcal{B}_J$. It follows from (51) that for any $\delta > 0$ one has the bound

$$U_t \ge \frac{e^{-(r(m_1-1)-\lambda)\delta}}{h(x)(1+\varepsilon)}\sum_{u\in\overline{\xi}_{n\delta}} e^{-(r(m_1-1)-\lambda)n\delta}h(u_{n\delta})\Xi^{(u)}_{B,\delta,\varepsilon} \tag{59}$$

for $t \in [n\delta,(n+1)\delta)$, where $\Xi^{(u)}_{B,\delta,\varepsilon}$ denotes the indicator function of the event $A^{(u),1}_{B,\delta,\varepsilon}\cap A^{(u),2}_{B,\delta,\varepsilon}$, where

$$A^{(u),1}_{B,\delta,\varepsilon} := \left\{v_t \in B \text{ and } h(v_t) > \frac{1}{1+\varepsilon}h(u_{n\delta}) \text{ for all } v\in\overline{\xi}^{(u)}_t \text{ and } t\in[n\delta,(n+1)\delta)\right\}$$

and

$$A^{(u),2}_{B,\delta,\varepsilon} := \big\{\text{The particle } u\in\overline{\xi}_{n\delta} \text{ does not branch in } [n\delta,(n+1)\delta)\big\}.$$

(Recall that $\xi^{(u)}$ is the sub-dynamics of ξ associated to the particle u.) Now, $\Xi^{(u)}_{B,\delta,\varepsilon}$ is not a function of $u_{n\delta}$ alone as it also depends on the evolution of $\xi^{(u)}$ on the time interval $[n\delta,(n+1)\delta)$, so that our previous analysis for lattice times does not directly apply. However, since $\Xi^{(u)}_{B,\delta,\varepsilon}$ is still bounded, one can adapt the previous analysis to our current setting to show that the expression

$$\sum_{u\in\overline{\xi}_{n\delta}} e^{-(r(m_1-1)-\lambda)n\delta}h(u_{n\delta})\Xi^{(u)}_{B,\delta,\varepsilon} - \mathbb{E}_x\left(\sum_{u\in\overline{\xi}_{n\delta}} e^{-(r(m_1-1)-\lambda)n\delta}h(u_{n\delta})\Xi^{(u)}_{B,\delta,\varepsilon}\middle|\mathcal{F}_{n\delta}\right)$$

tends to zero almost surely as $n\to+\infty$. Now, it is easy to see that

$$\mathbb{E}_x\left(\sum_{u\in\overline{\xi}_{n\delta}} e^{-(r(m_1-1)-\lambda)n\delta}h(u_{n\delta})\Xi^{(u)}_{B,\delta,\varepsilon}\middle|\mathcal{F}_{n\delta}\right) = \sum_{u\in\overline{\xi}_{n\delta}} e^{-(r(m_1-1)-\lambda)n\delta}h(u_{n\delta})\zeta_{B,\delta,\varepsilon}(u_{n\delta})$$

where for $y \in J$ we define

$$\zeta_{B,\delta,\varepsilon}(y) := e^{-r(m_1-1)\delta} P_y\left(X_t \in B \text{ and } h(X_t) > \frac{1}{1+\varepsilon} h(y) \text{ for all } t \in [0,\delta) \right).$$

But by our previous analysis conducted for lattice times we know that almost surely

$$\frac{1}{h(x)} \sum_{u \in \overline{\xi}_{n\delta}} e^{-(r(m_1-1)-\lambda)n\delta} h(u_{n\delta}) \zeta_{B,\delta,\varepsilon}(u_{n\delta}) \longrightarrow \nu(h\zeta_{B,\delta,\varepsilon}) \cdot D_\infty$$

so that by (59) we conclude that almost surely

$$\liminf_{t\to+\infty} U_t \geq \frac{e^{-(r(m_1-1)-\lambda)\delta}}{1+\varepsilon} \nu(h\zeta_{B,\delta,\varepsilon}) \cdot D_\infty. \tag{60}$$

Since B is an open set, h is continuous and X has cadlag trajectories, we have that for any $y \in B$

$$\lim_{\delta\to 0} P_y\left(X_t \in B \text{ and } h(X_t) > \frac{1}{1+\varepsilon} h(y) \text{ for all } t \in [0,\delta) \right) = 1$$

so that, by first letting $\delta \to 0$ and afterwards taking $\varepsilon \to 0$ in (60), we have that almost surely

$$\liminf_{t\to+\infty} U_t \geq \nu(h\mathbb{1}_B) \cdot D_\infty. \tag{61}$$

Now, since J is locally compact and separable, there exists a countable basis $\mathcal{G} := (G_k)_{k\in\mathbb{N}}$ for its topology which is closed under finite unions. In particular, there exists a full P-measure set $\tilde{\Omega}$ such that for any $\omega \in \tilde{\Omega}$ the inequality in (61) holds simultaneously for every $G_k \in \mathcal{G}$. Hence, since any open set $G \subseteq J$ can be written as the union of an increasing sequence of open sets $(G_{n_k})_{k\in\mathbb{N}} \subseteq \mathcal{G}$, we conclude that for any $\omega \in \tilde{\Omega}$ and $k \in \mathbb{N}$

$$\liminf_{t\to+\infty} U_t^{\mathbb{1}_G}(\omega) \geq \liminf_{t\to+\infty} U_t^{\mathbb{1}_{G_{n_k}}}(\omega) \geq \nu(h\mathbb{1}_{G_{n_k}}) \cdot D_\infty(\omega).$$

By taking $k \to +\infty$, we obtain that for any $\omega \in \tilde{\Omega}$

$$\liminf_{t\to+\infty} U_t^{\mathbb{1}_G}(\omega) \geq \nu(h\mathbb{1}_G) \cdot D_\infty(\omega) \tag{62}$$

holds simultaneously for all open sets $G \subseteq J$. Hence, if for $\omega \in \tilde{\Omega} \cap \{D_\infty > 0\} \cap \{D_t \to D_\infty\}$ and each $t \geq 0$ we define the probability measures $\mu_t(\omega)$ and μ on $(J, \mathcal{B}_J)$ by the formulas

$$\mu_t(\omega)(B) := \frac{U_t^{\mathbb{1}_B}(\omega)}{D_t(\omega)} \quad \text{and} \quad \mu(B) := \nu(h\mathbb{1}_B)$$

for each $B \in \mathcal{B}_J$ (observe that μ_t is well-defined because $D_t > 0$ on the event $\{D_\infty > 0\}$) then, by Portmanteau's theorem [37, Theorem 4.25] and (62), $\mu_t(\omega)$ converges weakly to μ. In particular, since h is strictly positive and continuous on J, if $B \in \mathcal{C}_X$ is such that $\nu(\partial B) = 0$ then $\frac{\mathbb{1}_B}{h} : J \to \mathbb{R}$ is a bounded and $\mathcal{B}_J$-measurable function whose set of discontinuities has μ-null measure, so that by the weak convergence we obtain

$$\lim_{t\to+\infty} \mu_t(\omega)\left(\frac{\mathbb{1}_B}{h}\right) = \mu\left(\frac{\mathbb{1}_B}{h}\right) = \nu(B).$$

Since $D_t(\omega) \to D_\infty(\omega)$ for $\omega \in \tilde{\Omega} \cap \{D_\infty > 0\} \cap \{D_t \to D_\infty\}$, the former implies that

$$\lim_{t\to+\infty} U^{\frac{\mathbb{1}_B}{h}}(\omega) = \nu(B) \cdot D_\infty(\omega).$$

On the other hand, if $\omega \in \tilde{\Omega} \cap \{D_\infty = 0\} \cap \{D_t \to D_\infty\}$ then clearly as $t \to +\infty$ we have

$$\left|U_t^{\frac{\mathbb{1}_B}{h}}(\omega)\right| \leq \left\| \frac{\mathbb{1}_B}{h} \right\|_\infty D_t(\omega) \longrightarrow 0 = \nu(B) \cdot D_\infty(\omega).$$

Thus, since $\Omega := \tilde{\Omega} \cap \{D_t \to D_\infty\}$ is a full P-measure set by Theorem 2.5, we obtain (18). Finally, (19) follows from (18) as in the proof of (10). This concludes the proof of Theorem 2.16.

Acknowledgements

We are very grateful to an anonymous referee for pointing out useful references for Lemma 6.2 and underlining the seminal work of Shinzo Watanabe.

M.J. would like to warmly thank Elie Aïdekon, Julien Berestycki, Simon Harris, Pablo Groisman and Pascal Maillard for many fruitful discussions on branching dynamics.

S.S. would like to also thank Maria Eulália Vares for useful suggestions on how to improve this manuscript and Bastien Mallein for very helpful discussions on these topics.

The work of M.J. was supported in part at the University of Buenos Aires by the PICT fellowship, and by the Bizkaia talent excellence grant at BCAM center. The work of S.S. was supported in part at the Technion by a fellowship from the Lady Davis Foundation and also by the Israeli Science Foundation grants no. 1723/14 and 765/18.

References

[1] E. Andjel, F. Ezanno, P. Groisman and L. T. Rolla. Subcritical contact process seen from the edge: Convergence to quasi-equilibrium. *Electron. J. Probab.* **20** (32) (2015) 1–16. MR3335823 https://doi.org/10.1214/EJP.v20-3881

[2] S. Asmussen and H. Hering. Strong limit theorems for general supercritical branching processes with applications to branching diffusions. *Z. Wahrsch. Verw. Gebiete* **36** (3) (1976) 195–212. MR0420889 https://doi.org/10.1007/BF00532545

[3] A. Asselah, P. Ferrari and P. Groisman. Quasistationary distributions and Fleming–Viot processes in finite spaces. *J. Appl. Probab.* **48** (2) (2011) 322–332. MR2840302 https://doi.org/10.1239/jap/1308662630

[4] A. Asselah, P. A. Ferrari, P. Groisman and M. Jonckheere. Fleming–Viot selects the minimal quasi-stationary distribution: The Galton–Watson case. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2) (2016) 647–668. MR3498004 https://doi.org/10.1214/14-AIHP635

[5] K. Athreya. Limit theorems for multitype continuous time Markov branching processes. II. The case of an arbitrary linear functional. *Z. Wahrsch. Verw. Gebiete* **13** (1969) 204–214. MR0254928 https://doi.org/10.1007/BF00539201

[6] J. Bérard and J. Gouéré. Brunet–Derrida behavior of branching-selection particle systems on the line. *Comm. Math. Phys.* **298** (2) (2010) 323–342. MR2669438 https://doi.org/10.1007/s00220-010-1067-y

[7] D. Bertacchi and F. Zucca. Strong local survival of branching random walks is not monotone. *Adv. in Appl. Probab.* **46** (2) (2014) 400–421. MR3215539 https://doi.org/10.1239/aap/1401369700

[8] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *Ann. Probab.* **18** (4) (1990) 1462–1482. MR1071804

[9] N. H. Bingham, C. M. Goldie and J. L. Teugels. *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge University Press, Cambridge, 1989. MR1015093

[10] E. Brunet and B. Derrida. Shift in the velocity of a front due to a cutoff. *Phys. Rev. E (3)* **56** (3, part A) (1997) 2597–2604. MR1473413 https://doi.org/10.1103/PhysRevE.56.2597

[11] É. Brunet and B. Derrida. Effect of microscopic noise on front propagation. *J. Stat. Phys.* **103** (1–2) (2001) 269–282. MR1828730 https://doi.org/10.1023/A:1004875804376

[12] K. Burdzy, R. Holyst and P. March. A Fleming–Viot particle representation of the Dirichlet Laplacian. *Comm. Math. Phys.* **214** (3) (2000) 679–703. MR1800866 https://doi.org/10.1007/s002200000294

[13] Z.-Q. Chen, Y.-X. Ren, R. Song and R. Zhang. Strong law of large numbers for supercritical superprocesses under second moment condition. *Front. Math. China* **10** (4) (2015) 807–838. MR3352888 https://doi.org/10.1007/s11464-015-0482-y

[14] Z.-Q. Chen, Y.-X. Ren and Y.-X. Wang. An almost sure scaling limit theorem for Dawson–Watanabe superprocesses. *J. Funct. Anal.* **254** (7) (2008) 1988–2019. MR2397881 https://doi.org/10.1016/j.jfa.2007.12.003

[15] 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) (2017) 898–931. MR3687244 https://doi.org/10.1007/s10959-016-0671-y

[16] Z.-Q. Chen and Y. Shiozawa. Limit theorems for branching Markov processes. *J. Funct. Anal.* **250** (2) (2007) 374–399. MR2352485 https://doi.org/10.1016/j.jfa.2007.05.011

[17] R. Durrett and D. Remenik. Brunet–Derrida particle systems, free boundary problems and Wiener–Hopf equations. *Ann. Probab.* **39** (6) (2011) 2043–2078. MR2932664 https://doi.org/10.1214/10-AOP601

[18] J. Engländer. Branching diffusions, superdiffusions and random media. *Probab. Surv.* **4** (2007) 303–364. MR2368953 https://doi.org/10.1214/07-PS120

[19] J. Engländer. Law of large numbers for superdiffusions: The non-ergodic case. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (1) (2009) 1–6. MR2500226 https://doi.org/10.1214/07-AIHP156

[20] J. Engländer. *Spatial Branching in Random Environments and with Interaction*. *Advanced Series on Statistical Science & Applied Probability* **20**. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015. MR3362353 https://doi.org/10.1142/8991

[21] 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) (2010) 279–298. MR2641779 https://doi.org/10.1214/09-AIHP203

[22] J. Engländer and A. E. Kyprianou. Local extinction versus local exponential growth for spatial branching processes. *Ann. Probab.* **32** (1A) (2004) 78–99. MR2040776 https://doi.org/10.1214/aop/1078415829

[23] J. Engländer and D. Turaev. A scaling limit theorem for a class of superdiffusions. *Ann. Probab.* **30** (2) (2002) 683–722. MR1905855 https://doi.org/10.1214/aop/1023481006

[24] J. Engländer and A. Winter. Law of large numbers for a class of superdiffusions. *Ann. Inst. Henri Poincaré B, Probab. Stat.* **42** (2006) 171–185. MR2199796 https://doi.org/10.1016/j.anihpb.2005.03.004

[25] P. Ferrari and N. Marić. Quasi stationary distributions and Fleming–Viot processes in countable spaces. *Electron. J. Probab.* **12** (24) (2007) 684–702 (electronic). MR2318407 https://doi.org/10.1214/EJP.v12-415

[26] P. A. Ferrari, H. Kesten and S. Martínez. R-positivity, quasi-stationary distributions and ratio limit theorems for a class of probabilistic automata. *Ann. Appl. Probab.* **6** (2) (1996) 577–616. MR1398060 https://doi.org/10.1214/aoap/1034968146

[27] P. Groisman and M. Jonckheere. Simulation of quasi-stationary distributions on countable spaces. *Markov Process. Related Fields* **19** (3) (2013) 521–542. MR3156964

[28] M. Hairer. Convergence of Markov processes. Minicourse notes, 2016.

[29] J. W. Harris, S. C. Harris and A. E. Kyprianou. Further probabilistic analysis of the Fisher–Kolmogorov–Petrovskii–Piscounov equation: One sided travelling-waves. *Ann. Inst. Henri Poincaré Probab. Stat.* **42** (1) (2006) 125–145. MR2196975 https://doi.org/10.1016/j.anihpb.2005.02.005

[30] S. C. Harris and M. I. Roberts. Measure changes with extinction. *Statist. Probab. Lett.* **79** (8) (2009) 1129–1133. MR2510780 https://doi.org/10.1016/j.spl.2008.12.025

[31] S. C. Harris and M. I. Roberts. A strong law of large numbers for branching processes: Almost sure spine events. *Electron. Commun. Probab.* **19** (28) (2014) 6. MR3208326 https://doi.org/10.1214/ECP.v19-2641

[32] S. C. Harris and M. I. Roberts. The many-to-few lemma and multiple spines. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** (1) (2017) 226–242. MR3606740 https://doi.org/10.1214/15-AIHP714

[33] N. Ikeda, M. Nagasawa and S. Watanabe. Branching Markov processes I. *J. Math. Kyoto Univ.* **8** (2) (1968) 233–278. MR0232439 https://doi.org/10.1215/kjm/1250524137

[34] N. Ikeda, M. Nagasawa and S. Watanabe. Branching Markov processes II. *J. Math. Kyoto Univ.* **8** (3) (1968) 365–410. MR0238401 https://doi.org/10.1215/kjm/1250524059

[35] N. Ikeda, M. Nagasawa and S. Watanabe. Branching Markov processes III. *J. Math. Kyoto Univ.* **9** (1) (1969) 95–160. MR0246376 https://doi.org/10.1215/kjm/1250524013

[36] M. Jonckheere and S. Saglietti. On the Kesten–Stigum theorem in L^2 beyond λ-positivity, 2017. Available at arXiv:1701.07634.

[37] O. Kallenberg. *Foundations of Modern Probability*, 2nd edition. *Probability and Its Applications (New York)*. Springer-Verlag, New York, 2002. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

[38] H. Kesten and Branching. Brownian motion with absorption. *Stochastic Process. Appl.* **7** (1) (1978) 9–47. MR0494543 https://doi.org/10.1016/0304-4149(78)90035-2

[39] T. Kurtz, R. Lyons, R. Pemantle and Y. Peres. A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes. In *Classical and Modern Branching Processes (Minneapolis, MN, 1994)* 181–185. *IMA Vol. Math. Appl.* **84**. Springer, New York, 1997. MR1601737 https://doi.org/10.1007/978-1-4612-1862-3\protect\T1\textunderscore14

[40] A. E. Kyprianou and Z. Palmowski. Quasi-stationary distributions for Lévy processes. *Bernoulli* **12** (4) (2006) 571–581. MR2248228 https://doi.org/10.3150/bj/1155735927

[41] R.-L. Liu, Y.-X. Ren and R. Song. Strong law of large numbers for a class of superdiffusions. *Acta Appl. Math.* **123** (2013) 73–97. MR3010225 https://doi.org/10.1007/s10440-012-9715-1

[42] M. Lladser and J. San Martín. Domain of attraction of the quasi-stationary distributions for the Ornstein–Uhlenbeck process. *J. Appl. Probab.* **37** (2) (2000) 511–520. MR1781008 https://doi.org/10.1017/s0021900200015692

[43] O. Louidor and S. Saglietti. A strong law of large numbers for super-critical branching Brownian motion with absorption. Preprint, 2017. Available at arXiv:1708.08440.

[44] P. Maillard. Speed and fluctuations of N-particle branching Brownian motion with spatial selection. *Probab. Theory Related Fields* **166** (3–4) (2016) 1061–1173. MR3568046 https://doi.org/10.1007/s00440-016-0701-9

[45] P. Maillard. The λ-invariant measures of subcritical Bienaymé–Galton–Watson processes. *Bernoulli* **24** (1) (2018) 297–315. MR3706758 https://doi.org/10.3150/16-BEJ877

[46] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25** (3) (1993) 518–548. MR1234295 https://doi.org/10.2307/1427522

[47] S. C. Moy. Extensions of a limit theorem of Everett, Ulam and Harris on multitype branching processes to a branching process with countably many types. *Ann. Math. Stat.* **38** (1967) 992–999. MR0214155 https://doi.org/10.1214/aoms/1177698767

[48] M. Polak and T. Rolski. A note on speed of convergence to the quasi-stationary distribution. *Demonstratio Math.* **45** (2) (2012) 385–397. MR2963076

[49] E. Seneta and D. Vere-Jones. On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* **3** (2) (1966) 403–434. MR0207047 https://doi.org/10.2307/3212128

[50] D. Villemonais. Interacting particle systems and Yaglom limit approximation of diffusions with unbounded drift. *Electron. J. Probab.* **16** (61) (2011) 1663–1692. MR2835250 https://doi.org/10.1214/EJP.v16-925

[51] S. Watanabe. Limit theorem for a class of branching processes. In *Markov Processes and Potential Theory (Proc. Sympos. Math. Res. Center, Madison, Wis., 1967)* 205–232. Wiley, New York, 1967. MR0237007