LEBESGUE APPROXIMATION OF SUPERPROCESSES

by

Xin He

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 03, 2013

Keywords: Super-Brownian motion, measure-valued branching processes,
deterministic distribution properties

Approved by

Olav Kallenberg, Chair, Professor of Mathematics
Ming Liao, Professor of Mathematics
Jerzy Szulga, Professor of Mathematics
Erkan Nane, Associate Professor of Mathematics

Abstract

Superprocesses are certain measure-valued Markov processes, whose distributions can be characterized by two components: the branching mechanism and the spatial motion. It is well known that some basic superprocesses are scaling limits of various random spatially distributed systems near criticality.

We consider the Lebesgue approximation of superprocesses. The Lebesgue approximation means that the processes at a fixed time can be approximated by suitably normalized restrictions of Lebesgue measure to the small neighborhoods of their support. From this, we see that the processes distribute their mass over their support in a deterministic and "uniform" manner. It is known that the Lebesgue approximation holds for the most basic Dawson–Watanabe superprocesses but fails for certain superprocesses with discontinuous spatial motion.

In this dissertation we first prove that the Lebesgue approximation holds for superprocesses with Brownian spatial motion and a stable branching mechanism. Then we generalize the Lebesgue approximation even further to superprocesses with Brownian spatial motion and a regularly varying branching mechanism. We believe that the Lebesgue approximation holds for superprocesses with Brownian spatial motion and any "reasonable" branching mechanism. Our present results may be regarded as some progress towards a complete proof of this very general conjecture.

Acknowledgments

I owe my deepest gratitude to my advisor Dr. Olav Kallenberg, who has put lots of his time and effort in trying to develop me into a proper researcher in probability. His taste, style, and work ethic have all influenced me immensely, and they will continue to influence me in the future. I also sincerely thank Dr. Ming Liao, Dr. Jerzy Szulga, and Dr. Erkan Nane for serving in my committee and providing me much help when needed. Finally, I am grateful for the moral support of my family: My parents Anqing He and Ruizhen Hao, my sister Ni He, and my wife Yajie Chu. I am extremely lucky to have them and I simply can not imagine my life without them.

Table of Contents

List of Notation

δ_x	Dirac measure at x
$\overset{=}{\frown}$	equality up to a constant factor
$\hat{\mathcal{M}}_d$	space of finite measures on R^d
λ^d	Lebesgue measure on R^d
$\lesssim$	inequality up to a constant factor
μf	integral of function f with respect to measure μ
μ^ε	neighborhood measure of μ, defined as the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp}\mu$
$\overline{A}$	closure of set A
$\Vert f \Vert$	supremum norm of function f
$\Vert \mu \Vert$	total variation of measure μ
$\Phi = \Phi(v)$	branching mechanism
$\asymp$	the combination of $\lesssim$ and $\gtrsim$
$\xi = (\xi_t)$	superprocess or discrete spatial branching process
A^ε	ε-neighborhood of set A, $A^\varepsilon = \{x : d(x, A) < \varepsilon\}$
A^c	compliment of set A
B_x^r	an open ball around x of radius r

$f \approx g$	$f - g \to 0$
$f \sim g$	$f/g \to 1$
$\mathcal{M}_d$	space of σ-finite measures on R^d
$\stackrel{v}{\to}$	vague convergence of measures
$\stackrel{w}{\to}$	weak convergence of measures

Chapter 1

Introduction

1.1 A short introduction to superprocesses

In this section we give a short introduction to superprocesses. Three characterizations of superprocesses will be given. They are the Laplace functional approach, the weak convergence approach, and the martingale problem approach. Superprocesses were introduced by Watanabe [49] in 1968 and Dawson [3] in 1975, and have been studied extensively ever since. General surveys of superprocesses include the following excellent monographs and lecture notes: Dawson [4, 5], Dynkin [15, 16], Etheridge [17], Le Gall [32], Li [35], and Perkins [42]. Two extremely informative yet concise and very accessible introductions of superprocesses are Perkins [43] and Slade [46].

First let us explain the two defining components of a superprocess: the branching mechanism and the spatial motion. We begin with branching processes, which contain only one component of superprocesses: the branching mechanism. Galton-Watson processes are discrete branching processes. They describe the evolution in discrete time of a population of individuals who reproduce according to an *offspring distribution*, which is a probability measure on the nonnegative integers with expectation 1 (we only consider the critical case in this introduction). The distribution of a Galton-Watson process is determined by this offspring distribution. Continuous-state branching processes are continuous analogues of the Galton-Watson branching processes. Roughly speaking, they describe the evolution in continuous time of a *"population"* with values in the positive real line R_+. The "population" consists of uncountably many *"individuals"*, if its value is not 0. The distribution of a continuous-state branching process is determined by a function Φ of the following type (again, we only consider the critical case in this introduction, so no drift term here)

$$\Phi(v) = av^2 + \int_0^\infty (e^{-rv} - 1 + rv)\pi(dr) \tag{1.1}$$

where $a \geq 0$ and π is a σ-finite measure on $(0, \infty)$ such that $\int_0^\infty (r \wedge r^2)\pi(dr) < \infty$. This function Φ is called the *branching mechanism*. Continuous-state branching processes may also be obtained as weak limits of rescaled Galton-Watson processes, see (1.7). This is closely related to the weak convergence approach to superprocesses, see (1.8).

Spatial branching processes are obtained by combining the branching phenomenon with a *spatial motion*, which is usually given by a Markov process X. In the discrete setting, the branching phenomenon is a Galton-Watson process, and the individuals move independently in space according to the law of X. More precisely, when an individual dies at position x, her children begin to move from the initial point x, and they move in space independently according to the law of X. Writing $Y_t^1, Y_t^2, \ldots$ for the positions of all individuals alive at time t, we may define

$$\xi_t = \sum_i \delta_{Y_t^i} \tag{1.2}$$

where δ_y denotes the Dirac measure at y. The process $\xi = (\xi_t, t \geq 0)$ is the spatial branching process corresponding to the branching phenomenon of a Galton-Watson process and the spatial motion X. Note that this is a measure-valued process, whose value at time t records the positions of all individuals alive at time t.

In the continuous setting, the branching phenomenon is a continuous-state branching process with branching mechanism Φ. The construction of the spatial motions is harder, and so here we proceed only heuristically. For mathematical support of these heuristics, refer to the weak convergence approach later in this section (see (1.7) and (1.8)), cluster representation in Section 2.2, and historical superprocesses and random snakes in Section 2.3. Here we let the "individuals" move independently in space according to the law of a Markov process X. Thus when an "individual" dies at the position x, her "children" begin to move from the initial point x, and they move in space independently according to the law of X. Again we get a measure-valued process $\xi = (\xi_t, t \geq 0)$, whose value at time t records the positions of all "individuals" alive at time t. This measure-valued process $\xi = (\xi_t)$ is called the (X, Φ)–*superprocess* (or (X, Φ)-*process*, for short).

Superprocesses are measure-valued Markov processes. We first use the Laplace functional approach to characterize their distributions. For a (X, Φ)-process on R^d, the spatial motion X is a Markov process in R^d. Use μf to denote the integral of the function f with respect to the measure μ. Write $P_\mu(\xi \in \cdot)$ for the distribution of the process ξ with initial measure μ, and E_μ for the expectation corresponding to P_μ. The Laplace functional $E_\mu \exp(-\xi_t f)$ satisfies

$$E_\mu[\exp(-\xi_t f)|\xi_s] = \exp(-\xi_s v_{t-s}) \tag{1.3}$$

where $\big(v_t(x), t \geq 0, x \in \mathsf{R}^d\big)$ is the unique nonnegative solution of the integral equation

$$v_t(x) + \Pi_x \left(\int_0^t \Phi\left(v_{t-s}(X_s)\right) ds \right) = \Pi_x\left(f(X_t)\right) \tag{1.4}$$

Here we write $\Pi_x(X \in \cdot)$ for the distribution of the process X starting from x. If X is a Feller process in R^d with generator L, the integral equation (1.4) is the integral form of the following PDE, the so-called *evolution equation*

$$\dot{v} = Lv - \Phi(v) \tag{1.5}$$

with initial condition $v_0 = f$. More explicitly, PDE (1.5) means

$$\frac{\partial v_t}{\partial t}(x) = (Lv_t)(x) - \Phi(v_t(x)).$$

For the equivalence of the integral equation (1.4) and the differential equation (1.5), see Section 7.1 in [35]. If X is a rotation invariant (or spherically symmetric, or isotropic) α-stable Lévy process in R^d for some $\alpha \in (0, 2]$ (see Definition 14.12 and Theorem 14.14 in

[45]) and $\Phi(v) = v^{1+\beta}$ for some $\beta \in (0, 1]$, we get a superprocess corresponding to the PDE

$$\dot{v} = \frac{1}{2}\gamma\Delta_\alpha v - v^{1+\beta}$$

where $\frac{1}{2}\gamma\Delta_\alpha$ is the generator of the rotation invariant α-stable process X (see Theorem 19.10 in [24]), and $\Delta_\alpha = -(-\Delta)^{\alpha/2}$ is the fractional Laplacian ($\Delta_2 = \Delta$ is the Laplacian, see Section 2.6 in [39]). Taking $\gamma = 1$ in the above PDE, we get a superprocess corresponding to the PDE

$$\dot{v} = \frac{1}{2}\Delta_\alpha v - v^{1+\beta}. \tag{1.6}$$

We call it the (α, β)-*superprocess* ((α, β)-*process* for short). For the most basic and most important superprocess, we take $\alpha = 2$ and $\beta = 1$ to get a $(2, 1)$-process, which is often called the *Dawson–Watanabe superprocess* (*DW-process* for short). Clearly a DW-process has Brownian spatial motion and branching mechanism $\Phi(v) = v^2$. We may abuse the notation further by referring to (α, Φ)-processes and (X, β)-processes. Specifically, an (α, Φ)-process has rotation invariant α-stable spatial motion and branching mechanism Φ, and an (X, β)-process has spatial motion X and branching mechanism $\Phi(v) = v^{1+\beta}$.

Next we move to the weak convergence approach, which is the most intuitive way to define superprocesses. Just as continuous-state branching processes may be obtained as weak limits of rescaled Galton-Watson processes (see (1.7)), superprocesses can be obtained as weak limits of rescaled discrete spatial branching processes (see (1.8)). Recall that we can get intuition about Brownian motion from rescaled random walks, similarly here we may get some intuition about superprocesses from rescaled discrete spatial branching processes.

We consider a sequence N^k, $k \geq 1$ of Galton-Watson processes such that as $k \to \infty$,

$$\left(\frac{1}{a_k} N^k_{[kt]}, t \geq 0\right) \xrightarrow{fd} (Z_t, t \geq 0) \tag{1.7}$$

where constants $a_k \uparrow \infty$, Z is a continuous-state branching process with branching mechanism Φ, and the symbol $\xrightarrow{fd}$ means weak convergence of finite-dimensional marginals. Then, according to (1.2), we consider a sequence ξ^k, $k \geq 1$ of spatial branching processes corresponding to the Galton-Watson processes N^k, $k \geq 1$ and the spatial motion X. Clearly ξ^k_t is a random element with values in the space of finite measures on R^d, equipped with the topology of weak convergence. Now, according to (1.7), we consider a sequence of rescaled spatial branching processes $\frac{1}{a_k}\xi^k_{[k\cdot]}$, $k \geq 1$. Suppose that the initial measures converge as $k \to \infty$ ($\stackrel{w}{\to}$ denotes weak convergence) :

$$\frac{1}{a_k}\xi^k_0 \stackrel{w}{\to} \mu,$$

where μ is a finite measure on R^d. Finally, under adequate regularity assumptions on the spatial motion X, there exists a measure-valued Markov process ξ such that

$$(\frac{1}{a_k}\xi^k_{[kt]}, t \geq 0) \xrightarrow{fd} (\xi_t, t \geq 0), \tag{1.8}$$

where ξ is an (X, Φ)-process with initial measure μ.

Finally, superprocesses can also be characterized as solutions to martingale problems. Chapter 7 in [35] is an excellent reference on martingale problems of very general superprocesses. We first discuss a martingale problem of $(X, 1)$-processes, where X is a Feller process in R^d with generator L. Write $\hat{\mathcal{M}}_d$ for the space of finite measures on R^d. Then write $(D([0,\infty), \hat{\mathcal{M}}_d), \xi_t, \mathcal{F}_t)$ for the space of rcll $\hat{\mathcal{M}}_d$-valued paths, the coordinate process, and the canonical completed right continuous filtration. For any $f \in D(L)$ (domain of generator L), define the process $M_t(f)$ by

$$M_t(f) = \xi_t f - \xi_0 f - \int_0^t \xi_s(Lf)ds. \tag{1.9}$$

For any $\mu \in \hat{\mathcal{M}}_d$, use $\mathcal{L}_\mu$ to denote the distribution of an $(X,1)$-process with initial measure μ. This is the unique distribution on $\mathcal{F} = \sigma(\bigcup_{t\geq 0} \mathcal{F}_t)$ such that the coordinate process satisfies the following martingale problem: $\xi_0 = \mu$, *and for any* $f \in D(L)$, *the process* $M_t(f)$ *defined in (1.9) is a continuous martingale with quadratic variation process*

$$[M(f)]_t = \int_0^t \xi_s(f^2)ds.$$

For a (X,Φ)-process, the corresponding martingale is not continuous in general. In this case, we may split the martingale into two parts: the continuous martingale $M_t^c(f)$ and the purely discontinuous martingale $M_t^d(f)$ (see Theorem 26.14 in [24]). Then we write

$$M_t(f) = M_t^c(f) + M_t^d(f) = \xi_t f - \xi_0 f - \int_0^t \xi_s(Lf)ds,$$

where $M_t^c(f)$ is a continuous martingale with quadratic variation process

$$[M^c(f)]_t = \int_0^t \xi_s(af^2)ds, \tag{1.10}$$

and $M_t^d(f)$ is a purely discontinuous martingale, which can be defined through a compensated random measure relating to the jumps of ξ. For details, see Section 7.2 in [35]. Note that the jumps of ξ are related to the measure π in the branching mechanisam Φ of (1.1), not the jumps of the spatial motion X (see Section 2.6 in [17]). We may also note that the continuous martingale $M_t^c(f)$ is related to the term av^2 in the branching mechanisam Φ of (1.1) through its quadratic variation process $[M^c(f)]_t$ in (1.10).

1.2 Summary of contents

The purpose of this dissertation is to discuss the Lebesgue approximation of superprocesses in details. In Section 1 we discussed the definitions of superprocesses. Here we give a summary of the contents of the following chapters.

In Chapter 2, we discuss some basic ingredients of superprocesses in the first three sections, which are crucial for the Lebesgue approximation of superprocesses. Then we discuss the background of Lebesgue approximation and some related known results in the last two sections. In Section 1 we discuss the first moment measure $E_\mu\xi_t$ and the second moment measure $E_\mu\xi_t^2$. In particular, the second moment measure does not exist in general, which causes a real difficulty for generalizing certain results. In Section 2 we discuss the very important cluster representation of superprocesses, which contains partial information of the whole genealogical evolution underlying superprocesses. This cluster representation is transparent in the discrete setting, however in the continuous setting it is not easy at all to obtain it rigorously. In Section 3 we discuss two approaches to encode the genealogical information and to obtain the cluster representation. They are Historical superprocesses approach and random snakes approach. In Section 4 we discuss some classical results about the Hausdorff dimensions and Hausdorff measures of superprocesses. The point is that the Hausdorff measure approach is a more traditional, more successful way to do what the Lebesgue approximation approach tries to do: Construct nontrivial measures on some random null sets. Finally in Section 5 we discuss basic ideas of Lebesgue approximation and review almost all known Lebesgue approximation results. At the end of this section we also discuss some related open problems.

In Chapter 3, we discuss the Lebesgue approximation of Dawson-Watanabe superprocesses of dimension $d \geq 3$, which is the most basic and most transparent case. This chapter is based on Kallenberg's proof of Lebesgue approximation of DW-processes of dimension $d \geq 3$ in [25], with some technical simplifications. Note that Tribe first proved this result in [48]. Extra efforts haae been made to explain Kallenberg's approach clearly and to make it more accessible. In Section 1 we explain some crucial components in the proof and review some terminology and notation. In Section 2 we first explain the crucial ideas about cluster representations, then state several lemmas which will not be used directly in the main proof of Lebesgue approximation, including the important upper bound of the hitting multiplicities. In Section 3 we state and prove the Lebesgue approximation for DW-processes of dimensions $d \geq 3$. In order to do so, we list several lemmas that are needed in the main proof. Finally, in Section 4, we prove all the lemmas in this chapter. We suggest that the reader read the first three sections in the linear order, then, when need arises, read the proofs of some lemmas in Section 4.

In Chapter 4, we discuss the Lebesgue approximation of $(2, \beta)$-superprocesses of dimension $d > 2/\beta$. This chapter is based on my 2013 paper [22]. In Section 1 we explain the additional difficulties for the Lebesgue approximation of $(2, \beta)$-processes and review our general approach, which overcomes these difficulties. In Section 2 we develop further a truncation of (α, β)-processes from [38]. We also characterize the local finiteness of any (α, β)-superprocess, which can be used to extend certain results to some superprocesses with σ-finite initial measures. In Section 3, we develop some lemmas about hitting bounds and neighborhood measures of $(2, \beta)$-processes, in particular, we improve the upper bounds of hitting probabilities. In Section 4, we derive some asymptotic results of these hitting probabilities. In particular, for the $(2, \beta)$-superprocess ξ we show that $\varepsilon^{2/\beta - d} P_\mu\{\xi_t B_x^\varepsilon > 0\} \to c_{\beta,d}\,(\mu * p_t)(x)$, which extends the corresponding result for DW-processes. Finally in Section 5 we state and prove the Lebesgue approximation of $(2, \beta)$-processes and their truncated processes. Whenever one feels the lack of details of some results in this chapter, refer back to appropriated places in Chapter 3.

In Chapter 5, we discuss the Lebesgue approximation of superprocesses with a regularly varying branching mechanism. The branching mechanisms we consider here include the stable branching mechanisms considered in Chapter 4 as special cases. In Section 5.1 we explain the new difficulties for the Lebesgue approximation of superprocesses with the more general branching mechanism and review our general approach, which overcomes these difficulties. In Section 2 we review the truncation of superprocesses in a more general setting. In Section 3, we develop some lemmas about hitting bounds and neighborhood measures of the more general superprocesses. In Section 4, we derive some asymptotic results of these hitting probabilities. Finally in Section 5 we state and prove the Lebesgue approximation of superprocesses with a regularly varying branching mechanism and their truncated processes. This general result contains all previous Lebesgue approximation of superprocesses as special cases.

Chapter 2

Some Basics of Superprocesses

2.1 Moment measures

Moment measures play an important role in the study of superprocesses. For the Markov process X, write $T_t f(x) = \Pi_x(f(X_t))$ for the semigroup of X, where $\Pi_x(X \in \cdot)$ denotes the distribution of the process X starting from x. Then the first moment measure of the (X, Φ)-process ξ (see (1.3) and (1.4)) is

$$E_\mu(\xi_t f) = \mu(T_t f). \tag{2.1}$$

Note that the branching mechanism Φ of (1.1) plays no role here. Write $p_t^\alpha(x)$ for the transition density of the rotation invariant α-stable Lévy process with generator $\frac{1}{2}\Delta_\alpha$ (see (1.6)). Then the first moment measure of the (α, Φ)-process ξ takes the equivalent measure form

$$E_\mu \xi_t = (\mu * p_t^\alpha) \cdot \lambda^d,$$

where $\mu * p_t^\alpha(x) = \int p_t^\alpha(x-y)\mu(dy)$ and $f \cdot \lambda^d$ denotes the measure defined by $(f \cdot \lambda^d)(B) = \int_B f d\lambda^d$.

The second moment measure depends crucially on the branching mechanism. In fact, second moments do not exist in general. However, they do exist when the measure $\pi = 0$ in the branching mechanism Φ of (1.1), that is, for the $(X, \gamma v^2)$-process ξ. The second moment measure of the $(X, \gamma v^2)$-process ξ is

$$E_\mu(\xi_t f)^2 = (\mu(T_t f))^2 + 2\gamma \int_0^t \mu\left(T_s(T_{t-s}f)^2\right) ds. \tag{2.2}$$

Refer to Section 2.4 in [32] for the proofs of (2.1) and (2.2). For the (X, β)-process ξ with $\beta < 1$, only moments of order less than $1 + \beta$ exist. A useful inequality along this line is Lemma 2.1 in [37]: For $0 < \theta < \beta < 1$,

$$E_\mu(\xi_t f)^{1+\theta} \leq 1 + c(\theta) \left[(\mu(T_t f))^{1+\beta} + \int_0^t \mu \left(T_s (T_{t-s} f)^{1+\beta} \right) ds \right],$$

where $c(\theta) \to \infty$ as $\theta \to \beta$. When we need to use the second moments, we may truncate ξ at any level $K > 0$ to get the truncated process ξ^K, which has finite second moments. For details about this truncation method, see pages 484 - 487 and Lemma 3 in [38].

Using series expansions of Laplace functionals, Dynkin [13] gives moment measure formulas for very general superprocesses. See Section 14.7 in [16] for a concise review of these formulas. Finally we mention that, for DW-processes, Theorem 4.2 of Kallenberg [27] contains a basic cluster decomposition of moment measures. Theorem 4.4 of that paper gives a fundamental connection between moment measures and certain uniform Brownian trees, first noted by Etheridge in Section 2.1 of [17]. It would be interesting to study this connection for more general superprocesses. For details about the cluster decomposition of moment measures, See Theorem 5.1 in Kallenberg [26].

2.2 Cluster representation

In this section we discuss the very important cluster representation of superprocesses. Note that although a superprocess records the positions of all "individuals" alive at time t, they do not keep track of all the genealogy of these "individuals". More precisely, let us pick an "individual" alive at time t, then try to identify her "ancestor" at an earlier time s. Although we know from ξ_s the positions of all "individuals" alive at time s, we don't know which specific "individual" at time s is the "ancestor" of the "individual" we picked at time t. However, in the study of some deep properties of superprocesses, the genealogical structure underlying the evolution can be extremely useful, even when the final results have nothing directly to do with the genealogy. The cluster representation of superprocesses, while containing only partial information of the genealogy, is enough for many purposes.

In order to discuss the cluster representation, let us first recall the definition of Poisson cluster processes. To define a cluster process, we start with a point process $\zeta = \sum_i \delta_{\tau_i}$ on some space T. For a suitable class $\mathcal{M}_S$ of measures on S, we consider a probability kernel ν from T to $\mathcal{M}_S$. Choosing the random measures η_i to be conditionally independent of the τ_i with distributions ν_{τ_i}, we may introduce a random measure $\xi = \sum_i \eta_i$ on S. This random measure ξ is called a ν-cluster process generated by ζ. If ζ is Poisson or Cox, we call ξ a Poisson or Cox cluster process.

Due to the underlying independence structure, superprocesses have the following branching property: *If ξ and ξ' are two independent (X, Φ)-processes with initial measures μ and μ' respectively, then $\xi + \xi'$ is an (X, Φ)-process with initial measure $\mu + \mu'$.* This can be verified by using any of the three characterizations in Section 1.1. From this branching property, we see that, for any t, the superprocess ξ_t is an infinitely divisible random measure. A random measure is infinitely divisible iff it is the sum of a Poisson cluster process and a deterministic measure (see Theorem 1.28 in [17]). Since $P\{\xi_t = 0\} > 0$, the superprocess ξ_t is just a Poisson cluster process.

The cluster representation of (X, Φ)-processes depends crucially on the branching mechanism Φ of (1.1). For convenience, we first discuss the cluster representation of $(X, 1)$-processes (see Section 3.2 and 6.1 in [17]). For a $(X, 1)$-process ξ, at time 0, there are actually uncountably many "individuals". All "individuals" produce "offspring" randomly. However almost all "individuals" have no "offspring" alive at time $t > 0$, except finitely many "lucky" ones. In other words, the superprocess at time t is actually "offspring" of finitely many "ancestors". The point process records the locations of these finite many "ancestors" is a Poisson process ζ_0 with intensity measure $t^{-1}\mu$. This is the generating process in the Poisson cluster representation of ξ_t. Each one of these finitely many "ancestors" generates a random cluster at time t. Clearly this cluster is just her "offspring" at time t. These clusters are "the same", means that they have the same distribution if we move their "ancestors" to a common point. In summary, ξ_t being a Poisson cluster process, is a finite sum of conditionally independent clusters, equally distributed apart from shifts and rooted at the points of a Poisson process ζ_0 of "ancestors" with intensity measure $t^{-1}\mu$. By the Markov property of ξ, we have a similar representation of ξ_t for every $s = t - h \in (0, t)$ as a countable sum of conditionally independent h-clusters (clusters of age h), rooted at the points of a Cox process ζ_s directed by $h^{-1}\xi_s$. In other words, ζ_s is conditionally Poisson given ξ_s with intensity measure $h^{-1}\xi_s$ (see page 226 in [24]).

Under some restrictions of the branching mechanism Φ of (1.1), (X, Φ)-processes also have a similar cluster representation (see Section 11.5 in [5] and Section 3 in [7]). The function t^{-1} in the above intensity measure $t^{-1}\mu$ should be replaced by another function of t, determined by the branching mechanism Φ. The cluster distributions are also different, determined by both X and Φ.

2.3 Historical superprocesses and random snakes

It is clear that the cluster representation of the previous section cannot be recovered from the superprocess ξ itself, since ξ_t records only the positions of all "individuals" alive at time t. A complete picture of the evolution underlying a superprocess is given by a random tree composed from the paths of all individuals. Two approaches to encode this picture are provided by historical superprocesses and by random snakes. Both approaches can be used to verify the cluster representation.

The basic idea of historical superprocesses is very simple (see Section 1.9 in [17]). Let us explain the idea in the discrete setting, to make it even more transparent. For a discrete spatial branching process ξ, pick two individuals alive at time $t > 0$, and assume that they have their last common ancestor at time $s \in (0, t)$. Based on the Markov properties of the spatial motion X and the independence structures of spatial branching processes, clearly we can think that these two individuals perform the spatial motion X together as a single individual before time s, then separate at time s and begin to perform independent spatial motion X ever since. In other words, we can think of these two individuals as a single path before time s, and this path splits into two independent paths at time s. The same idea still holds in the continuous setting, that is, for the superprocesses. Note that in the construction of superprocesses, the spatial motion X_t is only the location of an individual at time t. In order to remember the spatial locations of all the members in her genealogy line before her, we may just replace X_t by the corresponding path process $\hat{X}_t$, which is a path-valued process. The value of $\hat{X}_t$ is the path of X over the time interval $[0,t]$. Now we construct the $(\hat{X},\Phi)$-superprocess $\hat{\xi}$, which corresponds to the (X,Φ)-superprocess ξ. We call $\hat{\xi}$ the (X,Φ)-historical superprocess. Note that the way we define $\hat{\xi}$ from ξ is different from the naive way we define $\hat{X}$ from X. If we define $\hat{\xi}$ as the corresponding path process of ξ, then we would not be able to specify the ancestors of any individual.

Denote the space of all rcll paths over the time interval $[0,t]$ by $\mathcal{W}_t$. Then the state space of $\hat{X}_t$ is $\mathcal{W}_t$, and so $\hat{X}_t$ is a time-inhomogeneous Markov process. Write $\Pi_{r,w}$ for the probability measure under which $\hat{X}$ starts from the path w at time r. Clearly w is an rcll path over the time interval $[0,r]$. Let $\mathcal{M}_t$ be the space of all finite measures on $\mathcal{W}_t$. This is the state space of $\hat{\xi}_t$, and so $\hat{\xi}_t$ is also a time-inhomogeneous Markov process. The (X,Φ)-historical superprocess $\hat{\xi}$ can be characterized by all three approaches in Section 1.1. It is obvious how to carry out the weak convergence approach. For the Laplace functional approach, we have

$$E_{r,\mu}e^{-\hat{\xi}_t f} = e^{-\mu v_t^r},$$

where μ is a finite measure on $\mathcal{W}_r$, $\hat{\xi}_t$ is a finite measure on $\mathcal{W}_t$, f is a function on $\mathcal{W}_t$, and v_t^r is a function on $\mathcal{W}_r$. The function $v_t^r(w)$ with $r \le t$ and $w \in \mathcal{W}_r$ is uniquely determined by the integral equation

$$v_t^r(w) + \Pi_{r,w}\left(\int_r^t \Phi\left(v_t^s(\hat{X}_s)\right) ds\right) = \Pi_{r,w}\left(f(\hat{X}_t)\right).$$

This may be compared with (1.4), the integral equation of ξ. The only difference is that $\hat{\xi}$ is a superprocess of historical paths, while ξ is a superprocess of spatial positions.

The concept of historical superprocesses was developed in Dawson and Perkins [7] and Dynkin [14]. We may also refer to Chapter 12 in [5] and Section II.8 in [42].

From the construction of historical superprocesses it is clear that $\hat{\xi}_t$ encodes the genealogy information of all "individuals" alive at time t.

The random snake approach developed by Le Gall and his co-authors allows to give a complete description of the genealogy. Here we only focus on the basic ideas, since the technical details and notation can be overwhelming.

The basic idea of random snakes stems from an important fact of branching processes: *The genealogical structures of branching processes can be completely encoded by a (random) function on* R_+. More precisely, the genealogical structure of Galton-Watson processes can be completely encoded by a discrete (random) function defined on nonnegative integers. These nonnegative integers correspond to all the individuals and the function values are the *generations* of these individuals. For continuous-state branching processes, similarly the genealogical structure can be completely encoded by a continuous (random) function on R_+. Again, any number in R_+ corresponds to an "individual", and the function value is the "generation" or *lifetime* of this "individual". This continuous coding function is called the *lifetime process*. In fact, continuous coding function can be obtained from a sequence of rescaled discrete coding functions. Clearly this is closed related to the fact that continuous-state branching processes may be obtained as weak limits of rescaled Galton-Watson processes. For a continuous-state branching process with the branching mechanism $\Phi = v^2$, the lifetime process ς_t is actually just the reflected one dimensional Brownian motion. The time parameter of the lifetime process ς_t is a labeling of all individuals in a certain order.

For the complete evolution of $(X, 1)$-processes, we then need to somehow combine the paths of individuals with this coding continuous random function. This is done by the so called *Brownian snake* W_t, which is a path-valued Markov process evolving according to both the spatial motion X and the lifetime process ς_t. Note that the term "Brownian" refers to the branching mechanism, actually the lifetime process, not the spatial motion. The behavior of the Brownian snake is actually not hard to explain, at least informally. The value W_t at time t of the Brownian snake is a path of the underlying spatial motion X (started at a fixed initial point) with the random lifetime ς_t. Informally, when ς_t decreases, the path W_t is shortened from its tip, and when ς_t increases, the path W_t is extended by adding (independently of the past) small "pieces of paths" following the law of the spatial motion X. In this way, we can generate the full set of historical paths of a $(X, 1)$-process by running the Brownian snake according to the lifetime process, in this way we are visiting all the "individuals" one by one.

For superprocesses with a general branching mechanism Φ, similarly the so called *Lévy snakes* can be defined. The basic ideas are similar, but technically it is much more complicated. The main reason is that the corresponding lifetime process is not Markov and its definition is quite involved. Actually part of the beauty, and the power, of the Brownian snake is that the lifetime process is itself a Markov process. The standard reference of Brownian snake is the excellent lecture notes [32] by Le Gall in 1999. For Lévy snakes, refer to the excellent monograph [12] by Duquesne and Le Gall in 2002.

2.4 Hausdorff dimensions and Hausdorff measures

In this section we review some classical results about the Hausdorff dimensions and Hausdorff measures of superprocesses. First let us review the definitions of Hausdorff dimension and Hausdorff measure. For a nice introduction of this topic in a probabilistic setting, see Chapter 4 and Section 6.4 in [36].

We first define Hausdorff measure, then Hausdorff dimension. Assume A to be a metric space with the metric ρ. Use $|A|$ to denote the diameter of the set A, which is defined by

$$|A| = \sup\{\rho(x, y) : x, y \in A\}.$$

For every $\alpha \geq 0$ define

$$\mathcal{H}^\alpha_\delta(A) = \inf\{\sum_{i=1}^{\infty} |A_i|^\alpha : A \subset \bigcup_{i=1}^{\infty} A_i, |A_i| \leq \delta \text{ for all } i\}. \tag{2.3}$$

Easy to see that the quantity $\mathcal{H}^\alpha_\delta(A)$ is increasing as δ decreases, so that the limit

$$\mathcal{H}^\alpha(A) = \lim_{\delta\to 0} \mathcal{H}^\alpha_\delta(A)$$

is well-defined, although it could be infinite. We call the limit $\mathcal{H}^\alpha(A) = \lim_{\delta\to 0} \mathcal{H}^\alpha_\delta(A)$ the *α-Hausdorff measure* of A.

Since subsets of a metric space are metric spaces on their own, the α-Hausdorff measure $\mathcal{H}^\alpha$ can be defined for all subsets of the space A. Using the definition of $\mathcal{H}^\alpha$, we can check that the function $\mathcal{H}^\alpha$ defined for all subsets satisfies all the properties of a metric Carathéodory exterior measure (see Section 7.1 in [47], or Section 11.2 in [20]). Thus $\mathcal{H}^\alpha$ is a countably additive measure when restricted to the Borel sets of A. So indeed, the α-Hausdorff measure defined on all Borel sets is a measure.

Let us define the Hausdorff dimension. The α-Hausdorff measure $\mathcal{H}^\alpha(A)$ has the following natural properties: If $0 \leq \alpha < \beta$, and $\mathcal{H}^\alpha(A) < \infty$, then $\mathcal{H}^\beta(A) = 0$; If $0 \leq \alpha < \beta$, and $\mathcal{H}^\beta(A) > 0$, then $\mathcal{H}^\alpha(A) = \infty$. So there exists a unique number which is denoted by $\dim A$ such that $\mathcal{H}^\alpha(A) = \infty$ for $\alpha < \dim A$, and $\mathcal{H}^\alpha(A) = 0$ for $\alpha > \dim A$. We call this unique number the *Hausdorff dimension* of the set A, denoted by $\dim A$. Or in other words, we define the Hausdorff dimension of the set A by

$$\dim A = \sup\{\alpha : \mathcal{H}^\beta(A) = \infty\} = \inf\{\alpha : \mathcal{H}^\beta(A) = 0\}.$$

Using the Hausdorff dimension, we can associate a nonnegative number to any set, which generalizes the usual integer dimensions. For example, the classical Cantor set has Hausdorff dimension $\log 2/\log 3$. The graph of a one dimensional Brownian motion, which is a continuous (random) curve on R^2, has Hausdorff dimension 3/2 a.s. (see Theorem 16.4 in [19]). This is related to the fact that one dimensional Brownian path is a.s. locally Hölder continuous with exponent c for any $c \in (0, \frac{1}{2})$.

Now we turn to superprocesses. For a DW-process ξ in R^d, we denote the support of ξ_t by $\operatorname{supp} \xi_t$, which is a random closed set in R^d. Actually this is even a random compact set, assuming $\xi_0 = \mu$ is a finite measure (see Theorem 1.2 in [6]). For fixed $t > 0$, if $d \geq 2$, then a.s. this is a null set (means that it has Lebesgue measure 0). Here Hausdorff dimension is useful for us to get some more understanding of the size of $\operatorname{supp} \xi_t$. It is well-known that a.s.

$$\dim (\operatorname{supp} \xi_t) = 2 \wedge d, \quad \text{on } \{\xi_t \neq 0\}.$$

Note that if $\xi_t = 0$, then $\operatorname{supp} \xi_t = \emptyset$. More generally, if ξ is a $(2, \beta)$-process in R^d with $\beta \in (0, 1]$, then for fixed $t > 0$, if $d \geq 2/\beta$, a.s. $\operatorname{supp} \xi_t$ is a null set (again, this is a random compact set, adapt the proof of Theorem 1.2 in [6] to generalize Theorem 1.1 in [8], or see Section 4.3 in [1]) and

$$\dim (\operatorname{supp} \xi_t) = (2/\beta) \wedge d, \quad \text{on } \{\xi_t \neq 0\}.$$

For this result and even more, see Theorem 2.1 in [9]. The situation for $(2, \beta)$-processes is in stark contrast to (α, β)-processes ξ in R^d with $\alpha \in (0, 2)$ and $\beta \in (0, 1]$, where the spatial motion has jumps. In this case, Evans and Perkins [18, 40] showed that For fixed $t > 0$ a.s.

$$\operatorname{supp} \xi_t = \mathsf{R}^d, \quad \text{on } \{\xi_t \neq 0\}. \tag{2.4}$$

We can also discuss the Hausdorff dimension of the range of superprocesses. First for $I \subset \mathsf{R}_+$, define the range of ξ on I by

$$\mathcal{R}(I) = \bigcup_{t \in I} \operatorname{supp} \xi_t, \tag{2.5}$$

and the closed range of ξ on I by $\overline{\mathcal{R}}(I) = \overline{\mathcal{R}(I)}$, where $\overline{\mathcal{R}(I)}$ is the closure of $\mathcal{R}(I)$. Then the range of ξ is defined by

$$\mathcal{R} = \bigcup_{\varepsilon>0} \overline{\mathcal{R}}([\varepsilon, \infty)).$$

For a DW-process ξ in R^d, if $d \geq 4$, a.s. $\mathcal{R}$ is a null set and $\dim \mathcal{R} = 4 \wedge d$. More generally, if ξ is a $(2, \beta)$-process in R^d with $\beta \in (0, 1]$, if $d \geq (2/\beta) + 2$, a.s. $\mathcal{R}$ is a null set and

$$\dim \mathcal{R} = [(2/\beta) + 2] \wedge d,$$

see Corollary 2.2 in [9]. Again, for (α, β)-processes ξ in R^d with $\alpha \in (0, 2)$ and $\beta \in (0, 1]$, a.s. $\mathcal{R} = \mathsf{R}^d$. This is immediate from (2.4) and the definition of the range.

Let us turn back to the α-Hausdorff measures. Although for any nonnegative α the α-Hausdorff measure is a Borel measure, for some metric spaces it is always a trivial measure for any α, means that for any α, the α-Hausdorff measure $\mathcal{H}^\alpha(B)$ can only be 0 or ∞ for any $B \in \mathcal{B}(A)$. For example, if ξ is a DW-process in R^d with $d \geq 2$, then for a fixed $t > 0$, a.s. $\mathcal{H}^2(\operatorname{supp} \xi_t) = 0$ and

$$\mathcal{H}^\alpha(B \cap \operatorname{supp} \xi_t) = \infty \text{ or } 0$$

for any $\alpha < 2$ and $B \in \mathcal{B}(\mathsf{R}^d)$ (see (2.7), (2.8), and (2.9)). So we need to generalize the α-Hausdorff measures if we want to construct a nontrivial measure on $\operatorname{supp} \xi_t$.

The definition of Hausdorff dimension still makes sense if we evaluate coverings by applying, instead of a simple power, an arbitrary non-decreasing function to the diameters of the sets in a covering. We call this function a *gauge function*. By a gauge function we mean a non-decreasing function $\phi : [0, \varepsilon) \to [0, \infty)$ with $\phi(0) = 0$.

As before, we define

$$\mathcal{H}^\phi_\delta(A) = \inf\{\sum_{i=1}^{\infty} \phi(|A_i|) : A \subset \bigcup_{i=1}^{\infty} A_i, |A_i| \leq \delta \text{ for all } i\}. \tag{2.6}$$

Clearly the α-Hausdorff measure $\mathcal{H}_\delta^\alpha$ in (2.3) is just the special case of $\mathcal{H}_\delta^\phi$ with $\phi(x) = x^\alpha$. Then define the *ϕ-Hausdorff measure* of A by

$$\mathcal{H}^\phi(A) = \lim_{\delta \to 0} \mathcal{H}_\delta^\phi(A).$$

As before, $\mathcal{H}^\phi$ is a measure on Borel sets.

Under this more general framework, it is more likely to construct nontrivial measures on a metric space, although this is still not always possible. For a DW-superprocess ξ, this approach is extremely successful. Perkins and his co-authors proved the exact Hausdorff measure of the support at a fixed time or of the range of the process. First about the support $\operatorname{supp} \xi_t$. For a fixed $t > 0$, a.s. we have

$$\mathcal{H}^\phi(\cdot \cap \operatorname{supp} \xi_t) = \xi_t(\cdot), \tag{2.7}$$

where for $d \geq 3$ (see Theorem 5.2 in [7]),

$$\phi(x) = x^2 \log \log(1/x), \tag{2.8}$$

and for $d = 2$ (see Theorem 1.1 in [33]),

$$\phi(x) = x^2 \log(1/x) \log \log \log(1/x). \tag{2.9}$$

Next the range $\mathcal{R}(0, t]$. For a fixed $t > 0$, a.s. we have

$$\mathcal{H}^\phi(\cdot \cap \mathcal{R}(0, t]) = \int_0^t \xi_s ds(\cdot), \tag{2.10}$$

where for $d \geq 5$,

$$\phi(x) = x^4 \log \log(1/x), \tag{2.11}$$

and for $d = 4$,

$$\phi(x) = x^4 \log(1/x) \log\log\log(1/x). \tag{2.12}$$

Note that $\int_0^t \xi_s ds$ is a measure on $\mathcal{B}(\mathsf{R}^d)$, which is defined by

$$\int_0^t \xi_s ds(B) = \int_0^t \xi_s(B) ds, \quad \text{for any } B \in \mathcal{B}(\mathsf{R}^d). \tag{2.13}$$

It is easy to see that the Hausdorff measure results here contain the Hausdorff dimension results that we reviewed previously.

One obvious remaining question is the exact Hausdorff measure function of $(2,\beta)$-processes, but this may be technically too challenging. Then it is also interesting to try to get some good upper bound and lower bound of the exact Hausdorff measure function.

2.5 Lebesgue approximations

From the previous section, we see that by choosing carefully a suitable gauge function ϕ, we can define some nontrivial random measures on certain random null sets. Since from the beginning we know that there are some naturally defined nontrivial random measures on these random null sets (the DW-process ξ_t on $\operatorname{supp} \xi_t$, and the local time measure of one dimensional Brownian motion on its level set, see below), in fact the Hausdorff measure approach gives representations of these measures with respect to only their support. So in order to recover these measures, we can forget about the related stochastic processes, only the support of these measures is needed. In this regard, we also have the packing measure approach (see [11, 34]), which is similar to the Hausdorff measure approach generally speaking. A more different approach to do this is the so called Lebesgue approximation approach. Kingman [28] explained this approach in a very accessible manner and also used this approach to recover the local time measure of certain Markov processes intrinsically from the level set.

Let us first explain Kingman's idea. For the subset $A \in \mathsf{R}^d$, we use A^ε to denote the ε-neighborhood of A , that is,

$$A^\varepsilon = \{x : d(x, A) < \varepsilon\}.$$

It's easy to see that $A^\varepsilon = (\overline{A})^\varepsilon$. Recall that $\overline{A}$ is the closure of A. For their corresponding Lebesgue measures, clearly $\lambda^d E^\varepsilon \in [0, \infty]$ and $\lambda^d E^\varepsilon \to \lambda^d \overline{E}$. So when $\overline{E}$ is a null set, we get $\lambda^d E^\varepsilon \to 0$. Here the interesting thing is that, the rate at which $\lambda^d E^\varepsilon$ converges to zero is an indication of the size of $\overline{E}$. For example, if E is a part of a sufficiently smooth d'-dimensional surface in R^d, where $d' < d$, then

$$\varepsilon^{d'-d} \lambda^d E^\varepsilon \sim C \lambda^{d'} E,$$

where C is a constant depending only on d and d'.

Now let us explain a special case of Kingman's Lebesgue approximation result. Let $L(t, x)$ be the local time of a one dimensional Brownian motion $B_1(t)$. Let $Z(t, x) = \{s : s \leq t, B_1(s) = x\}$. Kingman showed that there exists a constant c such that for fixed t and x, a.s.

$$\varepsilon^{-1/2} \lambda^d [Z(t, x)]^\varepsilon \to cL(t, x).$$

Kingman argued that unlike other approximation results of local time (see Corollary 1.9 and Theorem 1.10 in Chapter VI of [44]), this Lebesgue approximation result only requires the knowledge of $Z(t, x)$, which is the support of the local time measure $L(t, x)$, to recover $L(t, x)$.

Next we discuss all known Lebesgue approximation results of superprocesses. For any measure μ on R^d and constant $\varepsilon > 0$, write μ^ε for the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp} \mu$. Note that using our notations, the ε-neighborhood of $\operatorname{supp} \mu$

is denoted by $(\operatorname{supp}\mu)^\varepsilon$. So we may write μ^ε explicitly by

$$\mu^\varepsilon(\cdot) = \lambda^d\left(\cdot \cap (\operatorname{supp}\mu)^\varepsilon\right).$$

For a DW-process ξ in R^d with $d \geq 3$, Tribe [48] showed that for any fixed $t > 0$ and any bounded Borel set B in R^d, a.s. as $\varepsilon \to 0$,

$$\varepsilon^{2-d}\,\xi_t^\varepsilon(B) \to c_d\,\xi_t(B),$$

where $c_d > 0$ is a constant depending on d.

Shortly after, in order to prove the strong Markov property of the support process $\operatorname{supp}\xi_t$, Perkins [41] showed that the Lebesgue approximation result holds simultaneously for all time $t > 0$. More precisely, for a DW-process ξ in R^d with $d \geq 3$, a.s. as $\varepsilon \to 0$

$$\varepsilon^{2-d}\,\xi_t^\varepsilon \stackrel{w}{\to} c_d\,\xi_t, \quad \text{for all } t > 0, \tag{2.14}$$

where $\stackrel{w}{\to}$ denotes the weak convergence of measures. The corresponding Lebesgue approximation of two dimensional DW-processes ξ was still open at that time, even for fixed $t > 0$. However, Perkins conjectured that for fixed $t > 0$ and bounded Borel set B in R^2, a.s. as $\varepsilon \to 0$,

$$|\log\varepsilon|\,\xi_t^\varepsilon(B) \to c\,\xi_t(B).$$

Later, Kallenberg [25] essentially confirmed the above conjecture. More precisely, for a DW-process ξ in R^2, Kallenberg showed that for fixed $t > 0$ a.s. as $\varepsilon \to 0$,

$$\tilde{m}(\varepsilon)\,|\log\varepsilon|\,\xi_t^\varepsilon \xrightarrow{w} \xi_t,$$

where $\tilde{m}$ is a suitable normalizing function bounded below and above by two positive constants. Note that both the conjecture of Perkins and the proof of Kallenberg depend crucially on the hitting bounds of DW-processes ξ in R^2 from Le Gall [31]. Kallenberg's approach also works for DW-processes ξ in R^d with $d \geq 3$, and results in a more probabilistic proof of Lebesgue approximation of ξ_t.

In [?], we adapted Kallenberg's probabilistic approach in [25] to prove the Lebesgue approximation of $(2, \beta)$-processes with $\beta < 1$, combined with a truncation method of superprocesses from Mytnik and Villa [38], in order to overcome the additional difficulty imposed by the infinite variance of $(2, \beta)$-processes. More precisely, for a $(2, \beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$, we proved that, for fixed $t > 0$ a.s. as $\varepsilon \to 0$,

$$\varepsilon^{2/\beta - d}\, \xi_t^\varepsilon \xrightarrow{w} c_{\beta,d}\, \xi_t,$$

where $c_{\beta,d} > 0$ is a constant depending on β and d.

In view of the Hausdorff measure results (2.10), (2.11), and (2.12), we may ask about the Lebesgue approximation of the range of a superprocess. Here, Delmas [10] proved the Lebesgue approximation of the range of DW-processes ξ in R^d with $d \geq 4$, using Le Gall's Brownian snake. More precisely, for a DW-process ξ in R^d with $d \geq 4$, Delmas showed that, for fixed $t > 0$ and bounded Borel set B in R^d, a.s. as $\varepsilon \to 0$,

$$\phi(\varepsilon)\, \overline{\mathcal{R}}_t^\varepsilon(B) \to c_d \int_t^\infty \xi_s ds(B), \tag{2.15}$$

where $\mathcal{R}_t$ is the $\mathcal{R}([t, \infty))$ defined in (2.5), and $\int_t^\infty \xi_s ds$ is defined as in (2.13). About the normalizing function ϕ, it is shown that

$$\phi(\varepsilon) = \varepsilon^{4-d} \quad \text{for } d > 5, \quad \text{and,} \quad \phi(\varepsilon) = |\log \varepsilon| \quad \text{for } d = 4.$$

These known results lead to a couple immediate open problems. First we may ask if the Lebesgue approximation of $(2, \beta)$-processes holds simultaneously for all time $t > 0$, in view of (2.14). Since intuitively the $(2, \beta)$-process ξ and its support $\operatorname{supp} \xi$ do not jump at the same time, an immediate guess should be no. We may then ask if it is possible to prove some results supporting this guess. What about the strong Markov property of the $(2, \beta)$-support process $\operatorname{supp} \xi_t$? It seems that we need to find new approaches to prove it (or disprove it). The second question is that, whether it is possible to prove the Lebesgue approximation of the range of $(2, \beta)$-processes, in view of (2.15).

More generally, we may try to "translate" all Hausdorff measure results into corresponding Lebesgue approximation ones. Here the challenge is that while there is a solid theory behind Hausdorff measures which one could rely on, there is no such support for Lebesgue approximation results. One has to "invent" some approaches when trying to establish Lebesgue approximation results. Still it is very interesting to see that whenever we can get the Lebesgue approximation results, the results are always shorter and cleaner then the corresponding Hausdorff measure results.

Chapter 3

Lebesgue Approximation of Dawson-Watanabe Superprocesses

3.1 Introduction

In this chapter we discuss the Lebesgue approximation of Dawson-Watanabe superprocesses in detail. The Lebesgue approximation of DW-processes of dimension $d \geq 3$ was first proved by Tribe [48], using both probabilistic and analytic techniques. The case of critical dimension $d = 2$ is more difficult. However, Kallenberg [25] obtained a similar result for DW-processes in R^2 using a more probabilistic approach. His approach can also be applied to DW-processes of dimension $d \geq 3$, and indeed this was done in [25]. The present chapter is based on Kallenberg's proof of Lebesgue approximation of DW-processes of dimension $d \geq 3$ in [25], with some technical simplifications. Extra efforts have been made to explain Kallenberg's approach clearly and to make it more accessible.

We use $\xi = (\xi_t)$ to denote the DW-process of dimension $d \geq 3$. Recall that ξ is a measure-valued Markov process, so for fixed t and ω, the value $\xi_t(\omega)$ is a measure on R^d. We write ξ_t^ε for the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp} \xi_t$, the support of the measure ξ_t, which for fixed t and ω, is a compact set in R^d (see Theorem 1.2 in [6]). The Lebesgue approximation of DW-processes of dimension $d \geq 3$, which is Theorem 3.5 in this chapter, states that for fixed $t > 0$, $\varepsilon^{2-d}\, \xi_t^\varepsilon \xrightarrow{w} c_d\, \xi_t$ a.s. as $\varepsilon \to 0$, where $\xrightarrow{w}$ denotes weak convergence of measures and $c_d > 0$ is a universal constant depending on d. In particular, this confirms that ξ_t "distributes its mass over $\operatorname{supp} \xi_t$ in a deterministic manner" (cf. [17], p. 115, or [42], p. 212), as previously inferred from some deep results involving the exact Hausdorff measure (cf. [7]).

The proof depends crucially on some basic hitting estimates, due to Dawson, Iscoe, and Perkins [6]. Here we need the lower bound and upper bound of $P_\mu\{\xi_t B_0^\varepsilon > 0\}$ (Theorem

3.1.(a) in [6]), and also the precise convergence result $\varepsilon^{2-d}P_\mu\{\xi_t B_0^\varepsilon > 0\} \to c_d\,\mu p_t$ for $d \geq 3$ as $\varepsilon \to 0$ (Theorem 3.1.(b) in [6]), where B_x^r denotes an open ball around x of radius r.

The proof also depends crucially on the representation of the DW-process as a countable sum of conditionally independent clusters. Precisely, each ξ_t can be expressed as a countable sum of conditionally independent clusters of age $h \in (0, t]$, where the generating ancestors at time $s = t - h$ form a Cox process ζ_s directed by $h^{-1}\xi_s$ (cf. [7, 30]). Typically we let $h \to 0$ at a suitable rate depending on ε. However, a technical complication when dealing with cluster representations is the possibility of multiple hits. More specifically, a single cluster may hit (charge) several of the ε-neighborhoods of n distinct points $x_1, \dots, x_n$, or one of those neighborhoods may be hit by several clusters. In particular, Lemma 2.4 deals with this multiple hitting of a single neighborhood by several clusters. To minimize the effect of such multiplicities, we need the cluster age h to be sufficiently small. On the other hand, it needs to be large enough for the mentioned hitting estimates to apply to the individual clusters. Notice that we can translate the hitting estimates for the superprocess ξ to the hitting estimates for the cluster η, based on the connection between the superprocess and its clusters.

The reason we don't cover the case of critical dimension $d = 2$ is that, although the two cases of $d = 2$ and $d \geq 3$ use the same general approach, technically the case of $d = 2$ is much more involved, since we then have to deal with the Logarithm normalizing function $|\log(\varepsilon)|$ rather than the power normalizing function ε^{2-d} as in the case of $d \geq 3$. Also when $d = 2$, a corresponding crucial result to the precise convergence result for $d \geq 3$, as $\varepsilon \to 0$, $\varepsilon^{2-d}P_\mu\{\xi_t B_0^\varepsilon > 0\} \to c_d\,\mu p_t$, is not readily available. So in this chapter, we restrict our attention to the case of $d \geq 3$.

We proceed with some general remarks on terminology and notation. A *random measure* ξ on R^d is defined as a measurable function from Ω to the space $\mathcal{M}_d$ of locally finite measures on R^d, equipped by the σ-field generated by all evaluation maps $\pi_B\colon\, \mu \mapsto \mu B$ with $B \in \mathcal{B}^d$, where $\mathcal{B}^d$ denotes the Borel σ-field on R^d. The subclasses of measures and bounded sets are denoted by $\hat{\mathcal{M}}_d$ and $\hat{\mathcal{B}}^d$, respectively. The *weak* topology in $\mathcal{M}_d$ is generated by all integration maps $\pi_f : \mu \mapsto \mu f = \int f d\mu$ with f belonging to the space C_b^d of bounded, continuous functions $\mathsf{R}^d \to \mathsf{R}_+$. Thus, $\mu_n \stackrel{w}{\to} \mu$ in $\mathcal{M}_d$ iff $\mu_n f \to \mu f$ for all $f \in C_b^d$.

Throughout the chapter we use relations such as $\underset{\frown}{=}$, $\underset{\frown}{<}$, $\underset{\frown}{\sim}$, and $\asymp$, where the first three mean equality, inequality, and asymptotic equality up to a constant factor, and the last one is the combination of $\underset{\frown}{<}$ and $\underset{\frown}{>}$. We often write $a \ll b$ to mean $a/b \to 0$. The double bars $\|\cdot\|$ denote the supremum norm when applied to functions and total variation when applied to signed measures. In any Euclidean space R^d, we write B_x^r for the open ball of radius $r > 0$ centered at $x \in \mathsf{R}^d$. The shift and scaling operators θ_x and S_r are given by $\theta_x y = x + y$ and $S_r x = rx$, respectively, and for measures μ on R^d we define $\mu\theta_x$ and μS_r by $(\mu\theta_x)B = \mu(\theta_x B)$ and $(\mu S_r)B = \mu(S_r B)$, respectively. In particular, $(\mu S_r)f = \mu(f \circ S_r^{-1})$ for measurable functions f on R^d. Convolutions of measures μ with functions f are given by $(\mu * f)(x) = \int f(x-u)\,\mu(du)$.

This chapter is organized as follows. In Section 2 we first explain the crucial ideas about cluster representations, then state several lemmas which will not be used directly in the main proof of Lebesgue approximation, including the important upper bound of the hitting multiplicities. In Section 3 we state and prove the Lebesgue approximation for DW-processes of dimensions $d \geq 3$. In order to do so, we list several lemmas that are needed in the main proof. Finally, in Section 4 we prove all the lemmas in this chapter. We suggest that the reader read the first three sections in the linear order, then, when need arises, read the proofs of some lemmas in Section 4.

3.2 Preliminaries

Let us first explain the cluster representations of DW-processes. We write $\mathcal{L}_\mu(\xi) = P_\mu\{\xi \in \cdot\}$ for the distribution of the process ξ with initial measure μ. For every fixed μ, the DW-process ξ is infinitely divisible under P_μ and admits a decomposition into a Poisson “forest” of conditionally independent *clusters*, corresponding to the excursions of the contour process in the ingenious “Brownian snake” representation of Le Gall [32]. In particular, this yields a cluster representation of ξ_t for every fixed $t > 0$. More generally, the “ancestors” of ξ_t at an earlier time $s = t - h$ form a Cox process ζ_s directed by $h^{-1}\xi_s$ (meaning that ζ_s is conditionally Poisson with intensity $h^{-1}\xi_s$, given ξ_s; cf. [24], p. 226), and the generated clusters η_h^i are conditionally independent and identically distributed apart from shifts. In this paper, a generic cluster of age $t > 0$ is denoted by η_t; we write $\mathcal{L}_x(\eta_t) = P_x\{\eta_t \in \cdot\}$ for the distribution of a t-cluster centered at $x \in \mathsf{R}^d$ and put $P_\mu\{\eta_t \in \cdot\} = \int \mu(dx)\, P_x\{\eta_t \in \cdot\}$.

The first lemma is about some basic scaling properties of DW-processes and their associated clusters.

Lemma 3.1 *Let ξ be a DW-process in R^d with associated clusters η_t. Then for any measure μ on R^d, and $r, t > 0$,*

(i) $\mathcal{L}_{\mu S_r}(r^2\xi_t) = \mathcal{L}_{r^2\mu}(\xi_{r^2 t} S_r)$,

(ii) $\mathcal{L}_{\mu S_r}(r^2\eta_t) = \mathcal{L}_{\mu}(\eta_{r^2 t} S_r)$.

Although the above two compact identities look nice, they may not be very intuitive for some people. In order to appreciate better these scaling properties, first we translate the $\mathcal{L}$ notation back to the P notation

$$\begin{aligned} P_{\mu S_r}\{r^2\xi_t \in \cdot\} &= P_{r^2\mu}\{\xi_{r^2 t} S_r \in \cdot\}, \\ P_{\mu S_r}\{r^2\eta_t \in \cdot\} &= P_{\mu}\{\eta_{r^2 t} S_r \in \cdot\}. \end{aligned}$$

Recall that the evaluation map $\pi_B\colon\ \mu \mapsto \mu B$ is a function defined on the space $\mathcal{M}_d$ of locally finite measures on R^d. According to the definition of σ-field on $\mathcal{M}_d$, the set $\{\pi_{B_0^{1/r}} > 0\}$ is a measurable set on $\mathcal{M}_d$. In the above two identities, take $r = 1/\varepsilon$, $\mu S_r = \delta_x$, and, $\cdot = \{\pi_{B_0^{1/r}} > 0\}$, we get

$$P_x\{\xi_t B_0^\varepsilon > 0\} = P_{(1/\varepsilon^2)\delta_{x/\varepsilon}}\{\xi_{t/\varepsilon^2} B_0^1 > 0\},$$

$$P_x\{\eta_t B_0^\varepsilon > 0\} \;=\; P_{x/\varepsilon}\{\eta_{t/\varepsilon^2} B_0^1 > 0\}.$$

Now these two identities should be intuitive enough for one to appreciate the scaling properties.

Next we state a well-known relationship between the hitting probabilities of ξ_t and η_t.

Lemma 3.2 *Let the DW-process ξ in R^d with associated clusters η_t be locally finite under P_μ, and fix any $B \in \hat{\mathcal{B}}^d$. Then*

$$\begin{aligned} P_\mu\{\eta_t B > 0\} &= -t \log\left(1 - P_\mu\{\xi_t B > 0\}\right), \\ P_\mu\{\xi_t B > 0\} &= 1 - \exp\left(-t^{-1} P_\mu\{\eta_t B > 0\}\right). \end{aligned}$$

In particular, $P_\mu\{\xi_t B > 0\} \sim t^{-1} P_\mu\{\eta_t B > 0\}$ as either side tends to 0.

The following lemma contains some slight variations of classical hitting estimates for DW-processes of dimension $d \geq 3$. By Lemma 3.2 it is enough to consider the corresponding clusters η_t, and by shifting it suffices to consider balls centered at the origin.

Lemma 3.3 *Let the η_t be clusters of a DW-process in R^d with $d \geq 3$, and consider a σ-finite measure μ on R^d. Then for $0 < \varepsilon \leq \sqrt{t}$, we have*

$$\mu p_t \precsim t^{-1} \varepsilon^{2-d} P_\mu\{\eta_t B_0^\varepsilon > 0\} \precsim \mu p_{2t},$$

The classical upper bound is $\mu p_{t+\varepsilon}$. Note that as $\varepsilon \to 0$, the upper bound $\mu p_{t+\varepsilon}$ is approaching the lower bound μp_t, however the constants before these two bounds are definitely different. Still this suggests that as $\varepsilon \to 0$, the normalized hitting probability $t^{-1}\varepsilon^{2-d} P_\mu\{\eta_t B_0^\varepsilon > 0\}$ converges to $c\,\mu p_t$ for some constant $c > 0$. This is indeed the case. Although the classical upper bound can give us this intuitive impression, for all practical purposes our upper bound μp_{2t} is as good, if not better. The reason is that mathematically speaking, p_{2t} is almost the same as p_t.

Next we need to estimate the probability that a small ball in R^d is hit by more than one subcluster of our DW-process ξ. This result will play a crucial role throughout the remainder of the chapter.

Lemma 3.4 *Let the DW-process ξ in R^d be locally finite under P_μ. For any $t \geq h > 0$ and $\varepsilon > 0$, let κ_h^ε be the number of h-clusters hitting B_0^ε at time t. Then for $d \geq 3$ and as $\varepsilon^2 \ll h \leq t$, we have*

$$E_\mu \kappa_h^\varepsilon (\kappa_h^\varepsilon - 1) \lesssim \varepsilon^{2(d-2)} \left(h^{1-d/2} \mu p_t + (\mu p_{2t})^2 \right).$$

Here the intuition is that, if compare to h, the radius ε is small enough, then most likely there will be only one cluster hitting this tiny ball, or no cluster at all. Actually what we want to control is the discrete quantity $(\kappa_h^\varepsilon - 1)_+$. However it seems that the only natural way to relate this quantity to the DW-process ξ_t is through the following simple inequality

$$(\kappa_h^\varepsilon - 1)_+ \leq \kappa_h^\varepsilon (\kappa_h^\varepsilon - 1).$$

Then we can relate $E_\mu \kappa_h^\varepsilon (\kappa_h^\varepsilon - 1)$ to $E_\mu \xi_t$ and $E_\mu \xi_t^2$, the first and second moment of the DW-process ξ_t. This is actually a very important point, especially in the next chapter when we are dealing with the $(2, \beta)$-superprocesses. Since the $(2, \beta)$-superprocesses have infinite second moment, to control $E_\mu (\kappa_h^\varepsilon - 1)_+$ we have to truncate the $(2, \beta)$-processes, in order to get the finite second moment.

3.3 Lebesgue approximation

In this section we first state the main result of this chapter, the Lebesgue approximation of DW-processes of dimension $d \geq 3$, which is Theorem 3.5. In order to give the proof of Theorem 3.5, we then state Lemma 3.6, 3.7, and 3.8, which will be used directly in the proof of Theorem 3.5. However we leave all proofs of lemmas in the next section. At the end of the present section we give the proof of Theorem 3.5.

For any measure μ on R^d and constant $\varepsilon > 0$, we define the associated *neighborhood measure* μ^ε as the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp}\mu$, so that μ^ε has Lebesgue density $1\{\mu B_x^\varepsilon > 0\}$. First note that μ^ε is a measure defined from the measure μ. Then recall that $\xi_t(\omega)$ is a measure for fixed t and ω, so $\xi_t^\varepsilon(\omega)$ is just the neighborhood measure of $\xi_t(\omega)$. Also recall that $\hat{\mathcal{M}}_d$ is the space of finite measures on R^d. For random measures ξ_n and ξ with values in $\hat{\mathcal{M}}_d$, the weak convergence in L^1, denoted by

$$\xi_n \stackrel{w}{\to} \xi \quad \text{in } L^1,$$

means that $\xi_n f \to \xi f$ in L^1 for all f in C_b^d. Write $\tilde{c}_d = 1/c_d$ for convenience, where c_d is such as in (3.1).

Now we are ready to state the main result of this chapter, the Lebesgue approximation of DW-processes of dimension $d \geq 3$.

Theorem 3.5 *Let ξ be the DW-process in R^d with $d \geq 3$. Fix any $\mu \in \hat{\mathcal{M}}_d$ and $t > 0$. Then under P_μ, we have as $\varepsilon \to 0$*

$$\tilde{c}_d\, \varepsilon^{2-d}\, \xi_t^\varepsilon \stackrel{w}{\to} \xi_t \text{ a.s. and in } L^1.$$

Here the a.s. convergence means that for every ω outside a null set,

$$\tilde{c}_d\, \varepsilon^{2-d}\, \xi_t^\varepsilon(\omega) \stackrel{w}{\to} \xi_t(\omega).$$

Note that for fixed t and ω, both $\xi_t^\varepsilon(\omega)$ and $\xi_t(\omega)$ are deterministic measures.

Next we are going to study $(\eta_h^i)^\varepsilon$, the neighborhood measures of the clusters. Since we will use the cluster decomposition $\xi_t = \Sigma_i \eta_h^i$ throughout the proof, naturally in order to prove $\tilde{c}_d\, \varepsilon^{2-d}\, \xi_t^\varepsilon \stackrel{w}{\to} \xi_t$ we also need to study $(\eta_h^i)^\varepsilon$. Write $(\eta_h^i)^\varepsilon = \eta_h^{i\varepsilon}$ for convenience.

Lemma 3.6 *Let the η_h^i be conditionally independent h-clusters in R^d, rooted at the points of a Poisson process ξ with $E\xi = \mu$. Fix any measurable function $f \geq 0$ on R^d. Then*

(i) $E_\mu \sum_i \eta_h^{i\varepsilon} = (\mu * p_h^\varepsilon) \cdot \lambda^d$,

(ii) $\mathrm{Var}_\mu \sum_i \eta_h^{i\varepsilon} f \lesssim h^2 \varepsilon^{2(d-2)} \|f\|^2 \|\mu\|$ *for* $\varepsilon^2 \leq h$.

In part (i), notice that $\sum_i \eta_h^{i\varepsilon}$ is a random measure, its expectation is the deterministic measure $(\mu * p_h^\varepsilon) \cdot \lambda^d$, which means that for any measurable $f \geq 0$

$$E_\mu \sum\nolimits_i \eta_h^{i\varepsilon} f = ((\mu * p_h^\varepsilon) \cdot \lambda^d) f$$

where $((\mu * p_h^\varepsilon) \cdot \lambda^d) f$ is the integral of the function f with respect to the measure $(\mu * p_h^\varepsilon) \cdot \lambda^d$. In part (ii), notice that $\sum_i \eta_h^{i\varepsilon} f$ is a real-valued random variable, its variance is bounded above by $h^2 \varepsilon^{2(d-2)} \|f\|^2 \|\mu\|$.

Next we compare ξ_t^ε and $\sum_i \eta_h^{i\varepsilon}$, and prove that asymptotically they are the same, so that we can just replace ξ_t^ε by $\sum_i \eta_h^{i\varepsilon}$. Intuitively this result is clear: Since the ages of clusters h and the parameter of neighborhood measures ε are both going to 0 at some suitable rates, asymptotically there are no overlaps between the neighborhood measures of clusters, so that asymptotically $\sum_i \eta_h^{i\varepsilon}$ and ξ_t^ε are the same.

Lemma 3.7 *Let ξ be a DW-process in R^d with $d \geq 3$, and for fixed $t > 0$, let η_h^i denote the subclusters in ξ_t of age $h > 0$. Fix a $\mu \in \hat{\mathcal{M}}_d$. Then as $\varepsilon^2 \leq h \to 0$,*

$$E_\mu \left\| \sum_i \eta_h^{i\varepsilon} - \xi_t^\varepsilon \right\| \lesssim (\varepsilon^2/\sqrt{h})^{d-2}.$$

Recall that for a signed measure μ on R^d with f as the density with respect to λ^d, the total variation $\|\mu\|$ satisfies

$$\|\mu\| = \lambda^d |f| = \int |f(x)| dx.$$

Note that $\sum_i \eta_h^{i\varepsilon}$ and ξ_t^ε have the density $\sum_i 1\{\eta_h^i B_x^\varepsilon > 0\}$ and $1\{\xi B_x^\varepsilon > 0\}$ respectively, so

$$E_\mu \left\| \sum_i \eta_h^{i\varepsilon} - \xi_t^\varepsilon \right\| = E_\mu \int \left| \sum_i 1\{\eta_h^i B_x^\varepsilon > 0\} - 1\{\xi B_x^\varepsilon > 0\} \right| dx.$$

Now clearly the integrand $|\sum_i 1\{\eta_h^i B_x^\varepsilon > 0\} - 1\{\xi B_x^\varepsilon > 0\}|$ is related to the multiple hitting of Lemma 3.4.

The last lemma is a precise convergence result about the hitting probability $P_x\{\eta_h B_0^\varepsilon > 0\}$.

For a DW-process ξ of dimension $d \geq 3$, we know from Theorem 3.1 of Dawson, Iscoe, and Perkins [6] (cf. Remark III.5.12 in [42]) that, for fixed $t > 0$, $x \in \mathsf{R}^d$, and finite μ, as $\varepsilon \to 0$

$$\varepsilon^{2-d} P_\mu\{\xi_t B_x^\varepsilon > 0\} \to c_d\,(\mu * p_t)(x), \tag{3.1}$$

where $c_d > 0$ is a constant depending only on d, and the convergence is uniform for $x \in \mathsf{R}^d$ and for bounded t^{-1} and $\|\mu\|$. Notice that in this classical result t can change, but it has to be bounded away from 0.

By using the scaling property of DW-processes, from the classical result above we can get a precise convergence result about $P_x\{\eta_h B_0^\varepsilon > 0\}$ as both h and ε are approaching 0 at some suitable rates. More precisely, after the scaling term $(c_d)^{-1} h^{-1} \varepsilon^{2-d}$ multiplied to $P_x\{\eta_h B_0^\varepsilon > 0\}$, the measure

$$(c_d)^{-1} h^{-1} \varepsilon^{2-d} P_x\{\eta_h B_0^\varepsilon > 0\} dx$$

converges in a certain sense to δ_0, the Dirac measure at 0 δ_0, as both h and ε are approaching 0 at some suitable rates. This result should be easy to understand since if $x \neq 0$, then for small enough h and ε, the η_h stated from x will not be able to reach B_0^ε before time h. Only the η_h stated from 0 will be able to reach B_0^ε before time h, although the probability is decreasing to 0. After the scaling term $h^{-1}\varepsilon^{2-d}$ multiplied to $P_0\{\eta_h B_0^\varepsilon > 0\}$, it converges to the constant c_d.

Lemma 3.8 *Write $p_h^\varepsilon(x) = P_x\{\eta_h B_0^\varepsilon > 0\}$, where the η_h are clusters of a DW-process in R^d, and fix a bounded, uniformly continuous function $f \geq 0$ on R^d. Then as $0 < \varepsilon^2 \ll h \to 0$, we have*

$$\left\| \, h^{-1}\varepsilon^{2-d}\,(p_h^\varepsilon * f) - c_d\, f \, \right\| \to 0.$$

The result holds uniformly over any class of uniformly bounded and equicontinuous functions $f \geq 0$ on R^d.

Here $\left\| \, h^{-1}\varepsilon^{2-d}\,(p_h^\varepsilon * f) - c_d\, f \, \right\|$ is the supremum norm of the function

$$h^{-1}\varepsilon^{2-d}\,(p_h^\varepsilon * f)(x) - c_d\, f(x),$$

as a function of x.

Now we are ready to prove Theorem 3.5, but before giving the proof let us discuss the main ideas in the proof carefully. First of all, we have two possible approaches to attack this theorem: one is to prove the L^1-convergence first, then use some interpolation to get the a.s. convergence from the L^1-convergence (this is indeed what Tribe did in [48]); the other is to prove the a.s. convergence first. In the first approach, we need to get the a.s. convergence from the L^1-convergence by the usual Borel-Cantelli argument: If $E\sum |f_n| < \infty$, then $f_n \to 0$ a.s. as $n \to \infty$. In order to do so, we need an upper bound of the approximating error

$$\varepsilon^{2-d} P_\mu\{\xi_t B_x^\varepsilon > 0\} - c_d\,(\mu * p_t)(x),$$

which we don't have here. So we will use the second approach: prove the a.s. convergence first.

In order to prove the a.s. convergence, we need to show that a.s. for all $f \in C_b^d$, we have that $\tilde{c}_d\,\varepsilon^{2-d}\,\xi_t^\varepsilon f \to \xi_t f$, where C_b^d is the class of bounded, continuous functions $\mathsf{R}^d \to \mathsf{R}_+$.

However since there exists a countable, convergence-determining class of functions f in C_b^d, we only need to prove for any fixed $f \in C_b^d$, we have $\tilde{c}_d\, \varepsilon^{2-d}\, \xi_t^\varepsilon f \to \xi_t f$ a.s.

In order to prove this, we write

$$\begin{aligned}\left|\, \varepsilon^{2-d}\, \xi_t^\varepsilon f - c_d\, \xi_t f \,\right| \;\; \leq \;\; & \varepsilon^{2-d} \left| \xi_t^\varepsilon f - \sum\nolimits_i \eta_h^{i\varepsilon} f \right| \\ & + \varepsilon^{2-d} \left| \sum\nolimits_i \eta_h^{i\varepsilon} f - h^{-1}\, \xi_s(p_h^\varepsilon * f) \right| \\ & + \|\xi_s\| \left\| \varepsilon^{2-d}\, h^{-1}\, (p_h^\varepsilon * f) - c_d f \right\| \\ & + c_d\, |\xi_s f - \xi_t f|. \end{aligned}$$

Notice that the last term converges to 0 by the a.s. weak continuity of ξ and the third term converges to 0 by Lemma 3.4. The first term is related to Lemma 3.3 and the second term is related to Lemma 3.2, however these two lemmas are about the expectations and variances of those terms.

In order to get a.s. convergence from results of expectations and variances, we use the usual Borel-Cantelli argument: take a sequence ε_n and get $f(\varepsilon_n) \to 0$ as $n \to \infty$ by showing that $E \sum |f(\varepsilon_n)| < \infty$. Finally we extend the a.s. convergence from the sequence ε_n to the whole interval $(0, 1)$ by interpolation.

As for the L^1-convergence, since by (1) we easily get

$$\varepsilon^{2-d} E_\mu \xi_t^\varepsilon f \to c_d\, E_\mu \xi_t f,$$

so the L^1-convergence follows from the a.s. convergence by an usual proposition.

Proof of Theorem 3.1:

Proof: (i) Let $d \geq 3$, and fix any $t > 0$, $\mu \in \hat{\mathcal{M}}_d$, and $f \in C_K^d$. Write η_h^i for the subclusters of ξ_t of age h. Since the ancestors of ξ_t at time $s = t - h$ form a Cox process directed by ξ_s/h, Lemma 3.6 (i) yields

$$E_\mu \left[\sum\nolimits_i \eta_h^{i\varepsilon} f \,\middle|\, \xi_s \right] = h^{-1} \xi_s(p_h^\varepsilon * f),$$

and so by Lemma 3.6 (ii)

$$\begin{aligned} E_\mu \Big| \sum\nolimits_i \eta_h^{i\varepsilon} f - h^{-1}\, \xi_s(p_h^\varepsilon * f) \Big|^2 &= E_\mu \mathrm{Var}\Big[\sum\nolimits_i \eta_h^{i\varepsilon} f \,\Big|\, \xi_s \Big] \\ &\lesssim \varepsilon^{2(d-2)}\, h^2\, \|f\|^2\, E_\mu \|\xi_s/h\| \\ &= \varepsilon^{2(d-2)}\, h\, \|f\|^2\, \|\mu\|. \end{aligned}$$

Combining with Lemma 3.7 gives

$$\begin{aligned} & E_\mu \big|\, \xi_t^\varepsilon f - h^{-1}\, \xi_s(p_h^\varepsilon * f) \,\big| \\ & \quad \le E_\mu \Big| \xi_t^\varepsilon f - \sum\nolimits_i \eta_h^{i\varepsilon} f \Big| + E_\mu \Big| \sum\nolimits_i \eta_h^{i\varepsilon} f - h^{-1}\, \xi_s(p_h^\varepsilon * f) \Big| \\ & \quad \lesssim \varepsilon^{2(d-2)}\, h^{1-d/2}\, \|f\| + \varepsilon^{d-2}\, h^{1/2}\, \|f\| \\ & \quad = \varepsilon^{d-2} \left(\sqrt{h} + (\varepsilon/\sqrt{h})^{d-2} \right) \|f\|. \end{aligned}$$

Taking $h = \varepsilon = r^n$ for a fixed $r \in (0,1)$ and writing $s_n = t - r^n$, we obtain

$$\begin{aligned} & E_\mu \sum\nolimits_n r^{n(2-d)} \,\big|\, \xi_t^{r^n} f - r^{-n}\, \xi_{s_n}(p_{r^n}^{r^n} * f) \,\big| \\ & \qquad \lesssim \sum\nolimits_n \big(r^{n/2} + r^{n(d-2)/2} \big)\, \|f\| < \infty, \end{aligned}$$

which implies

$$r^{n(2-d)} \,\big|\, \xi_t^{r^n} f - r^{-n}\, \xi_{s_n}(p_{r^n}^{r^n} * f) \,\big| \to 0 \ \text{ a.s. } P_\mu. \tag{3.2}$$

Now we write

$$\begin{aligned} \big|\, \varepsilon^{2-d}\, \xi_t^\varepsilon f - c_d\, \xi_t f \,\big| &\le \varepsilon^{2-d} \,\big|\, \xi_t^\varepsilon f - h^{-1}\, \xi_s(p_h^\varepsilon * f) \,\big| + c_d\, |\xi_s f - \xi_t f| \\ &\quad + \|\xi_s\| \,\big\|\, \varepsilon^{2-d}\, h^{-1}\, (p_h^\varepsilon * f) - c_d f \,\big\|. \end{aligned}$$

Using (3.2), Lemma 3.8, and the a.s. weak continuity of ξ (cf. Proposition 2.15 in [17]), we see that the right-hand side tends a.s. to 0 as $n \to \infty$, which implies $\varepsilon^{2-d} \xi_t^\varepsilon f - c_d \xi_t f$ a.s.

as $\varepsilon \to 0$ along the sequence (r^n) for any fixed $r \in (0,1)$. Since this holds simultaneously, outside a fixed null set, for all rational $r \in (0,1)$, the a.s. convergence extends by Lemma 2.3 in [25] to the entire interval $(0,1)$.

Now let $\mu \in \mathcal{M}_d$ be arbitrary with $\mu p_t < \infty$ for all $t > 0$. Write $\mu = \mu' + \mu''$ for bounded μ', and let $\xi = \xi' + \xi''$ be the corresponding decomposition of ξ into independent components with initial measures μ' and μ''. Fixing an $r > 1$ with $\operatorname{supp} f \subset B_0^{r-1}$ and using the result for bounded μ, we get a.s. on $\{\xi_t'' B_0^r = 0\}$

$$\varepsilon^{2-d}\, \xi_t^\varepsilon f = \varepsilon^{2-d}\, \xi_t'^\varepsilon f \to c_d\, \xi_t' f = c_d\, \xi_t f.$$

As $\mu' \uparrow \mu$, we get by Lemma 4.3 in [25]

$$P_\mu\{\xi_t'' B_0^r = 0\} = P_{\mu''}\{\xi_t B_0^r = 0\} \to 1,$$

and the a.s. convergence extends to μ. Applying this result to a countable, convergence-determining class of functions f (cf. Lemma 3.2.1 in [5]), we obtain the required a.s. vague convergence. If μ is bounded, then ξ_t has a.s. bounded support (cf. Corollary 6.8 in [17]), and the a.s. convergence remains valid in the weak sense.

To prove the convergence in L^1, we note that for any $f \in C_K^d$

$$\begin{aligned} \varepsilon^{2-d} E_\mu \xi_t^\varepsilon f &= \varepsilon^{2-d} \int P_\mu\{\xi_t B_x^\varepsilon > 0\}\, f(x)\, dx \\ &\to \int c_d\, (\mu * p_t)(x)\, f(x)\, dx = c_d\, E_\mu \xi_t f, \end{aligned} \tag{3.3}$$

by Theorem 5.3.(i) in [25]. Combining this with the a.s. convergence under P_μ and using Proposition 4.12 in [24], we obtain $E_\mu|\varepsilon^{2-d}\, \xi_t^\varepsilon f - c_d\, \xi_t f| \to 0$. For bounded μ, (5.14) extends to any $f \in C_b^d$ by dominated convergence based on Lemmas 4.1 and 4.2 (i) in [25], together with the fact that $\lambda^d(\mu * p_t) = \|\mu\| < \infty$ by Fubini's theorem.

3.4 Proofs of lemmas

Proof of Lemma 3.1:

(i) If v solves the evolution equation for ξ, that is,

$$\dot{v} = \frac{1}{2}\Delta v - v^2$$

then so does $\tilde{v}(t,x) = r^2 v(r^2 t, rx)$. Writing $\tilde{\xi}_t = r^{-2}\xi_{r^2 t}S_r$, $\tilde{\mu} = r^{-2}\mu S_r$, and $\tilde{f}(x) = r^2 f(rx)$, we get

$$E_\mu e^{-\tilde{\xi}_t \tilde{f}} = E_\mu e^{-\xi_{r^2 t} f} = e^{-\mu v_{r^2 t}} = e^{-\tilde{\mu}\tilde{v}_t} = E_{\tilde{\mu}} e^{-\xi_t \tilde{f}},$$

and so $\mathcal{L}_\mu(\tilde{\xi}) = \mathcal{L}_{\tilde{\mu}}(\xi)$, which is equivalent to (i).

(ii) Define the cluster kernel ν by $\nu_x = \mathcal{L}_x(\eta)$, $x \in \mathsf{R}^d$, and consider the cluster decomposition $\xi = \int m\,\zeta(dm)$, where ζ is a Poisson process with intensity $\mu\nu$ when $\xi_0 = \mu$. Here

$$r^{-2}\xi_{r^2 t}S_r = \int (r^{-2} m_{r^2 t} S_r)\,\zeta(dm), \quad r, t > 0.$$

Using (i) and the uniqueness of the Lévy measure, we obtain

$$(r^{-2}\mu S_r)\nu = \mu(\nu\{r^{-2}\hat{m}_{r^2} S_r \in \cdot\}),$$

which is equivalent to

$$r^{-2}\mathcal{L}_{\mu S_r}(\eta) = \mathcal{L}_{r^{-2}\mu S_r}(\eta) = \mathcal{L}_\mu(r^{-2}\hat{\eta}_{r^2} S_r). \qquad \Box$$

Proof of Lemma 3.2:

Under P_μ we have $\xi_t = \sum_i \eta_t^i$, where the η_t^i are conditionally independent clusters of age t rooted at the points of a Poisson process with intensity μ/t. For a cluster rooted at x, the hitting probability is $b_x = P_x\{\eta_t B > 0\}$. Hence (e.g. by Proposition 12.3 in [24]), the number of clusters hitting B is Poisson distributed with mean $\mu b/t$, and so $P_\mu\{\xi_t B = 0\} = \exp(-\mu b/t)$, which yields the asserted formulas. $\square$

Proof of Lemma 3.3:

$\square$

Proof of Lemma 3.4:

Let ζ_s be the Cox process of ancestors to ξ_t at time $s = t - h$, and write η_h^i for the associated h-clusters. Using Lemma 3.3, the conditional independence of the clusters, and the fact that $E_\mu \zeta_s^2 = h^{-2} E_\mu \xi_s^2$ outside the diagonal, we get with $p_h^\varepsilon(x) = P_x\{\eta_h B_0^\varepsilon > 0\}$

$$\begin{aligned}
E_\mu \kappa_h^\varepsilon(\kappa_h^\varepsilon - 1) &= E_\mu \sum\sum\nolimits_{i\neq j} 1\{\eta_h^i B_0^\varepsilon \wedge \eta_h^j B_0^\varepsilon > 0\} \\
&= \iint_{x\neq y} p_h^\varepsilon(x)\, p_h^\varepsilon(y)\, E_\mu \zeta_s^2(dx\,dy) \\
&\lesssim \varepsilon^{2(d-2)} \iint p_{h(\varepsilon)}(x)\, p_{h(\varepsilon)}(y)\, E_\mu \xi_s^2(dx\,dy).
\end{aligned}$$

By the formula of first moment, Fubini's theorem, and the semigroup property of (p_t), we get

$$\begin{aligned}
\int p_{h(\varepsilon)}(x)\, E_\mu \xi_s(dx) &= \int p_{h(\varepsilon)}(x)\, (\mu * p_s)(x)\, dx \\
&= \int \mu(du)\, (p_{h(\varepsilon)} * p_s)(u) = \mu p_{t(\varepsilon)}.
\end{aligned}$$

Next, we get by the formula of second moment, Fubini's theorem, the properties of (p_t), and the relations $t \le t_\varepsilon \le 2t - s$

$$\begin{aligned}
&\iint p_{h(\varepsilon)}(x)\, p_{h(\varepsilon)}(y)\, \mathrm{Cov}_\mu\, \xi_s(dx\,dy) \\
&\quad = 2 \iint p_{h(\varepsilon)}(x)\, p_{h(\varepsilon)}(y)\, dx\,dy \int \mu(du) \int_0^s dr \\
&\qquad\qquad \int p_r(v-u)\, p_{s-r}(x-v)\, p_{s-r}(y-v)\, dv
\end{aligned}$$

$$
\begin{aligned}
&= 2\int \mu(du)\int_0^s dr \int p_r(u-v)\left(p_{t(\varepsilon)-r}(v)\right)^2 dv \\
&\lesssim \int \mu(du)\int_0^s (t-r)^{-d/2}\,(p_r * p_{(t(\varepsilon)-r)/2})(u)\,dr \\
&= \int \mu(du)\int_0^s (t-r)^{-d/2}\,p_{(t(\varepsilon)+r)/2}(u)\,dr \\
&\lesssim \int p_t(u)\,\mu(du)\int_h^t r^{-d/2}\,dr \lesssim \mu p_t\, h^{1-d/2}.
\end{aligned}
$$

The assertion follows by combination of these estimates. □

Proof of Lemma 3.6:

(i) By Fubini's theorem and the definitions of η_h^ε and p_h^ε, we have

$$E_x\eta_h^\varepsilon f = E_x\int 1\{\eta_h B_u^\varepsilon > 0\}\, f(u)\,du = (p_h^\varepsilon * f)(x),$$

and so by independence

$$E\left[\sum\nolimits_i \eta_h^{i\varepsilon} f \,\Big|\, \xi\right] = \int \xi(dx)\, E_x\eta_h^\varepsilon f = \xi(p_h^\varepsilon * f). \tag{3.4}$$

Hence, by Fubini's theorem

$$E_\mu \sum\nolimits_i \eta_h^{i\varepsilon} f = E_\mu \xi(p_h^\varepsilon * f) = \mu(p_h^\varepsilon * f) = ((\mu * p_h^\varepsilon)\cdot\lambda^d)f.$$

(ii) First,

$$\mathrm{Var}_x(\eta_h^{K\varepsilon} f) \le E_x(\eta_h^{K\varepsilon} f)^2 \le E_x\|\eta_h^{K\varepsilon}\|^2\,\|f\|^2 = \|f\|^2 E_x\|\eta_h^{K\varepsilon}\|^2.$$

For $E_x\|\eta_h^{K\varepsilon}\|^2$, using Cauchy inequality and Lemma 3.3, we get

$$E_x\|\eta_h^{K\varepsilon}\|^2 \quad = \quad E_x\left(\int 1\{\eta_h^K B_y^\varepsilon > 0\}dy \int 1\{\eta_h^K B_z^\varepsilon > 0\}dz\right)$$

$$
\begin{aligned}
&= \int\int P_x\left(\{\eta_h^K B_y^\varepsilon > 0\} \cap \{\eta_h^K B_z^\varepsilon > 0\}\right) dydz \\
&\le \int\int (P_x\{\eta_h^K B_y^\varepsilon > 0\} P_x\{\eta_h^K B_z^\varepsilon > 0\})^{1/2} dydz \\
&\lesssim a_h \varepsilon^{d-2/\beta} \int\int (p_{2h}(y-x)p_{2h}(z-x))^{1/2} dydz \\
&\eqsim a_h \varepsilon^{d-2/\beta} h^{d/2} \int\int p_{4h}(y-x)p_{4h}(z-x) dydz \\
&= a_h \varepsilon^{d-2/\beta} h^{d/2}.
\end{aligned}
$$

Hence, by independence

$$
E_\mu \mathrm{Var}\left[\sum\nolimits_i \eta_h^{Ki\varepsilon} f | \zeta\right] = E_\mu \int \zeta(dx) \mathrm{Var}_x(\eta_h^{K\varepsilon} f) \lesssim a_h \varepsilon^{d-2/\beta} h^{d/2} \, \|f\|^2 \|\mu\|.
$$

□

Proof of Lemma 3.7:

Let $\kappa_h^\varepsilon(x)$ denote the number of subclusters of age h hitting B_x^ε at time t. Then Lemma 3.4 yields,

$$
\begin{aligned}
E_\mu \left\| \sum\nolimits_i \eta_h^{i\varepsilon} - \xi_t^\varepsilon \right\| &= E_\mu \int \left| \sum\nolimits_i 1\{\eta_h^i B_x^\varepsilon > 0\} - 1\{\xi B_x^\varepsilon > 0\} \right| dx \\
&= \int E_\mu (\kappa_h^\varepsilon(x) - 1)_+ \, dx \\
&\lesssim \varepsilon^{2(d-2)} \lambda^d \big(h^{1-d/2}(\mu * p_t) + (\mu * p_{2t})^2\big) \\
&\lesssim \varepsilon^{2(d-2)} \left(h^{1-d/2}\|\mu\| + t^{-d/2}\|\mu\|^2\right).
\end{aligned}
$$

□

Proof of Lemma 3.8:

Using (3.1) and Lemmas 3.1 (ii), 3.2, and 3.3, we get by dominated convergence

$$\lambda^d p_h^\varepsilon = h^{d/2} \lambda^d p_1^{\varepsilon/\sqrt{h}} \sim c_d \, h^{d/2} \, (\varepsilon/\sqrt{h})^{d-2} \lambda^d p_1 = c_d \, \varepsilon^{d-2} h. \tag{3.5}$$

Similarly, Lemma 3.3 yields for fixed $r > 0$ and a standard normal random vector γ in R^d

$$\begin{aligned} \varepsilon^{2-d} \, h^{-1} \int_{|x|>r} p_h^\varepsilon(x) \, dx \quad &\lesssim \quad \int_{|u|>r/\sqrt{h}} p_{l(\varepsilon)}(u) \, du \\ &= \quad P\Big\{ |\gamma| \, l_\varepsilon^{1/2} > r/\sqrt{h} \Big\} \to 0. \end{aligned} \tag{3.6}$$

By (3.5) it is enough to show that $\|\hat{p}_h^\varepsilon * f - f\| \to 0$ as h, $\varepsilon^2/h \to 0$, where $\hat{p}_h^\varepsilon = p_h^\varepsilon / \lambda^d p_h^\varepsilon$. Writing w_f for the modulus of continuity of f, we get

$$\begin{aligned} \|\hat{p}_h^\varepsilon * f - f\| \quad &= \quad \sup_x \left| \int \hat{p}_h^\varepsilon(u) \, (f(x-u) - f(x)) \, du \right| \\ &\le \quad \int \hat{p}_h^\varepsilon(u) \, w_f(|u|) \, du \\ &\le \quad w_f(r) + 2 \, \|f\| \int_{|u|>r} \hat{p}_h^\varepsilon(u) \, du, \end{aligned}$$

which tends to 0 as h, $\varepsilon^2/h \to 0$ and then $r \to 0$, by (3.6) and the uniform continuity of f.

$\square$

Chapter 4

Lebesgue Approximation of $(2, \beta)$-Superprocesses

4.1 Introduction

Throughout this chapter, we use μf to denote the integral of the function f with respect to the measure μ. By an (α, β)-*superprocess* (or (α, β)-*process*, for short) in R^d we mean a vaguely rcll, measure-valued strong Markov process $\xi = (\xi_t)$ in R^d satisfying $E_\mu e^{-\xi_t f} = e^{-\mu v_t}$ for suitable functions $f \geq 0$, where $v = (v_t)$ is the unique solution to the *evolution equation* $\dot{v} = \frac{1}{2}\Delta_\alpha v - v^{1+\beta}$ with initial condition $v_0 = f$. Here $\Delta_\alpha = -(-\Delta)^{\alpha/2}$ is the fractional Laplacian, $\alpha \in (0, 2]$ refers to the spatial motion, and $\beta \in (0, 1]$ refers to the branching mechanism. When $\alpha = 2$ and $\beta = 1$ we get the *Dawson–Watanabe superprocess* (*DW-process* for short), where the spatial motion is standard Brownian motion. General surveys of superprocesses include the excellent monographs and lecture notes [5, 15, 17, 32, 35, 42].

In this chapter we consider superprocesses with possibly infinite initial measures. Indeed, by the additivity property of superprocesses, we can construct the (α, β)-process ξ with any σ-finite initial measure μ. In Lemma 4.5 we show that ξ_t is a.s. locally finite for every $t > 0$ iff $\mu p_\alpha(t, \cdot) < \infty$ for all t, where $p_\alpha(t, x)$ denotes the transition density of a symmetric α-stable process in R^d. Note that when $\alpha = 2$, $p_2(t, x) = p_t(x)$ is the normal density in R^d.

For any measure μ on R^d and constant $\varepsilon > 0$, write μ^ε for the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp}\mu$. For a DW-process ξ in R^d with any finite initial measure, Tribe [48] showed that $\varepsilon^{2-d}\,\xi_t^\varepsilon \xrightarrow{w} c_d\,\xi_t$ a.s. as $\varepsilon \to 0$ when $d \geq 3$, where $\xrightarrow{w}$ denotes weak convergence and $c_d > 0$ is a constant depending on d. For a locally finite DW-process ξ in R^2, Kallenberg [25] showed that $\tilde{m}(\varepsilon)\,|\log \varepsilon|\,\xi_t^\varepsilon \xrightarrow{v} \xi_t$ a.s. as $\varepsilon \to 0$, where $\xrightarrow{v}$ denotes vague convergence and $\tilde{m}$ is a suitable normalizing function. Our main result in this chapter is Theorem 4.18, where we prove that, for a locally finite $(2, \beta)$-process ξ in R^d with $\beta < 1$ and

$d > 2/\beta$, $\varepsilon^{2/\beta-d}\,\xi_t^\varepsilon \xrightarrow{v} c_{\beta,d}\,\xi_t$ a.s. as $\varepsilon \to 0$, where $c_{\beta,d} > 0$ is a constant depending on β and d. In particular, the $(2,\beta)$-process ξ_t distributes its mass over $\operatorname{supp}\xi_t$ in a deterministic manner, which extends the corresponding property of DW-processes (cf. [17], page 115, or [42], page 212). See the end of the present chapter for a detailed explanation of this deterministic distribution property. For DW-processes, this property can also be inferred from some deep results involving the exact Hausdorff measure (cf. [7]). However, for any (α,β)-process ξ with $\alpha < 2$, $\operatorname{supp}\xi_t = \mathsf{R}^d$ or $\emptyset$ a.s. (cf. [18, 40]), and so the corresponding property fails. Our result shows that this property depends only on the spatial motion.

To prove our main result, we adapt the probabilistic approach for DW-processes from [25]. However, the finite variance of DW-processes plays a crucial role there. In order to deal with the infinite variance of $(2,\beta)$-processes with $\beta < 1$, we use a truncation of (α,β)-processes from [38], which will be further developed in Section 2 of the present chapter. By this truncation we may reduce our discussion to the truncated processes, where the variance is finite.

To adapt the probabilistic approach from [25] to study the truncated processes, we also need to develop some technical tools. Thus, in Section 3 we improve the upper bounds of hitting probabilities for $(2,\beta)$-processes with $\beta < 1$ and their truncated processes. As an immediate application, in Theorem 4.8 we improve some known extinction criteria of the $(2,\beta)$-process ξ by showing that the local extinction property $\xi_t \xrightarrow{d} 0$ and the seemingly stronger support property $\operatorname{supp}\xi_t \xrightarrow{d} \emptyset$ are equivalent. Then in Section 4 we derive some asymptotic results of these hitting probabilities. In particular, for the $(2,\beta)$-process ξ we show in Theorem 4.15 that $\varepsilon^{2/\beta-d} P_\mu\{\xi_t B_x^\varepsilon > 0\} \to c_{\beta,d}\,(\mu * p_t)(x)$, where B_x^r denotes an open ball around x of radius r, which extends the corresponding result for DW-processes (cf. Theorem 3.1(b) in [6]). Since the truncated processes do not have the scaling properties of the $(2,\beta)$-process, our general method is first to study the $(2,\beta)$-process, then to estimate the truncated processes by the $(2,\beta)$-process, in order to get the needed results for the truncated processes.

The extension of results of DW-processes to general (α, β)-processes is one of the major themes in the research of superprocesses. Since the spatial motion of the (α, β)-process is not continuous when $\alpha < 2$ and the (α, β)-process has infinite variance when $\beta < 1$, many extensions are not straightforward, and some may not even be valid. However, it turns out that several properties of the support of $(2, \beta)$-processes depend only on the spatial motion. These properties include short-time propagation of the support (cf. Theorem 9.3.2.2 in [5]) and Hausdorff dimension of the support (cf. Theorem 9.3.3.5 in [5]). Our result also belongs to that category.

In this chapter we are mainly using the notations in [25]. Recall that the double bars $\|\cdot\|$ denote the supremum norm when applied to functions and total variation when applied to signed measures. We also use relations such as $\overset{=}{\frown}$, $\overset{<}{\frown}$, and $\asymp$, where the first two mean equality and inequality up to a constant factor, and the last one is the combination of $\overset{<}{\frown}$ and $\overset{>}{\frown}$. Other notation will be explained whenever it occurs.

4.2 Truncated superprocesses and local finiteness

Although our main result of the present chapter is about $(2, \beta)$-processes, in this section we discuss the truncation and local finiteness of all (α, β)-processes, due to their independent interests.

It is well known that the $(\alpha, 1)$-process has weakly continuous sample paths. By contrast, the (α, β)-process ξ with $\beta < 1$ has only weakly rcll sample paths with jumps of the form $\Delta\xi_t = r\delta_x$, for some $t > 0$, $r > 0$, and $x \in \mathsf{R}^d$. Let

$$N_\xi(dt, dr, dx) = \sum_{(t,r,x):\ \Delta\xi_t = r\delta_x} \delta_{(t,r,x)}.$$

Clearly the point process N_ξ on $\mathsf{R}_+ \times \mathsf{R}_+ \times \mathsf{R}^d$ records all information about the jumps of ξ. By the proof of Theorem 6.1.3 in [5], we know that N_ξ has compensator measure

$$\hat{N}_\xi(dt, dr, dx) = c_\beta(dt)r^{-2-\beta}(dr)\xi_t(dx), \tag{4.1}$$

where c_β is a constant depending on β. Due to all the "big" jumps, ξ_t has infinite variance. Some methods for $(\alpha, 1)$-processes, which rely on the finite variance of the processes, are not directly applicable to (α, β)-processes with $\beta < 1$.

In [38], Mytnik and Villa introduced a truncation method for (α, β)-processes with $\beta < 1$, which can be used to study (α, β)-processes with $\beta < 1$, especially to extend results of $(\alpha, 1)$-processes to (α, β)-processes with $\beta < 1$. Specifically, for the (α, β)-process ξ with $\beta < 1$, we define the stopping time $\tau_K = \inf\{t > 0 : \|\Delta\xi_t\| > K\}$ for any constant $K > 0$, where $\inf \emptyset = \infty$ as usual. When $\Delta\xi_t = r\delta_x$, we see that $\|\Delta\xi_t\| = r$. Clearly τ_K is the time when ξ has the first jump greater than K. For any finite initial measure μ, they proved that one can define ξ and a weakly rcll, measure-valued Markov process ξ^K (which is Y^K on page 485 of [38]) on a common probability space such that $\xi_t = \xi_t^K$ for $t < \tau_K$. Intuitively, ξ^K euqals ξ minus all masses produced by jumps greater than K along with the future evolution of those masses. In this chapter, we call ξ^K the *truncated K-process* of ξ. Since all "big" jumps are omitted, ξ_t^K has finite variance. They also proved that ξ_t^K and ξ_t agree asymptotically as $K \to \infty$. We give a different proof of this result, since similar ideas will also be used at several crucial stages later. We write $P_\mu\{\xi \in \cdot\}$ for the distribution of ξ with initial measure μ.

Lemma 4.1 *Fix any finite μ and $t > 0$. Then $P_\mu\{\tau_K > t\} \to 1$ as $K \to \infty$.*

Proof: If $\tau_K \leq t$, then ξ has at least one jump greater than K before time t. Noting that $N_\xi([0, t], (K, \infty), \mathsf{R}^d)$ is the number of jumps greater than K before time t, we get by

Theorem 25.22 of [24] and (4.1),

$$\begin{aligned} P_\mu\{\tau_K \le t\} &\le E_\mu N_\xi\big([0,t],(K,\infty),\mathsf{R}^d\big) \\ &= E_\mu \hat{N}_\xi\big([0,t],(K,\infty),\mathsf{R}^d\big) \\ &\asymp K^{-1-\beta} E_\mu \int_0^t \|\xi_s\| ds = t\|\mu\| K^{-1-\beta} \to 0 \end{aligned}$$

as $K \to \infty$, where the last equation holds by $E_\mu\|\xi_s\| = \|\mu\|$. $\square$

Using Lemma 1 of [38] and a recursive construction, we can prove that $\xi_t^K(\omega) \le \xi_t(\omega)$ for any t and ω. So indeed, ξ^K is a "truncation" of ξ.

Lemma 4.2 *We can define ξ and ξ^K on a common probability space such that:*

(i) *ξ is an (α,β)-process with $\beta < 1$ and a finite initial measure μ, and ξ^K is its truncated K-process, which has no jumps greater than K,*

(ii) *$\xi_t(\omega) \ge \xi_t^K(\omega)$ for any t and ω,*

(iii) *$\xi_t(\omega) = \xi_t^K(\omega)$ for $t < \tau_K(\omega)$.*

Proof: Let $\xi_{m,n}(t)$ denote the process $\xi_{m,n}$ at time t. Use $D([0,\infty),\hat{\mathcal{M}}_d)$ as our Ω, the space of rcll functions from $[0,\infty)$ to $\hat{\mathcal{M}}_d$, where $\hat{\mathcal{M}}_d$ is the set of finite measures on R^d. We endow Ω with the Skorohod J_1-topology. Let $\mathcal{A} = B(\Omega)$.

Let $\zeta_1(t,\omega) = \omega(t)$ be an (α,β)-process defined on $(\Omega,\mathcal{A},P)$ with initial measure μ, and define $\tau_{K_1} = \inf\{t > 0 : \|\Delta\zeta_1(t)\| > K\}$. Then define a kernel u from $\hat{\mathcal{M}}_d$ to Ω such that $u(\nu,\cdot)$ is the distribution of an (α,β)-process with initial measure ν, and a kernel u^K from $\hat{\mathcal{M}}_d$ to Ω such that $u^K(\nu,\cdot)$ is the distribution of the truncated K-process of an (α,β)-process with initial measure ν. By Lemma 6.9 in [24], we can define $\zeta_{1,\infty}$ to be an (α,β)-process with initial measure $\zeta_1(\tau_{K_1})$ on an extension of $(\Omega,\mathcal{A},P)$, and $\zeta'_{1,\infty}$ to be the truncated K-process of an (α,β)-process with initial measure $\zeta_1(\tau_{K_1}^-)$. Now define ξ_1 and ξ_1^K by

$$\xi_1(t) = \begin{cases} \zeta_1(t), & t < \tau_{K_1}, \\ \zeta_{1,\infty}(t-\tau_{K_1}), & t \geq \tau_{K_1}, \end{cases}$$

$$\xi_1^K(t) = \begin{cases} \zeta_1(t), & t < \tau_{K_1}, \\ \zeta'_{1,\infty}(t-\tau_{K_1}), & t \geq \tau_{K_1}. \end{cases}$$

By the strong Markov property of (α,β)-processes and the above construction, we can verify that ξ_1 is an (α,β)-process. By Lemma 1 in [38], ξ_1^K is the truncated K-process of an (α,β)-process. Moreover, ξ_1 and ξ_1^K satisfy conditions (ii) and (iii) on $[0,\tau_{K_1})$.

Let u' be a kernel from $\hat{\mathcal{M}}_d \times \hat{\mathcal{M}}_d$ to $\mathcal{A}\times\mathcal{A}$ such that $u'(\nu,\nu',\cdot,\cdot)$ is the distribution of a pair of two independent (α,β)-processes with initial measures ν and ν' respectively. Define $(\zeta_{2,0},\zeta_{2,1})$ with distribution

$$u'\left(\xi_1^K(\tau_{K_1}^-), \xi_1(\tau_{K_1}) - \xi_1^K(\tau_{K_1}^-), \cdot, \cdot\right).$$

Let $\zeta_2 = \zeta_{2,0} + \zeta_{2,1}$, $\zeta'_2 = \zeta_{2,0}$, and $\tau_{K_2} = \inf\{t > 0 : \|\Delta\zeta_2(t)\| > K\}$. Let $\zeta_{2,\infty}$ be an (α,β)-process with initial measure $\zeta_2(\tau_{K_2})$, and let $\zeta'_{2,\infty}$ be the truncated K-process of an (α,β)-process with initial measure $\zeta'_2(\tau_{K_2}^-)$. Now define ξ_2 and ξ_2^K by

$$\xi_2(t) = \begin{cases} \xi_1(t), & t < \tau_{K_1}, \\ \zeta_2(t-\tau_{K_1}), & \tau_{K_1} \leq t < \tau_{K_1}+\tau_{K_2}, \\ \zeta_{2,\infty}(t-\tau_{K_1}-\tau_{K_2}), & t \geq \tau_{K_1}+\tau_{K_2}, \end{cases}$$

$$\xi_2^K(t) = \begin{cases} \xi_1^K(t), & t < \tau_{K_1}, \\ \zeta'_2(t-\tau_{K_1}), & \tau_{K_1} \leq t < \tau_{K_1}+\tau_{K_2}, \\ \zeta'_{2,\infty}(t-\tau_{K_1}-\tau_{K_2}), & t \geq \tau_{K_1}+\tau_{K_2}. \end{cases}$$

Similarly, ξ_2 is an (α, β)-process and ξ_2^K is the truncated K-process of an (α, β)-process. They satisfy conditions (ii) and (iii) on $[0, \tau_{K_1} + \tau_{K_2})$.

Continue the above construction: For every n, define ξ_n and ξ_n^K such that ξ_n is an (α, β)-process, ξ_n^K it the truncated K-process of an (α, β)-process, and they satisfy conditions (ii) and (iii) on $[0, \sum_{k=1}^n \tau_{K_k})$.

It suffices to prove that $\sum_{k=1}^\infty \tau_{K_k} = \infty$ a.s. Suppose $P(\sum_{k=1}^\infty \tau_{K_k} < \infty) > 0$. Then there exist t and a such that $P(\sum_{k=1}^\infty \tau_{K_k} < t) = a > 0$. Since for every n, ξ_n is an (α, β)-process with initial measure μ, we get

$$an \leq E_\mu \hat{N}_{\xi_n}\left([0, t], (K, \infty), \mathsf{R}^d\right).$$

Noting that by (4.1) $E_\mu \hat{N}_{\xi_n}([0, t], (K, \infty), \mathsf{R}^d)$ is the same finite constant for different n, we get a contradiction. So $\sum_{k=1}^\infty \tau_{K_k} = \infty$ a.s. $\square$

Just as the DW-process, the (α, β)-process ξ and its truncated K-process ξ^K also have cluster structures (cf. Corollary 11.5.3 in [5], or Section 3 in [7], especially page 41 there). Specifically, for any fixed t, ξ_t is a Cox cluster process, such that the "ancestors" of ξ_t at time $s = t - h$ form a Cox process directed by $(\beta h)^{-1/\beta} \xi_s$, and the generated h-clusters η_h^i are conditionally independent and identically distributed apart from shifts. For the truncated K-process ξ^K, the situation is similar, except that the clusters are different (because of the truncation) and the term $(\beta h)^{-1/\beta}$ for ξ needs to be replaced by $a_K(h)$ (or a_h, when K is fixed). Use $\eta_h^{K,i}$ (or η_h^{Ki}) to denote the generated h-clusters of ξ^K. Write $P_x\{\eta_t \in \cdot\}$ for the distribution of η_t centered at $x \in \mathsf{R}^d$, and define $P_\mu\{\eta_t \in \cdot\} = \int \mu(dx) P_x\{\eta_t \in \cdot\}$. The following comparison of $a_K(h)$ and $(\beta h)^{-1/\beta}$, although not used explicitly in the present chapter, should be useful in other applications of the truncation method.

Lemma 4.3 *Fix any $K > 0$. Then as $h \to 0$,*

$$(\beta h)^{1/\beta} \leq a_K(h) \leq 2(\beta h)^{1/\beta}.$$

Proof: From Lemma 3.4 of [7] we know that

$$(\beta h)^{1/\beta} = \lim_{\theta \to \infty} 1/v_0(h, \theta),$$

where $v_0(h, \theta)$ is the solution of $\dot{v} = -v^{1+\beta}$ with initial condition $v \equiv \theta$, and

$$a_K(h) = \lim_{\theta \to \infty} 1/v_1(h, \theta),$$

where $v_1(h, \theta)$ is the solution of (1.12) in [38] with initial condition $v \equiv \theta$. Define $M_K(\lambda) = C_\beta(K)\lambda + \Phi^K(\lambda)$, where $C_\beta(K)$ and Φ^K are such as in (1.12) of [38]. Then M_K satisfies

$$\lambda^{1+\beta} \leq M_K(\lambda) \text{ and } \lim_{\lambda \to \infty} \frac{M_K(\lambda)}{\lambda^{1+\beta}} = 1.$$

Clearly it is enough to show that $(1/2)v_0(h, \theta) \leq v_1(h, \theta) \leq v_0(h, \theta)$ as $h \to 0$ and $\theta \to \infty$. This follows from the above properties of M_K. $\square$

Unlike the normal densities, we have no explicit expressions for the transition densities of symmetric α-stable processes when $\alpha < 2$. However, a simple estimate of $p_\alpha(t, x)$ is enough for our needs.

Lemma 4.4 *Let $p_\alpha(t, x)$, $\alpha \in (0, 2]$, $t > 0$, and $x \in \mathsf{R}^d$, denote the transition densities of a symmetric α-stable process on R^d. Then for any fixed α and d,*

$$p_\alpha(t, x + y) \lesssim p_\alpha(2t, x), \quad |y|^\alpha \leq t.$$

Proof: First let $\alpha = 2$. Note that $p_2(t,x) = p_t(x)$ is the standard normal density on R^d. For $|x| \le 4\sqrt{t}$, trivially $p_t(x+y) \lesssim p_{2t}(x)$. For $|x| > 4\sqrt{t}$, it suffices to check that

$$-\frac{|x+y|^2}{2t} \le -\frac{|x|^2}{4t},$$

that is, $2|x+y|^2 \ge |x|^2$, which follows easily from $|x| \ge 4|y|$.

Now let $\alpha < 2$. By the arguments after Remark 5.3 of [2],

$$p_\alpha(t,x) \asymp \left(t^{-d/\alpha} \wedge \frac{t}{|x|^{d+\alpha}}\right). \tag{4.2}$$

Choose $K > 2^{1/\alpha}$ to satisfy $1 \le 2(1-1/K)^{d+\alpha}$. Since $|y| \le t^{1/\alpha}$, we have for $|x| > Kt^{1/\alpha}$,

$$\frac{t}{|x+y|^{d+\alpha}} \le \frac{2t}{|x|^{d+\alpha}}.$$

Noticing also that $(2t)/|x|^{d+\alpha} < (2t)^{-d/\alpha}$ for $|x| > Kt^{1/\alpha}$, we get $p_\alpha(t,x+y) \lesssim p_\alpha(2t,x)$ for $|y| \le t^{1/\alpha}$ and $|x| > Kt^{1/\alpha}$. The same inequality holds trivially for $|y| \le t^{1/\alpha}$ and $|x| \le Kt^{1/\alpha}$. □

Using Lemma 4.2 and Lemma 4.4, we can generalize Lemma 3.2 in [25] to any (α,β)-process.

Lemma 4.5 *Let ξ be an (α,β)-process in R^d, $\alpha \in (0,2]$ and $\beta \in (0,1]$, and fix any σ-finite measure μ. Then for any fixed $t > 0$, the following two conditions are equivalent:*

(i) *ξ_t is locally finite a.s. P_μ,*

(ii) *$E_\mu \xi_t$ is locally finite.*

Furthermore, (i) *and* (ii) *hold for every $t > 0$ iff*

(iii) *$\mu p_\alpha(t,\cdot) < \infty$ for all $t > 0$,*

and if $\alpha < 2$, then (iii) *is equivalent to*

(iv) $\mu p_\alpha(t,\cdot) < \infty$ *for some* $t > 0$.

Proof: The formulas for $E_\mu \xi_t$ and $E_\mu \xi_t^2$ (when $\beta < 1$), well known for finite μ, as well as the formulas in Lemma 3 of [38], extend by monotone convergence to any σ-finite measure μ. We also need the simple inequality that for any fixed $\alpha < 2$, s, and t,

$$p_\alpha(s,x) \asymp p_\alpha(t,x). \tag{4.3}$$

To prove it, use (4.2) and consider three cases: $|x| \le (s \wedge t)^{1/\alpha}$, $|x| \ge (s \vee t)^{1/\alpha}$, and $(s \wedge t)^{1/\alpha} < |x| < (s \vee t)^{1/\alpha}$.

If $\alpha = 2$ and $\beta = 1$, then this is Lemma 3.2 of [25]. For $\alpha < 2$ and $\beta = 1$, using Lemma 4.4 and (4.3) we can proceed as in Lemma 3.2 of [25]. For example, for any fixed $t > 0$ and $x \in \mathsf{R}^d$, $p_\alpha(t, x-u) \lesssim p_\alpha(|x|^{1/\alpha}, x-u) \lesssim p_\alpha(2|x|^{1/\alpha}, -u) = p_\alpha(2|x|^{1/\alpha}, u)$ yields $\mu * p_\alpha(t,\cdot)(x) < \infty$.

Now assume $\beta < 1$. Condition (ii) clearly implies (i). Conversely, suppose that $E_\mu \xi_t B = \infty$ for some B. Then $E_\mu \xi_t^K B = \infty$ for any fixed $K > 0$ by Lemma 3 of [38]. Also, we get by Lemma 3 of [38],

$$P_\mu\left\{\frac{\xi_t^K B}{E_\mu \xi_t^K B} > r\right\} \ge (1-r)^2 \frac{(E_\mu \xi_t^K B)^2}{E_\mu (\xi_t^K B)^2} \ge \frac{(1-r)^2}{1 + ct\,(E_\mu \xi_t^K B)^{-1}}$$

for any $r \in (0,1)$. Arguing as in the proof of Lemma 3.2 in [25], we get $\xi_t^K B = \infty$ a.s., and so $\xi_t B = \infty$ a.s. by Lemma 4.2. In particular, this shows that (i) implies (ii). To prove the equivalence of (ii) and (iii), again using Lemma 4.4 and (4.3) we can proceed as in Lemma 3.2 of [25]. The last assertion is obvious from (4.3). □

4.3 Hitting bounds and neighborhood measures

From now on we consider only $(2,\beta)$-processes. The Lebesgue approximation depends crucially on estimates of the hitting probability $P_\mu\{\xi_t B_0^\varepsilon > 0\}$. In this section, we first estimate $P_\mu\{\xi_t B_0^\varepsilon > 0\}$ and $P_\mu\{\xi_t^K B_0^\varepsilon > 0\}$. Then we use these estimates to study multiple hitting and neighborhood measures of the clusters η_h^K associated with the truncated K-process ξ^K. We begin with a well-known relationship between the hitting probabilities of ξ_t and η_t, which can be proved as in Lemma 4.1 of [25].

Lemma 4.6 *Let the $(2,\beta)$-process ξ in R^d with associated clusters η_t be locally finite under P_μ, let ξ^K be its truncated K-process with associated clusters η_t^K, and fix any $B \in \hat{\mathcal{B}}^d$. Then*

$$\begin{aligned}
P_\mu\{\eta_t B > 0\} &= -(\beta t)^{1/\beta} \log\left(1 - P_\mu\{\xi_t B > 0\}\right), \\
P_\mu\{\xi_t B > 0\} &= 1 - \exp\left(-(\beta t)^{-1/\beta} P_\mu\{\eta_t B > 0\}\right), \\
P_\mu\{\eta_t^K B > 0\} &= -a_t \log\left(1 - P_\mu\{\xi_t^K B > 0\}\right), \\
P_\mu\{\xi_t^K B > 0\} &= 1 - \exp\left(-a_t^{-1} P_\mu\{\eta_t^K B > 0\}\right).
\end{aligned}$$

In particular, $P_\mu\{\xi_t B > 0\} \sim (\beta t)^{-1/\beta} P_\mu\{\eta_t B > 0\}$ and $P_\mu\{\xi_t^K B > 0\} \sim a_t^{-1} P_\mu\{\eta_t^K B > 0\}$ as either side tends to 0.

Upper and lower bounds of $P_\mu\{\xi_t B_0^\varepsilon > 0\}$ have been obtained by Delmas [9], using a subordinated Brownian snake approach. However, in this chapter we need the following improved upper bound.

Lemma 4.7 *Let η_t be the clusters of a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$, let η_t^K be the clusters of ξ^K, the truncated K-process of ξ, and consider a σ-finite measure μ on R^d. Then for $0 < \varepsilon \le \sqrt{t}$,*

(i) $\mu p_{t'} \lesssim \varepsilon^{2/\beta - d} (\beta t)^{-1/\beta} P_\mu\{\eta_t B_0^\varepsilon > 0\} \lesssim \mu p_{2t}$, *where $t' = \beta t/(1+\beta)$,*

(ii) $\varepsilon^{2/\beta - d} a_t^{-1} P_\mu\{\eta_t^K B_0^\varepsilon > 0\} \lesssim \mu p_{2t}$.

Proof: (i) From the proof of Theorem 2.3 in [9] we know that

$$P_x\{\xi_t B_0^\varepsilon > 0\} = 1 - \exp(-N_x\{Y_t B_0^\varepsilon > 0\}),$$

where N_x and Y_t are defined in Section 4.2 of [9]. Comparing this with Lemma 5.4 yields

$$(\beta t)^{-1/\beta} P_x\{\eta_t B_0^\varepsilon > 0\} = N_x\{Y_t B_0^\varepsilon > 0\}.$$

By Proposition 6.2 in [9] we get the lower bound. For our upper bound, we will now improve the upper bound in Proposition 6.1 of [9].

For $0 < \varepsilon/2 < \sqrt{t}$, define

$$\begin{aligned} \Delta &= \{(r,y) \in \mathsf{R}^+ \times \mathsf{R}^d,\ r < t, |y| > \varepsilon/2\} \\ &\quad \bigcup \{(r,y) \in \mathsf{R}^+ \times \mathsf{R}^d,\ r < t - \varepsilon^2/16, |y| \le \varepsilon/2\}. \end{aligned}$$

Following the proof of Proposition 6.1 in [9], we have

$$(\beta t)^{-1/\beta} P_x\{\eta_t B_0^\varepsilon > 0\} \precsim \varepsilon^{-2/\beta} P_0\{\gamma_s \in B_x^{\varepsilon/2} \text{ for some } s \in [t - \varepsilon^2/16, t)\},$$

where γ is a standard Brownian motion. Define

$$T = \inf\{s \ge t - \varepsilon^2/16 : \gamma_s \in B_x^{\varepsilon/2}\},$$

where $\inf \emptyset = \infty$ as usual. Then $\{T < t\} = \{\gamma_s \in B_x^{\varepsilon/2} \text{ for some } s \in [t - \varepsilon^2/16, t)\}$. To get our upper bound, it remains to show that

$$P_0\{T < t\} \precsim \varepsilon^d p_{2t}(x).$$

To prove this, we need the elementary fact that for any $x \in \mathsf{R}^d$, $\varepsilon > 0$, $y \in B_x^{\varepsilon/2}$, and $s \le s' = \varepsilon^2/16$,

$$P_y\{\gamma_s \notin B_x^\varepsilon\} \le P_z\{\gamma_{s'} \notin B_x^\varepsilon\} \precsim P_z\{\gamma_{s'} \in B_x^\varepsilon\} \le P_y\{\gamma_s \in B_x^\varepsilon\},$$

where z is a point on the surface of $B_x^{\varepsilon/2}$, and the second relation holds since $P_z\{\gamma_{s'} \notin B_x^\varepsilon\}$ and $P_z\{\gamma_{s'} \in B_x^\varepsilon\}$ are both positive constants. Now return to $P_0\{T < t\}$. Noting $t - T \le \varepsilon^2/16$ on $\{T < t\}$, we get

$$\begin{aligned}
P_0\{T < t\} &= P_0\{T < t, \gamma_t \in B_x^\varepsilon\} + P_0\{T < t, \gamma_t \notin B_x^\varepsilon\} \\
&= P_0\{T < t, \gamma_t \in B_x^\varepsilon\} + E_0\{P_{\gamma_T}\{\gamma_{t-T} \notin B_x^\varepsilon\}, T < t\} \\
&\precsim P_0\{T < t, \gamma_t \in B_x^\varepsilon\} + E_0\{P_{\gamma_T}\{\gamma_{t-T} \in B_x^\varepsilon\}, T < t\} \\
&= P_0\{T < t, \gamma_t \in B_x^\varepsilon\} + P_0\{T < t, \gamma_t \in B_x^\varepsilon\} \\
&\precsim P_0\{\gamma_t \in B_x^\varepsilon\} \precsim \varepsilon^d p_{2t}(x),
\end{aligned}$$

where the second and fourth relations hold by the strong Markov property of Brownian motion and the last relation holds by Lemma 4.4.

(ii) This is obvious from (i), Lemma 4.2, and Lemma 5.4. □

As an immediate application of the improved upper bound, we may improve some known extinction criteria for $(2, \beta)$-processes in R^d with $\beta < 1$ and $d > 2/\beta$. This extends Theorem 4.5 of [25] for DW-processes of dimension $d \ge 2$. Note that the special case of convergence of random measures $\xi_t \xrightarrow{d} 0$ is equivalent to $\xi_t B \xrightarrow{P} 0$ for any bounded Borel set B. Convergence of closed random sets is defined as usual with respect to the Fell topology (cf. [24], pp. 324, 566). However, in this chapter we need only the special case of convergence to the empty set $\operatorname{supp} \xi_t \xrightarrow{d} \emptyset$, which is equivalent to $1\{\xi_t B > 0\} \xrightarrow{P} 0$ for any bounded Borel set B.

Theorem 4.8 *Let ξ be a locally finite $(2,\beta)$-process in R^d, $\beta < 1$ and $d > 2/\beta$, with arbitrary initial distribution. Then these conditions are equivalent as $t \to \infty$:*

(i) $\xi_t \xrightarrow{d} 0$,

(ii) $\mathrm{supp}\, \xi_t \xrightarrow{d} \emptyset$,

(iii) $\xi_0 p_t \xrightarrow{P} 0$.

Proof: By Lemma 5.4 and Lemma 5.5(i) we get for any fixed r

$$P_\mu\{\xi_t B_0^r > 0\} \le (\beta t)^{-1/\beta} P_\mu\{\eta_t B_0^r > 0\} \lesssim \mu p_{2t},$$

and so $P_\mu\{\xi_t B_0^r > 0\} \lesssim \mu p_{2t} \wedge 1$. For a general initial distribution,

$$P\{\xi_t B_0^r > 0\} \lesssim E(\xi_0 p_{2t} \wedge 1),$$

which shows that (iii) implies (ii). Since clearly (ii) implies (i), it remains to prove that (i) implies (iii).

Let ξ be locally finite under P_μ. We first choose $f \in C_c^{++}(\mathsf{R}^d)$ with $\mathrm{supp} f \in B_0^1$, where $C_c^{++}(\mathsf{R}^d)$ is such as in Proposition 2.6 of [29]. Clearly $\xi_t f \xrightarrow{P} 0$ if $\xi_t B_0^1 \xrightarrow{P} 0$. By dominated convergence

$$\exp(-\mu v_t) = E_\mu \exp(-\xi_t f) \to 1,$$

and so $\mu v_t \to 0$. By Proposition 2.6 of [29], we have for t large enough

$$p_{t/2}(x) \asymp \phi(t/2, x) \lesssim v_t(x),$$

where ϕ is defined in (1.15) of [29] (on page 1061, see also (1.17) and (1.18) there). So $\mu p_{t/2} \to 0$. For general ξ_0, we may proceed as in the proof of Theorem 4.5 in [25]. □

The following simple fact is often useful to extend results for finite initial measures μ to the general case. Here $\hat{\mathcal{B}}^d$ denotes the space of bounded sets in the Borel σ-algebra $\mathcal{B}^d$.

Lemma 4.9 *Let the $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$ be locally finite under P_μ, and suppose that $\mu \geq \mu_n \downarrow 0$. Then $P_{\mu_n}\{\xi_t B > 0\} \to 0$ as $n \to \infty$ for any fixed $t > 0$ and $B \in \hat{\mathcal{B}}^d$.*

Proof: Follow the proof of Lemma 4.3 in [25], then use Lemma 4.5, Lemma 5.4, and Lemma 5.5(i). $\square$

As in [25] we need to estimate the probability that a ball in R^d is hit by more than one subcluster of the truncated K-process ξ^K. This is where the truncation of ξ is needed.

Lemma 4.10 *Fix any $K > 0$. Let ξ^K be the truncated K-process of a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$. For any $t \geq h > 0$ and $\varepsilon > 0$, let κ_h^ε be the number of h-clusters of ξ_t^K hitting B_0^ε at time t. Then for $\varepsilon^2 \leq h \leq t$,*

$$E_\mu \kappa_h^\varepsilon(\kappa_h^\varepsilon - 1) \lesssim \varepsilon^{2(d-2/\beta)} \left(h^{1-d/2} \mu p_t + (\mu p_{2t})^2 \right).$$

Proof: Follow Lemma 4.4 in [25], then use Lemma 3 of [38] and Lemma 5.5(ii). $\square$

Now we consider the neighborhood measures of the clusters η_h^K associated with the truncated K-process ξ^K. For any measure μ on R^d and constant $\varepsilon > 0$, we define the associated *neighborhood measure* μ^ε as the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp}\mu$, so that μ^ε has Lebesgue density $1\{\mu B_x^\varepsilon > 0\}$. Let $p_h^{K,\varepsilon}(x) = P_x\{\eta_h^K B_0^\varepsilon > 0\}$, where the η_h^K are clusters of ξ^K. Write $p_h^{K,\varepsilon}(x) = p_h^{K\varepsilon}(x)$ and $(\eta_h^{K,i})^\varepsilon = \eta_h^{Ki\varepsilon}$ for convenience.

Lemma 4.11 *Let ξ^K be the truncated K-process of a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$. Let the η_h^{Ki} be conditionally independent h-clusters of ξ^K, rooted at the points of a Poisson process ζ with $E\zeta = \mu$. Fix any measurable function $f \geq 0$ on R^d. Then,*

(i) $E_\mu \sum_i \eta_h^{Ki\varepsilon} = (\mu * p_h^{K\varepsilon}) \cdot \lambda^d$,

(ii) $E_\mu \mathrm{Var}\left[\sum_i \eta_h^{Ki\varepsilon} f|\zeta\right] \lesssim a_h \varepsilon^{d-2/\beta} h^{d/2}\, \|f\|^2 \|\mu\|$ *for* $\varepsilon^2 \le h$.

Proof: (i) Follow the proof of Lemma 6.2 (i) in [25].

(ii) First,

$$\mathrm{Var}_x(\eta_h^{K\varepsilon} f) \le E_x(\eta_h^{K\varepsilon} f)^2 \le E_x \|\eta_h^{K\varepsilon}\|^2\, \|f\|^2 = \|f\|^2 E_x \|\eta_h^{K\varepsilon}\|^2.$$

For $E_x\|\eta_h^{K\varepsilon}\|^2$, using Cauchy inequality and Lemma 5.5(ii), we get

$$\begin{aligned}
E_x\|\eta_h^{K\varepsilon}\|^2 &= E_x\left(\int 1\{\eta_h^K B_y^\varepsilon > 0\} dy \int 1\{\eta_h^K B_z^\varepsilon > 0\} dz\right) \\
&= \int\int P_x\left(\{\eta_h^K B_y^\varepsilon > 0\} \cap \{\eta_h^K B_z^\varepsilon > 0\}\right) dydz \\
&\le \int\int (P_x\{\eta_h^K B_y^\varepsilon > 0\} P_x\{\eta_h^K B_z^\varepsilon > 0\})^{1/2} dydz \\
&\lesssim a_h \varepsilon^{d-2/\beta} \int\int (p_{2h}(y-x) p_{2h}(z-x))^{1/2} dydz \\
&\eqsim a_h \varepsilon^{d-2/\beta} h^{d/2} \int\int p_{4h}(y-x) p_{4h}(z-x) dydz \\
&= a_h \varepsilon^{d-2/\beta} h^{d/2}.
\end{aligned}$$

Hence, by independence

$$E_\mu \mathrm{Var}\left[\sum\nolimits_i \eta_h^{Ki\varepsilon} f|\zeta\right] = E_\mu \int \zeta(dx) \mathrm{Var}_x(\eta_h^{K\varepsilon} f) \lesssim a_h \varepsilon^{d-2/\beta} h^{d/2}\, \|f\|^2 \|\mu\|.$$

□

We also need to estimate the overlap between subclusters.

Lemma 4.12 *Let* ξ^K *be the truncated* K*-process of a* $(2,\beta)$*-process* ξ *in* R^d *with* $\beta < 1$ *and* $d > 2/\beta$*. For any fixed* $t > 0$*, let* η_h^{Ki} *denote the subclusters in* ξ^K *of age* $h > 0$*. Fix any*

$\mu \in \hat{\mathcal{M}}_d$. *Then as* $\varepsilon^2 \le h \to 0$,

$$E_\mu \Big\| \sum\nolimits_i \eta_h^{Ki\varepsilon} - \xi_t^{K\varepsilon} \Big\| \lesssim \varepsilon^{2(d-2/\beta)} h^{1-d/2}.$$

Proof: Follow the proof of Lemma 6.3(i) in [25], then use Lemma 5.5(ii). □

4.4 Hitting asymptotics

For a DW-process ξ of dimension $d \ge 3$, we know from Theorem 3.1(b) of Dawson, Iscoe, and Perkins [6] that, as $\varepsilon \to 0$,

$$\varepsilon^{2-d} P_\mu\{\xi_t B_x^\varepsilon > 0\} \to c_d\,(\mu * p_t)(x),$$

uniformly for bounded $\|\mu\|$, bounded t^{-1}, and $x \in \mathsf{R}^d$. A similar result for DW-processes of dimension $d = 2$ is Theorem 5.3(ii) of [25]. In this section, using Lemma 5.5(i), we can prove the corresponding result for $(2, \beta)$-processes in R^d with $\beta < 1$ and $d > 2/\beta$.

First we fix a continuous function f on R^d such that $0 < f(x) \le 1$ for $x \in B_0^1$ and $f(x) = 0$ otherwise. Let v_λ be the solution of $\dot{v} = \frac{1}{2}\Delta v - v^{1+\beta}$ with initial condition $v(0) = \lambda f$. Since v_λ is increasing in λ, we can define $v_\infty = \lim_{\lambda\to\infty} v_\lambda$. Using Lemma 5.5(i), we can get an upper bound of v_∞, similar to Lemma 3.2 in [6].

Lemma 4.13 *For any* $t \ge 1$ *and* $x \in \mathsf{R}^d$, $v_\infty(t, x) \lesssim p(2t, x)$.

Proof: Letting $\lambda \to \infty$ in $E_x \exp(-\xi_t \lambda f) = \exp[-v_\lambda(t, x)]$, we get

$$P_x\{\xi_t B_0^1 > 0\} = 1 - \exp[-v_\infty(t, x)].$$

Comparing this with Lemma 5.4 yields

$$v_\infty(t, x) = (\beta t)^{-1/\beta} P_x\{\eta_t B_0^1 > 0\}. \tag{4.4}$$

Now Lemma 4.13 follows from Lemma 5.5(i). □

As in Lemma 3.3 of [6], we can apply a PDE result to get the uniform convergence of v_∞. Notice that the improved upper bound in Lemma 5.5(i) is crucial here.

Lemma 4.14 *There exists a constant $c_{\beta,d} > 0$ such that*

$$\lim_{\varepsilon\to 0} \varepsilon^{-d} v_\infty(\varepsilon^{-2}t, \varepsilon^{-1}x) = c_{\beta,d}\cdot p(t,x).$$

The convergence is uniform for bounded t^{-1} and $x \in \mathsf{R}^d$.

Proof: We follow the proof of Lemma 3.3 in [6]. By Lemma 4.13, $v_\infty(t,x)$ is finite for any $t \geq 1$ and $x \in \mathsf{R}^d$. Then by a standard regularity argument in PDE theory,

$$\dot{v}_\infty = \frac{1}{2}\Delta v_\infty - v_\infty^{1+\beta} \tag{4.5}$$

on $[1,\infty)\times\mathsf{R}^d$. By Lemma 4.13, $v_\infty(1) \in L^1(\mathsf{R}^d)$. Set

$$w_\varepsilon(t,x) = \varepsilon^{-d} v_\infty(1+\varepsilon^{-2}t, \varepsilon^{-1}x).$$

Then by (4.5), $\dot{w}_\varepsilon = \frac{1}{2}\Delta w_\varepsilon - \varepsilon^{\beta d-2} w_\varepsilon^{1+\beta}$ with initial condition $w_\varepsilon(0,x) = \varepsilon^{-d} v_\infty(1, \varepsilon^{-1}x)$.

Applying Proposition 3.1 in [21] gives

$$\lim_{\varepsilon\to 0} \varepsilon^{-d} v_\infty(1+\varepsilon^{-2}t, \varepsilon^{-1}x) = c_{\beta,d}\cdot p(t,x),$$

uniformly on compact subsets of $(0,\infty)\times\mathsf{R}^d$. Together with Lemma 4.13 this yields the uniform convergence on $[a,\infty)\times\mathsf{R}^d$ for any $a>0$. Moreover, letting $t = t' - \varepsilon^2$, we get

$$\lim_{\varepsilon\to 0} \varepsilon^{-d} v_\infty(\varepsilon^{-2}t', \varepsilon^{-1}x) = c_{\beta,d}\cdot p(t',x),$$

uniformly on $[a,\infty)\times \mathsf{R}^d$ for any $a>0$.

It remains to prove that $c_{\beta,d}>0$. Using (4.4) and the lower bound in Lemma 5.5(i), we obtain

$$\begin{aligned} \varepsilon^{-d}v_\infty(\varepsilon^{-2}t,\varepsilon^{-1}x) &= \varepsilon^{-d}(\beta t)^{-1/\beta}P_{\varepsilon^{-1}x}\{\eta_{\varepsilon^{-2}t}B_0^1>0\} \\ &\gtrsim \varepsilon^{-d}p\left(\frac{\beta\varepsilon^{-2}t}{1+\beta},\varepsilon^{-1}x\right)=p\left(\frac{\beta t}{1+\beta},x\right), \end{aligned}$$

and so $c_{\beta,d}>0$. □

Now we can derive the asymptotic hitting rate for a $(2,\beta)$-process.

Theorem 4.15 *Let the $(2,\beta)$-process ξ in R^d with $\beta<1$ and $d>2/\beta$ be locally finite under P_μ. Fix any $t>0$ and $x\in\mathsf{R}^d$. Then as $\varepsilon\to 0$,*

$$\varepsilon^{2/\beta-d}P_\mu\{\xi_t B_x^\varepsilon>0\}\to c_{\beta,d}(\mu * p_t)(x).$$

The convergence is uniform for bounded $\|\mu\|$, bounded t^{-1}, and $x\in\mathsf{R}^d$. Similar results hold for the clusters η_t with p_t replaced by $(\beta t)^{1/\beta}p_t$.

Proof: We first prove that as $\varepsilon\to 0$,

$$\varepsilon^{2/\beta-d}(\beta t)^{-1/\beta}P_\mu\{\eta_t B_x^\varepsilon>0\}\to c_{\beta,d}(\mu * p_t)(x), \tag{4.6}$$

uniformly for bounded $\|\mu\|$, bounded t^{-1}, and $x\in\mathsf{R}^d$.

Use $\mu-x$ to denote the measure μ shifted by $-x$. If μ is finite, then by the scaling of η, (4.4), and Lemma 4.14, we can get the following chain of relations, which proves the uniform convergence of (4.6):

$$\varepsilon^{2/\beta-d}(\beta t)^{-1/\beta}P_\mu\{\eta_t B_x^\varepsilon>0\}$$

$$
\begin{aligned}
&= \varepsilon^{2/\beta-d}(\beta t)^{-1/\beta}\int P_y\{\eta_t B_0^\varepsilon > 0\}(\mu - x)(dy) \\
&= \varepsilon^{2/\beta-d}(\beta t)^{-1/\beta}\int P_{y/\varepsilon}\{\eta_{t/\varepsilon^2} B_0^1 > 0\}(\mu - x)(dy) \\
&= \varepsilon^{-d}\int v_\infty(\varepsilon^{-2}t, \varepsilon^{-1}y)(\mu - x)(dy) \to c_{\beta,d}(\mu * p_t)(x).
\end{aligned}
$$

Let μ be an infinite σ-finite measure satisfying $\mu p_t < \infty$ for all t. From the proof of Lemma 4.5, we know that $(\mu * p_{2t})(x) < \infty$ for any $x \in \mathsf{R}^d$. Then by dominated convergence based on Lemma 5.5(i), we can still get (4.6).

Now we turn to ξ_t. First note that by Lemma 5.4, as $\varepsilon \to 0$,

$$
\varepsilon^{2/\beta-d} P_\mu\{\xi_t B_x^\varepsilon > 0\} \to c \;\;\Leftrightarrow\;\; \varepsilon^{2/\beta-d}(\beta t)^{-1/\beta} P_\mu\{\eta_t B_x^\varepsilon > 0\} \to c, \tag{4.7}
$$

$$
\varepsilon^{2/\beta-d} P_\mu\{\xi_t^K B_x^\varepsilon > 0\} \to c \;\;\Leftrightarrow\;\; \varepsilon^{2/\beta-d} a_t^{-1} P_\mu\{\eta_t^K B_x^\varepsilon > 0\} \to c. \tag{4.8}
$$

It remains to prove the uniform convergence for ξ_t. Since $(\mu * p_t)(x) \le t^{-d/2}\|\mu\|$, we know that by (4.6), $(\beta t)^{-1/\beta} P_\mu\{\eta_t B_x^\varepsilon > 0\} \to 0$, uniformly for bounded $\|\mu\|$, bounded t^{-1}, and $x \in \mathsf{R}^d$. Then we may use Lemma 5.4 to get the uniform convergence for ξ_t. $\square$

The following result, especially part (ii), will play a crucial role in Section 5. Here we approximate the hitting probabilities $p_h^{K\varepsilon}$ by suitably normalized Dirac functions. This will be used in Lemma 4.17 to prove the Lebesgue approximation of ξ^K.

Lemma 4.16 *Let $p_h^\varepsilon(x) = P_x\{\eta_h B_0^\varepsilon > 0\}$, where the η_h are clusters of a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$. Recall that $p_h^{K\varepsilon}(x) = P_x\{\eta_h^K B_0^\varepsilon > 0\}$, where the η_h^K are clusters of ξ^K, the truncated K-process of ξ. Fix any bounded, uniformly continuous function $f \ge 0$ on R^d.*

(i) *As $0 < \varepsilon^2 \ll h \to 0$,*

$$
\left\| \varepsilon^{2/\beta-d}(\beta h)^{-1/\beta}\,(p_h^\varepsilon * f) - c_{\beta,d}\, f \right\| \to 0.
$$

(ii) *Fix any* $b \in (0, 1/2)$*. Then as* $0 < \varepsilon^2 \ll h \to 0$ *with* $\varepsilon^{2/\beta-d} h^{1+bd} \to 0$,

$$\left\| \varepsilon^{2/\beta-d} \, a_h^{-1} \, (p_h^{K\varepsilon} * f) - c_{\beta,d} \, f \, \right\| \to 0.$$

Both results hold uniformly over any class of uniformly bounded and equicontinuous functions $f \geq 0$ *on* R^d.

Proof: (i) We follow the proof of Lemma 5.2(i) in [25]. By scaling of η and (4.6),

$$\varepsilon^{2/\beta-d} (\beta h)^{-1/\beta} \lambda^d p_h^\varepsilon = (\varepsilon/\sqrt{h})^{2/\beta-d} (\beta)^{-1/\beta} \lambda^d p_1^{\varepsilon/\sqrt{h}} \to c_{\beta,d}. \tag{4.9}$$

Defining $\hat{p}_h^\varepsilon = p_h^\varepsilon / \lambda^d p_h^\varepsilon$, we need to show that $\|\hat{p}_h^\varepsilon * f - f\| \to 0$. Write w_f for the modulus of continuity of f, that is, a function $w_f = w(f, \cdot)$ defined by

$$w_f(r) = \sup\{|f(s) - f(t)|; s, t \in \mathsf{R}^d, |s - t| \leq r\}, \qquad r > 0.$$

Clearly $w_f(r) \to 0$ as $r \to 0$ since f is uniformly continuous. Now we get

$$\begin{aligned} \|\hat{p}_h^\varepsilon * f - f\| &= \sup_x \left| \int \hat{p}_h^\varepsilon(u) \; (f(x-u) - f(x)) \; du \right| \\ &\leq \int \hat{p}_h^\varepsilon(u) \, w_f(|u|) \, du \\ &\leq w_f(r) + 2 \, \|f\| \int_{|u|>r} \hat{p}_h^\varepsilon(u) \, du. \end{aligned}$$

It remains to show that $\int_{|u|>r} \hat{p}_h^\varepsilon(u) \, du \to 0$ for any fixed $r > 0$. Then notice that for any fixed $r > 0$ by Lemma 5.5(i),

$$\varepsilon^{2/\beta-d} (\beta h)^{-1/\beta} \int_{|u|>r} p_h^\varepsilon(u) \, du \lesssim \int_{|u|>r} p_{2h}(u) \, du \to 0.$$

(ii) For $p_h^{K\varepsilon}$, Lemma 5.5(ii) yields for any fixed $r > 0$,

$$\varepsilon^{2/\beta-d} a_h^{-1} \int_{|u|>r} p_h^{K\varepsilon}(u)\, du \lesssim \int_{|u|>r} p_{2h}(u)\, du \to 0.$$

Following the steps of the previous proof, it is enough to show that

$$\varepsilon^{2/\beta-d} a_h^{-1} \lambda^d p_h^{K\varepsilon} \to c_{\beta,d}. \tag{4.10}$$

Since $\int_{|u|>h^b} p_{2h}(u)\, du \to 0$, Lemma 5.5 yields

$$\varepsilon^{2/\beta-d}(\beta h)^{-1/\beta} 1\{(B_0^{h^b})^c\}\lambda^d p_h^{\varepsilon} \to 0,\ \ \varepsilon^{2/\beta-d} a_h^{-1} 1\{(B_0^{h^b})^c\}\lambda^d p_h^{K\varepsilon} \to 0.$$

By (5.12), to prove (4.10) it suffices to show that

$$\varepsilon^{2/\beta-d}(\beta h)^{-1/\beta} 1\{B_0^{h^b}\}\lambda^d p_h^{\varepsilon} - \varepsilon^{2/\beta-d} a_h^{-1} 1\{B_0^{h^b}\}\lambda^d p_h^{K\varepsilon} \to 0,$$

or equivalently (by (4.7) and (4.8)),

$$\varepsilon^{2/\beta-d}\left(P_{1\{B_0^{h^b}\}\lambda^d}\{\xi_h B_0^{\varepsilon} > 0\} - P_{1\{B_0^{h^b}\}\lambda^d}\{\xi_h^K B_0^{\varepsilon} > 0\}\right) \to 0.$$

By Theorem 25.22 of [24] and (4.1),

$$\begin{aligned}
&\varepsilon^{2/\beta-d}\left(P_{1\{B_0^{h^b}\}\lambda^d}\{\xi_h B_0^{\varepsilon} > 0\} - P_{1\{B_0^{h^b}\}\lambda^d}\{\xi_h^K B_0^{\varepsilon} > 0\}\right)\\
\leq\ \ & \varepsilon^{2/\beta-d} E_{1\{B_0^{h^b}\}\lambda^d} N_\xi\left([0,h],(K,\infty),\mathsf{R}^d\right)\\
=\ \ & \varepsilon^{2/\beta-d} E_{1\{B_0^{h^b}\}\lambda^d} \hat{N}_\xi\left([0,h],(K,\infty),\mathsf{R}^d\right)\\
\eqsim\ \ & \varepsilon^{2/\beta-d} E\int_0^h \|\xi_s\| ds \eqsim \varepsilon^{2/\beta-d} h^{1+bd} \to 0. \qquad \square
\end{aligned}$$

4.5 Lebesgue approximations

To prove the Lebesgue approximation for a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$, we begin with the Lebesgue approximation for ξ^K, the truncated K-process of ξ. Since ξ and ξ^K agree asymptotically as $K \to \infty$, we have thus proved the Lebesgue approximation for ξ. Write $\tilde{c}_{\beta,d} = 1/c_{\beta,d}$ for convenience, where $c_{\beta,d}$ is such as in Lemma 4.14. Recall that $\xi_t^{K\varepsilon} = (\xi_t^K)^\varepsilon$, the ε-neighborhood measure of ξ_t^K.

Lemma 4.17 *Let ξ^K be the truncated K-process of a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$. Fix any $\mu \in \hat{\mathcal{M}}_d$ and $t > 0$. Then under P_μ, we have as $\varepsilon \to 0$:*

$$\tilde{c}_{\beta,d}\, \varepsilon^{2/\beta - d}\, \xi_t^{K\varepsilon} \xrightarrow{w} \xi_t^K \quad a.s.$$

Proof: We follow the proof of Theorem 7.1 in [25]. Fix any $f \in C_K^d$. Write η_h^{Ki} for the subclusters of ξ_t^K of age h. Since the ancestors of ξ_t^K at time $s = t - h$ form a Cox process directed by ξ_s^K / a_h, Lemma 5.7(i) yields

$$E_\mu\Big[\sum\nolimits_i \eta_h^{Ki\varepsilon} f \,\Big|\, \xi_s^K\Big] = a_h^{-1} \xi_s^K (p_h^{K\varepsilon} * f),$$

and so by Lemma 5.7(ii)

$$\begin{aligned} E_\mu\Big| \sum\nolimits_i \eta_h^{Ki\varepsilon} f - a_h^{-1}\, \xi_s^K (p_h^{K\varepsilon} * f)\Big|^2 &= E_\mu \mathrm{Var}\Big[\sum\nolimits_i \eta_h^{Ki\varepsilon} f \,\Big|\, \xi_s^K\Big] \\ &\lesssim a_h\, \varepsilon^{d-2/\beta}\, h^{d/2}\, \|f\|^2\, E_\mu \|\xi_s^K / a_h\| \\ &\leq \varepsilon^{d-2/\beta}\, h^{d/2}\, \|f\|^2\, \|\mu\|, \end{aligned}$$

where the last inequality follows from $E_\mu \|\xi_s^K\| \leq \|\mu\|$. Combining with Lemma 5.8 gives

$$\begin{aligned} & E_\mu \big|\, \xi_t^{K\varepsilon} f - a_h^{-1}\, \xi_s^K (p_h^{K\varepsilon} * f)\,\big| \\ &\quad \leq \; E_\mu \Big|\, \xi_t^{K\varepsilon} f - \sum\nolimits_i \eta_h^{Ki\varepsilon} f\,\Big| + E_\mu \Big| \sum\nolimits_i \eta_h^{Ki\varepsilon} f - a_h^{-1}\, \xi_s^K (p_h^{K\varepsilon} * f)\,\Big| \end{aligned}$$

$$
\begin{aligned}
&\lesssim \quad \varepsilon^{2(d-2/\beta)}\, h^{1-d/2}\, \|f\| + \varepsilon^{1/2(d-2/\beta)}\, h^{d/4}\, \|f\| \\
&= \quad \varepsilon^{d-2/\beta} \left(\varepsilon^{d-2/\beta} h^{1-d/2} + \varepsilon^{-1/2(d-2/\beta)} h^{d/4}\right) \|f\|.
\end{aligned}
$$

Let c satisfy

$$
(d - 2/\beta) + (-d/2 + 1/2)c = 0. \tag{4.11}
$$

Clearly $c \in (0, 2)$. Taking $\varepsilon = r^n$ for a fixed $r \in (0, 1)$ and $h = \varepsilon^c = r^{cn}$, and writing $s_n = t - h = t - r^{cn}$, we obtain

$$
\begin{aligned}
&E_\mu \sum_n r^{n(2/\beta - d)} \left| \xi_t^{Kr^n} f - a_{r^{cn}}^{-1}\, \xi_{s_n}^K (p_{r^{cn}}^{Kr^n} * f) \right| \\
&\qquad \lesssim \sum_n \left(r^{[(d-2/\beta)+(-d/2+1)c]n} + r^{[-1/2(d-2/\beta)+(d/4)c]n} \right) \|f\| < \infty,
\end{aligned}
$$

since $(d - 2/\beta) + (-d/2 + 1)c > 0$ and $-1/2(d - 2/\beta) + (d/4)c > 0$ by (4.11). This implies

$$
r^{n(2/\beta - d)} \left| \xi_t^{Kr^n} f - a_{r^{cn}}^{-1}\, \xi_{s_n}^K (p_{r^{cn}}^{Kr^n} * f) \right| \to 0 \ \text{ a.s. } P_\mu. \tag{4.12}
$$

Now we write

$$
\begin{aligned}
&\left| \varepsilon^{2/\beta - d}\, \xi_t^{K\varepsilon} f - c_{\beta,d}\, \xi_t^K f \right| \\
&\quad \leq \ \varepsilon^{2/\beta - d} \left| \xi_t^{K\varepsilon} f - a_h^{-1}\, \xi_s^K (p_h^{K\varepsilon} * f) \right| + c_{\beta,d}\, |\xi_s^K f - \xi_t^K f| \\
&\qquad + \|\xi_s^K\| \left\| \varepsilon^{2/\beta - d}\, a_h^{-1}\, (p_h^{K\varepsilon} * f) - c_{\beta,d} f \right\|.
\end{aligned}
$$

For the last term, we first fix $b = 1/2 - 1/d$, then apply Lemma 4.16. Noting that by (4.11)

$$
(2/\beta - d) + (1 + bd)c = (2/\beta - d) + (d/2)c > 0,
$$

we get by Lemma 4.16

$$
\left\| \varepsilon^{2/\beta - d}\, a_h^{-1}\, (p_h^{K\varepsilon} * f) - c_{\beta,d} f \right\| \to 0
$$

along the sequence (r^n). Using (5.13) and the a.s. weak continuity of ξ^K at the fixed time t, we see that the right-hand side tends a.s. to 0 as $n \to \infty$, which implies $\varepsilon^{2/\beta - d}\, \xi_t^{K\varepsilon} f \to c_{\beta,d}\, \xi_t^K f$ a.s. as $\varepsilon \to 0$ along the sequence (r^n) for any fixed $r \in (0,1)$. Since this holds simultaneously, outside a fixed null set, for all rational $r \in (0,1)$, the a.s. convergence extends by Lemma 2.3 in [25] to the entire interval $(0,1)$.

Applying this result to a countable, convergence-determining class of functions f (cf. Lemma 3.2.1 in [5]), we obtain the required a.s. vague convergence. Since μ is finite, the $(2,\beta)$-process ξ_t has a.s. compact support (cf. Theorem 9.3.2.2 of [5] and the proof of Theorem 1.2 in [6]). By Lemma 4.2, ξ_t^K also has a.s. compact support, and so the a.s. convergence remains valid in the weak sense. $\square$

Now we may prove our main result, the Lebesgue approximation of $(2,\beta)$-processes. Again, we write $\tilde{c}_{\beta,d} = 1/c_{\beta,d}$ for convenience, where $c_{\beta,d}$ is such as in Lemma 4.14. Also recall that $\xi_t^\varepsilon = (\xi_t)^\varepsilon$ denotes the ε-neighborhood measure of ξ_t. For random measures ξ_n and ξ on R^d, $\xi_n \xrightarrow{v} \xi$ (or $\xrightarrow{w}$) in L^1 means that $\xi_n f \to \xi f$ in L^1 for all f in C_K^d (or C_b^d).

Theorem 4.18 *Let the $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$ be locally finite under P_μ, and fix any $t > 0$. Then under P_μ, we have as $\varepsilon \to 0$:*

$$\tilde{c}_{\beta,d}\, \varepsilon^{2/\beta - d}\, \xi_t^\varepsilon \xrightarrow{v} \xi_t \ \text{ a.s. and in } L^1.$$

This remains true in the weak sense when μ is finite. The weak version holds even for the clusters η_t when $\|\mu\| = 1$.

Proof: For a finite initial measure μ, by Lemma 4.17 and Lemma 4.1 we get as $\varepsilon \to 0$

$$\tilde{c}_{\beta,d}\, \varepsilon^{2/\beta - d}\, \xi_t^\varepsilon \xrightarrow{w} \xi_t \ \text{a.s.}$$

For a general σ-finite measure μ on R^d with $\mu p_t < \infty$ for all $t > 0$, write $\mu = \mu' + \mu''$ for a finite μ', and let $\xi = \xi' + \xi''$ be the corresponding decomposition of ξ into independent components with initial measures μ' and μ''. Fixing an $r > 1$ with $\operatorname{supp} f \subset B_0^{r-1}$ and using the result for finite μ, we get a.s. on $\{\xi_t'' B_0^r = 0\}$

$$\varepsilon^{2/\beta-d}\,\xi_t^\varepsilon f = \varepsilon^{2/\beta-d}\,\xi_t'^\varepsilon f \to c_{\beta,d}\,\xi_t' f = c_{\beta,d}\,\xi_t f.$$

As $\mu' \uparrow \mu$, we get by Lemma 4.9

$$P_\mu\{\xi_t'' B_0^r = 0\} = P_{\mu''}\{\xi_t B_0^r = 0\} \to 1,$$

and the a.s. convergence extends to μ. As in the proof of Lemma 4.17, we can obtain the required a.s. vague convergence.

To prove the convergence in L^1, we note that for any $f \in C_K^d$

$$\begin{aligned} \varepsilon^{2/\beta-d} E_\mu \xi_t^\varepsilon f &= \varepsilon^{2/\beta-d} \int P_\mu\{\xi_t B_x^\varepsilon > 0\}\, f(x)\, dx \\ &\to \int c_{\beta,d}\,(\mu * p_t)(x)\, f(x)\, dx = c_{\beta,d}\, E_\mu \xi_t f, \end{aligned} \tag{4.13}$$

by Theorem 4.15. Combining this with the a.s. convergence under P_μ and using Proposition 4.12 in [24], we obtain $E_\mu|\varepsilon^{2/\beta-d}\,\xi_t^\varepsilon f - c_{\beta,d}\,\xi_t f| \to 0$. For finite μ, (4.13) extends to any $f \in C_b^d$ by dominated convergence based on Lemmas 5.4 and 5.5(i), together with the fact that $\lambda^d(\mu * p_t) = \|\mu\| < \infty$ by Fubini's theorem.

To extend the Lebesgue approximation to the individual clusters η_t, let ζ_0 denote the process of ancestors of ξ_t at time 0, and note that

$$P_x\{\eta_t \in \cdot\} = P_{\delta_x}[\xi_t \in \cdot \,|\, \|\zeta_0\| = 1],$$

where $P_{\delta_x}\{\|\zeta_0\| = 1\} = (\beta t)^{-1/\beta} e^{-(\beta t)^{-1/\beta}} > 0$. The a.s. convergence then follows from the corresponding statement for ξ_t. Since

$$P_\mu\{\eta_t \in \cdot\} = \int \mu(dx) P_x\{\eta_t \in \cdot\},$$

the a.s. convergence under any P_μ with $\|\mu\| = 1$ also follows. To obtain the weak L^1-convergence in this case, we note that for $f \in C_b^d$,

$$\begin{aligned} \varepsilon^{2/\beta-d} E_\mu \eta_t^\varepsilon f &= \varepsilon^{2/\beta-d} \int P_\mu\{\eta_t B_x^\varepsilon > 0\}\, f(x)\, dx \\ &\to c_{\beta,d}\, (\beta t)^{1/\beta} \int (\mu * p_t)(x)\, f(x)\, dx = c_{\beta,d}\, E_\mu \eta_t f, \end{aligned}$$

by dominated convergence based on Lemma 5.5(i) and Theorem 4.15. □

As in Corollary 7.2 of [25], for the intensity measures in Theorem 4.18, we have even convergence in total variation.

Corollary 4.19 *Let ξ be a $(2,\beta)$-process in R^d with $\beta < 1$ and $d > 2/\beta$. Then for any finite μ and $t > 0$, we have as $\varepsilon \to 0$:*

$$\left\| \varepsilon^{2/\beta-d} E_\mu \xi_t^\varepsilon - c_{\beta,d}\, E_\mu \xi_t \right\| \to 0.$$

This remains true for the clusters η_t, and it also holds locally for ξ_t whenever ξ is locally finite under P_μ.

Finally let us give a detailed explanation of the deterministic distribution property of $(2,\beta)$-processes. Here the deterministic distribution property has two aspects. Define deterministic functions Φ_ε, Φ similar to those defined on page 309 of [41], Theorem 4.18 shows that a.s.

$$\Phi(\operatorname{supp} \xi_t) = \xi_t,$$

so a.s. ξ_t is a deterministic function of its support $\operatorname{supp}\xi_t$. This is the first aspect of the deterministic distribution property. Now the second aspect. Since $\lambda^d(\partial B_x^r) = 0$, we get $\xi_t(\partial B_x^r) = 0$ a.s. by noting $E_\mu \xi_t = (\mu * p_t) \cdot \lambda^d$. With the help of Portmanteau Theorem for finite measures, Theorem 4.18 shows that a.s. for all open balls B with rational centers and rational radius,

$$\lim_{\varepsilon \to 0} \Phi_\varepsilon(\operatorname{supp}\xi_t)(B) = \xi_t(B),$$

so the construction of $\xi_t(\omega)$ from its support $\operatorname{supp}\xi_t(\omega)$ is the same everywhere for any fixed ω outside a null set.

Chapter 5

Lebesgue Approximation of Superprocesses with a Regularly Varying Branching Mechanism

5.1 Introduction

Superprocesses are certain measure-valued Markov processes $\xi = (\xi_t)$, whose distributions can be characterized by two components: the *branching mechanism* specified by a function $\Phi(v)$, and the *spatial motion* usually given by a Markov process X. If X is a Feller process in R^d with generator L, then the laplace functional $E_\mu \exp(-\xi_t f)$ satisfies $E_\mu[\exp(-\xi_t f)|\xi_s] = \exp(-\xi_s v_{t-s})$ where $v_t(x)$ is the unique nonnegative solution of the so-called *evolution equation* $\dot{v} = Lv - \Phi(v)$ with initial condition $v_0 = f$. We call this superprocess an (L, Φ)-superprocess (or (L, Φ)*-process* for short). For $\alpha \in (0, 2]$ and $\beta \in (0, 1]$, if X is a rotation invariant α-stable Lévy process in R^d with generator $\frac{1}{2}\Delta_\alpha$ and $\Phi(v) = v^{1+\beta}$, we get a superprocess corresponding to the PDE $\dot{v} = \frac{1}{2}\Delta_\alpha v - v^{1+\beta}$. We call it an (α, β)*-superprocess* ((α, β)*-process* for short), which is just a $(\frac{1}{2}\Delta_\alpha, v^{1+\beta})$-superprocess in our previous notation. General surveys of superprocesses include the excellent monographs and lecture notes [5, 15, 17, 32, 35, 42].

For any measure μ on R^d and constant $\varepsilon > 0$, write μ^ε for the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp}\mu$. For a (2,1)-process ξ in R^d, Tribe [48] showed that $\varepsilon^{2-d}\,\xi_t^\varepsilon \xrightarrow{w} c_d\,\xi_t$ a.s. as $\varepsilon \to 0$ for fixed time $t > 0$ when $d \geq 3$, where $\xrightarrow{w}$ denotes weak convergence. Perkins [41] improved Tribe's result by showing that the Lebesgue approximation actually holds for all time $t > 0$ simultaneously. Kallenberg [25] proved the Lebesgue approximation of 2-dimensional (2,1)-processes. In [22], we showed that, for any $(2, \beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$, $\varepsilon^{2/\beta-d}\,\xi_t^\varepsilon \xrightarrow{w} c_{\beta,d}\,\xi_t$ a.s. as $\varepsilon \to 0$ for fixed time $t > 0$. In particular, the Lebesgue approximation result implies that the superprocess ξ_t distributes its mass over $\operatorname{supp}\xi_t$ in a deterministic manner. See the end of [22] for a detailed explanation of this deterministic distribution property. However, for any (α,β)-process ξ with $\alpha < 2$, $\operatorname{supp}\xi_t = \mathsf{R}^\mathsf{d}$ or $\emptyset$ a.s. (cf. [18, 40]), and so the corresponding property fails. From all these Lebesgue approximation results, we raise the natural conjecture: Lebesgue approximation holds for superprocesses with Brownian spatial motion and any "reasonable" branching mechanism.

As a first step to prove this general conjecture, in this chapter we study the Lebesgue approximation of superprocesses with Brownian spatial motion and a regularly varying branching mechanism. For a precise description of the branching mechanism we consider in this chapter, refer to the beginning of Section 3. The stable branching mechanism $\Phi(v) = v^{1+\beta}$ with $\beta \in (0,1]$ is a special case of the regularly varying branching mechanism we consider here. Our main result in this chapter is Theorem 5.5, where we prove that the Lebesgue approximation still holds for these more general superprocesses. Specifically, $\tilde{m}(\varepsilon)\,\xi_t^\varepsilon \xrightarrow{w} \xi_t$ a.s. as $\varepsilon \to 0$ for fixed time $t > 0$, where $m(\varepsilon)$ is a suitable normalizing function. In particular, if the branching mechanism is the stable one, we may recover all previous Lebesgue approximation results for fixed time $t > 0$.

Although the previous conjecture may seems very natural, technically we have limited tools to support some rigorous arguments needed. One such boundary is imposed by the availability of the very important cluster representation of superprocesses. Luckily the superprocesses we consider here do have the cluster representation. Another boundary is imposed by the availability of the lower and upper bounds of the hitting probabilities $P_\mu\{\xi_t B_x^\varepsilon > 0\}$, which is fundamental for the Lebesgue approximation. The restriction on the branching mechanism we consider actually follows from Theorem 2.3 in [9], which is exactly the lower and upper bounds of the hitting probabilities.

Armed with the hitting estimates, then we are able to overcome the main difficulty in this chapter, that is, to obtain an asymptotic result of the hitting probabilities $P_\mu\{\xi_t B_x^\varepsilon > 0\}$, which is Theorem 5.11. Note that for a $(2,\beta)$-process, such a result is obtained by using the strong scaling property. Since the regularly varying branching mechanism we consider here has much weaker scaling property, we then have to rely on only the cluster representation and the hitting estimations. Also the form of the asymptotic result of the hitting probabilities is not clear in our general setting. By adapting an idea in Section 5 of [25], we can get the correct form of our asymptotic result, which determines the form of the Lebesgue approximation.

This chapter is organized as follows. In Section 2 we review the truncation of superprocesses in a more general setting. In Section 3, we develop some lemmas about hitting bounds and neighborhood measures of the more general superprocesses. In Section 4, we derive some asymptotic results of these hitting probabilities. Finally in Section 5 we state and prove the Lebesgue approximation of superprocesses with a regularly varying branching mechanism and their truncated processes. This general result contains all previous Lebesgue approximation of superprocesses as special cases.

5.2 Truncation of superprocesses

In this section we discuss the truncation of superprocesses with a general branching mechanism, due to their independent interests.

We consider a general branching mechanism function Φ defined on R_+ as

$$\Phi(v) = av + bv^2 + \int_{(0,\infty)} (e^{-rv} - 1 + rv)\pi(dr),$$

where $b \geq 0$ and π is a measure on $(0, \infty)$ such that $\int_0^\infty (r \wedge r^2)\pi(dr) < \infty$.

It is well known that the $(L, 1)$-process has weakly continuous sample paths. By contrast, when $\pi \neq 0$, the corresponding superprocess ξ has only weakly rcll sample paths with jumps of the form $\Delta\xi_t = r\delta_x$, for some $t > 0$, $r > 0$, and $x \in \mathsf{R}^d$. Let

$$N_\xi(dt, dr, dx) = \sum_{(t,r,x):\ \Delta\xi_t = r\delta_x} \delta_{(t,r,x)}.$$

Clearly the point process N_ξ on $\mathsf{R}_+ \times \mathsf{R}_+ \times \mathsf{R}^d$ records all information about the jumps of ξ. By the proof of Theorem 6.1.3 in [5], we know that N_ξ has compensator measure

$$\hat{N}_\xi(dt, dr, dx) = (dt)\pi(dr)\xi_t(dx). \tag{5.1}$$

Due to all the "big" jumps, ξ_t has infinite variance. Some methods for $(L,1)$-processes, which rely on the finite variance of the processes, are not directly applicable to superprocesses with a branching mechanism having $\pi \neq 0$.

$$M_t(f) = M_t^c(f) + M_t^d(f) = \xi_t f - \xi_0 f - \int_0^t \xi_s(Lf)ds,$$

$$\xi_t f = \xi_0 f + \int_0^t \xi_s(Lf)ds + M_t^c(f) + M_t^d(f)$$

where $M_t^c(f)$ is a continuous martingale with quadratic variation process

$$[M^c(f)]_t = \int_0^t \xi_s(bf^2)ds, \tag{5.2}$$

and $M_t^d(f)$ is a purely discontinuous martingale, which can be written as follows

$$\begin{aligned}
M_t^d(f) &= \int_0^t \int_{(0,\infty)} \int_{\mathsf{R}^d} rf(x)\hat{N}_\xi(dt, dr, dx) \\
&= \int_0^t \int_{(0,K]} \int_{\mathsf{R}^d} rf(x)\hat{N}_\xi(dt, dr, dx) + \int_0^t \int_{(K,\infty)} \int_{\mathsf{R}^d} rf(x)\hat{N}_\xi(dt, dr, dx) \\
&= \int_0^t \int_{(0,K]} \int_{\mathsf{R}^d} rf(x)\hat{N}_\xi(dt, dr, dx) \\
&\quad + \int_0^t \int_{(K,\infty)} \int_{\mathsf{R}^d} rf(x)N_\xi(dt, dr, dx) - \pi(K,\infty)\int_0^t \xi_s f ds
\end{aligned}$$

$$E_\mu[\exp(-\xi_t^K f)|\xi_s^K] = \exp(-\xi_s^K v_{t-s}) \tag{5.3}$$

$$\dot{v} = Lv - \Phi^K(v), \tag{5.4}$$

where $\Phi^K = (a - \pi(K,\infty))v + bv^2 + \int_{(0,K]}(e^{-rv} - 1 + rv)\pi(dr)$

$$\xi_t^K f = \xi_0^K f + \int_0^t \xi_s^K(Lf)ds + M_t^c(f) + M_t^d(f) - \pi[K,\infty)\int_0^t \xi_s^K f ds$$

where $M_t^c(f)$ is a continuous martingale with quadratic variation process

$$[M^c(f)]_t = \int_0^t \xi_s^K(bf^2)ds, \tag{5.5}$$

and $M_t^d(f)$ is a purely discontinuous martingale, which can be written as follows

$$\begin{aligned} M_t^d(f) &= \int_0^t \int_{(0,\infty)} \int_{\mathsf{R}^d} rf(x)\hat{N}_{\xi^K}(dt,dr,dx) \\ &= \int_0^t \int_{(0,K)} \int_{\mathsf{R}^d} rf(x)\hat{N}_{\xi^K}(dt,dr,dx) \end{aligned}$$

$$N_{\xi^K}(dt,dr,dx) = \sum_{(t,r,x):\ \Delta\xi_t^K = r\delta_x} \delta_{(t,r,x)}.$$

$$\hat{N}_{\xi^K}(dt,dr,dx) = (dt)1_{(0,K)}(r)\pi(dr)\xi_t^K(dx). \tag{5.6}$$

In [38], Mytnik and Villa introduced a truncation method for (α,β)-processes with $\beta < 1$, which can be used to study (α,β)-processes with $\beta < 1$, especially to extend results of $(\alpha,1)$-processes to (α,β)-processes with $\beta < 1$. Specifically, for the (α,β)-process ξ with $\beta < 1$, we define the stopping time $\tau_K = \inf\{t > 0 : \|\Delta\xi_t\| > K\}$ for any constant $K > 0$. Clearly τ_K is the time when ξ has the first jump greater than K. For any finite initial measure μ, they proved that one can define ξ and a weakly rcll, measure-valued Markov process ξ^K on a common probability space such that $\xi_t = \xi_t^K$ for $t < \tau_K$. Intuitively, ξ^K euqals ξ minus all masses produced by jumps greater than K along with the future evolution of those masses. In this paper, we call ξ^K the *truncated K-process* of ξ. Since all "big" jumps are omitted, ξ_t^K has finite variance. They also proved that ξ_t^K and ξ_t agree asymptotically as $K \to \infty$. We give a different proof of this result, since similar ideas will also be used at several crucial stages later. We write $P_\mu\{\xi \in \cdot\}$ for the distribution of ξ with initial measure μ.

Using the same proof of Lemma 1 in [38], we can construct ξ and ξ^K on a common probability space such that $\xi_t(\omega) = \xi_t^K(\omega)$ for $t < \tau_K(\omega)$. This confirms our intuition that ξ^K euqals ξ minus all masses produced by jumps greater than K along with the future evolution of those masses.

Lemma 5.1 *We can define ξ and ξ^K on a common probability space such that:*

(i) *ξ is an (α, β)-process with $\beta < 1$ and a finite initial measure μ, and ξ^K is its truncated K-process,*

(ii) *$\xi_t(\omega) = \xi_t^K(\omega)$ for $t < \tau_K(\omega)$.*

Now we can prove that ξ_t^K and ξ_t agree asymptotically as $K \to \infty$. We choose to give a complete proof of this result, since similar ideas will also be used at several crucial stages later. We write $P_\mu\{\xi \in \cdot\}$ for the distribution of ξ with initial measure μ.

Lemma 5.2 *Fix any finite μ and $t > 0$. Then $P_\mu\{\tau_K > t\} \to 1$ as $K \to \infty$.*

Proof: If $\tau_K \le t$, then ξ has at least one jump greater than K before time t. Noting that $N_\xi([0,t],(K,\infty),\mathsf{R}^d)$ is the number of jumps greater than K before time t, we get by Theorem 25.22 of [24] and (4.1),

$$
\begin{aligned}
P_\mu\{\tau_K \le t\} &\le E_\mu N_\xi\big([0,t],(K,\infty),\mathsf{R}^d\big) \\
&= E_\mu \hat{N}_\xi\big([0,t],(K,\infty),\mathsf{R}^d\big) \\
&\underset{\frown}{=} \pi[K,\infty) E_\mu \int_0^t \|\xi_s\| ds = t\|\mu\|\pi[K,\infty) \to 0
\end{aligned}
$$

as $K \to \infty$, where the last equation holds by $E_\mu \|\xi_s\| = \|\mu\|$. $\square$

Using the same proof of Lemma 2.2 in [22], we can prove that $\xi_t^K(\omega) \leq \xi_t(\omega)$ for any t and ω. So indeed, ξ^K is a "truncation" of ξ.

Lemma 5.3 *We can define ξ and ξ^K on a common probability space such that:*

(i) *ξ is an (α, β)-process with $\beta < 1$ and a finite initial measure μ, and ξ^K is its truncated K-process,*

(ii) *$\xi_t(\omega) \geq \xi_t^K(\omega)$ for any t and ω,*

(iii) *$\xi_t(\omega) = \xi_t^K(\omega)$ for $t < \tau_K(\omega)$.*

5.3 Hitting bounds

First we specify the regularly varying branching mechanism we consider for the Lebesgue approximation. We consider the increasing function Φ defined on R_+ by

$$\Phi(v) = bv^2 + \int_{(0,\infty)} \frac{2rv^2}{1+2rv} \pi'(dr),$$

where $b \geq 0$ and π' is a measure on $(0, \infty)$ such that $\int_{(0,\infty)} (1 \wedge r)\pi'(dr) < \infty$. To avoid trivial cases, we assume either $b > 0$ or $\pi'((0, \infty)) = \infty$. The function Φ can be expressed in the usual form for branching mechanism functions,

$$\Phi(v) = bv^2 + \int_{(0,\infty)} (e^{-rv} - 1 + rv)\pi(dr),$$

where $\pi(dr) = [\int_{(0,\infty)} e^{-r/(2u)}/(4u^2)\pi'(du)]dr$ satisfies $\int_{(0,\infty)} (r \wedge r^2)\pi(dr) < \infty$. Notice that if we take $b = 0$ and $\pi'(dr) = c'r^{-(1+\beta)}dr$ then we get the stable case $\Phi(v) = cv^{1+\beta}$.

We consider the following two assumptions:

(A1) The function Φ is regularly varying at ∞ with index $1+\beta$ where $\beta \in (0,1]$; that is to say,

$$\lim_{u\to\infty} \frac{\Phi(ru)}{\Phi(u)} = r^{1+\beta} \text{ for every } r > 0$$

(A2) $\limsup_{r\to 0+} r^{-(1+\beta)}\Phi(r) < \infty$.

The stable case $\Phi(v) = v^{1+\beta}$ satisfies all these assumptions.

The Lebesgue approximation depends crucially on estimates of the hitting probability $P_\mu\{\xi_t B_0^\varepsilon > 0\}$. In this section, we first estimate $P_\mu\{\xi_t B_0^\varepsilon > 0\}$ and $P_\mu\{\xi_t^K B_0^\varepsilon > 0\}$. Then we use these estimates to study multiple hitting and neighborhood measures of the clusters η_h^K associated with the truncated K-process ξ^K. We begin with a well-known relationship between the hitting probabilities of superprocesses and their clusters, which can be proved as in Lemma 4.1 of [25].

Lemma 5.4 *Let the (α,β)-process ξ in R^d with associated clusters η_t be locally finite under P_μ, let ξ^K be its truncated K-process with associated clusters η_t^K, and fix any $B \in \mathcal{B}^d$. Then*

$$\begin{aligned}
P_\mu\{\eta_t B > 0\} &= -a_t \log(1 - P_\mu\{\xi_t B > 0\}), \\
P_\mu\{\xi_t B > 0\} &= 1 - \exp\left(-a_t^{-1} P_\mu\{\eta_t B > 0\}\right), \\
P_\mu\{\eta_t^K B > 0\} &= -a_t^K \log(1 - P_\mu\{\xi_t^K B > 0\}), \\
P_\mu\{\xi_t^K B > 0\} &= 1 - \exp(-(a_t^K)^{-1} P_\mu\{\eta_t^K B > 0\}).
\end{aligned}$$

In particular, $P_\mu\{\xi_t B > 0\} \sim a_t^{-1} P_\mu\{\eta_t B > 0\}$ and $P_\mu\{\xi_t^K B > 0\} \sim (a_t^K)^{-1} P_\mu\{\eta_t^K B > 0\}$ as either side tends to 0.

Upper and lower bounds of $P_\mu\{\xi_t B_0^\varepsilon > 0\}$ have been obtained by Delmas [9], using the Brownian snake. However, in this paper we need the following improved upper bound.

Lemma 5.5 *Let η_t be the clusters of a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$, let η_t^K be the clusters of ξ^K, the truncated K-process of ξ, and consider a σ-finite measure μ on R^d. Then for $0 < \varepsilon \le \sqrt{t}$,*

(i) $l_2(\varepsilon)\mu p_{t'} \lesssim \varepsilon^{2/\beta-d} a_t^{-1} P_\mu\{\eta_t B_0^\varepsilon > 0\} \lesssim l_1(\varepsilon)\mu p_{2t}$, *where* $t' = \beta t/(1+\beta)$,

(ii) $\varepsilon^{2/\beta-d}(a_t^K)^{-1} P_\mu\{\eta_t^K B_0^\varepsilon > 0\} \lesssim l_1(\varepsilon)\mu p_{2t}$.

Proof: (i) Follow the proof of Lemma 6.3(i) in [25], then use Lemma 5.5(ii).

(ii) This is obvious from (i), Lemma 4.2, and Lemma 5.4. $\square$

As in [25] we need to estimate the probability that a ball in R^d is hit by more than one subcluster of the truncated K-process ξ^K. This is where the truncation of ξ is needed.

Lemma 5.6 *Fix any $K > 0$. Let ξ^K be the truncated K-process of a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$. For any $t \geq h > 0$ and $\varepsilon > 0$, let $\kappa_h^{K\varepsilon}$ be the number of h-clusters of ξ_t^K hitting B_0^ε at time t. Then for $\varepsilon^2 \leq h \leq t$,*

$$E_\mu \kappa_h^{K\varepsilon}(\kappa_h^{K\varepsilon} - 1) \lesssim l_1^2(\varepsilon)\varepsilon^{2(d-2/\beta)} \left(h^{1-d/2}\mu p_t + (\mu p_{2t})^2 \right).$$

Proof: Follow Lemma 4.4 in [25], then use Lemma 3 of [38] and Lemma 5.5(ii). $\square$

Now we consider the neighborhood measures of the clusters η_h^K associated with the truncated K-process ξ^K. For any measure μ on R^d and constant $\varepsilon > 0$, we define the associated *neighborhood measure* μ^ε as the restriction of Lebesgue measure λ^d to the ε-neighborhood of $\operatorname{supp}\mu$, so that μ^ε has Lebesgue density $1\{\mu B_x^\varepsilon > 0\}$. Let $p_h^{K,\varepsilon}(x) = P_x\{\eta_h^K B_0^\varepsilon > 0\}$, where the η_h^K are clusters of ξ^K. Write $p_h^{K,\varepsilon}(x) = p_h^{K\varepsilon}(x)$ and $(\eta_h^{K,i})^\varepsilon = \eta_h^{Ki\varepsilon}$ for convenience.

Lemma 5.7 *Let ξ^K be the truncated K-process of a $(2,\beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$. Let the η_h^{Ki} be conditionally independent h-clusters of ξ^K, rooted at the points of a Poisson process ζ with $E\zeta = \mu$. Fix any measurable function $f \geq 0$ on R^d. Then,*

(i) $E_\mu \sum_i \eta_h^{Ki\varepsilon} = (\mu * p_h^{K\varepsilon}) \cdot \lambda^d$,

(ii) $E_\mu \mathrm{Var}\left[\sum_i \eta_h^{Ki\varepsilon} f | \zeta\right] \lesssim l_1(\varepsilon) a_h^K \varepsilon^{d-2/\beta} h^{d/2} \|f\|^2 \|\mu\|$ *for* $\varepsilon^2 \leq h$.

Proof: (i) Follow the proof of Lemma 6.2(i) in [25].

(ii)Follow the proof of Lemma 4.4(ii) in [22]. $\square$

We also need to estimate the overlap between subclusters.

Lemma 5.8 *Let ξ^K be the truncated K-process of a $(2, \beta)$-process ξ in R^d with $\beta < 1$ and $d > 2/\beta$. For any fixed $t > 0$, let η_h^{Ki} denote the subclusters in ξ^K of age $h > 0$. Fix any $\mu \in \hat{\mathcal{M}}_d$. Then as $\varepsilon^2 \leq h \to 0$,*

$$E_\mu \left\| \sum\nolimits_i \eta_h^{Ki\varepsilon} - \xi_t^{K\varepsilon} \right\| \precsim l_1^2(\varepsilon)\varepsilon^{2(d-2/\beta)}h^{1-d/2}.$$

Proof: Follow the proof of Lemma 6.3(i) in [25], then use Lemma 5.5(ii). $\square$

5.4 Hitting asymptotics

Write $p_h^\varepsilon(x) = P_x\{\eta_h B_0^\varepsilon > 0\}$ and $p_h^{K,\varepsilon}(x) = P_x\{\eta_h^K B_0^\varepsilon > 0\}$, where η_h and η_h^K denote an h-cluster associated with the superprocess ξ in R^d and its truncated K-process ξ^K respectively. Recall that $\lambda^d p_h^\varepsilon = P_{\lambda^d}\{\eta_h B_0^\varepsilon > 0\}$. Write $p_h^{K\varepsilon} = p_h^{K,\varepsilon}$ for convenience. For the functions p_h^ε and $p_h^{K\varepsilon}$, we have the following basic asymptotic property. Since we do not have a lower bound for $P_x\{\eta_h^K B_0^\varepsilon > 0\}$ in Lemma, this asymptotic property is crucial to us by showing that essentially $P_x\{\eta_h B_0^\varepsilon > 0\}$ and $P_x\{\eta_h^K B_0^\varepsilon > 0\}$ share the same lower bound.

Lemma 5.9 *As $0 < \varepsilon^2 \ll h \to 0$ with $\varepsilon^{2/\beta-d-b'}h^{1+bd} \to 0$ for some $b' > 0$ and $b \in (0, 1/2)$,*

$$a_h^{-1}\lambda^d p_h^\varepsilon \sim (a_h^K)^{-1}\lambda^d p_h^{K\varepsilon}.$$

Proof: We just need to show that

$$\frac{a_h^{-1}\lambda^d p_h^\varepsilon - (a_h^K)^{-1}\lambda^d p_h^{K\varepsilon}}{a_h^{-1}\lambda^d p_h^\varepsilon} \to 0.$$

By Lemma 5.5(i), we get

$$a_h^{-1}\lambda^d p_h^\varepsilon \geq l_2(\varepsilon)\lambda^d p_{h'}\varepsilon^{d-2/\beta} = l_2(\varepsilon)\varepsilon^{d-2/\beta},$$

$$\begin{aligned} a_h^{-1}P_{1\{(B_0^{h^b})^c\}\lambda^d}\{\eta_h B_0^\varepsilon > 0\} &\leq l_1(\varepsilon)1\{(B_0^{h^b})^c\}\lambda^d p_h \varepsilon^{d-2/\beta} &(5.7)\\ &= l_1(\varepsilon)\varepsilon^{d-2/\beta}\int_{|x|\geq h^b} p_h(x)dx &(5.8)\\ &\leq l_1(\varepsilon)\varepsilon^{d-2/\beta}h^c, &(5.9)\end{aligned}$$

for some $c > 0$. Since $\varepsilon^{2/\beta-d-b'}h^{1+bd} \to 0$, we get

$$\frac{a_h^{-1}P_{1\{(B_0^{h^b})^c\}\lambda^d}\{\eta_h B_0^\varepsilon > 0\}}{a_h^{-1}\lambda^d p_h^\varepsilon} \to 0.$$

Similarly, we get

$$\frac{(a_h^K)^{-1}P_{1\{(B_0^{h^b})^c\}\lambda^d}\{\eta_h^K B_0^\varepsilon > 0\}}{a_h^{-1}\lambda^d p_h^\varepsilon} \to 0.$$

By Lemma 5.4, finally it suffices to show that

$$\frac{P_{1\{B_0^{h^b}\}\lambda^d}\{\xi_h B_0^\varepsilon > 0\} - P_{1\{B_0^{h^b}\}\lambda^d}\{\xi_h^K B_0^\varepsilon > 0\}}{a_h^{-1}\lambda^d p_h^\varepsilon} \to 0 \tag{5.10}$$

By Theorem 25.22 of [24] and (4.1),

$$\begin{aligned} &\varepsilon^{2/\beta-d}\left(P_{1\{B_0^{h^b}\}\lambda^d}\{\xi_h B_0^\varepsilon > 0\} - P_{1\{B_0^{h^b}\}\lambda^d}\{\xi_h^K B_0^\varepsilon > 0\}\right)\\ \leq\ & \varepsilon^{2/\beta-d}E_{1\{B_0^{h^b}\}\lambda^d}N_\xi\left([0,h],(K,\infty),\mathsf{R}^\mathsf{d}\right)\\ =\ & \varepsilon^{2/\beta-d}E_{1\{B_0^{h^b}\}\lambda^d}\hat{N}_\xi\left([0,h],(K,\infty),\mathsf{R}^\mathsf{d}\right)\\ \underset{\frown}{=}\ & \varepsilon^{2/\beta-d}E\int_0^h \|\xi_s\|ds \underset{\frown}{=} \varepsilon^{2/\beta-d}h^{1+bd} \to 0.\end{aligned}$$

□

Define normalizing functions $m(\varepsilon)$ by

$$m(\varepsilon) = a_{\varepsilon^c}^{-1} \lambda^d p_{\varepsilon^c}^{\varepsilon} \tag{5.11}$$

with a fixed c satisfying

$$(d - 2/\beta) + (-d/2 + 1/2)c = 0. \tag{5.12}$$

Clearly $c \in (0, 2)$.

Lemma 5.10 *Fix any bounded, uniformly continuous function $f \geq 0$ on R^d. As $\varepsilon \to 0$,*

$$\left\| \tilde{m}(\varepsilon)(a_{\varepsilon^r}^K)^{-1} (p_{\varepsilon^r}^{K\varepsilon} * f) - f \right\| \to 0.$$

The result holds uniformly over any class of uniformly bounded and equicontinuous functions $f \geq 0$ on R^d.

Proof: By Lemma 5.9, we get

$$\tilde{m}(\varepsilon)(a_{\varepsilon^r}^K)^{-1} \lambda_d \, p_{\varepsilon^r}^{K\varepsilon} \to 1.$$

Defining $\hat{p}_h^{K\varepsilon} = p_h^{K\varepsilon} / \lambda^d p_h^{K\varepsilon}$, we then only need to show that $\|\hat{p}_h^{K\varepsilon} * f - f\| \to 0$. Now follow the proof of Lemma 4.4(i) in [22] and use Lemma 5.5(ii). $\Box$

Theorem 5.11 *Let ξ be a superprocess in R^d. Then for any $t > 0$ and bounded μ, we have as $\varepsilon \to 0$*

$$\left\| \tilde{m}(\varepsilon) P_\mu \{\xi_t^K B_\cdot^\varepsilon > 0\} - e^{-b_K t} (\mu * p_t) \right\| \to 0,$$

$$\| \tilde{m}(\varepsilon) P_\mu \{\xi_t B_\cdot^\varepsilon > 0\} - \mu * p_t \| \to 0.$$

Proof:

$$P_\mu \{\xi_t^K B_\cdot^\varepsilon > 0\} \approx E_\mu(\zeta_s^K * p_h^{K\varepsilon}) = (a_h^K)^{-1} E_\mu(\xi_s^K * p_h^{K\varepsilon})$$

$$
\begin{aligned}
&= e^{-b_K s}(a_h^K)^{-1}(\mu * p_s * p_h^{K\varepsilon}) \approx e^{-b_K s} m(\varepsilon)(\mu * p_s) \\
&\approx e^{-b_K t} m(\varepsilon)(\mu * p_t).
\end{aligned}
$$

$$P_\mu\{\xi_t B_x^\varepsilon > 0\} \le E_\mu(\zeta_s * p_h^\varepsilon)$$

$$\|\tilde{m}(\varepsilon) E_\mu(\zeta_s * p_h^\varepsilon) - \mu * p_t\| \to 0,$$

$$P_\mu\{\xi_t B_x^\varepsilon > 0\} \ge P_\mu\{\xi_t^K B_x^\varepsilon > 0\}$$

$$\|e^{-b_K t}(\mu * p_t) - \mu * p_t\| \to 0$$

as $K \to \infty$ since $b_K \to 0$ $\square$

5.5 Lebesgue approximations

Same as in Section 5 of [22], here we begin with the Lebesgue approximation for ξ^K, the truncated K-process of ξ. Then we get the Lebesgue approximation for ξ immediately by Lemma 5.2. Write $\tilde{m}(\varepsilon) = 1/m(\varepsilon)$ for convenience, where $m(\varepsilon)$ is defined in (5.11). Recall that $\xi_t^{K\varepsilon} = (\xi_t^K)^\varepsilon$, the ε-neighborhood measure of ξ_t^K. For random measures ξ_n and ξ on R^d, $\xi_n \stackrel{w}{\to} \xi$ in L^1 means that $\xi_n f \to \xi f$ in L^1 for all f in C_b^d.

Theorem 5.12 *Let ξ^K be the truncated K-process of a superprocess ξ in R^d satisfying assuptions (A1) and (A2) with $\beta < 1$ and $d > 2/\beta$. Fix any $\mu \in \hat{\mathcal{M}}_d$ and $t > 0$. Then under P_μ, we have as $\varepsilon \to 0$:*

$$\tilde{m}(\varepsilon)\, \xi_t^{K\varepsilon} \stackrel{w}{\to} \xi_t^K \text{ a.s. and in } L^1$$

Proof: Fix any $f \in C_K^d$. We first prove that $\tilde{m}(\varepsilon)\,\xi_t^{K\varepsilon} f \to \xi_t^K f$ a.s. as $\varepsilon \to 0$. In order to do that, we only need to show that for any sequence $\varepsilon_n \to 0$ as $n \to \infty$, we can pick a subsequence (still denoted by ε_n) such that $\tilde{m}(\varepsilon_n)\,\xi_t^{K\varepsilon_n} f \to \xi_t^K f$ a.s. To do this we fix an $r \in (0,1)$. Then for any given sequence $\varepsilon_n \to 0$ as $n \to \infty$, we pick the subsequence ε_n satisfying $\varepsilon_n \le r^n$.

We follow the proof of Lemma 5.1 in [22]. Write η_h^{Ki} for the subclusters of ξ_t^K of age h. Since the ancestors of ξ_t^K at time $s = t - h$ form a Cox process directed by ξ_s^K / a_h^K, Lemma 5.7(i) yields

$$E_\mu\Big[\sum\nolimits_i \eta_h^{Ki\varepsilon} f \,\Big|\xi_s^K\Big] = (a_h^K)^{-1} \xi_s^K (p_h^{K\varepsilon} * f),$$

and so by Lemma 5.7(ii)

$$\begin{aligned} E_\mu\Big|\sum\nolimits_i \eta_h^{Ki\varepsilon} f - (a_h^K)^{-1}\,\xi_s^K(p_h^{K\varepsilon} * f)\Big|^2 &= E_\mu \mathrm{Var}\Big[\sum\nolimits_i \eta_h^{Ki\varepsilon} f \,\Big|\xi_s^K\Big] \\ &\precsim l_1(\varepsilon) a_h^K\, \varepsilon^{d-2/\beta}\, h^{d/2}\, \|f\|^2\, E_\mu \|\xi_s^K / a_h^K\| \\ &\le l_1(\varepsilon)\varepsilon^{d-2/\beta}\, h^{d/2}\, \|f\|^2\, \|\mu\|, \end{aligned}$$

where the last inequality follows from $E_\mu\|\xi_s^K\| \le \|\mu\|$. Combining with Lemma 5.8 gives

$$\begin{aligned} &E_\mu\big|\,\xi_t^{K\varepsilon} f - (a_h^K)^{-1}\,\xi_s^K(p_h^{K\varepsilon} * f)\,\big| \\ &\quad\le E_\mu\Big|\,\xi_t^{K\varepsilon} f - \sum\nolimits_i \eta_h^{Ki\varepsilon} f\,\Big| + E_\mu\Big|\sum\nolimits_i \eta_h^{Ki\varepsilon} f - (a_h^K)^{-1}\,\xi_s^K(p_h^{K\varepsilon} * f)\Big| \\ &\quad\precsim l_1^2(\varepsilon)\varepsilon^{2(d-2/\beta)}\, h^{1-d/2}\,\|f\| + l_1^{1/2}(\varepsilon)\varepsilon^{1/2(d-2/\beta)}\, h^{d/4}\,\|f\| \\ &\quad= \varepsilon^{d-2/\beta}\left(l_1^2(\varepsilon)\varepsilon^{d-2/\beta} h^{1-d/2} + l_1^{1/2}(\varepsilon)\varepsilon^{-1/2(d-2/\beta)} h^{d/4}\right)\|f\|. \end{aligned}$$

Taking $h_n = \varepsilon_n^c$, where c is defined in (5.12) and writing $s_n = t - h_n = t - \varepsilon_n^c$, we obtain

$$\begin{aligned} &E_\mu \sum_n \tilde{m}^K(\varepsilon_n)\,\big|\,\xi_t^{K\varepsilon_n} f - a_{h_n}^{-1}\,\xi_{s_n}^K(p_{h_n}^{K\varepsilon_n} * f)\,\big| \\ &\quad\precsim \sum_n \Big(l_2(r^n) l_1^2(r^n) r^{[(d-2/\beta)+(-d/2+1)c]n} + l_2(r^n) l_1^{1/2}(r^n) r^{[-1/2(d-2/\beta)+(d/4)c]n}\Big)\,\|f\| < \infty, \end{aligned}$$

since $(d-2/\beta)+(-d/2+1)c>0$ and $-1/2(d-2/\beta)+(d/4)c>0$ by (5.12). Note that in the previous inequality we also used the fact that the subsequence ε_n satisfying $\varepsilon_n \le r^n$. The inequality above about the expectations clearly implies

$$\tilde{m}^K(\varepsilon_n)\,\big|\,\xi_t^{K\varepsilon_n}f - a_{h_n}^{-1}\,\xi_{s_n}^K(p_{h_n}^{K\varepsilon_n} * f)\,\big| \to 0 \text{ a.s. } P_\mu. \tag{5.13}$$

Now we write

$$\begin{aligned} &\big|\,\tilde{m}^K(\varepsilon)\,\xi_t^{K\varepsilon}f - \xi_t^K f\,\big| \\ &\le\ \tilde{m}^K(\varepsilon)\,\big|\,\xi_t^{K\varepsilon}f - (a_h^K)^{-1}\,\xi_s^K(p_h^{K\varepsilon} * f)\,\big| + |\xi_s^K f - \xi_t^K f| \\ &\quad + \|\xi_s^K\|\,\big\|\,\tilde{m}^K(\varepsilon)\,a_h^{-1}\,(p_h^{K\varepsilon} * f) - f\,\big\|\,. \end{aligned}$$

For the last term, we first fix $b = 1/2 - 1/d$, then apply Lemma 4.16. Noting that by (4.11)

$$(2/\beta - d) + (1+bd)c = (2/\beta - d) + (d/2)c > 0,$$

we get by Lemma 4.16

$$\big\|\tilde{m}^K(\varepsilon)(a_h^K)^{-1}\,(p_h^{K\varepsilon} * f) - f\,\big\| \to 0$$

along the sequence (r^n). Using (5.13) and the a.s. weak continuity of ξ^K at the fixed time t, we see that the right-hand side tends a.s. to 0 as $n\to\infty$, which implies $\varepsilon^{2/\beta-d}\,\xi_t^{K\varepsilon}f \to c_{\beta,d}\,\xi_t^K f$ a.s. as $\varepsilon\to 0$ along the sequence (r^n) for any fixed $r\in(0,1)$. Since this holds simultaneously, outside a fixed null set, for all rational $r\in(0,1)$, the a.s. convergence extends by Lemma 2.3 in [25] to the entire interval $(0,1)$.

Applying this result to a countable, convergence-determining class of functions f (cf. Lemma 3.2.1 in [5]), we obtain the required a.s. vague convergence. Since μ is finite, the $(2,\beta)$-process ξ_t has a.s. compact support (cf. Theorem 9.3.2.2 of [5] and the proof of Theorem 1.2 in [6]). By Lemma 4.2, ξ_t^K also has a.s. compact support, and so the a.s. convergence remains valid in the weak sense. $\square$

Now we may prove our main result, the Lebesgue approximation of superprocesses with a regularly varying branching mechanism. Again, we write $\tilde{m}(\varepsilon) = 1/m(\varepsilon)$ for convenience, where $m(\varepsilon)$ is defined in (5.11). Also recall that $\xi_t^{K\varepsilon} = (\xi_t^K)^\varepsilon$, the ε-neighborhood measure of ξ_t^K. For random measures ξ_n and ξ on R^d, $\xi_n \xrightarrow{w} \xi$ in L^1 means that $\xi_n f \to \xi f$ in L^1 for all f in C_b^d.

Theorem 5.13 *Let the superprocess ξ in R^d satisfy assuptions (A1) and (A2) with $\beta < 1$ and $d > 2/\beta$. Fix any $\mu \in \hat{\mathcal{M}}_d$ and $t > 0$. Then under P_μ, we have as $\varepsilon \to 0$:*

$$\tilde{m}(\varepsilon)\, \xi_t^\varepsilon \xrightarrow{w} \xi_t \text{ a.s. and in } L^1$$

Proof: by Theorem 5.12 and Lemma 5.2 we get as $\varepsilon \to 0$

$$\tilde{m}(\varepsilon)\, \xi_t^\varepsilon \xrightarrow{w} \xi_t \text{ a.s..}$$

To prove the convergence in L^1, we note that for any $f \in C_b^d$

$$\begin{aligned} \tilde{m}(\varepsilon)\, E_\mu \xi_t^\varepsilon f &= \tilde{m}(\varepsilon) \int P_\mu\{\xi_t B_x^\varepsilon > 0\}\, f(x)\, dx \\ &\to \int (\mu * p_t)(x)\, f(x)\, dx = E_\mu \xi_t f, \end{aligned} \tag{5.14}$$

by Theorem 5.11. Combining this with the a.s. convergence under P_μ and using Proposition 4.12 in [24], we obtain $E_\mu|\tilde{m}(\varepsilon)\, \xi_t^\varepsilon f - \xi_t f| \to 0$. $\square$

If ξ is a $(2,1)$-process in R^d with $d \geq 3$, then $a_t = t$. By (4) in [25], we get

$$m(\varepsilon) = \varepsilon^{-r} \lambda^d p_{\varepsilon^r}^\varepsilon \sim \varepsilon^{-r} c_d\, \varepsilon^{d-2} \varepsilon^r = c_d\, \varepsilon^{d-2}.$$

So we recover the Lebesgue approximation of $(2,1)$-processes, that is,

$$\tilde{c}_d\, \varepsilon^{2-d}\, \xi_t^\varepsilon \xrightarrow{w} \xi_t \text{ a.s. and in } L^1.$$

Similarly, if ξ is a $(2,\beta)$-process in R^d with $\beta < 1$ and $d > 2/\beta$, then $a_t = (\beta t)^{1/\beta}$. By (9) in [22], we get

$$m(\varepsilon) = (\beta\varepsilon^r)^{-1/\beta}\lambda^d p^\varepsilon_{\varepsilon^r} \sim c_{\beta,d}\, \varepsilon^{d-2/\beta}.$$

Again, we recover the Lebesgue approximation of $(2,\beta)$-processes, that is,

$$\tilde{c}_{\beta,d}\, \varepsilon^{2/\beta-d}\, \xi_t^\varepsilon \xrightarrow{w} \xi_t \text{ a.s. and in } L^1.$$

Bibliography

[1] Bertoin, J., Le Gall, J. F. and Le Jan, Y. (1999). Spatial branching processes and subordination. *Canad. J. Math.* **49**, 24–54.

[2] Bass, R.F., Levin, D.A. (2002). Transition probabilities for symmetric jump processes. *Trans. Amer. Math. Soc.* **354**, 2933–2953.

[3] Dawson, D. A. (1975). Stochastic evolution equations and related measure processes. *J. Multivariate Anal.* **5** 1–52.

[4] Dawson, D.A. (1992). Infinitely divisible random measures and superprocesses In: *Stochastic analysis and related topics. Progr. Probab.* **31**, 1—129. Birkhäuser, Boston, MA.

[5] Dawson, D.A. (1993). Measure-valued Markov processes. In: *École d'Été de Probabilités de Saint-Flour XXI–1991. Lect. Notes in Math.* **1541**, 1—260. Springer, Berlin.

[6] Dawson, D.A., Iscoe, I., Perkins, E.A. (1989). Super-Brownian motion: Path properties and hitting probabilities. *Probab. Th. Rel. Fields* **83**, 135–205.

[7] Dawson, D.A., Perkins, E.A. (1991). Historical processes. *Mem. Amer. Math. Soc.* **93**, #454.

[8] Dawson, D. A., Vinogradov, V. (1994). Almost-sure path properties of $(2, d, \beta)$-superprocesses. *Stochastic Process. Appl.* **51** 221–258.

[9] Delmas, J.-F. (1999). Path Properties of Superprocesses with a General Branching Mechanism. *Ann. Probab.* **27**, 1099-1134.

[10] Delmas, J.-F. (1999). Some properties of the range of super-Brownian motion. *Probab. Th. Rel. Fields* **114**, 505–547.

[11] Duquesne, T. (2009). The packing measure of the range of Super-Brownian motion. *Ann. Probab.* **37**, 2431–2458.

[12] Duquesne, T., Le Gall, J.-F. (2002). Random trees, Lévy processes and spatial branching processes. *Astérisque* **281**, vi-147.

[13] Dynkin, E.B. (1991). Branching particle systems and superprocesses. *Ann. Probab.* **19**, 1157–1194.

[14] Dynkin, E.B. (1991). Path processes and historical superprocesses. *Probab. Th. Rel. Fields* **90**, 1–36.

[15] Dynkin, E.B. (1994). *An Introduction to Branching Measure-Valued Processes*. CRM Monograph Series **6**. AMS, Providence, RI.

[16] Dynkin, E.B. (2002). *Diffusions, Superdiffusions and Partial Differential Equations*. Colloquium Publications **50**. AMS, Providence, RI.

[17] Etheridge, A.M. (2000). *An Introduction to Superprocesses*. University Lecture Series **20**. AMS, Providence, RI.

[18] Evans, S.N., Perkins, E. (1991). Absolute continuity results for superprocesses with some applications. *Trans. Amer. Math. Soc.* **325**, 661–681.

[19] Falconer, K. (2007). *Fractal geometry: mathematical foundations and applications*. Wiley.

[20] Folland, G. B. (1999). *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, New York.

[21] Gmira, A., Veron, L. (1984). Large time behaviour of the solutions of a semilinear parabolic equation in R^N. *J. Differ. Equations* **53**, 258–276.

[22] He, X. (2013). Lebesgue approximation of $(2, \beta)$-superprocesses. *Stochastic Process. Appl.* **123** 1802–1819.

[23] He, X. (2013). Lebesgue approximation of superprocesses with a regularly varying branching mechanism. Unpublished manuscript.

[24] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York.

[25] Kallenberg, O. (2008). Some local approximations of Dawson-Watanabe superprocesses. *Ann. Probab.* **36**, 2176–2214.

[26] Kallenberg, O. (2011). Iterated Palm Conditioning and Some Slivnyak-Type Theorems for Cox and Cluster Processes. *J. Theor. Probab.* **24**, 875–893.

[27] Kallenberg, O. (2013). Local conditioning in Dawson-Watanabe superprocesses. *Ann. Probab.* **41**, 385–443.

[28] Kingman, J.F.C. (1973). An intrinsic description of local time. *J. London Math. Soc.* (2) **6**, 725–731.

[29] Klenke, A. (1998). Clustering and invariant measures for spatial branching models with infinite variance. *Ann. Probab.* **26**, 1057–1087.

[30] Le Gall, J.F. (1991). Brownian excursions, trees and measure-valued branching processes. *Ann. Probab.* **19**, 1399–1439.

[31] Le Gall, J. F. (1994). A lemma on super-Brownian motion with some applications. *The Dynkin Festschrift. Progr. Probab.* **24** 237–251. Birkhäuser, Boston, MA.

[32] Le Gall, J.F. (1999). *Spatial Branching Processes, Random Snakes and Partial Differential Equations.* Lectures in Mathematics, ETH Zürich. Birkhäuser, Basel.

[33] Le Gall, J.F., Perkins, E. (1995). The Hausdorff Measure of the Support of Two-Dimensional Super-Brownian Motion. *Ann. Probab.* **23**, 1719–1747.

[34] Le Gall, J.F., Perkins, E.A., Taylor S.J. (1995). The packing measure of the support of super-Brownian motion. *Stochastic Process. Appl.* **59** 1–20.

[35] Li, Z. (2011). *Measure-Valued Branching Markov Processes.* Springer, New York.

[36] Mörters, P, Peres, Y (2010). *Brownian Motion.* Cambridge University Press.

[37] Mytnik, L., Perkins, E. (2003). Regularity and irregularity of $(1+\beta)$-stable super-Brownian motion. *Ann. Probab.* **31**, 1413–1440.

[38] Mytnik, L., Villa, J. (2007). Self-intersection local time of (α, d, β)-superprocess. *Ann. Inst. Henri Poincar Probab. Stat.* **43**, 481–507.

[39] Pazy, A. (1983). *Semigroups of linear operators and applications to partial differential equations.* **198**. Springer, New York.

[40] Perkins, E. (1990). Polar sets and multiple points for super-Brownian motion. *Ann. Probab.* **18**, 453–491.

[41] Perkins, E. (1994). The strong Markov property of the support of super-Brownian motion. *The Dynkin Festschrift. Progr. Probab.* **24** 307–326. Birkhäuser, Boston, MA.

[42] Perkins, E. (2002). Dawson-Watanabe superprocesses and measure-valued diffusions. *École d'Été de Probabilités de Saint-Flour XXIX–1999. Lect. Notes in Math.* **1781**, 125–329. Springer, Berlin.

[43] Perkins, E. (2004). Super-Brownian motion and criticial spatial stochastic systems. *Bull. Can. Math. Soc.* **47**, 280–297.

[44] Revuz, D., Yor, M (1999). *Continuous martingales and Brownian motion.* **293**. Springer, Berlin.

[45] Sato, K. I. (1999). *Lévy processes and infinitely divisible distribution.* Cambridge University Press.

[46] Slade, G. (2002). Scaling limits and super-Brownian motion. *Notices Amer. Math. Soc.* **49**, 1056–1067.

[47] Stein, E.M., Shakarchi, R. (2005). *Real analysis: measure theory, integration, and Hilbert spaces.* Princeton Lectures in Analysis **3**. Princeton University Press.

[48] Tribe, R. (1994). A representation for super Brownian motion. *Stochastic Process. Appl.* **51** 207–219.

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