Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2010, Vol. 46, No. 1, 279–298
DOI: 10.1214/09-AIHP203

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

Strong Law of Large Numbers for branching diffusions

János Engländer[a], Simon C. Harris[b] and Andreas E. Kyprianou[c]

[a] *Department of Statistics and Applied Probability University of California, Santa Barbara, CA 93106-3110, USA. E-mail: englander@pstat.ucsb.edu*
[b] *Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK. E-mail: S.C.Harris@bath.ac.uk*
[c] *Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK. E-mail: a.kyprianou@bath.ac.uk*

Received 11 February 2008; revised 29 October 2008; accepted 20 March 2009

Abstract. Let X be the branching particle diffusion corresponding to the operator $Lu+\beta(u^2-u)$ on $D\subseteq\mathbb{R}^d$ (where $\beta\ge 0$ and $\beta\not\equiv 0$). Let λ_c denote the generalized principal eigenvalue for the operator $L+\beta$ on D and assume that it is finite. When $\lambda_c>0$ and $L+\beta-\lambda_c$ satisfies certain spectral theoretical conditions, we prove that the random measure $\exp\{-\lambda_c t\}X_t$ converges almost surely in the vague topology as t tends to infinity. This result is motivated by a cluster of articles due to Asmussen and Hering dating from the mid-seventies as well as the more recent work concerning analogous results for superdiffusions of [*Ann. Probab.* **30** (2002) 683–722, *Ann. Inst. H. Poincaré Probab. Statist.* **42** (2006) 171–185]. We extend significantly the results in [*Z. Wahrsch. Verw. Gebiete* **36** (1976) 195–212, *Math. Scand.* **39** (1977) 327–342, *J. Funct. Anal.* **250** (2007) 374–399] and include some key examples of the branching process literature. As far as the proofs are concerned, we appeal to modern techniques concerning martingales and "spine" decompositions or "immortal particle pictures."

Résumé. Soit X le processus de diffusion avec branchement correspondant à l'operateur $Lu+\beta(u^2-u)$ sur $D\subseteq\mathbb{R}^d$ (où $\beta\ge 0$ et $\beta\not\equiv 0$). La valeur propre principale généralisée de l'operateur $L+\beta$ sur D est dénotée par λ_c et on la suppose finie. Quand $\lambda_c>0$ et $L+\beta-\lambda_c$ satisfait certaines conditions spectrales théoriques, on montre que la mesure aléatoire $\exp\{-\lambda_c t\}X_t$ converge presque sûrement pour la topologie vague quand t tend vers l'infini. Ce résultat est motivé par un ensemble d'articles par Asmussen et Hering datant du milieu des années soixante-dix, ainsi que par des travaux plus récents [*Ann. Probab.* **30** (2002) 683–722, *Ann. Inst. H. Poincaré Probab. Statist.* **42** (2006) 171–185] concernant des résultats analogues pour les superdiffusions. Nous généralisons de manière significative les résultats de [*Z. Wahrsch. Verw. Gebiete* **36** (1976) 195–212, *Math. Scand.* **39** (1977) 327–342, *J. Funct. Anal.* **250** (2007) 374–399] et nous donnons quelques exemples clés de la théorie des processus de branchement. En ce qui concerne les démonstrations, nous faisons appel aux techniques modernes de martingales et aux "spine decompositions" ou "immortal particle pictures."

MSC: Primary 60J60; secondary 60J80

Keywords: Law of Large Numbers; Spine decomposition; Spatial branching processes; Branching diffusions; Measure-valued processes; h-transform; Criticality; Product-criticality; Generalized principal eigenvalue

1. Introduction and statement of results

1.1. *Model*

Let $D\subseteq\mathbb{R}^d$ be a nonempty domain and write $C^{i,\eta}(D)$ to denote the space of i times ($i=1,2$) continuously differentiable functions with all their ith order derivatives belonging to $C^{\eta}(D)$. [Here $C^{\eta}(D)$ denotes the usual Hölder space.] Consider $Y=\{Y_t;t\ge 0\}$, the diffusion process with probabilities $\{\mathbb{P}_x, x\in D\}$ corresponding to the operator

$$L=\frac{1}{2}\nabla\cdot a\nabla+b\cdot\nabla\quad\text{on } D, \tag{1}$$

where the coefficients $a_{i,j}$ and b_i belong to $C^{1,\eta}(D)$, $i, j = 1, \ldots, d$, for some η in (0, 1], and the symmetric matrix $a(x) = \{a_{i,j}(x)\}$ is positive definite for all $x \in D$. At this point, we do not assume that Y is conservative, that is, the exit time from D may be finite with positive probability. Intuitively, this means that Y may get killed at the Euclidean boundary of D or "run out to infinity" in finite time.

Furthermore let us first assume that $0 \le \beta \in C^{\eta}(D)$ is bounded from above on D and $\beta \not\equiv 0$. The (strictly dyadic) $(L, \beta; D)$-*branching diffusion* is the Markov process with motion component Y and with spatially dependent rate β, replacing particles by precisely two offspring when branching and starting from a finite configuration of individuals. At each time $t > 0$, the state of the process is denoted by $X_t \in \mathcal{M}(D)$ where

$$\mathcal{M}(D) := \left\{ \sum_{i=1}^{n} \delta_{x_i} \colon n \in \mathbb{N} \text{ and } x_i \in D \text{ for } i = 1, \ldots, n \right\}.$$

We will also use the following notation: $X = \{X_t \colon t \ge 0\}$ has probabilities $\{P_\mu \colon \mu \in \mathcal{M}(D)\}$, and E_μ is expectation with respect to P_μ. As usual, $\langle f, \mu \rangle := \int_D f(x)\mu(\mathrm{d}x)$ and $\langle f, g \rangle := \int_D f(x)g(x)\,\mathrm{d}x$, where $\mathrm{d}x$ is Lebesgue measure, and so $\langle f, g\,\mathrm{d}x \rangle = \langle fg, \mathrm{d}x \rangle = \langle f, g \rangle$.

When β is not bounded from above, one may wonder if the $(L, \beta; D)$-branching diffusion is still well defined, in particular, whether the global (or even local) mass may blow up in finite time. Let $\mathcal{M}_{\mathrm{loc}}(D)$ denote the family of locally finite measures on D and consider the family of locally finite discrete measures on D:

$$\mathcal{M}^*_{\mathrm{loc}}(D) := \left\{ \mu \in \mathcal{M}_{\mathrm{loc}}(D) \colon \sum_i \delta_{x_i} \colon x_i \in D \right\}.$$

Let

$$\lambda_c = \lambda_c(L + \beta, D) := \inf\bigl\{\lambda \in \mathbb{R} \colon \exists u > 0 \text{ satisfying } (L + \beta - \lambda)u = 0 \text{ in } D\bigr\}$$

denote the *generalized principal eigenvalue* for $L + \beta$ on D. It is known (see Chapter 4 in [27]) that $\lambda_c < \infty$ whenever β is upper bounded and that, for general β, there exists an $h > 0$ satisfying that

$$(L + \beta - \lambda_c)h = 0, \tag{2}$$

whenever $\lambda_c < \infty$. From (2) it follows by standard approximation arguments (either by truncating the domain or by truncating β) that the $(L, \beta; D)$-branching diffusion X is well defined and is $\mathcal{M}^*_{\mathrm{loc}}(D)$-valued, and, when weighted by h, it is even finite measure-valued. ($W^h_t := \mathrm{e}^{-\lambda_c t} \sum_i h(X^i_t)$ is a supermartingale for $X_t = \sum_i \delta_{X^i_t}$.) Therefore, from now on, *we relax the assumption that* $\sup_D \beta < \infty$ *and replace it with the much milder assumption* $\lambda_c < \infty$.

1.2. *Motivation*

This paper concerns the local growth of mass for branching particle diffusions. In doing so we address a gap in the literature dating back to the mid-seventies when the study of growth of typed branching processes on compact domains of the type space was popularized by Asmussen and Hering. Also we complement a recent revival in this field which has appeared amongst the superprocess community.

Before discussing main results, we shall introduce the topic in detail.

Definition 1 (Local extinction). *Fix $\mu \in \mathcal{M}(D)$. We say that X exhibits local extinction under P_μ if for every Borel set $B \subset\subset D$, there exists a random time τ_B such that*

$$P_\mu\bigl(\tau_B < \infty,\ \text{and } X_t(B) = 0 \text{ for all } t \ge \tau_B\bigr) = 1.$$

[*Here $B \subset\subset D$ means that B is bounded and its closure is a subset of D.*]

Local extinction has been studied by [15], [28] (for superprocesses) and [14] (for branching diffusions). To explain their results, recall that we assume that the generalized principal eigenvalue for $L + \beta$ on D is finite. In fact, $\lambda_c \le 0$ if and only if there exists a function $h > 0$ satisfying $(L + \beta)h = 0$ on D – see Section 4.4 in [27]. Following the papers [15,28] where similar issues were addressed for superprocesses, in [14] the following was shown.

Theorem 2 (Local extinction versus local exponential growth). *Let $\mathbf{0} \neq \mu \in \mathcal{M}(D)$.*

(i) *X under P_μ exhibits local extinction if and only if there exists a function $h > 0$ satisfying $(L + \beta)h = 0$ on D, that is, if and only if $\lambda_c \leq 0$.*
(ii) *When $\lambda_c > 0$, for any $\lambda < \lambda_c$ and $\varnothing \neq B \subset\subset D$ open,*

$$P_\mu\Big(\limsup_{t\uparrow\infty} \mathrm{e}^{-\lambda t} X_t(B) = \infty\Big) > 0,$$

$$P_\mu\Big(\limsup_{t\uparrow\infty} \mathrm{e}^{-\lambda_c t} X_t(B) < \infty\Big) = 1.$$

In particular, local extinction/local exponential growth does not depend on the initial measure $\mathbf{0} \neq \mu \in \mathcal{M}(D)$.

(In [14] it is assumed that β is upper bounded, whereas in [15] only the finiteness of λ_c is assumed. The proofs of [14] go through for this latter case too.) On closer inspection this last theorem says that when $\lambda_c \leq 0$ mass "escapes out of B" even though the entire process may survive with positive probability. (If Y is conservative in D for example then it survives with probability one). Further, when $\lambda_c > 0$ mass accumulates on all nonempty bounded open domains and in such a way that with positive probability this accumulation grows faster than any exponential rate $\lambda < \lambda_c$. On the other hand, mass will not grow faster than at the exponential rate λ_c. It is natural then to ask whether in fact λ_c gives an exact growth rate or not. That is to say, for each $\varnothing \neq B \subset\subset D$ do the random measures $\{\exp\{-\lambda_c t\}X_t \colon t \geq 0\}$ converge in the vague topology almost surely? (The latter we henceforth refer to as the SLLN, the use of the word "strong" here pertains to a.s. convergence.) Further, can one identify the limit? This is precisely the object of interest of a variety of previous studies for both branching diffusions and superprocesses which we shall now review.

We note already here that the process in expectation is given by the *linear kernel* corresponding to the operator $L + \beta$ on D. Therefore, trusting in the Law of Large Numbers for branching processes, one should expect that the process itself grows like the linear kernel too. On the other hand, it is easy to see that the linear kernel does not in general scale precisely with $\exp\{-\lambda_c t\}$ but rather with $f(t)\exp\{-\lambda_c t\}$, where f grows to infinity as $t \to \infty$ and at the same time is subexponential. (Take, for example, $L = \Delta/2$ and $\beta > 0$ constant on $\mathbb{R}^d$, then $f(t) = t^{d/2}$.) In fact the growth is pure exponential if and only if $L + \beta$ is product-critical (see the definition later in this subsection). Proving SLLN seems to be significantly harder in the general case involving the subexponential term f.

In the late seventies Asmussen and Hering wrote a series of papers concerning weak and strong laws of large numbers for a reasonably general class of branching processes which included branching diffusions. See [1] and [2]. In the context of the branching diffusions we consider here one can summarize briefly their achievements by saying that, when D is *bounded*, for a special class of operators $L + \beta$, the rescaled process $\{\exp\{-\lambda_c t\}X_t \colon t \geq 0\}$ converges in the vague topology, almost surely for branching diffusions. Further, for the same class of $L + \beta$ when D is unbounded they proved that there exists the limit *in probability* of $\exp\{-\lambda_c t\}X_t$ as $t \uparrow \infty$ (in the vague topology). The class of $L + \beta$ alluded to they called "positively regular." The latter is a subclass of the class $\mathcal{P}_p^*(D)$ (the class that we shall work with) given below.

A more detailed comparison with [1,2] as well as the discussion on related results on superprocesses is deferred to Section 2.

Before we give the definition of the basic classes of operators that we shall use, $\mathcal{P}_p(D)$ and $\mathcal{P}_p^*(D)$, we need to recall certain concepts of the so-called *criticality theory* of second-order operators. The operator $L + \beta - \lambda_c$ is called *critical* if the associated space of positive harmonic functions is nonempty but the operator does not possess a (minimal positive) Green's function. In this case the space of positive harmonic functions is in fact one-dimensional. Moreover, the space of positive harmonic functions of the adjoint of $L + \beta - \lambda_c$ is also one dimensional.

Assumption 3. *Suppose we choose representatives of these two spaces to be ϕ and $\widetilde{\phi}$ respectively. Throughout the paper and without further reference, we will always assume that $L + \beta - \lambda_c$ is product-critical, i.e. $\langle \phi, \widetilde{\phi} \rangle < \infty$, and in this case we pick ϕ and $\widetilde{\phi}$ with the normalization $\langle \phi, \widetilde{\phi} \rangle = 1$.*

We now define the classes $\mathcal{P}_p(D)$ and $\mathcal{P}_p^*(D)$. Since we want to talk about spatial spread on a generic domain D, we fix, for the rest of the paper, an arbitrary family of domains $\{D_t, t \geq 0\}$ with $D_t \subset\subset D$, $D_t \uparrow D$. (For $D = \mathbb{R}^d$, D_t can be the t-ball, but we can take any other family with $D_t \subset\subset D$, $D_t \uparrow D$ too.)

Definition 4 ($\mathcal{P}_p(D)$ and $\mathcal{P}_p^*(D)$). *For $p \geq 1$, we write $L+\beta \in \mathcal{P}_p(D)$ if*

(i) $\lambda_c = \lambda_c(L+\beta; D) > 0$.
(ii) $\langle \phi^p, \widetilde{\phi} \rangle < \infty$, *in which case we say that $L+\beta-\lambda_c$ is product p-critical.*

Let $q(x,y,t)$ be transition density of $L+\beta$ and

$$Q(x,y,t) := q(x,y,t) - \mathrm{e}^{\lambda_c t}\widetilde{\phi}(y)\phi(x).$$

We write $L+\beta \in \mathcal{P}_p^*(D)$ when the following additional conditions hold for each given $x \in D$ and $\varnothing \neq B \subset\subset D$.

(iii) *There exists a function $a : [0,\infty) \to [0,\infty)$ such that for all $\delta > 0$,*

$$P_{\delta_x}\big(\exists n_0, \forall n > n_0\colon \operatorname{supp}(X_{n\delta}) \subset D_{a_{n\delta}}\big) = 1.$$

(iv) *There exists a function $\zeta : [0,\infty) \to [0,\infty)$ such that, as $t \uparrow \infty$,*

(1) $\zeta(t) \uparrow \infty$,
(2) $\zeta(a_t) = \mathcal{O}(t)$,
(3) $$\alpha_t := \sup_{z\in D_t, y\in B} \frac{|Q(z,y,\zeta(t))|}{\widetilde{\phi}(y)\phi(z)} = \mathrm{o}\big(\mathrm{e}^{\lambda_c t}\big).$$

Let $p(x,y,t)$ denote the transition density of the diffusion corresponding to the operator $(L+\beta-\lambda_c)^{\phi}$. Then $p(x,y,t) = \mathrm{e}^{-\lambda_c t}\phi(y)\phi^{-1}(x)q(x,y,t)$, and thus, (iv) *is equivalent to*

(iv*) *With the same ζ as in* (iv),

$$\lim_{t\to\infty} \sup_{z\in D_t, y\in B} \left| \frac{p(z,y,\zeta(t))}{\phi\widetilde{\phi}(y)} - 1 \right| = 0.$$

Note that a depends on x and ζ, α depend on x and B through (2) and (3). For notational efficiency and on account of the fact that in our proofs no uniformity in x and B is required, we have chosen not to emphasize this dependency. Moreover, it is often the case that ζ and α in fact *do not depend* on x or B, as we shall see in the examples of Section 3 where explicit cases of these quantities are discussed.

Remark 5 (Ergodicity). *Note that criticality is invariant under h-transforms. Moreover, an easy computation shows that ϕ and $\widetilde{\phi}$ transforms into 1 and $\phi\widetilde{\phi}$ respectively when turning from $(L+\beta-\lambda_c)$ to the h-transformed ($h=\phi$) operator $(L+\beta-\lambda_c)^{\phi} = L + a\phi^{-1}\nabla\phi\cdot\nabla$. Therefore, product criticality is invariant under h-transforms too* (*this is not the case with product p-criticality when $p>1$*). *Further, for operators with no zeroth-order term, it is equivalent to positive recurrence* (*ergodicity*) *of the corresponding diffusion process. In particular, $(L+\beta-\lambda_c)^{\phi}$ corresponds to an ergodic diffusion process provided $(L+\beta-\lambda_c)$ is product critical* (*see* [27], *Section* 4.9).

1.3. *Main results*

With the following theorem we wish to address the issue of almost sure convergence in the vague topology of $\{\exp\{-\lambda_c t\}X_t\colon t \geq 0\}$ for branching diffusions with $L+\beta$ belonging to $\mathcal{P}_p^*(D)$, $p>1$ thus generalizing the results of Asmussen and Hering.

Note that since $L+\beta-\lambda_c$ is critical, ϕ is the unique (up to constant multiples) *invariant positive function* for the linear semigroup corresponding to $L+\beta-\lambda_c$ (Theorem 4.8.6. in [27]). Let $\{S_t\}_{t\geq 0}$ denote the semigroup corresponding to $L+\beta$. It is a standard fact (sometimes called "the one particle picture") that

$$S_t(g)(x) = E_{\delta_x}\langle g, X_t\rangle \tag{3}$$

for all nonnegative bounded measurable g's. Even though ϕ is not necessarily bounded from above, $S_t(\phi)$ makes sense and (3) remains valid when g is replaced by ϕ, because ϕ can be approximated with a monotone increasing sequence of g's and the finiteness of the limit is guaranteed precisely by the invariance property of ϕ. By the invariance of ϕ,

$E_{\delta_x} \mathrm{e}^{-\lambda_c t}\langle \phi, X_t\rangle = \mathrm{e}^{-\lambda_c t} S_t(\phi)(x) = \phi(x)$, which, together with the branching property, is sufficient to deduce that W^{ϕ} is a martingale where

$$W_t^{\phi} := \mathrm{e}^{-\lambda_c t}\langle \phi, X_t\rangle, \quad t \geq 0.$$

Indeed note that

$$E_{\delta_x}\big(\mathrm{e}^{-\lambda_c (t+s)}\langle \phi, X_{t+s}\rangle | \mathcal{F}_t\big) = \mathrm{e}^{-\lambda_c t} E_{X_t}\big(\mathrm{e}^{-\lambda_c s}\langle \phi, X_s\rangle\big) = \mathrm{e}^{-\lambda_c t}\langle \phi, X_t\rangle.$$

Being a positive martingale, P_{δ_x}-almost sure convergence is guaranteed, and the a.s. martingale limit $W_\infty^{\phi} := \lim_{t\to\infty} W_t^{\phi}$ appears in the following main conclusion.

Let $C_c^+(D)$ denote the space of nonnegative, continuous and compactly supported functions on D.

Theorem 6 (SLLN). *Assume that $L + \beta \in \mathcal{P}_p^*(D)$ for some $p \in (1, 2]$ and $\langle \beta\phi^p, \widetilde{\phi}\rangle < \infty$. Then,*

$$\lim_{t\uparrow\infty} \mathrm{e}^{-\lambda_c t}\langle g, X_t\rangle = \langle g, \widetilde{\phi}\rangle W_\infty^{\phi}, \quad g \in C_c^+(D) \tag{4}$$

holds P_{δ_x}-a.s. for all $x \in D$, and $E_{\delta_x}(W_\infty^{\phi}) = \phi(x)$. Moreover, if $\sup_D \beta < \infty$ then the restriction $p \in (1, 2]$ can be replaced by $p > 1$.

We close this subsection with the Weak Law of Large Numbers. Here we change the class $\mathcal{P}_p^*(D)$ to the larger class $\mathcal{P}_p(D)$ and get $L^1(P_{\delta_x})$-convergence instead of a.s. convergence (hence the use of the word "weak"). It is important to point out, however, that the class $\mathcal{P}_p^*(D)$ is already quite large – see Section 3, where we verify that key examples from the literature are in fact in $\mathcal{P}_p^*(D)$ and thus obey the SLLN.

Theorem 7 (WLLN). *Suppose that $L + \beta \in \mathcal{P}_p(D)$ for some $p \in (1, 2]$ and $\langle \beta\phi^p, \widetilde{\phi}\rangle < \infty$. Then for all $x \in D$, (4) holds in the $L^1(P_{\delta_x})$ sense and $E_{\delta_x}(W_\infty^{\phi}) = \phi(x)$. Moreover, if $\sup_D \beta < \infty$ then the restriction $p \in (1, 2]$ can be replaced by $p > 1$.*

In closing, we would like to mention an important concept that will play a central role in the proof of Theorems 6 and 7: the so called "spine" will intuitively represent a "typical" particle within the branching diffusion whose motion is governed by the operator $(L + \beta - \lambda_c)^{\phi} = L + a\phi^{-1}\nabla\phi \cdot \nabla$. In Section 4, we give the spine construction and probabilistic interpretation of the branching diffusion under a change of measure using martingale W^{ϕ}. The reader may find it helpful to familiarize themselves with this notion even before reading the proofs. Also see, for example, Section 4.1 of [12] and [22].

1.4. *Outline*

The rest of this paper is organized as follows. In Section 2 we embed our results into the literature, while in Section 3 we discuss some key examples for the SLLN. The proofs are given in Section 4.

2. Detailed comparison with some older results

The methods of Asmussen and Hering were based for the most part on classical techniques of truncation and applications of the Borel–Cantelli lemma. Using this method, they proved the convergence of $\mathrm{e}^{-\lambda_c t}\langle X_t, g\rangle$ for all $0 \leq g \in L^1(\widetilde{\phi}(x)\,\mathrm{d}x)$. It is also worth noting that the generic strength of their method extended to many other types of branching processes; discrete time, discrete space and so on.

Interestingly, preceding all work of Asmussen and Hering is the single article [29] (later improved upon by [4]). Watanabe demonstrates that when a suitable Fourier analysis is available with respect to the operator $L + \beta$, then by spectrally expanding any $g \in C_c^+(D)$, one can show that $\{\langle g, X_t\rangle\colon t \geq 0\}$ is almost surely asymptotically equivalent to its mean. From this the classic Strong Law of Large Numbers for dyadic branching Brownian motion in $\mathbb{R}^d$ is recovered. Namely that when $L = \triangle/2$ and $\beta > 0$ is a constant,

$$\lim_{t\uparrow\infty} t^{d/2} e^{-\beta t} X_t(B) = (2\pi)^{d/2}|B| \times N_\mu,$$

where B is any Borel set, $|B|$ is its Lebesgue measure and N_μ is a strictly positive random variable depending on the initial configuration $\mu \in \mathcal{M}(\mathbb{R}^d)$. The operator $1/2\Delta + \beta$ does not fall into the class $\mathcal{P}_1(D)$. (This is important since [16,17] and also [8] assumes that the operator is in $\mathcal{P}_1(D)$.) For an analogous result on supercritical super-Brownian motion see [13].

Let us discuss now how our assumptions relate to the assumptions imposed in the article [1].

In [1] the domain is bounded (and even one dimensional) when the Strong Law of Large Numbers is stated for branching diffusions; on general domains, only convergence in probability was obtained. Furthermore, in [1] the notion of *positively regular* operators was introduced. In our context it first means that:

(A) $\lambda_c > 0$ (in [1] this property is called "supercriticality"),
(B) ϕ is bounded from above,
(C) $\langle\widetilde{\phi}, 1\rangle < \infty$.

Obviously, (B–C) is stronger than the assumption $\langle\phi, \widetilde{\phi}\rangle < \infty$ (product-criticality).

Secondly, $\{S_t\}_{t\geq 0}$, the semigroup corresponding to $L + \beta$ (the so called "expectation semigroup") satisfies the following condition. If η is a nonnegative, bounded measurable function on $\mathbb{R}^d$, then

$$\text{(D)} \quad S_t(\eta)(x) = \langle\eta, \widetilde{\phi}\rangle\phi(x)\big[e^{\lambda_c t} + o\big(e^{\lambda_c t}\big)\big] \quad \text{as } t \uparrow \infty, \text{ uniformly in } \eta.$$

Let T_t be the semigroup defined by $T_t(f) := S_t^{\phi}(f) = \frac{1}{\phi} S_t(\phi f)$, for all $0 \leq f$ measurable with ϕf being bounded. Then T_t correspond to the h-transformed ($h = \phi$) operator L_0^{ϕ}. Recall that L_0^{ϕ} corresponds to a positive recurrent diffusion process. Then, assuming that ϕ is bounded, it is easy to check that the following condition would suffice for (D) to hold:

$$\lim_{t\uparrow\infty} \sup_{x\in D} \sup_{\|g\|\leq 1} \langle g, \phi\widetilde{\phi}\rangle^{-1} \big|T_t(g) - \langle g, \phi\widetilde{\phi}\rangle\big| = 0, \tag{5}$$

where $\|\cdot\|$ denotes sup-norm. However this is not true in most cases on unbounded domains (or even on bounded domains with general unbounded coefficients) because of the requirement on the uniformity in x. (See our examples in Section 3 – neither of the examples on $\mathbb{R}^d$ satisfy (5).)

Turning to superprocesses, there would seem to be considerably fewer results of this kind in the literature (see the references [10,11,18] for superprocesses in general). The most recent and general work in this area we are aware of are [9,13,16,17].

In [16] it was proved that (in the vague topology) $\{\exp\{-\lambda_c t\}X_t \colon t \geq 0\}$ converges in law where X is the so called $(L, \beta, \alpha, \mathbb{R}^d)$-superprocess (with α being the "activity parameter") satisfying that $L + \beta \in \mathcal{P}_1(D)$ and that $\alpha\phi$ is bounded from above. (An additional requirement was that $\langle\phi, \mu\rangle < \infty$ where $\mu = X_0$ is the deterministic starting measure.) The long and technical proof relied heavily on the theory of dynamical systems applied to the Laplace transforms of $\{e^{-\lambda_c t}\phi X_t, t \geq 0\}$.

In [17] the convergence in law was replaced by convergence in probability. Furthermore, instead of $\mathbb{R}^d$ a general Euclidean domain $D \subseteq \mathbb{R}^d$ was considered. The heavy analytic method of [16] was replaced by a different, simpler and more probabilistic one. The main tool was the introduction of a "weighted superprocess" obtained by a "space–time H-transform."

Very recently, in [8], almost sure limits were proven for a class of Markov branching processes,[1] using mostly functional analytic methods. The main difference between the setup in [8] and our setup are that

(i) [8] used an L^2-approach and therefore the setting had to be restricted to symmetric operators, whereas our results are applicable for non-symmetric operators as well, as long as the other conditions are satisfied (a concrete class of such operators is given in Example 13 of the next section).

[1] A similar result for superprocesses has been obtained in [9].

(ii) Even within the symmetric case, our milder spectral assumptions include e.g. Examples 10 and 11 of the next section, which do not satisfy the assumptions in [8]. To be more specific, Example 10 does not satisfy the assumption that the ground state ϕ is upper bounded. In Example 11, since β is constant, the assumption that β belongs to the class $K_\infty(Y)$ is not satisfied. (The class $K_\infty(Y)$ depends on the motion process Y, and is defined in [8] with the help of standard Kato classes; it contains rapidly decaying functions.)
(iii) While [8] uses a functional analytic (Hilbert space) approach, our method is more probabilistic and is based on a "spine-decomposition."

3. Examples

In this section we give examples which satisfy all the assumptions we have, and thus, according to Theorem 6, obey the SLLN. (Those examples do not fall into the setting in [1,2] and two of them are not covered by [8] either.)

Before we turn to the specific examples, we give some heuristics. Although these are not actually needed for understanding the examples, we feel that the reader "gets a more complete picture" by first reading them.

Remark 8 (Expectation calculations and local vs. global growth rates). *From* (3), *we have*

$$E_x\langle \mathbb{1}_{\{\cdot\in \mathrm{d}y\}}, X_t\rangle = \mathrm{e}^{\lambda_c t}\frac{\phi(x)}{\phi(y)}p(t,x,\mathrm{d}y)$$

and then, by ergodicity,

$$\mathrm{e}^{-\lambda_c t}E_x\langle 1, X_t\rangle = \phi(x)\int_{\mathbb{R}^d}\frac{p(t,x,y)}{\phi(y)}\,\mathrm{d}y \to \phi(x)\int_{\mathbb{R}^d}\widetilde{\phi}(y)\,\mathrm{d}y, \quad \text{as } t\to\infty.$$

Hence, if $\langle\widetilde{\phi},1\rangle<\infty$, *then the global population growth is the same as the local population growth* (*in expectation*), *whereas, if* $\langle\widetilde{\phi},1\rangle=\infty$ *the global growth rate exceeds the local growth rate.*

Remark 9 (Heuristics for a and ζ). *One may wonder how one can find the functions a and* ζ *as in Definition* 4(iii)–(iv). *In fact, it will often be straightforward to find them.*

Fix $x\in D$. *If, for example, we can pick a deterministic increasing function a such that, for all* $\delta>0$,

$$\sum_{n=1}^{\infty}P_{\delta_x}\big(\mathrm{supp}(X_{n\delta})\not\subset D_{a_{n\delta}}\big)<\infty,$$

then Borel–Cantelli says that the function a is an appropriate choice, if also $\zeta(a_t)=\mathcal{O}(t)$ *holds. Since the probability one particle is present in a set is trivially dominated by the expected particle number in that set, it will be much easier to check that*

$$\sum_{n=1}^{\infty}E_{\delta_x}\langle \mathbb{1}_{D^c_{a_{n\delta}}}, X_{n\delta}\rangle<\infty.$$

If $D=\mathbb{R}^d$ *and* $D_t=B_t$ *and we can choose* a_t *such that, for some* $\varepsilon>0$,

$$\int_{|y|>a_t}\frac{p(x,y,t)}{\phi(y)}\,\mathrm{d}y<\mathrm{e}^{-(\lambda_c+\varepsilon)t},$$

then we will have satisfied

$$\sum_{n=1}^{\infty}E_x\langle \mathbb{1}_{D^c_{a_{n\delta}}}, X_{n\delta}\rangle<\infty.$$

Heuristically, at least for a nicely decaying $\widetilde{\phi}$, if the transition density[2] $p(t,x,y)$ *converges to its equilibrium* $\phi(y)\widetilde{\phi}(y)$ *sufficiently quickly even for very large* y, *we might hope to take*

$$a_t \approx F^{-1}\big(\mathrm{e}^{\lambda_c t}\big), \quad \text{where } F(\alpha) = \int_{|y|>\alpha} \widetilde{\phi}(y)\,\mathrm{d}y.$$

If the spine starts at a very large position, since it is ergodic it will tend to move back toward the origin, albeit taking a potentially large time, and Ventcel-Freidlin large deviation theory suggests that it will "closely" follow the path of a deterministic particle with the same drift function. We can use this to guess for a suitable form for $\zeta(t)$. *At least heuristically, to find out how far away the spine particle may start in order that it both returns to the vicinity of the origin and then ergodizes towards its invariant measure before large time* t, *we can consider the deterministic differential equation*

$$\dot{f}(t) = \mu\big(f(t)\big) - \sigma^2\big(f(t)\big)\frac{\nabla\phi(f(t))}{\phi(f(t))}$$

when $L = \frac{1}{2}\sigma^2(x)\Delta - \mu(x)\cdot\nabla$, *and, for example, often take* $\zeta(t)$ *a little larger than* $|f^{-1}(t)|$ *in one dimensional settings.*

Indeed, these heuristics appear to the correct form for both a_t *and* $\zeta(t)$ *in the examples considered below.*

Example 10 (OU process with quadratic breeding rate). *Let* $\sigma, \mu, a, b > 0$ *and consider*

$$L := \frac{1}{2}\sigma^2\Delta - \mu x\cdot\nabla \quad \textit{on } \mathbb{R}^d$$

corresponding to an (inward) Ornstein–Uhlenbeck process and let $\beta(x) := bx^2 + a$. *Since* L *corresponds to a recurrent diffusion and* β *is a smooth function with* $\beta \geq 0$ *and* $\beta \not\equiv 0$, *it follows that* $\lambda_c > 0$ *(see Chapter 4 in [27]). The equilibrium distribution for* L *is given by a normal density,*

$$\pi(x) = \left(\frac{\mu}{\pi\sigma^2}\right)^{d/2}\exp\left\{-\frac{\mu}{\sigma^2}x^2\right\}.$$

Suppose that $\mu > \sigma\sqrt{2b}$. *Defining* $\gamma^{\pm} := \frac{1}{2\sigma^2}(\mu \pm \sqrt{\mu^2 - 2b\sigma^2})$, *for the principal eigenvalue problem with* $(L+\beta)\phi = \lambda_c\phi$ *we can take*

$$\lambda_c := \sigma^2\gamma^- + a, \qquad \phi(x) := c^-\exp\{\gamma^- x^2\} \quad \textit{and} \quad \widetilde{\phi}(x) = c^+\exp\{-\gamma^+ x^2\},$$

where $c^- := (1-(2b\sigma^2/\mu^2))^{d/8}$, $c^+ := c^-(\mu/(\pi\sigma^2))^{d/2}$. *Note that* $L+\beta$ *is a self-adjoint operator with respect to* π, *and for the* h*-transform of a second-order operator* A,

$$\widetilde{\phi}_{A^h} = h\widetilde{\phi}_A.$$

Calculations using the "one-particle picture" (Eq. (3)) reveal that, in expectation, the support of the process grows like $\sqrt{\lambda_c t/\gamma^+}$ *and one can pick* $a_t = \sqrt{\lambda t/\gamma^+}$ *for any* $\lambda > \lambda_c$ *and condition* (iii) *in Definition 4 will hold.*

The spine is also an (inward) Ornstein–Uhlenbeck process with parameter $\alpha := \mu - 2\gamma^-\sigma^2 = \sqrt{\mu^2 - 2b\sigma^2}$ *with*

$$(L+\beta-\lambda_c)^{\phi} = L + \sigma^2\frac{\nabla\phi}{\phi}\cdot\nabla = \frac{1}{2}\sigma^2\Delta - \alpha x\cdot\nabla \quad \textit{on } \mathbb{R}^d,$$

and transition density

$$p(x,y,t) = \left(\frac{\alpha}{\pi\sigma^2(1-\mathrm{e}^{-2(\alpha/\sigma^2)t})}\right)^{d/2}\exp\left[-\frac{\alpha\sum_{i=1}^d (y_i - x_i\mathrm{e}^{-(\alpha/\sigma^2)t})^2}{\sigma^2(1-\mathrm{e}^{-2(\alpha/\sigma^2)t})}\right].$$

[2] Later we will see that this transition density corresponds to the "spine."

We see that the drift of the inward OU reduces the influence of any starting position exponentially in time. Indeed, one can take $\zeta(t) = (1+\varepsilon)(\sigma^2/2\alpha)\log t$ *for any* $\varepsilon > 0$ *for condition* (iv*) *in Definition* 4 *to hold. Finally, we trivially note that* $\zeta(a_t) = \mathcal{O}(t)$ (*in fact, only* $\log t$ *growth*), *hence, for* p *sufficiently close to* 1, *all necessary conditions are satisfied for our strong law theorem to hold.*

Note that a strong law for a generalization of this model can be found in [23] *where the convergence is proved using a martingale expansion for continuous functions* $g \in L^2(\pi)$ (*rather than compactly supported* g). *Almost sure asymptotic growth rates* (*and a.s. support*) *for the same model are studied in* [20].

This is certainly a non-trivial model and it highlights the strength of our general result. In particular, a quadratic breeding rate is critical in the sense that a BBM with breeding rate $\beta(x) = const \cdot x^p$ *explodes in a finite time a.s. if and only if* $p > 2$, *with explosion in the expected population size when* $p = 2$. *When a branching inward OU process with quadratic breeding is considered here, a strong enough drift with* $\mu > \sigma\sqrt{2b}$ *can balance the high breeding, whereas any lower drift would lead to a dramatically different behavior.*

Example 11 (Outward OU process with constant breeding rate). *Let* $\sigma, \mu > 0$; $b > d\mu$ *and consider*

$$L := \frac{1}{2}\sigma^2\Delta + \mu x \cdot \nabla \quad \text{on } \mathbb{R}^d$$

corresponding to an "outward" Ornstein–Uhlenbeck process and let $\beta(\cdot) \equiv b$. *As the spatial movements have no affect on the branching, the global population grows like* $e^{\beta t}$ *and this is achieved "naturally" with particles moving freely. This corresponds to* $(L+\beta)\widetilde{\phi} = b\widetilde{\phi}$ *with* $\widetilde{\phi} \equiv 1$. *On the other hand, the principal eigenvalue is* $\lambda_c = b - d\mu < b$ *with* $\phi(x) = const \cdot \exp\{-(\mu/\sigma^2)x^2\}$, *it being associated with the* local, *as opposed to global, growth rate.*

After some similar expectation calculations to the inward OU in quadratic potential, an upper bound on the process' spread is roughly the same as for an individual outward OU particle, that is, we can take $a_t = \exp\{(1+\delta)(\mu/\sigma^2)t\}$ *for any* $\delta > 0$.

Despite the transient nature of the original motion, the spine is an inward OU process

$$(L+\beta-\lambda_c)^{\phi} = L + \sigma^2\frac{\nabla\phi}{\phi}\cdot\nabla = \frac{1}{2}\sigma^2\Delta - \mu x\cdot\nabla \text{ on } \mathbb{R}^d,$$

with equilibrium $\phi\widetilde{\phi}(x) \propto \exp\{-(\mu/\sigma^2)x^2\}$. *Intuitively, this is the motion that maximizes the local growth rate at* λ_c (*here it is the original motion "conditioned to keep returning to the origin"*). *We can therefore take* $\zeta(t) = (1+\varepsilon)(\sigma^2/\mu)\log t$ *for any* $\varepsilon > 0$ *and hence still find that* $\zeta(a_t) = (1+\varepsilon)(1+\delta)t = \mathcal{O}(t)$. *All the conditions required for the strong law to hold are again satisfied.*

Example 12 (BBM with $\beta \in C_c^+(\mathbb{R}^d)$ ***and*** $\beta \not\equiv 0$ ***for*** $d = 1, 2$***).*** *Consider the* $(\frac{1}{2}\Delta + \beta)$*-branching diffusion where* $\beta \in C_c^+(\mathbb{R}^d)$ *and* $\beta \not\equiv 0$ *for* $d = 1, 2$. *Since Brownian motion is recurrent in dimension* $d = 1, 2$, *it follows that* $\lambda_c > 0$ *and in fact, the operator* $\frac{1}{2}\Delta + \beta - \lambda_c$ *is product-critical and even product-*p*-critical for all* $p > 1$ (*see Example* 22 *in* [16]).

We now show how to find a ζ *that satisfies* (iv*) *in Definition* 4. *We do it for* $d = 1$, *the* $d = 2$ *case is similar.*

Let $b > 0$ *be so large that supp*$(\beta) \subset [-b, b]$ *and let* $M := \max_{\mathbb{R}} \beta$. *Recall that* $p(t, x, y)$ *denotes the* (*ergodic*) *kernel corresponding to* $(\frac{1}{2}\Delta + \beta - \lambda_c)^{\phi}$. *In this example* P *will denote the corresponding probability. By comparison with the constant branching rate case, it is evident that* $a_t := \sqrt{2M}\cdot t$ *is an appropriate choice, because it is well known that a BBM with constant rate* M *has velocity* $\sqrt{2M}$. *Therefore, we have to find a* ζ *which satisfies that for any fixed ball* B,

$$\lim_{t\to\infty}\sup_{|z|\le t}\left|\frac{p(z, B, \zeta(t))}{\int_B \phi\widetilde{\phi}(y)dy} - 1\right| = 0$$

together with the condition that $\zeta(a_t) = \zeta(\sqrt{2M}\cdot t) = \mathcal{O}(t)$ *as* $n \to \infty$.

An easy computation (*see again Example* 22 *in* [16]) *shows that on* $\mathbb{R}\setminus[-b, b]$,

$$\left(\frac{1}{2}\Delta + \beta - \lambda_c\right)^{\phi} = \frac{1}{2}\Delta - \operatorname{sgn}(x)\cdot\sqrt{2\lambda_c}\frac{d}{dx},$$

where $\operatorname{sgn}(x) := x/|x|, x \neq 0$. *Fix an* ε *and let* $\tau_{\pm b}$ *and* τ_0 *denote the first hitting time (by a single Brownian particle) of* $[-b, b]$ *and of* 0, *respectively. We first show that as* $t \to \infty$,

$$\sup_{b<|x|\le t} P_x\left[\tau_{\pm b} > \frac{t(1+\varepsilon)}{\sqrt{2\lambda_c}}\right] \to 0. \tag{6}$$

Obviously, it is enough to show that, for example,

$$\mathbf{P}_t\left[\tau_0 > \frac{t(1+\varepsilon)}{\sqrt{2\lambda_c}}\right] \to 0,$$

where $\mathbf{P}$ *corresponds to* $\frac{1}{2}\Delta - \sqrt{2\lambda_c}\frac{\mathrm{d}}{\mathrm{d}x}$ *on* $[0, \infty)$. *Indeed, if* $\mathcal{W}$ *denotes standard Brownian motion starting at the origin with probability* Q, *then*

$$\begin{aligned}\mathbf{P}_t\left[\tau_0 > \frac{t(1+\varepsilon)}{\sqrt{2\lambda_c}}\right] &\le \mathbf{P}_t[Y_{t(1+\varepsilon)/\sqrt{2\lambda_c}} > 0] = Q\left[t - \sqrt{2\lambda_c}\frac{t(1+\varepsilon)}{\sqrt{2\lambda_c}} + \mathcal{W}_{t(1+\varepsilon)/\sqrt{2\lambda_c}} > 0\right]\\ &= Q[\mathcal{W}_{t(1+\varepsilon)/\sqrt{2\lambda_c}} > \varepsilon t] \to 0\end{aligned}$$

(*the last term tends to zero by the SLLN for* $\mathcal{W}$).

We now claim that $\zeta(t) := \frac{t(1+2\varepsilon)}{\sqrt{2\lambda_c}}$ *satisfies*

$$\lim_{t\to\infty} \sup_{|z|\le t} \left|\frac{p(z, B, \zeta(t))}{\int_B \phi\widetilde{\phi}(y)\,\mathrm{d}y} - 1\right| = 0.$$

(*The condition* $\zeta(a_t) = \mathcal{O}(t)$ *is obviously satisfied.*) *By the ergodicity of* $p(t, x, y)$, *it is sufficient to show that* ζ *satisfies*

$$\lim_{t\to\infty} \sup_{b<|z|\le t} \left|\frac{p(z, B, \zeta(t))}{\int_B \phi\widetilde{\phi}(y)\,\mathrm{d}y} - 1\right| = 0.$$

Let, for example, $b < x \le t$. *By the strong Markov property at* τ_b (*the hitting time of* b) *and by* (6),

$$\frac{p(x, B, \zeta(t))}{\int_B \phi\widetilde{\phi}(y)\,\mathrm{d}y} = \frac{p(b, B, \zeta(t) - t(1+\varepsilon)/\sqrt{2\lambda_c})}{\int_B \phi\widetilde{\phi}(y)\,\mathrm{d}y} P_x\left[\tau_b \le \frac{t(1+\varepsilon)}{\sqrt{2\lambda_c}}\right] + \mathrm{o}(1),$$

uniformly in $b < x \le t$.

Finally,

$$\lim_{t\to\infty} \frac{p(b, B, \zeta(t) - t(1+\varepsilon)/\sqrt{2\lambda_c})}{\int_B \phi\widetilde{\phi}(y)\,\mathrm{d}y} = 1$$

because $p(t, x, y)$ *is an ergodic kernel and*

$$\lim_{t\to\infty}\left[\zeta(t) - \frac{t(1+\varepsilon)}{\sqrt{2\lambda_c}}\right] = \lim_{t\to\infty} \frac{t\varepsilon}{\sqrt{2\lambda_c}} = \infty,$$

completing the proof of our claim about ζ.

Example 13 (Non-symmetric operator). *For the sake of concreteness, we give here a simple example for a non-symmetric operator that satisfies our assumptions, by slightly modifying the setting of Example* 12. (*For more on symmetric operators, see Section* 4.10 *in* [27].)

In Example 12 *set d=2. Now add a drift* $b(x, y)$ *as follows. Let* $b = (b_1, b_2)^T$, *where* $b_1(x, y) := m(x)n(y)$ *and* $b_2(x, y) := p(x)q(y)$. *If* m, n, p, q *are smooth compactly supported functions, then so is* b, *and the same argument* *as in the previous example shows that the conditions are satisfied, but if $m(x)n'(y)$ is not equal to $p'(x)q(y)$ for all (x, y), that is, if $m/p'(x)$ is not equal to $q/n'(y)$, then the operator is not symmetric, because then b is not a gradient vector.*

Hence, whenever q/n' is not a constant or m/p' is not a constant, this setting constitutes a non-symmetric example for the SLLN.

Example 14 (Bounded domain). *First note that when D is bounded, an important subset of $\mathcal{P}_p(D)$, $p > 1$ is formed by the operators $L + \beta$ which are uniformly elliptic on D with bounded coefficients which are smooth up to the boundary of D and with $\lambda_c > 0$. That is, in this case $L + \beta - \lambda_c$ is critical (see [27], Section 4.7), and since ϕ and $\widetilde{\phi}$ are Dirichlet eigenfunctions (are zero at the boundary of D), it is even product-p-critical for all $p > 1$. Theorem 7 thus applies.*

Although in this case Y is not conservative in D, in fact even Theorem 6 will be applicable whenever (iv) can be strengthened to the following uniform convergence on D:*

$$\lim_{t\to\infty} \sup_{z\in D, y\in B} \left| \frac{p(z, y, \zeta(t))}{\phi\widetilde{\phi}(y)} - 1 \right| = 0. \tag{7}$$

(Note that [1] has a similar global uniformity assumption – see the paragraph after (5).) Indeed, then the proof of Theorem 6 (which can be found later, in Section 4) can be simplified, because the function a is not actually needed: D_{a_n} can be replaced by D for all $n \geq 1$.

As far as (7) is concerned, it is often relatively easy to check. For example, assume that $d = 1$ (the method can be extended for radially symmetric settings too) and so let $D = (r, s)$. Then the drift term of the spine is $b + a(\log\phi)'$. Now, if this is negative and bounded away from zero at $s - \varepsilon < x < s$ and positive and bounded away from zero at $r < x < r + \varepsilon$ with some $\varepsilon \in (0, s - r)$, then (7) can be verified by a method similar to the one in the previous example. The above condition on the drift is not hard to check in a concrete example (note that since ϕ satisfies the Dirichlet boundary condition, $\log\phi$ tends to $-\infty$ at the boundary).

If we relax the regularity assumptions on $L + \beta$ then, for example, ϕ is not necessarily upper bounded, and so we are leaving the family of operators handled in [2] (see the four paragraphs preceding (5)); nevertheless our method still works as long as $L + \beta \in \mathcal{P}_p^*(D)$, $p > 1$ (for the SLLN) or $L + \beta \in \mathcal{P}_p(D)$, $p > 1$ (for the WLLN).

4. Proofs

4.1. *A spine approach*

To establish the $L^p(P_{\delta_x})$ convergence of W^ϕ for $p > 1$ we appeal to a, by now, standard technique that have been introduced to the literature by [26] and by the references given there (in particular [25]) and involves a change of measure inducing a "spine" decomposition. Similar applications can be found in [3,5,14,21] to name but a few. See, for example, [18,19] as well as the discussion in [14] for yet further references.

It is important to point out that we will need the spine decomposition not only to establish the L^p-convergence mentioned above but also in the key lemma (Lemma 18) in the proof of Theorem 6. In both cases, we found the spine method to be indispensable and we were not able to replace it by other L^p methods.

Before we can state our spine decomposition, we need to recall some facts concerning changes of measures for diffusions and Poisson processes.

Girsanov change of measure
Suppose that Y is adapted to some filtration $\{\mathcal{G}_t\colon t \geq 0\}$. Under the change of measure

$$\left.\frac{\mathrm{d}\mathbb{P}_x^\phi}{\mathrm{d}\mathbb{P}_x}\right|_{\mathcal{G}_t} = \frac{\phi(Y_t)}{\phi(x)} \mathrm{e}^{-\int_0^t (\lambda_c - \beta(Y_s))\,\mathrm{d}s} \tag{8}$$

the process $(Y, \mathbb{P}_x^\phi)$ corresponds to the h-transformed ($h = \phi$) generator $(L + \beta - \lambda_c)^\phi = L + a\phi^{-1}\nabla\phi \cdot \nabla$. Note now in particular that since $L + \beta \in \mathcal{P}_1(D)$, it follows that $(Y, \mathbb{P}_x^\phi)$ is an ergodic diffusion with transition density $p(x, y, t)$ and an invariant density $\phi\widetilde{\phi}$.

Change of measure for Poisson processes
Suppose that given a non-negative continuous function $g(t), t \geq 0$, the Poisson process $(n, \mathbb{L}^g)$ where $n = \{\{\sigma_i\colon i = 1, \ldots, n_t\}\colon t \geq 0\}$ has instantaneous rate $g(t)$. Further, assume that n is adapted to $\{\mathcal{G}_t\colon t \geq 0\}$. Then under the change of measure

$$\left.\frac{\mathrm{d}\mathbb{L}^{2g}}{\mathrm{d}\mathbb{L}^g}\right|_{\mathcal{G}_t} = 2^{n_t} \exp\left\{-\int_0^t g(s)\,\mathrm{d}s\right\}$$

the process $(n, \mathbb{L}^{2g})$ is also a Poisson process with rate $2g$. See Chapter 3 in [24].

Theorem 15 (The spine construction). *Let $\{\mathcal{F}_t\colon t \geq 0\}$ be the natural filtration generated by X. Define the change of measure*

$$\left.\frac{\mathrm{d}\widetilde{P}_{\delta_x}}{\mathrm{d}P_{\delta_x}}\right|_{\mathcal{F}_t} = \mathrm{e}^{-\lambda_c t}\frac{\langle \phi, X_t\rangle}{\phi(x)} = \frac{W_t^\phi}{\phi(x)}.$$

Then, under $\widetilde{P}_{\delta_x}$, X can be constructed as follows:

- *a single particle, $Y = \{Y_t\}_{t\geq 0}$, referred to as the spine, initially starts at x and moves as a diffusion corresponding to the h-transformed operator $L + a\phi^{-1}\nabla\phi \cdot \nabla$;*
- *the spine undergoes fission into two particles at an accelerated rate $2\beta(Y_t)$ at time t, one of which is selected uniformly at random to continue the spine motion Y;*
- *the remaining child gives rise to an independent copy of a P-branching diffusion started at its space–time point of creation.*

A similar construction for BBM was established in Chauvin and Rouault [7]. See Theorem 5 in [14] on how to prove it.[3]

Remark 16 (The spine decomposition). *Theorem 15 says that $(X, \widetilde{P}_{\delta_x})$ has the same law as a process constructed in the following way. A $(Y, \mathbb{P}_x^\phi)$-diffusion is initiated along which $(L, \beta; D)$-branching processes immigrate at space–time points $\{(Y_{\sigma_i}, \sigma_i)\colon i \geq 1\}$ where, given Y, $n = \{\{\sigma_i\colon i = 1, \ldots, n_t\}\colon t \geq 0\}$ is a Poisson process with law $\mathbb{L}^{2\beta(Y)}$. It will often be very useful to think of $(X, \widetilde{P}_{\delta_x})$ as being constructed in this richer way and it will be convenient to define the natural filtration of the spine and the birth process along the spine as $\mathcal{G}_t := \sigma(Y_s, n_s\colon s \leq t)$. Note that using the "spine construction" of $(X, \widetilde{P})$, we can write*

$$W_t^\phi = \mathrm{e}^{-\lambda_c t}\phi(Y_t) + \sum_{i=1}^{n_t} \mathrm{e}^{-\lambda_c \sigma_i} W_i,$$

where, conditional on the spine filtration $\mathcal{G}_t$, W_i is an independent copy of the martingale W_t^ϕ started from position Y_{σ_i} and run for time $t - \sigma_i$ where σ_i is the ith fission time along the spine for $i = 1, \ldots, n_t$. Remembering that particles off the spine behave the same as if under the original measure P and that the martingale property gives $E_{\delta_x}(W_t^\phi) = \phi(x)$, we then have the so called "spine decomposition":

$$\widetilde{E}\left(W_t^\phi | \mathcal{G}_t\right) = \mathrm{e}^{-\lambda_c t}\phi(Y_t) + \sum_{i=1}^{n_t} \mathrm{e}^{-\lambda_c \sigma_i}\phi(Y_{\sigma_i}). \tag{9}$$

---

[3] Although the construction in [14] was for a branching process with killing on the boundary of a compact domain, the analysis applies almost verbatim with obvious changes to take account of the fact that there are no boundary conditions; this straightforward exercise is left to the reader.

The L^p-convergence of the martingale
The a.s. convergence of W^ϕ can be complemented with the following result.

Lemma 17. *Assume that $L+\beta$ belongs to $\mathcal{P}_p(D)$ and that $\langle \beta\phi^p, \widetilde{\phi}\rangle < \infty$ for some $p \in (1,2]$. Then, for $x \in D$, W^ϕ is an $L^p(P_{\delta_x})$-convergent martingale. Moreover, if $\sup_D \beta < \infty$, then the same conclusion holds if we only assume that $p > 1$.*

Proof. Pick p so that $q = p-1 \in (0,1]$, $\langle \phi^p, \widetilde{\phi}\rangle < \infty$ and $\langle \beta\phi^p, \widetilde{\phi}\rangle < \infty$. (If $K := \sup_D \beta < \infty$ and we only assume that $p > 1$, then $\langle \phi^p, \widetilde{\phi}\rangle = \langle \phi^q, \phi\widetilde{\phi}\rangle < \infty$ implies $\langle \phi^r, \widetilde{\phi}\rangle = \langle \phi^{r-1}, \phi\widetilde{\phi}\rangle < \infty$ and $\langle \beta\phi^r, \widetilde{\phi}\rangle \le K\langle \phi^{r-1}, \phi\widetilde{\phi}\rangle < \infty$ for all $r \in (0,p)$, and so we can assume that in fact $p \in (1,2]$.)

We adopt an approach similar to the one in [22]. Using the conditional form of Jensen's inequality, the spine decomposition (9) and that $(u+v)^q \le u^q + v^q$ for $u, v > 0$ when $q \in (0,1]$, we find

$$\begin{aligned}
\phi(x)^{-1} E_{\delta_x}\big[\big(W_t^\phi\big)^p\big] &= \widetilde{E}_{\delta_x}\big[\big(W_t^\phi\big)^q\big] = \widetilde{E}_{\delta_x}\big\{\widetilde{E}\big[\big(W_t^\phi\big)^q|\mathcal{G}_t\big]\big\} \\
&\le \widetilde{E}_{\delta_x}\big\{\big[\widetilde{E}\big(W_t^\phi|\mathcal{G}_t\big)\big]^q\big\} \\
&\le \mathbb{E}_x^\phi \mathbb{L}^{2\beta(Y)}\left(\mathrm{e}^{-\lambda_c q t}\phi(Y_t)^q + \sum_{i=1}^{n_t} \mathrm{e}^{-\lambda_c q \sigma_i}\phi(Y_{\sigma_i})^q\right) \\
&= \mathrm{e}^{-\lambda_c q t}\mathbb{E}_x^\phi\big[\phi(Y_t)^q\big] + \mathbb{E}_x^\phi\left[\int_0^t \mathrm{e}^{-\lambda_c q s} 2\beta(Y_s)\phi(Y_s)^q \,\mathrm{d}s\right].
\end{aligned}$$

Call the two expressions on the right-hand side the *spine term*, $A(x,t)$, and the *sum term*, $B(x,t)$, respectively. Since Y has generator $L + a\phi^{-1}(\nabla\phi)\cdot\nabla$ and $\langle \phi, \widetilde{\phi}\rangle = 1$, Y is ergodic and $\mathbb{E}_x^\phi(f(Y_t)) \to \langle f\phi, \widetilde{\phi}\rangle$ for every measurable f satisfying $\langle f\phi, \widetilde{\phi}\rangle < \infty$. Then

$$\lim_{t\uparrow\infty} \mathrm{e}^{\lambda_c q t} A(t,x) = \lim_{t\uparrow\infty} \mathbb{E}_x^\phi\big(\phi(Y_t)^q\big) = \big\langle \phi^p, \widetilde{\phi}\big\rangle < \infty$$

for all $x \in D$. For the sum term note that $\lim_{s\uparrow\infty} \mathbb{E}_x^\phi(\beta(Y_s)\phi(Y_s)^q) = \langle \beta\phi^p, \widetilde{\phi}\rangle < \infty$ and so $\lim_{t\uparrow\infty} B(t,x) < \infty$ for all $x \in D$. By Doob's inequality, W^ϕ is therefore an L^p-convergent (uniformly integrable) martingale, as required. □

4.2. *Proof of Theorem 6 along lattice times*

The statement that $E_{\delta_x}(W_\infty^\phi) = \phi(x)$ as well as the one after it follow from Lemma 17 and the first paragraph of its proof respectively.

The rest of the proof will be based on the following key lemma.

Lemma 18. *Fix $\delta > 0$ and let $B \subset\subset D$. Define*

$$U_t = \mathrm{e}^{-\lambda_c t}\langle \phi|_B, X_t\rangle,$$

where $\phi|_B(x) = \phi(x)\mathbb{1}_{(x\in B)}$. Then for any non-decreasing sequence $\{m_n\}_{n\ge1}$,

$$\lim_{n\uparrow\infty} \big|U_{(m_n+n)\delta} - E(U_{(m_n+n)\delta}|\mathcal{F}_{n\delta})\big| = 0, \quad P_{\delta_x}\text{-a.s.}$$

Proof. We will suppress the dependence in n in our notation below and simply write m instead m_n. Suppose that $\{X_i\colon i = 1,\dots,N_{n\delta}\}$ describes the configuration of particles at time $n\delta$. Note that we may always write

$$U_{(m+n)\delta} = \sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-n\delta\lambda_c} U_{m\delta}^{(i)}, \tag{10}$$

where given $\mathcal{F}_{n\delta}$, the collection $\{U_{m\delta}^{(i)}: i=1,\dots,N_{n\delta}\}$ are mutually independent and equal in distribution to $U_{m\delta}$ under $P_{\delta_{X_i}}$ respectively.

By the Borel–Cantelli lemma, it is sufficient to prove that for $x\in D$ and for all $\varepsilon>0$,

$$\sum_{n\ge 1} P_{\delta_x}\big(\big|U_{(m+n)\delta}-E(U_{(m+n)\delta}|\mathcal{F}_{n\delta})\big|>\varepsilon\big)<\infty.$$

To this end we first note that,

$$P_{\delta_x}\big(\big|U_{(m+n)\delta}-E(U_{(m+n)\delta}|\mathcal{F}_{n\delta})\big|>\varepsilon\big)\le\frac{1}{\varepsilon^p}E_{\delta_x}\big(\big|U_{(m+n)\delta}-E(U_{(m+n)\delta}|\mathcal{F}_{n\delta})\big|^p\big).$$

Now recall the following very useful result, for example, see [4] or [6]: if $p\in(1,2)$ and Z_i are independent random variables with $E(Z_i)=0$ (or they are martingale differences), then

$$E\left|\sum_{i=1}^n Z_i\right|^p\le 2^p\sum_{i=1}^n E|Z_i|^p. \tag{11}$$

(This is trivially true for $p=2$.) Jensen's inequality also implies that for each $n\ge1$, $|\sum_{i=1}^n u_i|^p\le n^{p-1}\sum_{i=1}^n(|u_i|^p)$ and, in particular, $|u+v|^p\le 2^{p-1}(|u|^p+|v|^p)$.

Note that

$$U_{s+t}-E(U_{s+t}|\mathcal{F}_t)=\sum_{i=1}^{N_t}\mathrm{e}^{-\lambda_c t}\big(U_s^{(i)}-E\big(U_s^{(i)}|\mathcal{F}_t\big)\big),$$

where conditional on $\mathcal{F}_t$, $Z_i:=U_s^{(i)}-E(U_s^{(i)}|\mathcal{F}_t)$ are independent with $E(Z_i)=0$. Thus, by (11) and Jensen,

$$\begin{aligned}
&E\big(\big|U_{s+t}-E(U_{s+t}|\mathcal{F}_t)\big|^p|\mathcal{F}_t\big)\\
&\quad\le 2^p\mathrm{e}^{-p\lambda_c t}\sum_{i=1}^{N_t}E\big(\big|U_s^{(i)}-E\big(U_s^{(i)}|\mathcal{F}_t\big)\big|^p|\mathcal{F}_t\big)\\
&\quad\le 2^p\mathrm{e}^{-p\lambda_c t}\sum_{i=1}^{N_t}E\big(2^{p-1}\big(\big|U_s^{(i)}\big|^p+\big|E\big(U_s^{(i)}|\mathcal{F}_t\big)\big|^p\big)|\mathcal{F}_t\big)\\
&\quad\le 2^p\mathrm{e}^{-p\lambda_c t}\sum_{i=1}^{N_t}2^{p-1}E\big(\big|U_s^{(i)}\big|^p+E\big(\big|U_s^{(i)}\big|^p|\mathcal{F}_t\big)|\mathcal{F}_t\big)\\
&\quad\le 2^{2p}\mathrm{e}^{-p\lambda_c t}\sum_{i=1}^{N_t}E\big(\big|U_s^{(i)}\big|^p|\mathcal{F}_t\big).
\end{aligned}$$

Then, as a consequence of the previous estimate, we have that

$$\begin{aligned}
&\sum_{n\ge1}E_{\delta_x}\big(\big|U_{(m+n)\delta}-E(U_{(m+n)\delta}|\mathcal{F}_{n\delta})\big|^p\big)\\
&\quad\le 2^{2p}\sum_{n\ge1}\mathrm{e}^{-\lambda_c n\delta p}E_{\delta_x}\left(\sum_{i=1}^{N_{n\delta}}E_{\delta_{X_i}}\big[(U_{m\delta})^p\big]\right).
\end{aligned}\tag{12}$$

Recalling the definition of the terms $A(x,t)$ and $B(x,t)$ from the proof of Lemma 17 and trivially noting that $U_t \le W_t^{\phi}$, we have

$$\begin{aligned}
&\sum_{n\ge 1} E_{\delta_x}\big(\big|U_{(m+n)\delta} - E(U_{(m+n)\delta}|\mathcal{F}_{n\delta})\big|^p\big)\\
&\quad\le 2^{2p}\sum_{n\ge 1} \mathrm{e}^{-\lambda_c n\delta p} E_{\delta_x}\left(\sum_{i=1}^{N_{n\delta}} E_{\delta_{X_i}}\big[\big(W_{m\delta}^{\phi}\big)^p\big]\right)\\
&\quad\le 2^{2p}\sum_{n\ge 1} E_{\delta_x}\left(\sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-p\lambda_c n\delta}\phi(X_i)\big(A(X_i,m\delta)+B(X_i,m\delta)\big)\right)\\
&\quad= 2^{2p}\sum_{n\ge 1}\phi(x)\mathrm{e}^{-q\lambda_c\delta n}\mathbb{E}_x^{\phi}\big(A(Y_{n\delta},m\delta)+B(Y_{n\delta},m\delta)\big),
\end{aligned}$$

where we have used the "one-particle picture" (Eq. (3)) and the spine change of measure at (8). Since the spine Y is Markovian and ergodic under $\mathbb{P}_x^{\phi}$, we know that

$$\mathbb{E}_x^{\phi}\big[A(Y_{n\delta},m\delta)\big] = \mathrm{e}^{-\lambda_c q m\delta}\mathbb{E}_x^{\phi}\big(\phi(Y_{(m+n)\delta})^q\big).$$

Denoting $m_\infty := \lim_{n\to\infty} m_n$, the latter converges to $\mathrm{e}^{-q\lambda_c m_\infty\delta}\langle\phi^p,\widetilde{\phi}\rangle$ (which will be zero if $m_\infty=\infty$) as $n\uparrow\infty$. Recall the assumption that $\langle\beta\phi^p,\widetilde{\phi}\rangle<\infty$. Similarly to previously, we have that

$$\mathbb{E}_x^{\phi}\big[B(Y_{n\delta},m\delta)\big] = 2\int_0^{m\delta}\mathrm{e}^{-\lambda_c q s}\mathbb{E}_x^{\phi}\big(\beta(Y_{s+n\delta})\phi(Y_{s+n\delta})^q\big)\,\mathrm{d}s$$

which has a finite limit equal to $2\int_0^{m_\infty\delta}\mathrm{e}^{-\lambda_c q s}\langle\beta\phi^p,\widetilde{\phi}\rangle\,\mathrm{d}s$ as $n\uparrow\infty$. These facts are enough to conclude that the last sum remains finite to complete the Borel–Cantelli argument. □

We now complete the proof of Theorem 6 along lattice times. Assume that $L+\beta\in\mathcal{P}_p^*$ for some $p>1$. Recall now that $I(B) := \int_B \phi(y)\widetilde{\phi}(y)\,\mathrm{d}y < 1$ and that $\{X_i\colon i=1,\dots,N_t\}$ describes the configuration of particles in the process at time $t>0$. Note that, similarly to (10),

$$\begin{aligned}
E(U_{t+s}|\mathcal{F}_t) &= \sum_{i=1}^{N_t}\mathrm{e}^{-\lambda_c t}\phi(X_i)p(X_i,B,s) = \sum_{i=1}^{N_t}\mathrm{e}^{-\lambda_c(t+s)}\int_B\phi(y)q(X_i,y,s)\,\mathrm{d}y\\
&= \sum_{i=1}^{N_t}\mathrm{e}^{-\lambda_c t}\phi(X_i)I(B) + \sum_{i=1}^{N_t}\mathrm{e}^{-\lambda_c(t+s)}\int_B\phi(y)Q(X_i,y,s)\,\mathrm{d}y\\
&= I(B)W_t^{\phi} + \sum_{i=1}^{N_t}\mathrm{e}^{-\lambda_c(t+s)}\int_B\phi(y)Q(X_i,y,s)\,\mathrm{d}y =: I(B)W_t^{\phi} + \Theta(t,s).
\end{aligned}$$

Let us replace now t by $n\delta$ and s by $m_n\delta$, where

$$m_n := \zeta(a_{n\delta})/\delta,$$

and a,ζ are the functions[4] appearing in the definition of $\mathcal{P}_p^*$. (Although we do not need it yet, we note that, according to (iv) in Definition 4, one has $m_n \le Kn$, where $K>0$ does not depend on δ.) Then

$$E(U_{(n+m_n)\delta}|\mathcal{F}_{n\delta}) = I(B)W_{n\delta}^{\phi} + \Theta(n\delta, m_n\delta).$$

[4]Note that since x, B are fixed we re-iterate our earlier comment that it is not necessary to indicate the dependency of ζ and α on B or the dependency of ζ and a on x.

Define the event

$$A_n := \{\mathrm{supp}(X_{n\delta}) \not\subset D_{a_{n\delta}}\}.$$

Using the second part of Definition 4(iv) along with the choice of m_n and that $I(B) < 1$, we have

$$\begin{aligned}\left|\Theta(n\delta, m_n\delta)\right| &\leq \sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-\lambda_c n\delta}\phi(X_i)\mathrm{e}^{-\lambda_c m_n\delta}\alpha_{m_n\delta} + \left|\Theta(n\delta, m_n\delta)\right|\mathbb{1}_{A_n} \\ &= \mathrm{e}^{-\lambda_c m_n\delta}\alpha_{m_n\delta}W^{\phi}_{n\delta} + \left|\Theta(n\delta, m_n\delta)\right|\mathbb{1}_{A_n}.\end{aligned}$$

Since, according to Definition 4(iii), $\lim_{n\to\infty}\mathbb{1}_{A_n} = 0$, P-a.s., therefore

$$\limsup_{n\uparrow\infty}\left|\Theta(n\delta, m_n\delta)\right| \leq \lim_{n\uparrow\infty}\mathrm{e}^{-\lambda_c m_n\delta}\alpha_{m_n\delta}W^{\phi}_{n\delta} = 0 \quad P_{\delta_x}\text{-a.s.},$$

and so

$$\lim_{n\uparrow\infty}\left|E_{\delta_x}(U_{(n+m_n)\delta}|\mathcal{F}_{n\delta}) - \langle\phi|_B, \widetilde{\phi}\,\mathrm{d}x\rangle W^{\phi}_{\infty}\right| = 0 \quad P_{\delta_x}\text{-a.s.} \tag{13}$$

Since Span$\{\phi|_B, B \subset\subset D\}$ is dense in C_c^+, the result for lattice times follows by standard arguments along with Lemma 18.

4.3. *Replacing lattice times with continuous time*

The following lemma is enough to conclude the convergence in Theorem 6 (see the remark after the lemma). It upgrades convergence along lattice times to the full sequence of times and is based on the idea to be found in Lemma 8 of [1].

Lemma 19. *Suppose that for some $p > 1$, $\langle\phi^p, \widetilde{\phi}\rangle < \infty$ and for all $\delta > 0$ it is true that for all $g \in C_c^+(D)$ and $x \in D$*

$$\lim_{n\uparrow\infty}\mathrm{e}^{-\lambda_c n\delta}\langle g, X_{n\delta}\rangle = \langle g, \widetilde{\phi}\rangle W^{\phi}_{\infty} \quad P_{\delta_x}\text{-a.s.},$$

then the same result holds when $n\delta$ is replaced by t and $\lim_{n\uparrow\infty}$ by $\lim_{t\uparrow\infty}$.

Remark 20. *Recall that we assumed that $\zeta(a_t) = \mathcal{O}(t)$ as $t \to \infty$, and so referring to the previous subsection, $m_n = \zeta(a_{n\delta})/\delta \leq Kn$ with some $K > 0$ which does not depend on δ. In fact, by possibly further increasing the function a, we can actually take $\zeta(a_t) = Kt$ and $m_n = Kn$. Then, from the previous subsection we already know that*

$$\lim_{n\uparrow\infty}\mathrm{e}^{-\lambda_c(K+1)n\delta}\langle g, X_{(K+1)n\delta}\rangle = \langle g, \widetilde{\phi}\rangle W^{\phi}_{\infty} \quad P_{\delta_x}\text{-a.s.}$$

Thus the assumption in Lemma 19 is indeed satisfied (write $\delta' := \delta(K+1)$).

Proof of Lemma 19. First suppose that $B \subset\subset D$ and for each $x \in D$ and $\varepsilon > 0$, define

$$B^{\varepsilon}(x) = \{y \in B\colon \phi(y) > (1+\varepsilon)^{-1}\phi(x)\}.$$

Note in particular that $x \in B^{\varepsilon}(x)$ if and only if $x \in B$. Next define for each $\delta > 0$

$$\varXi_B^{\delta,\varepsilon}(x) = \mathbb{1}_{\{\mathrm{supp}(X_t)\subset B^{\varepsilon}(x) \text{ for all } t\in[0,\delta]\}},$$

where $X_0 = x$, and let $\xi_B^{\delta,\varepsilon}(x) = E_{\delta_x}(\Xi_B^{\delta,\varepsilon}(x))$. An important feature of the latter quantity in the forthcoming proof is that $\xi_B^{\delta,\varepsilon}(x) \to \mathbb{1}_B(x)$ as $\delta \downarrow 0$. With this notation we now note the crucial estimate

$$\mathrm{e}^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq \frac{\mathrm{e}^{-\lambda_c \delta}}{(1+\varepsilon)} \sum_{i=1}^{N_{n\delta}} \mathrm{e}^{-\lambda_c n\delta}\phi(X_i)\Xi_i \quad \text{a.s., } t \in \big[n\delta, (n+1)\delta\big],$$

where, given $\mathcal{F}_{n\delta}$, the random variables $\{\Xi_i\colon i = 1, \dots, N_{n\delta}\}$ are independent and Ξ_i is equal in distribution to $\Xi_B^{\delta,\varepsilon}(x)$ with $x = X_i$ for $i = 1, \dots, N_{n\delta}$ respectively. Note that the sum on the right-hand side is of the form (10) where now the role of $U_{(m+n)\delta}$ is played by the right-hand side above and the role of $U_{m\delta}^{(i)}$ is played by

$$\phi(X_i)\Xi_i \cdot \mathrm{e}^{-\lambda_c \delta}.$$

Similar L^p estimates to those found in Lemma 18 show us that an estimate of the type of (12) is still valid in the setting here and hence

$$\begin{aligned}
&\sum_{n\geq 1} E_{\delta_x}\big(\big|U_{(m+n)\delta} - E(U_{(m+n)\delta}|\mathcal{F}_{n\delta})\big|^p\big) \\
&\quad \leq 2^{2p}\sum_{n\geq 1}\mathrm{e}^{-\lambda_c n\delta p} E_{\delta_x}\left(\sum_{i=1}^{N_{n\delta}} \phi(X_i)^p \xi_B^{\delta,\varepsilon}(X_i)\right).
\end{aligned}$$

However, with $q = p - 1$, the right-hand side can again be upper estimated by

$$2^{2p}\sum_{n\geq 1}\mathrm{e}^{-\lambda_c n\delta p} E_{\delta_x}\langle \phi^p, X_{n\delta}\rangle = 2^{2p}\sum_{n\geq 1}\mathrm{e}^{-\lambda_c n\delta q}\mathbb{E}_x^{\phi}\big(\phi(Y_{n\delta})^q\big) < \infty,$$

where the equality follows by Eq. (3), and the fact that the final sum is finite, follows by the ergodicity of $\mathbb{P}_x^{\phi}$ and the assumption that $\langle \phi^p, \widetilde{\phi}\rangle < \infty$.

We may now appeal to the Borel–Cantelli lemma to deduce that

$$\lim_{n\uparrow\infty}\left|\sum_{i=1}^{N_{n\delta}}\mathrm{e}^{-\lambda_c n\delta}\phi(X_i)\Xi_i - \mathrm{e}^{-\lambda_c n\delta}\langle \phi\xi_B^{\delta,\varepsilon}, X_{n\delta}\rangle\right| = 0$$

P_{δ_x}-almost surely and hence, using the fact that the Strong Law of Large Numbers has been proved already for $n\delta$-sequences,

$$\liminf_{t\uparrow\infty}\mathrm{e}^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq \frac{\mathrm{e}^{-\lambda_c\delta}}{(1+\varepsilon)}\langle \phi\xi_B^{\delta,\varepsilon}, \widetilde{\phi}\rangle W_\infty^{\phi}.$$

Since $\xi_B^{\delta,\varepsilon} \in [0, 1]$, therefore taking $\delta \downarrow 0$, by dominated convergence we have that $\langle \phi\xi_B^{\delta,\varepsilon}, \widetilde{\phi}\rangle \to \langle \phi|_B, \widetilde{\phi}\rangle$ in the lower estimate above; hence subsequently taking $\varepsilon \downarrow 0$ gives us

$$\liminf_{t\uparrow\infty}\mathrm{e}^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq \langle \phi|_B, \widetilde{\phi}\rangle W_\infty^{\phi}.$$

Recall that this estimate was computed for the case that $B \subset\subset D$. Suppose now that $B \subseteq D$ (not necessarily bounded). Then there exists an increasing sequence of compactly embedded domains in B, say $\{B_n\colon n \geq 1\}$, such that $\bigcup_{n\geq 1} B_n = B$. We may then note that for each $n \geq 1$

$$\liminf_{t\uparrow\infty}\mathrm{e}^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq \liminf_{t\uparrow\infty}\mathrm{e}^{-\lambda_c t}\langle \phi|_{B_n}, X_t\rangle \geq \langle \phi|_{B_n}, \widetilde{\phi}\rangle W_\infty^{\phi},$$

and hence, as n is arbitrary,

$$\liminf_{t\uparrow\infty} \mathrm{e}^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq \langle \phi|_B, \widetilde{\phi}\rangle W^{\phi}_{\infty}, \quad P_{\delta_x}\text{-a.s.}$$

Now that we have a tight lower estimate for the liminf for arbitrary Borel $B \subseteq D$, we shall look at the limsup, also for arbitrary Borel $B \subseteq D$. Using the normalization $\langle \phi, \widetilde{\phi}\rangle = 1$, one has

$$\limsup_{t\uparrow\infty} \mathrm{e}^{-\lambda_c t}\langle \phi|_B, X_t\rangle = W^{\phi}_{\infty} - \liminf_{t\uparrow\infty} \mathrm{e}^{-\lambda_c t}\langle \phi|_{D\setminus B}, X_t\rangle \leq \langle \phi|_B, \widetilde{\phi}\rangle W^{\phi}_{\infty}, \quad P_{\delta_x}\text{-a.s.}$$

This, together with the liminf result, yields

$$\lim_{t\uparrow\infty} \mathrm{e}^{-\lambda_c t}\langle \phi|_B, X_t\rangle = \langle \phi|_B, \widetilde{\phi}\rangle W^{\phi}_{\infty}, \quad P_{\delta_x}\text{-a.s.}$$

Then, just like for lattice times, a straightforward measure theoretical consideration shows that $\phi|_B$ can be replaced by an arbitrary test function $g \in C_c^+(D)$, completing the proof. □

4.4. *Proof of Theorem 7*

Proof. The last part of the theorem is merely a consequence of the second paragraph of the proof of Lemma 17. For any $g \in C_c^+(D)$ define for each $x \in D$ the function $h_s(x) = \mathbb{E}^{\phi}_x[g(Y_s)]$, and note that, uniformly in x and s, the function $h_s(x)$ is bounded. Now define $U_t[g] = \mathrm{e}^{-\lambda_c t}\langle g\phi, X_t\rangle$ and observe that, just as in Theorem 6, one has

$$U_{t+s}[g] = \sum_{i=1}^{N_t} \mathrm{e}^{-\lambda_c t} U^{(i)}_s[g],$$

where by (3),

$$E\big(U^{(i)}_s[g]|\mathcal{F}_t\big) = \phi\big(X_i(t)\big)h_s\big(X_i(t)\big).$$

Next, note from the Markov property at t and the proof[5] of Theorem 6 (along lattice times) that for fixed s and $x \in D$,

$$\lim_{t\uparrow\infty} \mathbb{E}_{\delta_x}\big(\big|U_{t+s}[g] - E\big(U_{t+s}[g]|\mathcal{F}_t\big)\big|^p\big) = 0$$

and hence by the monotonicity of norms

$$\lim_{t\uparrow\infty} \mathbb{E}_{\delta_x}\big(\big|U_{t+s}[g] - E\big(U_{t+s}[g]|\mathcal{F}_t\big)\big|\big) = 0. \tag{14}$$

Next, making use of the one particle picture in Eq. (3) and the spine change of measure in (8), we have that

$$\begin{aligned}
&E_{\delta_x}\big|E\big(U_{t+s}[g]|\mathcal{F}_t\big) - \langle \phi g, \widetilde{\phi}\rangle W^{\phi}_t\big| \\
&\quad\leq E_{\delta_x}\left(\sum_{i=1}^{N_t} \mathrm{e}^{-\lambda_c t}\phi\big(X_i(t)\big)\big|h_s\big(X_i(t)\big) - \langle \phi g, \widetilde{\phi}\rangle\big|\right) \\
&\quad= \phi(x)\mathbb{E}^{\phi}_x\big|h_s(Y_t) - \langle \phi g, \widetilde{\phi}\rangle\big|.
\end{aligned}$$

Hence taking limits as $t \uparrow \infty$, and using ergodicity of the spine as well as the uniform boundedness of $h_s(x)$, we have

$$\lim_{t\uparrow\infty} E_{\delta_x}\big|E\big(U_{t+s}[g]|\mathcal{F}_t\big) - \langle \phi g, \widetilde{\phi}\rangle W^{\phi}_t\big| \leq \phi(x)\big\langle\big|h_s - \langle \phi g, \widetilde{\phi}\rangle\big|, \phi\widetilde{\phi}\big\rangle.$$

---

[5] Note that even though U_t is defined differently, we still have martingale differences and the key upper estimate of $U_t \leq \text{const}\cdot W^{\phi}_t$ still holds.

Finally, noting that $\lim_{s\uparrow\infty} h_s(x) = \langle \phi g, \widetilde{\phi}\rangle$ and that the $h_s(x)$ is uniformly bounded, we have by dominated convergence that

$$\lim_{s\uparrow\infty}\lim_{t\uparrow\infty} E_{\delta_x}\left|E\left(U_{t+s}[g]|\mathcal{F}_t\right) - \langle \phi g, \widetilde{\phi}\rangle W_t^{\phi}\right| \le \phi(x)\left\langle \lim_{s\uparrow\infty}\left|h_s - \langle \phi g, \widetilde{\phi}\rangle\right|, \phi\widetilde{\phi}\right\rangle = 0. \tag{15}$$

Now recall from Lemma 17 that the martingale W^{ϕ} converges in the L^p norm and hence

$$\lim_{t\uparrow\infty} \mathbb{E}_{\delta_x}\left(\left|W_t^{\phi} - W_{\infty}^{\phi}\right|\right) = 0. \tag{16}$$

The proof is completed by a simple application of the triangle inequality together with (14)–(16) and taking $g = \kappa/\phi$ for any $\kappa \in C_c^+(D)$. □

Acknowledgments

This research has been completed during the first author's visit to Bath and was funded by an EPSRC grant (EP/E05448X/1). The generous support of EPSRC is gratefully acknowledged. The first author also expresses his thanks to the University of Bath for the hospitality during his stay.

The authors owe many thanks to Z.-Q. Chen for his valuable comments on the manuscript. Having read the manuscript of this paper, he has outlined for us a possible way to extend the methods of this paper and that of [8], in order to include a richer class in the symmetric case. In a planned future joint project, the authors of [8] and of this article will make an attempt to unify the methodologies.

Finally, we are grateful to an anonymous referee for his/her close reading of the manuscript.

References

[1] S. Asmussen and H. Hering. Strong limit theorems for general supercritical branching processes with applications to branching diffusions. *Z. Wahrsch. Verw. Gebiete* **36** (1976) 195–212. MR0420889

[2] S. Asmussen and H. Hering. Strong limit theorems for supercritical immigration-branching processes. *Math. Scand.* **39** (1977) 327–342. MR0438498

[3] K. Athreya. Change of measures for Markov chains and the LlogL theorem for branching processes. *Bernoulli* **6** (2000) 323–338. MR1748724

[4] J. Biggins. Uniform convergence in the branching random walk. *Ann. Probab.* **20** (1992) 137–151. MR1143415

[5] J. D. Biggins and A. E. Kyprianou. Measure change in multitype branching. *Adv. in Appl. Probab.* **36** (2004) 544–581. MR2058149

[6] A. Champneys, S. C. Harris, J. Toland, J. Warren and D. Williams. Algebra, analysis and probability for a coupled system of reaction-diffusion equations. *Philos. Trans. R. Soc. Lond. Ser. A* **350** (1995) 69–112. MR1325205

[7] B. Chauvin and A. Rouault. KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees. *Probab. Theory Related Fields* **80** (1988) 299–314. MR0968823

[8] Z.-Q. Chen and Y. Shiozawa. Limit theorems for branching Markov processes. *J. Funct. Anal.* **250** (2007) 374–399. MR2352485

[9] Z.-Q. Chen, Y. Ren and H. Wang. An almost sure scaling limit theorem for Dawson–Watanabe superprocesses *J. Funct. Anal.* **254** (2008) 1988–2019. MR2397881

[10] D. A. Dawson. Measure-valued Markov processes. In *Ecole d'Eté Probabilités de Saint Flour XXI* 1–260. *Lecture Notes in Math.* **1541**. Springer, Berlin, 1993. MR1242575

[11] E. B. Dynkin. *An Introduction to Branching Measure-Valued Processes*. *CRM Monograph Series* **6**. Amer. Math. Soc., Providence, RI, 1994. MR1280712

[12] J. Engländer. Branching diffusions, superdiffusions and random media. *Probab. Surv.* **4** (2007) 303–364. MR2368953

[13] J. Engländer. Law of large numbers for superdiffusions: The non-ergodic case. *Ann. Inst. H. Poincare Probab. Statist.* **45** (2009) 1–6.

[14] J. Engländer and A. Kyprianou. Local extinction versus local exponential growth for spatial branching processes. *Ann. Probab.* **32** (2003) 78–99. MR2040776

[15] J. Engländer and R. Pinsky. On the construction and support properties of measure-valued diffusions on $D \subseteq R^d$ with spatially dependent branching. *Ann. Probab.* **27** (1999) 684–730. MR1698955

[16] J. Engländer and D. Turaev. A scaling limit theorem for a class of superdiffusions. *Ann. Probab.* **30** (2002) 683–722. MR1905855

[17] J. Engländer and A. Winter. Law of large numbers for a class of superdiffusions. *Ann. Inst. H. Poincare Probab. Statist.* **42** (2006) 171–185. MR2199796

[18] A. Etheridge. *An Introduction to Superprocesses*. *University Lecture Series* **20**. Amer. Math. Soc., Providence, RI, 2000. MR1779100

[19] S. N. Evans. Two representations of a superprocess. *Proc. Roy. Soc. Edinburgh Sect. A* **123** (1993) 959–971. MR1249698

[20] Y. Git, J. W. Harris and S. C. Harris. Exponential growth rates in a typed branching diffusion. *Ann. Appl. Probab.* **17** (2007) 609–653. MR2308337

[21] R. Hardy and S. C. Harris. A conceptual approach to a path result for branching Brownian motion. *Stochastic Process Appl.* **116** (2006) 1992–2013. MR2307069

[22] R. Hardy and S. C. Harris. A spine approach to branching diffusions with applications to *Lp*-convergence of martingales. In *Séminaire de Probabilités XLII*. C. Donati-Martin, M. Émery, A. Rouault and C. Stricker (Eds). **1979**, 2009.

[23] S. C. Harris. Convergence of a "Gibbs–Boltzman" random measure for a typed branching diffusion. In *Séminaire de Probabilités XXXIV* 239–256. *Lecture Notes in Math.* **1729**. Springer, Berlin, 2000. MR1768067

[24] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, 2nd edition. *Grundlehren der Mathematischen Wissenschaften* **288**. Springer, Berlin, 2003. MR1943877

[25] O. Kallenberg. Stability of critical cluster fields. *Math. Nachr.* **77** (1977) 7–43. MR0443078

[26] R. Lyons, R. Pemantle and Y. Peres. Conceptual proofs of $L \log L$ criteria for mean behaviour of branching processes. *Ann. Probab.* **23** (1995) 1125–1138. MR1349164

[27] R. G. Pinsky. *Positive Harmonic Functions and Diffusion*. Cambridge Univ. Press, Cambridge, 1995. MR1326606

[28] R. G. Pinsky. Transience, recurrence and local extinction properties of the support for supercritical finite measure-valued diffusions. *Ann. Probab.* **24** (1996) 237–267. MR1387634

[29] S. Watanabe. A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.* **8** (1968) 141–167. MR0237008