Superprocesses and Large-Scale Networks

submitted by

Maren Eckhoff

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

November 2014

SUMMARY

The main theme of this thesis is the use of the branching property in the analysis of random structures. The thesis consists of two self-contained parts.

In the first part, we study the long-term behaviour of supercritical superdiffusions and prove the strong law of large numbers. The key tools are spine and skeleton decompositions, and the analysis of the corresponding diffusions and branching particle diffusions.

In the second part, we consider preferential attachment networks and quantify their vulnerability to targeted attacks. Despite the very involved global topology, locally the network can be approximated by a multitype branching random walk with two killing boundaries. Our arguments exploit this connection.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my supervisors, Andreas Kyprianou and Peter Mörters, for their constant support and encouragement during the last three years. Their enthusiasm, the numerous enjoyable meetings, and the inspiring discussions made my PhD a very enriching experience.

I also wish to thank my other coauthors Jesse Goodman, Remco van der Hofstad, Francesca Nardi, Silke Rolles, and Matthias Winkel from whom I have learned so much. Moreover, I would like to thank all members of the Prob-L@b at Bath and, generally, the Department of Mathematical Sciences for providing a very friendly and stimulating working environment. I thank my fantastic office mates, Matt, Acyr, Ben, and Hüseyin, for going with me through all the highs and lows, and for the Kinder surprises.

I have been extremely fortunate to be taught by very knowledgeable and inspiring people during my time at Gymnasium Meckelfeld, the Technical University of Munich, the Technical University of Eindhoven and, of course, here in Bath. Besides Andreas and Peter, I would particularly like to thank Remco, Silke, Herrn Schied, and Herrn Beutler.

A big thank you to all my friends, especially to Janina, Marina, Nina, Quynh, Michi, Domi, Schmidie, Mark, Sven and Luke, for always being there and for numerous wonderful holidays, weekends and evenings, and Istvan, Andrew and Nathan for all the entertaining and enlightening conversations. I would like to thank in particular Luke for many adventures and laughs, for his patience, encouragement and love.

Finally, I wish to thank Mom, Peter, Julia and Thomas for their endless support and invaluable reflections. None of this would have been possible without you.

CONTENTS

INTRODUCTION

Branching processes are one of the fundamental objects studied in probability theory. The reason is twofold: one is the striking beauty of the modern mathematical analysis of these processes, the other is the fact that they appear naturally in the study of many real-world phenomena. Examples range from population genetics, over large-scale networks, to the search for the optimal route between two locations.

Historically, branching processes have been considered as stochastic population models. At every given time, there is a family of individuals, and each individual reproduces independently of all others, living or dead. The population is observed either continuously at every time $t \geq 0$, or only at discrete time steps, $t \in \mathbb{N}_0$, say. The reproduction law and lifetime can depend on attributes of the individual that are usually encoded by a type or spatial location. In the latter case, one may also consider spatial movement during the lifetime of individuals.

The key characteristic of branching processes is the *branching property* that states the following. If the population is split into two subpopulations at a fixed time t, then the families descending from the members of the two subpopulations form independent copies of the original process with initial distributions given by the respective configurations at time t. Starting from this property, a good understanding of branching processes has been developed since the 1950s, and many key concepts like criticality and the asymptotic behaviour of finite-type processes have been well understood for some time now [83, 93, 7]. Recent advances led to deep structural insights, and made branching processes once again one of the hot topics in probability theory; see [1] and the references therein.

In this thesis we aim on the one hand to add to the theoretical understanding of branching processes and, on the other, to highlight their strength in the analysis of applied problems.

Motivation and perspective

This thesis consists of two self-contained parts. The first part is concerned with the asymptotic behaviour of spatial branching processes; the second analyses the vulnerability of large-scale networks to targeted attacks.

Part I. Superprocesses appeared first in work of Watanabe [122] as a high-density limit of branching particle processes.

A branching particle process $Y = (Y_t)_{t\geq 0}$ describes a cloud of particles that evolve over time as follows[1]. Start with a finite number of particles in $\mathbb{R}^d$. Each particle moves according to the same distribution as a Markov process $(\xi = (\xi_t)_{t\geq 0} \colon (\mathbb{P}_x)_{x\in\mathbb{R}^d})$, and dies after an exponential time of rate $q > 0$. Upon its death, the particle gives birth to a random number of offspring according to a distribution $(p_k)_{k\in\mathbb{N}_0}$. Each offspring independently repeats the behaviour of its parent starting from its location of birth.

The configuration of particles at time t can be expressed as an integer-valued measure $Y_t = \sum_i \delta_{\xi_i(t)}$, where $\xi_i(t)$ denotes the location of the i-th particle at time t. Write P_ν for the distribution of Y when started with initial configuration ν, and $\langle f, \mu\rangle = \int f \, d\mu$ for the integral of a function f with respect to a measure μ. The branching property yields, for any bounded, nonnegative function f,

$$P_{\sum_i \delta_{x_i}}\big[e^{-\langle f, Y_t\rangle}\big] = \prod_i P_{\delta_{x_i}}\big[e^{-\langle f, Y_t\rangle}\big] = \exp\Big(- \Big\langle v_f(\cdot, t), \sum_i \delta_{x_i} \Big\rangle\Big),$$

where $v_f(x,t) = -\log P_{\delta_x}[e^{-\langle f, Y_t\rangle}]$. We denote by $\varphi(z) = \sum_{k=0}^\infty p_k z^k$, $z \in [0,1]$, the probability generating function of the offspring distribution. The function v_f can be characterised by splitting at the time where the first particle dies to obtain the integral equation

$$e^{-v_f(x,t)} = \mathbb{P}_x\big[e^{-qt} e^{-f(\xi_t)}\big] + \int_0^t \mathbb{P}_x\big[q e^{-qs} \varphi\big(e^{-v_f(\xi_s, t-s)}\big)\big] \, ds. \tag{$*$}$$

Here the first summand corresponds to the case that the initial particle survives at least until time t. The second term accounts for the case that the initial particle dies at time $s \in [0,t]$ in location ξ_s. With probability p_k it then has k offspring that initialise independent copies of Y that run for the remaining $t - s$ time.

If ξ is a diffusion with generator L (c.f. [110, 119]) and if suitable regularity conditions are satisfied, then the unique solution $u(x,t) = e^{-v_f(x,t)}$ to ($*$) is equal to the unique, $[0,1]$-valued solution to the differential equation

$$\begin{aligned} \partial_t u &= Lu - q(\varphi(u) - u) && \text{on } \mathbb{R}^d \times (0,\infty), \\ u(\cdot, 0) &= e^{-f} && \text{on } \mathbb{R}^d; \end{aligned} \tag{$**$}$$

[1] We describe only a simple case here. Greater generality and more details can be found in Part I of the thesis.

see, for example, Chapter 3 in [119] or Chapter 5 in [44]. This close relationship between branching particle processes and partial differential equations (PDEs) facilitates a spectacular interplay of probabilistic and analytic methods in the analysis of both, stochastic processes and PDEs. See [10] for a good literature review in the case that ξ is a Brownian motion.

For small populations, a branching particle process is a suitable model that tracks the evolution of types (recorded as a location in the above description). However, for very large populations the concentration of particles can become extremely high, and it is preferable to record the proportion of particles of a certain type instead of absolute numbers. A continuous approximation then leads to a finite measure-valued Markov process – a superprocess – that captures many of the key properties of interest. Besides the motion ξ, the main parameter of a superprocess is the branching mechanism $z \mapsto \psi(z)$ which plays the role of $z \mapsto q(\varphi(z)-z)$ from ($**$) in the continuum limit [43, 59, 66]. If ξ is again a diffusion with generator L, and if suitable regularity assumptions hold, then the superprocess $X = (X_t)_{t\geq 0}$ satisfies

$$P_\mu\big[e^{-\langle f, X_t\rangle}\big] = e^{-\langle u_f(\cdot,t),\mu\rangle}$$

for all initial configurations μ, and nonnegative, bounded, continuous test functions f, where $(x,t) \mapsto u_f(x,t)$ is the unique, nonnegative solution to

$$\begin{aligned} \partial_t u &= Lu - \psi(u) && \text{on } \mathbb{R}^d \times (0,\infty),\\ u(\cdot,0) &= f && \text{on } \mathbb{R}^d. \end{aligned}$$

A simple example for a branching mechanism is given by the quadratic function $\psi(z) = -\beta z + \alpha z^2$ with $\beta \in \mathbb{R}$ and $\alpha > 0$. The particles in an approximating branching particle process reproduce at rate n/c according to an offspring distribution with mean $1+c\beta/n$ and variance roughly $c\alpha$, where $c > 0$ is a normalising constant and $n \to \infty$ parametrises the limiting procedure for the approximation. However, the range of possible branching mechanisms goes far beyond quadratic functions or polynomials and therefore offers a probabilistic interpretation for PDEs not covered by ($**$). In particular, superprocesses with branching mechanism

$$\psi(z) = -\beta z + \gamma z^p, \qquad \text{where } p \in (1,2],$$

$\beta \in \mathbb{R}$, $\gamma > 0$, received a lot of attention in the literature; see [67, 54, 41] and the references therein.

Further motivation for the study of superprocesses comes from stochastic differential equations and statistical mechanics, and we refer the reader to Le Gall's lecture notes [100] for an excellent overview and some examples.

The first powerful tools for the analysis of superprocesses were developed by Dawson [34]. Following his pioneering work, a large variety of techniques have been established by Dynkin, Kuznetsov, Perkins, Le Gall, Donnelly, Kurtz and many others; see [44, 108, 100, 66] for comprehensive overviews, detailed discussions, and further references. Recent developments are reviewed in Part I of this thesis, and by Engländer in [53] and the forthcoming book [56].

A fundamental question regarding spatial branching processes concerns how they distribute mass in space. For a given set of types $B \subseteq \mathbb{R}^d$, the goal is to understand how many particles are located in B, i.e. $Y_t(B)$, or what proportion of the population is of a type in B, i.e. $X_t(B)$ in the above notation. There is a substantial body of literature analysing the asymptotic behaviour of the total mass assigned to compact sets B. The results usually state that, for suitable starting measures μ and test functions f,

$$\langle f, X_t \rangle \sim P_\mu\big[\langle f, X_t \rangle\big] W_\infty \qquad \text{as } t \to \infty,$$

where W_∞ is a finite, non-trivial random variable, and $a_t \sim b_t$ if $a_t/b_t \to 1$ in some sense as $t \to \infty$. (The statement for branching particle processes is analogous where X_t is replaced by Y_t.) However, establishing such a convergence almost surely for superprocesses proved to be challenging [54]. Our approach to the problem is based on the skeleton decomposition for supercritical superprocesses that constitutes a pathwise representation of the superprocess as immigration along a branching particle process. In Part I of this thesis, we exploit that representation, and carry results for particle processes into the superprocess setup to establish the first strong law of large numbers for superprocesses that covers many of the key processes studied in the literature.

This research exposes an intriguing view on the skeleton decomposition that turns out to be the right approach to the study and understanding of the asymptotic behaviour of measure-valued processes. It would be of interest to extend the ideas to study superprocesses based on discontinuous motions, branching mechanisms satisfying weaker moment assumptions, and to cover the so-called "non-ergodic case" (see Section 1.4 and page 28 below for discussions in this direction). Moreover, it is desirable to deepen the understanding of the relationship between the measure-valued process and its skeleton, for example through a joint spine decomposition.

Part II Complex networks play a fundamental role in our everyday life. The Internet, the power grid, protein interaction networks and social networks are only a few examples. For the mathematical analysis, members of the network are interpreted as vertices of a graph and links are modelled as edges. To deal with the presence of uncertainty in the network formation, the edges are allocated in a probabilistic fashion.

The first and simplest random graph model was formalised and studied by Gilbert [76], and Erdős and Rényi [63, 64]. It arises by taking n vertices and placing an edge between each pair of vertices independently with a fixed probability p. The model was named after Erdős and Rényi who showed that for $p = c/n$ the largest component in the network undergoes a phase transition at $c = 1$. In fact, asymptotically as $n \to \infty$, if $c < 1$, the graph consists only of small components of order at most $\log n$; if $c > 1$, there is a giant component of order n; and if $c = 1$, then it has a largest component with of order $n^{2/3}$ vertices. This striking result caught the attention of scientists from many fields and may be viewed as the starting point of the field of random graphs.

In the analysis of the Erdős–Rényi random graph (ERRG), a wide range of methods have been developed [18, 89]. However, already in the earlier papers it was mentioned that the model is too simplistic to describe real-world networks. The advancement of modern data analysis in the last 15-20 years allowed scientists to get a more accurate description of complex structures and led to the formulation of many new random graph models.

One, by now, well-established feature of most real-world networks is that they are scale-free, which means that their degree distribution does not depend on the network size, and the proportion of nodes with degree k has a decay of order $k^{-\tau}$ for a power law exponent $\tau > 1$. A scale-free random graph model can be obtained as a variation of the ERRG where edges are independent but the probability for the edge between vertices i and j, denoted by p_{ij}, varies in i and j. Models of this type are called *inhomogeneous random graphs*. For suitable choices of $(p_{ij} : i, j \in \{1, \dots, n\})$ the network is scale-free [20, 29, 85]. Another way to achieve an asymptotic degree distribution with a power law is to construct the network so that it has a given degree sequence. This approach leads to the *configuration model* [2, 106, 85]. One criticism of these models is that the power law is introduced artificially by choosing the edge probabilities p_{ij} or degree distributions.

In a highly influential publication, Barabási and Albert [9] argued that the main topological structure of networks like the World Wide Web or social networks can be explained by the fact that they are built dynamically, and new vertices prefer to connect to vertices which already have a high degree. Models that obey these principles are called *preferential attachment models*. Bollobás and Riordan [22] gave the first mathematically rigorous formulation, insisting that new vertices come with a fixed number of edges and that the probability for an edge between the new and an old vertex are proportional to the degree of the old vertex. Jointly with Spencer and Tusnády [24], they showed that the resulting model, called the LCD-model, is scale-free with power law exponent $\tau = 3$. Van der Hofstad [85] demonstrated that any power law exponent can be achieved when the edge probabilities are chosen proportional to an affine function of the degree.

The analysis of preferential attachment models proved difficult. Bollobás, Riordan and co-authors developed highly technical tools to get a grip on the LCD-model [114, 23, 22]; van der Hofstad and co-authors used a wide range of techniques to characterise the generalised version [85, 38]. However, the understanding of the configuration model and certain inhomogeneous random graphs advanced much faster to an impressive level of detail. The key to this deep understanding is a branching process approximation.

The use of branching processes in the analysis of networks is a long success story.

It is based on two fundamental observations. Firstly, many global features of large-scale complex networks are determined by the local neighbourhoods of the vertices, and secondly, the local neighbourhood of a typical vertex is similar to a tree. How strong these two features are depends crucially on the model. To see the connection between a network and a branching process, one explores the neighbourhood of a vertex in the graph step-by-step, and interprets the discovery of a new member of the cluster as the birth of a new individual in the branching process. The stronger the inhomogeneity in the graph, the more attributes of individuals have to be recorded to describe a suitable offspring distribution. For example, the homogeneous ERRG with edge probability $p = c/n$ is approximated by a Poisson(c)-Galton–Watson process [85], which is a process with indistinguishable individuals. By contrast, some sophisticated inhomogeneous random graphs are approximated by branching processes with uncountably many types [19].

To show a property of a random graph model using a branching process approximation, one must first establish a suitable coupling, and then study the relevant property in the branching process. The simpler dependency structure and the detailed knowledge of branching processes available make the second step often a much simpler task than the original problem.

For preferential attachment models a branching process approximation was found only recently by Dereich and Mörters [37]. Their model differs from the previously mentioned preferential attachment models in the sense that the number of edges a new vertex attaches to existing vertices is not fixed. This is not only more desirable from a modelling perspective, but also makes the model more tractable, whilst keeping all the characteristic features popularised by Barabási and Albert. Quantitatively, the model behaves in the same way as the LCD-model and its generalised version. The approximating branching process is considerably more involved than previously found approximations: it could be viewed as a multitype Galton–Watson process, but more intuition can be drawn from an interpretation as a multitype branching random walk with killing boundary (see Section 5.6 for a discussion). The discovery of the branching process approximation opens up the path to a much deeper understanding of preferential attachment models, which has already led to Dereich and Mörters identifying the exact location of the phase transition of the size of the largest component.

In Part II of this thesis, we continue that work and adapt the branching process approximation to investigate the vulnerability of preferential attachment networks to a targeted attack on highly-connected vertices. For power law exponent $\tau \in (2, 3)$, the network is robust to the random removal of vertices but a targeted attack changes its topology dramatically. Our analysis constitutes the first mathematically rigorous, systematic study of this phenomenon.

The fascination behind the asymptotic power law degree distribution comes not only from the fact that it has been observed experimentally in most real-world networks, but also from the discovery that most key properties of random graphs are determined by the power law exponent τ. While a different mathematical analysis may be required for different models, the scalings agree when the power law exponents agree.

This folklore knowledge has been established over the last 10-15 years and has been demonstrated for various models and properties; see [85] for an excellent overview. One contribution of this thesis is to highlight qualitatively *different behaviour* of models that have the *same* power law exponent.

We call models of *rank one* if the edge probabilities (roughly) factorise, i.e.,

$$p_{i,j} \approx \chi(i)\chi(j) \qquad (***)$$

for some function χ. Examples include the configuration model and inhomogeneous random graphs that satisfy $(***)$.

A comparison of the two model classes given by scale-free graphs with power law exponent τ that are either of rank one or preferential attachment networks shows the following. In rank one models with power law exponent $\tau \in (2,3)$, the shortest path between two typical vertices consists of two paths that go through layers of increasingly well-connected vertices and meet at a highly connected vertex. In contrast, typical paths in preferential attachment models alternate between well-connected vertices and ordinary vertices, and are therefore twice as long as the typical paths in rank one models [107, 35]. In Part II of this thesis we show that this different behaviour leads to different critical exponents for the question of vulnerability. The analysis exposes a striking difference between the model classes, structurally and quantitatively.

For power law exponents $\tau > 3$, the difference is even more compelling. Consider the supercritical regime where a positive fraction of all vertices lies in one component, the giant component. While in rank one models the relative size of the giant component experiences a polynomial decay close to criticality, the decay is exponentially fast in preferential attachment models [52, 46, 114]. The analysis of this phenomenon offers another beautiful application of modern branching process techniques in the study of complex networks.

Based on these results, it would be of great interest to explore the difference between rank one and preferential attachment models further. A particularly interesting property to study would be the size of the largest component in the subcritical regime, where a remarkable difference between the two model classes can be expected.

Another intriguing open problem for preferential attachment models would be to understand the size of the giant component close to criticality in the regime where the second moment of the asymptotic degree distribution is infinite.

Further applications The range of problems where branching processes play an integral part is vast. Besides the fields of population genetics and random graphs, computer sciences, epidemiology, and queuing theory are usually mentioned [26, 8, 92].

An example related to our study of random graphs is that of flows on networks:

many real-world networks transport information, passengers or commodities, and links require a certain time or have a certain economic cost attached to them. Mathematically, these networks can be modelled by edge-weighted graphs. The aim is to understand properties of the weight and length of the optimal path between two vertices. For a more global picture, the topology of the graph consisting of the shortest paths from one vertex to all other vertices, the smallest-weight graph, is of importance. These questions lead to first passage percolation on random or deterministic graphs, a field that has attracted considerable attention in recent years; c.f. Chapter 12 in [85].

Similar to the study of random graphs, the local neighbourhood of a vertex in the smallest-weight graph can be approximated by a branching process. However, here the edge weights should be interpreted as birth times of particles in the branching process, leading to continuous-time instead of discrete-time processes [14]. For edge weight distributions with a heavy tail at zero, the approximation leads to a continuous-time branching process that depends crucially on the graph size. Consequently, a study of the double asymptotics, large time and large graph, is required [48, 49].

This, and the discussed preferential attachment networks, are prime examples in which the application of branching process theory paves the way to a deep understanding of very involved structures. However, the discovered branching processes are very sophisticated, and their analysis requires state-of-the-art techniques and new results. The consequence is a fruitful interaction between the branching process community and other fields that leads to the development of powerful methods and remarkable insights.

Publication and collaboration details

Part I of this thesis is joint work with Andreas E. Kyprianou and Matthias Winkel. It has been accepted for publication in the Annals of Probability, and forms reference [50].

Part II is joint work with Peter Mörters. It has been published in the Electronic Journal of Probability, and forms reference [51].

During my time as a PhD student in Bath I have worked on further projects that are not included in this thesis. For completeness, I will mention that joint work with Jesse Goodman, Remco van der Hofstad and Francesca Nardi on first passage percolation has been published [47] and is being prepared for publication [48, 49]. Further work on preferential attachment networks with Peter Mörters is currently being prepared for publication [52].

Part I

Spines, skeletons and the strong law of large numbers for superdiffusions

Consider a supercritical superdiffusion $(X_t)_{t\geq 0}$ on a domain $D \subseteq \mathbb{R}^d$ with branching mechanism

$$(x, z) \mapsto -\beta(x)z + \alpha(x)z^2 + \int_{(0,\infty)} \left(e^{-zy} - 1 + zy\right) \Pi(x, dy).$$

The skeleton decomposition provides a pathwise description of the process in terms of immigration along a branching particle diffusion. We use this decomposition to derive the strong law of large numbers (SLLN) for a wide class of superdiffusions from the corresponding result for branching particle diffusions. That is, we show that, for suitable test functions f and starting measures μ,

$$\frac{\langle f, X_t \rangle}{P_\mu[\langle f, X_t \rangle]} \to W_\infty \qquad P_\mu\text{-almost surely as } t \to \infty,$$

where W_∞ is a finite, non-trivial random variable characterised as a martingale limit. Our method is based on skeleton and spine techniques and offers structural insights into the driving force behind the SLLN for superdiffusions. The result covers many of the key examples of interest and, in particular, proves a conjecture by Fleischmann and Swart [74] for the super-Wright–Fisher diffusion.

CHAPTER 1

INTRODUCTION

The asymptotic behaviour of the total mass assigned to a compact set by a superprocess was first characterised by Pinsky [111] at the level of the first moment. Motivated by this study, Engländer and Turaev [61] proved weak convergence of the ratio of the total mass in a compact set and its expectation. Others have further improved the mode of convergence; specifically, several authors conjectured an almost sure convergence result for a wide class of superprocesses [54, 62, 74, 102]. However, up to now it has not been possible to deal with many of the classical examples of interest. In the existing literature, for almost sure convergence, either motion and branching mechanism have to obey restrictive conditions [27] or the domain is assumed to be of finite Lebesgue measure [102]. In this thesis, we make a significant step towards closing the gap and establish the strong law of large numbers (SLLN) for a large class of superdiffusions on arbitrary domains. In particular, we prove a conjecture by Fleischmann and Swart for the super-Wright–Fisher diffusion.

Methodologically, previous results concerned with almost sure limit behaviour of superprocesses relied on Fourier analysis, functional analytic arguments or used the martingale formulation for superprocesses combined with stochastic analysis. We take a different approach. The core of our proof is the skeleton decomposition that represents the superprocess as an immigration process along a branching particle process, called the skeleton, where immigration occurs in a Poissonian way along the space-time trajectories and at the branch points of the skeleton. The skeleton may be interpreted as immortal particles that determine the long-term behaviour of the process. We exploit this fact and carry the SLLN from the skeleton over to the superprocess. Apart from the result itself, this approach provides insights into the driving force behind the law of large numbers for superprocesses.

A more detailed literature review and discussion of the ideas of proof is deferred to Sections 1.4 and 1.5. Before, we introduce the model in Section 1.1, our assumptions are stated in Section 1.2, and the main results are collected in Section 1.3.

1.1 Model and notation

Let $d \in \mathbb{N}$ and let $D \subseteq \mathbb{R}^d$ be a nonempty domain. For $k \in \mathbb{N}_0$, $\eta > 0$, we write $C^{k,\eta}(D)$ for the space of real-valued functions on D, whose k-th order partial derivatives are locally η-Hölder continuous, $C^{\eta}(D) := C^{0,\eta}(D)$. We denote by $\mathcal{B}(D)$ the Borel σ-algebra on D. The notation $B \subset\subset D$ means that $B \in \mathcal{B}(D)$ is bounded and there is an open set B_1 such that $B \subseteq B_1 \subseteq \overline{B}_1 \subseteq D$. The Lebesgue measure on $\mathcal{B}(D)$ is denoted by ℓ; the set of finite (and compactly supported) measures on $\mathcal{B}(D)$ is denoted by $\mathcal{M}_f(D)$ (and $\mathcal{M}_c(D)$ resp.). When μ is a measure on $\mathcal{B}(D)$ and $f\colon D \to \mathbb{R}$ measurable, let $\langle f, \mu \rangle := \int_D f(x)\,\mu(dx)$, whenever the right-hand side makes sense. If μ has a density ρ with respect to ℓ, we write $\langle f, \rho \rangle = \langle f, \mu \rangle$. For any metric space E, we denote by $p(E)$ and $b(E)$ the sets of Borel measurable functions on E that are nonnegative and bounded, respectively. Let $bp(E) = b(E) \cap p(E)$.

Let $(\xi = (\xi_t)_{t\geq 0}; (\mathbb{P}_x)_{x\in D})$ be a diffusion process on D with generator

$$L(x) = \frac{1}{2}\nabla \cdot a(x)\nabla + b(x) \cdot \nabla \qquad \text{on } D.$$

The diffusion matrix $a\colon D \to \mathbb{R}^{d\times d}$ takes values in the set of symmetric, positive definite matrices. Moreover, all components of a and $b\colon D \to \mathbb{R}^d$ belong to $C^{1,\eta}(D)$ for some $\eta \in (0, 1]$ (the parameter η remains fixed throughout Part I of this thesis). In other words, ξ denotes the unique solution to the generalised martingale problem associated with L on $D \cup \{\dagger\}$, the one-point compactification of D with cemetery state $\dagger$; see Chapter I in [110]. We write $\tau_D = \inf\{t \geq 0\colon \xi_t \notin D\}$.

Let $\beta \in C^{\eta}(D)$ be bounded and

$$\psi_0(x, z) := \alpha(x)z^2 + \int_{(0,\infty)} \left(e^{-zy} - 1 + zy\right) \Pi(x, dy), \tag{1.1}$$

where $\alpha \in bp(D)$ and Π is a kernel from D to $(0, \infty)$ such that $x \mapsto \int_{(0,\infty)}(y \wedge y^2)\,\Pi(x, dy)$ belongs to $bp(D)$. The function $\psi_\beta(x, z) := -\beta(x)z + \psi_0(x, z)$ is called the *branching mechanism*. If $\Pi \equiv 0$, we say that the branching mechanism is quadratic. In Section 4.2, we explain that our results carry over to a class of quadratic branching mechanisms with unbounded α and β.

The main process of interest is the $(L, \psi_\beta; D)$-*superdiffusion*, which we denote by $X = (X_t)_{t\geq 0}$. Its distribution is denoted by P_μ if the process is started in $\mu \in \mathcal{M}_f(D)$. That is, X is an $\mathcal{M}_f(D)$-valued time-homogeneous Markov process such that, for all $\mu \in \mathcal{M}_f(D)$, $f \in bp(D)$ and $t \geq 0$,

$$P_\mu[e^{-\langle f, X_t \rangle}] = e^{-\langle u_f(\cdot,t), \mu \rangle}, \tag{1.2}$$

where u_f is the unique nonnegative solution to the *mild equation*

$$u(x,t) = S_t f(x) - \int_0^t S_s[\psi_0(\cdot, u(\cdot, t-s))](x)\, ds \quad \text{for all } (x,t) \in D \times [0,\infty). \tag{1.3}$$

Here $S_t g(x) := \mathbb{P}_x[e^{\int_0^t \beta(\xi_s)\, ds} g(\xi_t)\mathbb{1}_{\{t<\tau_D\}}]$ for all $g \in p(D)$, i.e. $(S_t)_{t\geq 0}$ denotes the semigroup of the differential operator $L + \beta$. Every function g on D is automatically extended to $D \cup \{\dagger\}$ by $g(\dagger) := 0$. Hence,

$$S_t g(x) = \mathbb{P}_x\left[e^{\int_0^t \beta(\xi_s)\, ds} g(\xi_t)\right].$$

We refer to ξ as the *underlying motion* or just the *motion* in the *space* D. Informally, the $\mathcal{M}_f(D)$-valued process $X = (X_t)_{t\geq 0}$ describes a cloud of infinitesimal particles independently evolving according to the motion ξ and branching in a spatially dependent way according to the branching mechanism ψ_β. The existence of the superprocess X is guaranteed by [43, 72], and it satisfies the branching property (see (1.1) in [72] for a definition). By Theorem 3.1 in [42] or Theorem 2.11 in [72], there is a version of X such that $t \mapsto \langle f, X_t\rangle$ is almost surely right-continuous for all continuous $f \in bp(D)$. We will always work with this version. In most texts the mild equation (1.3) is written in a slightly different form: instead of $(S_t)_{t\geq 0}$, the semigroup of L is used and ψ_0 is replaced by ψ_β. However, since the first moment of the superprocess is determined by $(S_t)_{t\geq 0}$, while ψ_0 influences only higher moments (c.f. (2.11) below), the mild equation in form (1.3) leads to simpler moment estimates. Moreover, (1.3) has a unique nonnegative solution even when β is only bounded from above, not necessarily from below; see Appendix B and the discussion around (2.7). Using Feynman–Kac arguments (see Lemma A.1 (i) in the appendix below) and Gronwall's lemma, one easily checks that (for bounded β) the two representations are equivalent.

Our main goal is to determine the large-time behaviour of

$$\frac{\langle f, X_t\rangle}{P_\mu[\langle f, X_t\rangle]} \tag{1.4}$$

for suitable test functions f and starting measures μ. We say that X satisfies the *strong law of large numbers* (SLLN) if, for all test functions $f \in C_c^+(D)$, $f \neq \mathbf{0}$, the ratio in (1.4) converges to a finite, non-trivial random variable which is independent of f. Here $C_c^+(D)$ denotes the space of nonnegative, continuous functions of compact support, and $\mathbf{0}$ is the constant function with value 0.

1.2 Statement of assumptions

A probabilistic view on supercritical superprocesses is offered by the *skeleton decomposition*. This, by now classical [70, 59, 40, 13, 11, 98], decomposition has been studied under a variety of names. It provides a pathwise representation of the superprocess as an immigration process along a supercritical branching particle process that we call the *skeleton*. The skeleton captures the global behaviour of the superprocess and its discrete nature makes it much more tractable than the superprocess itself. We exploit these facts to establish the SLLN for superdiffusions. Specifically, our fundamental aim it to show that the SLLN for superdiffusions follows as soon as an appropriate SLLN holds for its skeleton. Given the existing knowledge for branching particle processes, this will lead us to a large class of superprocesses for which the SLLN can be stated.

Classically, the skeleton was constructed using the event $\mathcal{E}_{\text{fin}} = \{\exists t \geq 0\colon X_t(D) = 0\}$ of extinction after finite time to guide the branching particle process into regions where extinction of the superprocess is unlikely. The key property of $\mathcal{E}_{\text{fin}}$ exploited in the skeleton decomposition is that the function $x \mapsto w(x) = -\log P_{\delta_x}(\mathcal{E}_{\text{fin}})$ gives rise to the multiplicative martingale $((e^{-\langle w, X_t\rangle})_{t\geq 0}; P_\mu)$. In the more general setup of this thesis, we assume only the existence of such a *martingale function* w.

Assumption 1 (Skeleton assumption)**.** There exists a function $w \in p(D)$ that satisfies $w(x) > 0$ for all $x \in D$,

$$\sup_{x\in B} w(x) < \infty \qquad \text{for all } B \subset\subset D, \tag{1.5}$$

$$P_\mu\big[e^{-\langle w, X_t\rangle}\big] = e^{-\langle w,\mu\rangle} \qquad \text{for all } \mu \in \mathcal{M}_c(D), t \geq 0. \tag{1.6}$$

The martingale function w allows us to define the skeleton as a branching particle diffusion Z, where the spatial movement of each particle is equal in distribution to $(\xi = (\xi_t)_{t\geq 0}; (\mathbb{P}^w_x)_{x\in D})$ with

$$\frac{d\mathbb{P}^w_x}{d\mathbb{P}_x}\bigg|_{\sigma(\xi_s\colon s\in[0,t])} = \frac{w(\xi_t)}{w(x)} \exp\Big(-\int_0^t \frac{\psi_\beta(\xi_s, w(\xi_s))}{w(\xi_s)}\, ds\Big) \qquad \text{on } \{t < \tau_D\} \tag{1.7}$$

for all $t \geq 0$. We will see in Lemma 2.2 that $\mathbb{P}^w_x$ is well-defined. Each particle dies at spatially dependent rate $q \in p(D)$ and is replaced by a random number of offspring with distribution $(p_k(x))_{k\geq 2}$, where x is the location of its death. The branching rate q and the offspring distribution $(p_k)_{k\geq 2}$ are uniquely identified by

$$G(x,s) := q(x)\sum_{k=2}^{\infty} p_k(x)(s^k - s) = \frac{1}{w(x)}\Big(\psi_0\big(x, w(x)(1-s)\big) - (1-s)\psi_0\big(x, w(x)\big)\Big) \tag{1.8}$$

for all $s \in [0,1]$ and $x \in D$. The fact that q and $(p_k)_{k\geq 2}$ are well-defined by (1.8) is contained in Theorem 2.3 (i) below. In Section 2.1.1, we define Z on a rich probability space with probability measures $\mathbf{P}_\mu$, $\mu \in \mathcal{M}_f(D)$, where the initial configuration of Z under $\mathbf{P}_\mu$ is given by a Poisson random measure with intensity $w(x)\mu(dx)$.

As noted earlier, we are interested in the situation where the skeleton itself satisfies a SLLN. There is a substantial body of literature available that analyses the long-term behaviour of branching particle diffusions. To delimit the regime we want to study, we make two regularity assumptions. A detailed discussion of all assumptions can be found in Section 2.1. The first condition ensures that the semigroup $(S_t)_{t\geq 0}$ of $L+\beta$ grows precisely exponentially on compactly supported, continuous functions.

Assumption 2 (Criticality assumption)**.**

(i) The second order differential operator $L+\beta$ has positive *generalised principal eigenvalue*

$$\lambda_c := \lambda_c(L+\beta) := \inf\left\{\lambda \in \mathbb{R} \colon \exists u \in C^{2,\eta}(D), u > 0, (L+\beta-\lambda)u = 0\right\} > 0. \quad (1.9)$$

(ii) The operator $L+\beta-\lambda_c$ is *critical*, that is, it does not possess a Green's function but there exists $\phi \in C^{2,\eta}(D)$, $\phi > 0$, such that $(L+\beta-\lambda_c)\phi = 0$.

Given (ii), ϕ is unique up to constant multiples and is called the *ground state*. With $L+\beta-\lambda_c$ also its formal adjoint is critical and the corresponding ground state is denoted by $\widetilde{\phi}$.

(iii) $L+\beta-\lambda_c$ is *product L^1-critical*, i.e. $\langle \phi, \widetilde{\phi}\rangle < \infty$. We normalize to obtain $\langle \phi, \widetilde{\phi}\rangle = 1$.

Corollary 2.7 below shows that under Assumptions 1 and 2, the process

$$W_t^{\phi/w}(Z) = e^{-\lambda_c t}\langle \phi/w, Z_t\rangle, \quad t \geq 0,$$

is a nonnegative $\mathbf{P}_\mu$-martingale for all $\mu \in \mathcal{M}_f^\phi(D) := \{\mu \in M_f(D) \colon \langle \phi, \mu\rangle < \infty\}$, and $(W_t^{\phi/w}(Z))_{t\geq 0}$ has an almost sure limit. To have the notation everywhere, we define $W_\infty^{\phi/w}(Z) := \liminf_{t\to\infty} W_t^{\phi/w}(Z)$.

Our second regularity assumption consists essentially of moment conditions.

Assumption 3 (Moment assumption)**.** There exists $p \in (1,2]$ such that

$$\sup_{x\in D} \phi(x)\alpha(x) < \infty, \quad (1.10)$$

$$\sup_{x\in D} \phi(x) \int_{(0,1]} y^2\, \Pi(x,dy) < \infty, \quad (1.11)$$

$$\sup_{x\in D} \phi(x)^{p-1} \int_{(1,\infty)} y^p\, \Pi(x,dy) < \infty, \quad (1.12)$$

$$\langle \phi^{p-1}, \phi\widetilde{\phi}\rangle < \infty, \quad (1.13)$$

$$\left\langle \int_{(1,\infty)} y^2 e^{-w(\cdot)y}\, \Pi(\cdot,dy), \phi\widetilde{\phi}\right\rangle < \infty. \quad (1.14)$$

The parameter p remains fixed throughout Part I of this thesis. Assumption 3 is satisfied, for example, when ϕ is bounded and $\sup_{x\in D}\int_{(1,\infty)} y^2\,\Pi(x,dy) < \infty$. These second moment conditions appeared in the literature (cf. Section 2.1.3), and we will see several examples in Chapter 4. However, our results are valid under the weaker conditions of Assumption 3. In Sections 2.1.3 and 4.2, we explain that in the case of a quadratic branching mechanism only (1.10) is needed.

The SLLN has been proved for a large class of branching particle diffusions. Where it has not been established, yet, we assume a SLLN for the skeleton Z. It will be sufficient to assume convergence along lattice times.

Assumption 4 (Strong law assumption)**.** For all $\mu \in \mathcal{M}_c(D)$, $\delta > 0$ and continuous $f \in p(D)$ with fw/ϕ bounded,

$$\lim_{n\to\infty} e^{-\lambda_c n\delta}\langle f, Z_{n\delta}\rangle = \langle f, w\widetilde{\phi}\rangle W_\infty^{\phi/w}(Z) \qquad \mathbf{P}_\mu\text{-almost surely.}$$

At first, Assumption 4 may look like a strong assumption. However, given Assumptions 1–3, the SLLN for the skeleton has been proved under two additional conditions. The first condition controls the spread of the support of the skeleton when started from a single particle; the second condition is a uniformity assumption on the convergence of an associated ergodic motion (the "spine") to its stationary distribution. See Theorem 2.13 for details. These conditions hold for a wide class of processes, and we demonstrate this for several key examples in Chapter 4.

1.3 Statement of the main results

Before stating the SLLN for superdiffusion X, we relate the limiting random variable of (1.4) to the limit that appears in Assumption 4. In Corollary 2.7 below, we show that under Assumption 2 the process

$$W_t^\phi(X) = e^{-\lambda_c t}\langle \phi, X_t\rangle, \quad t \geq 0,$$

is a nonnegative P_μ-martingale for all $\mu \in \mathcal{M}_f^\phi(D) = \{\mu \in \mathcal{M}_f(D) \colon \langle \phi, \mu\rangle < \infty\}$, and $(W_t^\phi(X))_{t\geq 0}$ has an almost sure limit. To have the notation everywhere, we define $W_\infty^\phi(X) := \liminf_{t\to\infty} W_t^\phi(X)$.

Proposition 1.1. *Suppose Assumptions 1, 2, (1.10)–(1.12) hold. For all $\mu \in \mathcal{M}_f^\phi(D)$, the martingales $(W_t^\phi(X))_{t\geq 0}$ and $(W_t^{\phi/w}(Z))_{t\geq 0}$ are bounded in $L^p(\mathbf{P}_\mu)$, and*

$$W_\infty^\phi(X) = W_\infty^{\phi/w}(Z) \qquad \mathbf{P}_\mu\text{-almost surely.} \tag{1.15}$$

Recall that ℓ denotes the Lebesgue measure on the domain D. Our main theorem is the following.

Theorem 1.2. *Suppose Assumptions 1–4 hold. For every $\mu \in \mathcal{M}_f^\phi(D)$, there exists a measurable set Ω_0 such that $P_\mu(\Omega_0) = 1$ and, on Ω_0, for all ℓ-almost everywhere continuous functions $f \in p(D)$ with f/ϕ bounded,*

$$\lim_{t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle = \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X). \tag{1.16}$$

The convergence in (1.16) also holds in $L^1(P_\mu)$. In particular, $P_\mu[W_\infty^\phi(X)] = \langle \phi, \mu\rangle$.

Even though our main interest is almost sure convergence, Theorem 1.2 gives also new results for convergence in probability; see the examples in Chapter 4. We record the following corollary of Theorem 1.2 to present the result in possibly more familiar terms.

Corollary 1.3. *Suppose Assumptions 1–4 hold. In the vague topology, $e^{-\lambda_c t}X_t \to W^{\phi}_{\infty}(X)\widetilde{\phi}\ell$ P_μ-almost surely as $t \to \infty$. If, in addition, ϕ is bounded away from zero, then the convergence holds in the weak topology P_μ-almost surely.*

Finally, we present the SLLN as announced in (1.4). This makes the comparison between $\langle f, X_t\rangle$ and its mean explicit.

Corollary 1.4. *Suppose that Assumptions 1–4 hold. For all $\mu \in \mathcal{M}^{\phi}_f(D)$, $\mu \not\equiv 0$, $f \in C^+_c(D)$, $f \neq \mathbf{0}$,*

$$\lim_{t\to\infty} \frac{\langle f, X_t\rangle}{P_\mu[\langle f, X_t\rangle]} = \frac{1}{\langle \phi, \mu\rangle} W^{\phi}_{\infty}(X) \qquad P_\mu\text{-almost surely and in } L^1(P_\mu).$$

The weak law of large numbers (WLLN), and even the L^1-convergence in (1.16), can be obtained without assuming the SLLN for the skeleton as the next theorem reveals.

Theorem 1.5. *Suppose Assumptions 1, 2, (1.10)–(1.13) hold. For all $\mu \in \mathcal{M}^{\phi}_f(D)$ and $f \in p(D)$ with f/ϕ bounded, the convergence in (1.16) holds in $L^1(P_\mu)$.*

1.4 Literature review

Terminology in the literature is not always consistent, so let us clarify that we refer to branching particle processes and superprocesses as branching diffusions and superdiffusions, respectively, if the underlying motion is a diffusion. Similar wording is used for other classes of underlying motions.

The limit theory of supercritical branching processes has been studied since the 1960s when sharp statements were established for classical finite-type processes [93, 6]. The first result for branching diffusions was due to Watanabe [121] in 1967, who proved an almost sure convergence result for branching Brownian motion and certain one-dimensional motions. The key ingredient to the proof was Fourier analysis, a technique recently used by Wang [120] and Kouritzin and Ren [95] to establish the SLLN for super-Brownian motion. Super-Brownian motion on $\mathbb{R}^d$ with a spatially independent branching mechanism does not fall into the framework of this thesis since $L + \beta - \lambda_c$ is not product L^1-critical in that case. Rather, $\phi = \widetilde{\phi} = \mathbf{1}$, where $\mathbf{1}$ denotes the constant function with value 1, and $e^{-\lambda_c t}P_\mu[\langle f, X_t\rangle]$ converges to zero for all $f \in C^+_c(D)$. The missing scaling factor is $t^{d/2}$ and $P_\mu[\langle f, X_t\rangle] \sim (2\pi t)^{-d/2}e^{\lambda_c t}\langle f, \mathbf{1}\rangle \mu(\mathbb{R}^d)$ for $\mu \in \mathcal{M}_c(\mathbb{R}^d)$. Wang's [120] SLLN for super-Brownian motion takes the form

$$\lim_{t\to\infty} \frac{\langle f, X_t\rangle}{P_\mu[\langle f, X_t\rangle]} = \frac{W^{\phi}_{\infty}(X)}{\mu(\mathbb{R}^d)} \qquad P_\mu\text{-almost surely}$$

for all nontrivial nonnegative continuous functions f with compact support, for all $\mu = \delta_x$, $x \in \mathbb{R}^d$, and with martingale limit $W^\phi_\infty(X) = \lim_{t\to\infty} e^{-\lambda_c t}\langle \mathbf{1}, X_t\rangle$. Watanabe's argument is thought to be incomplete because the regularity for his argument is not proven; see [120]. Biggins [16] developed a method to show uniform convergence of martingales for branching random walks. Wang combined these arguments with the compact support property of super-Brownian motion started from $\mu \in \mathcal{M}_c(D)$. Kouritzin and Ren [95] proved the SLLN for super-stable processes of index $\alpha \in (0, 2]$ with spatially independent quadratic branching mechanism. The correct scaling factor in this case is $t^{d/\alpha}e^{-\lambda_c t}$. The authors allow any finite starting measure with finite mean and a class of continuous test functions that decrease sufficiently fast at infinity. Fourier-analytic methods were also used by Grummt and Kolb [79] to prove the SLLN for the two-dimensional super-Brownian motion with a single point source (see [73] for the definition and a proof of existence of this process). Earlier, Engländer [55] established convergence in probability for a class of superdiffusions that do not necessarily satisfy Assumption 2 using a time-dependent h-transform developed in [62].

In the product L^1-critical case, the dominant method to prove almost sure limit theorems is due to Asmussen and Hering [5]. (Kaplan and Asmussen use a similar method in [90].) The main idea is as follows. For $s, t \geq 0$, write $\mathcal{F}_t = \sigma(X_r : r \leq t)$ and

$$\begin{aligned}
&e^{-\lambda_c(s+t)}\langle f, X_{s+t}\rangle \\
&= e^{-\lambda_c t}P_\mu[e^{-\lambda_c s}\langle f, X_{s+t}\rangle|\mathcal{F}_t] + \Big(e^{-\lambda_c(s+t)}\langle f, X_{s+t}\rangle - e^{-\lambda_c t}P_\mu[e^{-\lambda_c s}\langle f, X_{s+t}\rangle|\mathcal{F}_t]\Big) \\
&= \mathrm{CE}_f(s,t) + \mathrm{D}_f(s,t).
\end{aligned}$$

Here CE stands for "conditional expectation" and D for "difference". The first step is to show that $\mathrm{D}_f(s,t) \to 0$ as $t \to \infty$. This is usually done via a Borel–Cantelli argument, and therefore, requires a restriction to lattice times $t = n\delta$. The second step is to show that $\mathrm{CE}_f(s,t)$ behaves like the desired limit for s and t large. This is the hardest part of the proof and usually causes most of the assumptions. The third and last step is to extend the result from lattice to continuous time.

Asmussen and Hering control $\mathrm{CE}_f(s,t)$ for branching particle processes by a uniform Perron–Frobenius condition on the semigroup $(S_t)_{t\geq 0}$. Passage to continuous time is obtained under additional continuity assumptions on process and test functions. Recently, their method was generalised by Engländer et al. [57] to establish the SLLN for a class of branching diffusions on arbitrary domains. The authors control $\mathrm{CE}_f(s,t)$ by an assumption that restricts the speed at which particles spread in space and a condition on the rate at which a certain ergodic motion (the "spine") converges to its stationary distribution.

While Asmussen and Hering's idea for the proof of SLLNs along lattice times is rather robust and (under certain assumptions) feasible also for superprocesses, the argument used for the transition from lattice to continuous time relies heavily on the finite number of particles in the branching diffusion.

A new approach to almost sure limit theorems for branching processes was introduced by Chen and Shiozawa [28] in the setup of branching symmetric Hunt processes. Amongst other assumptions, a spectral gap condition was used to obtain a Poincaré inequality which constitutes the main ingredient in the proof along lattice times. For the transition to continuous times the argument from Asmussen and Hering was adapted. Chen et al. [27] proved the first SLLN for superprocesses and relied on the same Poincaré inequality and functional analytic methods for the result along lattice times. For the transition to continuous time, Perkins' Itô formula for superprocesses [108] was used. Even though their SLLN holds on the full domain $\mathbb{R}^d$, the assumptions on motion and branching mechanism are restrictive in the following way: the motion has to be symmetric (and in the diffusive case must have a uniformly elliptic generator) and the coefficients of the branching mechanism have to satisfy a strict Kato class condition.

The idea to use stochastic analysis was brought much further by Liu et al. [102]. The authors gave a proof which is based entirely on the martingale problem for superprocesses, and decomposed the process into three martingale measures. Moreover, they introduced a new technique for the transition from lattice to continuous times based on the resolvent operator and estimates for the hitting probabilities of diffusions. The proof by Liu et al. follows again the three steps of Asmussen and Hering. To control the conditional expectation $\mathrm{CE}_f(s,t)$, they assume that the transition density of the underlying motion is intrinsically ultracontractive and that the domain D is of finite Lebesgue measure. This assumption excludes most of the classical examples; see Chapter 4.

To complete our review, we mention that the first law of large numbers for superdiffusions was proven by Engländer and Turaev [61] on the domain $D = \mathbb{R}^d$. The authors use analytic tools from the theory of dynamical systems, in particular properties of invariant curves, to show the convergence in distribution. Besides classical superdiffusions, the 1-dimensional super-Brownian motion with a single point source is studied.

1.5 Outline of the proof of Theorem 1.2

The key to our argument is the skeleton decomposition for the supercritical superprocess X. Intuitively, this representation result states that the superprocess is a cloud of subcritical superdiffusive mass immigrating off a supercritical branching diffusion, the skeleton, which governs the large-time behaviour of X. It is important to note that we use the skeleton to make a connection between the asymptotic behaviour of a branching diffusion and of the superdiffusion, and we do not use any classical approximation of the superprocess by branching particle systems in a high-density limit regime.

Broadly speaking, our proof of Theorem 1.2 follows the three steps of Asmussen and Hering outlined in Section 1.4. However, instead of the full process X we consider only the immigration occurring after time t in the decomposition into conditional expectation CE_f and difference D_f. This immigration is a subprocess of X and we show that the stated convergence for the full process follows when the subprocess converges to the claimed limit.

Using the tree structure of the skeleton, we can split the immigration that occurs after time t according to the different branches of the skeleton at time t. This fact allows us to appeal to discrete techniques for the analysis of the immigration process. To analyse the conditional expectation CE_f for the immigration after time t, we use the SLLN for the skeleton. After exponential rescaling, the immigration along different branches up to a fixed time s is of constant order and the SLLN for the skeleton describes the asymptotic behaviour for large t. Taking the observed time frame s to infinity then adjusts only the constants. To replace the limiting random variable $W_\infty^{\phi/w}(Z)$, coming from the SLLN for the skeleton, by $W_\infty^{\phi}(X)$, we can, as it turns out, reverse the order in which these limits are taken. Taking first the observed time horizon s to infinity for test function ϕ, we recover the martingale for the skeleton as a consequence of the same invariance property of ϕ that makes $(W_t^\phi(X))_{t\geq 0}$ a martingale.

The analysis of D_f for the immigration after time t is fairly standard, and for the transition from lattice to continuous times we adapt the argument by Liu et al. [102] relying again on the skeleton decomposition. The moment estimates needed for our analysis are obtained using a spine decomposition for the superprocess.

1.6 Overview

The outline of Part I of this thesis is as follows. We start in Section 2.1.1 with an analysis of the skeleton assumption (Assumption 1) and give a detailed description of the skeleton decomposition. In the remainder of Section 2.1, we discuss further basic properties of superprocesses and our other three main assumptions, and we compare them to conditions that appeared in the literature. Section 2.2 contains a spine decomposition for the superprocess X and the proof that the martingale $(W_t^\phi(X))_{t\geq 0}$ is bounded in L^p.

The proofs of the main results are collected in Chapter 3. First, in Section 3.1, we reduce the SLLN to a statement that focuses on the main technical difficulty. In Section 3.2, we show that the martingale limits for superprocess and skeleton agree, and in Section 3.3, we prove the WLLN stated in Theorem 1.5. The asymptotic behaviour of the immigration process is studied in Section 3.4, and the SLLN along lattice times is established. The transition from lattice to continuous times is performed in Section 3.5, and we conclude our main results.

In Chapter 4 we provide several examples to illustrate our results. Spatially independent branching mechanisms are discussed in Section 4.1; quadratic branching mechanisms are considered in Section 4.2. In Section 4.3 we study the super-Wright–Fisher diffusion and prove a conjecture by Fleischmann and Swart [74].

Some minor statements needed along the way are proved in the appendix: Appendix A contains Feynman–Kac-type arguments, and Appendix B discusses a generalised version of the mild equation (1.3) and monotonicity of its solution in domain and test function.

CHAPTER 2

PRELIMINARIES

This chapter is split into two parts. In the first part, we discuss our four main assumptions; in the second, we prove that the martingale $(W_t^\phi(X))_{t\geq 0}$ converges in L^p.

2.1 Basic properties

2.1.1 Skeleton decomposition

In this section, we work under Assumption 1. The skeleton decomposition for supercritical superprocesses offers a pathwise description of the superprocess in terms of a supercritical branching particle process dressed with an immigration process. Heuristically, one can think of the skeleton as the prolific individuals of the branching process, i.e. individuals belonging to infinite lines of descent. The martingale function w assigns a small value to regions that prolific individuals should avoid. If $w(x) = -\log P_{\delta_x}(\mathcal{E})$ for some event $\mathcal{E}$, then the skeleton particles avoid the behaviour specified by $\mathcal{E}$. Classical examples are the event of extinction in finite time $\mathcal{E}_{\text{fin}} = \{\exists t \geq 0 \colon \langle \mathbf{1}, X_t\rangle = 0\}$, cf. [70, 59], and the event of weak extinction $\mathcal{E}_{\text{lim}} = \{\lim_{t\to\infty}\langle \mathbf{1}, X_t\rangle = 0\}$; cf. [13, 11]. In Chapter 4, we discuss classes of superprocesses where these two events give a suitable martingale function. To allow for other events that may be more suitable for a certain process of interest, we assume only Assumption 1.

We proceed by deriving properties of w from Assumption 1. For $f \in p(D)$, let $\widetilde{f}(x,t) = f(x)$ for all $(x,t) \in D \times [0,\infty)$. Dynkin [43] derives the superprocess X from *exit measures* that describe the evolution of mass not only in time but also in space. He showed that, for any domain $B \subseteq D$ and $t \geq 0$, there exists a random, finite measure $\widetilde{X}_t^B$ on $D \times [0,\infty)$ such that for all $\mu \in \mathcal{M}_f(D)$ and $f \in bp(D)$,

$$P_\mu\big[e^{-\langle \widetilde{f}, \widetilde{X}_t^B\rangle}\big] = e^{-\langle \widetilde{u}_f^B(\cdot,t), \mu\rangle}, \tag{2.1}$$

where $\widetilde{u}_f^B$ is the unique, nonnegative solution to the integral equation

$$u(x,t) = \mathbb{P}_x\big[f(\xi_{t\wedge\tau_B})\big] - \mathbb{P}_x\Big[\int_0^{t\wedge\tau_B} \psi_\beta(\xi_s, u(\xi_s, t-s))\,ds\Big] \tag{2.2}$$

for all $(x,t) \in D \times [0,\infty)$ and $\tau_B = \inf\{t \geq 0\colon \xi_t \notin B\}$. For $f \in p(D)$, there exists a sequence of functions $f_k \in bp(D)$ such that $f_k \uparrow f$ pointwise. By (2.1), $\widetilde{u}_{f_k}^B(x,t)$ is monotonically increasing in k, and we denote the limit by $\widetilde{u}_f^B(x,t) \in [0,\infty]$. With this notation, the monotone convergence theorem implies that (2.1) is valid for all $f \in p(D)$. The same argument shows that (1.2) holds for all $f \in p(D)$, and (1.6) implies $u_w = w$. Hence, (1.6) holds for all $\mu \in \mathcal{M}_f(D)$. The superprocess X_t is obtained as a projection of $\widetilde{X}_t^D$ restricted to $D \times \{t\}$. Writing $\widetilde{w}(x,t) = w(x)$ for $(x,t) \in D \times [0,\infty)$, the Markov property (cf. Theorem I.1.3 [43]) and (1.6) yield, for all $\mu \in \mathcal{M}_f(D)$ and all domains $B \subseteq D$,

$$P_\mu\big[e^{-\langle \widetilde{w}, \widetilde{X}_t^B\rangle}\big] = P_\mu\big[e^{-\langle w, X_t\rangle}\big] = e^{-\langle w,\mu\rangle}. \tag{2.3}$$

Comparing (2.3) to (2.1), we deduce that $\widetilde{u}_w^B = \widetilde{w}$. Now let $B \subset\subset D$. If the support of μ, $\mathrm{supp}(\mu)$, is a subset of B, then $\widetilde{X}_t^B$ is supported on the boundary of $B \times [0,t)$; if $\mathrm{supp}(\mu) \subseteq D \setminus B$, then $\widetilde{X}_t^B = \mu$ almost surely (cf. Theorem I.1.2 in [43]). In particular, (1.5) implies that, for $\mu = \delta_x$, $x \in D$, $\widetilde{w}$ in $\langle \widetilde{w}, \widetilde{X}_t^B\rangle$ can be interpreted as a bounded function. We combine (2.3) and (2.2) to obtain for all $(x,t) \in \overline{B} \times [0,\infty)$,

$$w(x) = \mathbb{P}_x\big[w(\xi_{t\wedge\tau_B})\big] - \mathbb{P}_x\Big[\int_0^{t\wedge\tau_B} \psi_\beta(\xi_s, w(\xi_s))\,ds\Big]. \tag{2.4}$$

Since w is bounded on $\overline{B}$, the continuity of the diffusion ξ yields that w is continuous on B (see the argument in the last paragraph of page 708 in [59]). Because B was arbitrary, we conclude:

Lemma 2.1. *The martingale function w is continuous on D.*

Lemma A.1 (i) in the appendix shows that (2.4) can be transformed into

$$w(x) = \mathbb{P}_x\Big[w(\xi_{t\wedge\tau_B})\exp\Big(-\int_0^{t\wedge\tau_B} \frac{\psi_\beta(\xi_s, w(\xi_s))}{w(\xi_s)}\,ds\Big)\Big] \quad \text{for all } (x,t) \in \overline{B} \times [0,\infty).$$

Hence, for any domain $B \subset\subset D$, $x \in B$,

$$w(\xi_{t\wedge\tau_B})\exp\Big(-\int_0^{t\wedge\tau_B} \frac{\psi_\beta(\xi_s, w(\xi_s))}{w(\xi_s)}\,ds\Big), \quad t \geq 0, \quad \text{is a } \mathbb{P}_x\text{-martingale.} \tag{2.5}$$

Since every nonnegative local martingale is a supermartingale, we conclude that for all $x \in D$,

$$\frac{w(\xi_t)}{w(x)}\exp\Big(-\int_0^t \frac{\psi_\beta(\xi_s, w(\xi_s))}{w(\xi_s)}\,ds\Big), \quad t \geq 0, \quad \text{is a } \mathbb{P}_x\text{-supermartingale.}$$

In particular, the definition of $\mathbb{P}_x^w$ in (1.7) is valid:

Lemma 2.2. *For every $x \in D$, $\mathbb{P}_x^w$ is a well-defined (sub-)probability measure and $(\xi = (\xi_t)_{t\geq 0}; (\mathbb{P}_x^w)_{x\in D})$ is a (possibly non-conservative) Markov process, which we consider as a Markov process in $D \cup \{\dagger\}$.*

If w is bounded, the argument leading to (2.5) is valid for $B = D$, and $(\xi; \mathbb{P}^w)$ is conservative.

To give a description of the skeleton decomposition, we construct an auxiliary $\mathcal{M}_f(D)$-valued Markov process using the martingale function w. Let for all $x \in D$, $z \geq 0$ and $f \in p(D)$, $\Pi^*(x, dy) := e^{-w(x)y}\,\Pi(x, dy)$,

$$\beta^*(x) := \beta(x) - 2\alpha(x)w(x) - \int_{(0,\infty)} (1 - e^{-w(x)y})y\,\Pi(x, dy),$$

$$\psi_0^*(x, z) := \alpha(x)z^2 + \int_{(0,\infty)} (e^{-zy} - 1 + zy)\,\Pi^*(x, dy).$$

Since $\beta^*(x) \leq \beta(x)$ for all $x \in D$, β^* is bounded from above. However, it is not clear whether β^* is bounded from below. Hence, the branching mechanism $\psi_{\beta^*}^*(x, z) = -\beta^*(x)z + \psi_0^*(x, z)$ might not satisfy the assumptions from Section 1.1. To overcome this problem, set $\beta_+^* = \max\{\beta^*, 0\}$ and $\beta_-^* = \max\{-\beta^*, 0\}$ so that $\beta^* = \beta_+^* - \beta_-^*$ with β_+^* bounded and β_-^* nonnegative. We write for all $f \in p(D)$,

$$S_t^* f(x) := \mathbb{P}_x\big[e^{-\int_0^t \beta_-^*(\xi_s)\,ds} e^{\int_0^t \beta_+^*(\xi_s)\,ds} f(\xi_t)\big] = \mathbb{P}_x\big[e^{\int_0^t \beta^*(\xi_s)\,ds} f(\xi_t)\big], \tag{2.6}$$

where $(\xi, (\mathbb{P}_x)_{x\in D})$ is the original diffusion process on D with generator L defined in Section 1.1. Dynkin [43, Theorem I.1.1] proved the existence and uniqueness of the superprocess $X^* = (X_t^*)_{t\geq 0}$ whose motion is given by the diffusion with generator L killed at spatially dependent rate β_-^*, branching mechanism $\psi_{\beta_+^*}^*(x, z) = -\beta_+^*(x)z + \psi_0^*(x, z)$ and domain D. In particular, X^* is an $\mathcal{M}_f(D)$-valued, time-homogeneous Markov process such that, for all $\mu \in \mathcal{M}_f(D)$, $f \in bp(D)$ and $t \geq 0$,

$$P_\mu\big[e^{-\langle f, X_t^*\rangle}\big] = e^{-\langle u_f^*(\cdot,t), \mu\rangle},$$

where u_f^* is the unique nonnegative solution to

$$u(x, t) = S_t^* f(x) - \int_0^t S_s^*[\psi_0^*(\cdot, u(\cdot, t - s))](x)\,ds \quad \text{for all } (x, t) \in D \times [0, \infty). \tag{2.7}$$

Comparing (2.7) and (1.3), we refer to X^* as the $(L, \psi_{\beta^*}^*; D)$-superprocess. In Appendix B, we show that, alternatively, X^* can be obtained as a monotone, distributional limit of superprocesses whose motion is given by the diffusion with generator L and no additional killing. If $w(x) = -\log P_{\delta_x}(\mathcal{E})$ for a tail event $\mathcal{E}$ with $P_\mu(\mathcal{E}) = e^{-\langle w,\mu\rangle}$ for all $\mu \in \mathcal{M}_f(D)$, then X^* can be obtained from X by conditioning on $\mathcal{E}$, i.e., the distribution of X_t^* is given by $P_\mu(X_t \in \cdot\,|\,\mathcal{E})$; cf. [70, 59, 11, 98]. For our analysis it will be enough to know that on compactly supported, continuous functions the semigroup $(S_t^*)_{t\geq 0}$ grows more slowly than the semigroup $(S_t)_{t\geq 0}$, and we prove this fact in Lemma 3.5.

The following theorem is a concise version of the skeleton decomposition at the level of detail that is useful to us. It is based on a result from Kyprianou et al. [98]. We denote by $\mathcal{M}_a^{\text{loc}}$ the set of locally finite integer-valued measures on $\mathcal{B}(D)$.

Theorem 2.3 (Kyprianou et al. [98]). *There exists a probability space with probability measures $\mathbf{P}_{\mu,\nu}$, $\mu \in \mathcal{M}_f(D)$, $\nu \in \mathcal{M}_a^{\text{loc}}(D)$, that carries the following processes:*

(i) *$(Z = (Z_t)_{t\geq 0}; \mathbf{P}_{\mu,\nu})$ is a branching diffusion with motion $(\xi; \mathbb{P}^w)$ defined in (1.7), and branching rate q, and offspring distribution $(p_k)_{k\geq 2}$ defined by (1.8), and $\mathbf{P}_{\mu,\nu}(Z_0 = \nu) = 1$.*

(ii) *$(X^* = (X_t^*)_{t\geq 0}; \mathbf{P}_{\mu,\nu})$ is an $\mathcal{M}_f(D)$-valued time-homogeneous Markov process such that, for every $\mu \in \mathcal{M}_f(D)$, $f \in bp(D)$ and $t \geq 0$,*

$$\mathbf{P}_{\mu,\nu}\big[e^{-\langle f, X_t^*\rangle}\big] = e^{-\langle u_f^*(\cdot,t),\mu\rangle},$$

where u_f^* is the unique solution to (2.7). Moreover, X^* is independent of Z under $\mathbf{P}_{\mu,\nu}$.

(iii) *$(I = (I_t)_{t\geq 0}; \mathbf{P}_{\mu,\nu})$ is an $\mathcal{M}_f(D)$-valued process such that:*

(a) *$\mathbf{P}_{\mu,\sum_i \delta_{x_i}}[e^{-\langle f, I_t\rangle}] = \prod_i \mathbf{P}_{\mu,\delta_{x_i}}\big[e^{-\langle f, I_t\rangle}\big]$ for all $\mu \in \mathcal{M}_f(D), x_i \in D, f \in p(D)$. Moreover, $\mathbf{P}_{\mu,\nu}(I \in \cdot)$ does not depend on μ, $\mathbf{P}_{\mu,\nu}(I_0 = 0) = 1$, and, under $\mathbf{P}_{\mu,\nu}$, (Z, I) is independent of X^*.*

(b) *$((X, Z) := (X^* + I, Z); \mathbf{P}_{\mu,\nu})$ is a Markov process.*

(c) *$(X = X^* + I; \mathbf{P}_\mu)$ is equal in distribution to $(X; P_\mu)$, where $\mathbf{P}_\mu$ denotes the measure $\mathbf{P}_{\mu,\nu}$ with ν replaced by a Poisson random measure with intensity $w(x)\mu(dx)$.*

(d) *Under $\mathbf{P}_\mu$, conditionally given X_t, the measure Z_t is a Poisson random measure with intensity $w(x)X_t(dx)$.*

We call the probability space from Theorem 2.3 the *skeleton space*. The process I is called *immigration process* or simply *immigration*. As the processes $(X; \mathbf{P}_\mu)$ on the skeleton space and $(X; P_\mu)$ on the generic space have the same distribution, we may, without loss of generality, work on the skeleton space whenever it is convenient. Since the distributions of X^* and I under $\mathbf{P}_{\mu,\nu}$ do not depend on ν and μ, respectively, we sometimes write $\mathbf{P}_{\mu,\bullet}$ or $\mathbf{P}_{\bullet,\nu}$.

Kyprianou et al. [98] identify the immigration process explicitly. We need only the properties listed in Theorem 2.3, but for definiteness, we now give a full characterisation of the immigration process.

Dynkin and Kuznetsov [45] showed that on the canonical space of measure-valued càdlàg functions, $\mathbb{D}([0,\infty),\mathcal{M}_f(D))$, for every $x \in D$, there is a unique measure $\mathbb{N}_x$ such that, for all $f \in bp(D)$, $t \geq 0$,

$$-\log P_{\delta_x}[e^{-\langle f,X_t\rangle}] = \mathbb{N}_x[1-e^{-\langle f,X_t\rangle}]. \tag{2.8}$$

The corresponding measures associated with the superprocess X^* are denoted by $\mathbb{N}^*_x$, $x \in D$.

To describe the immigration processes, we use the classical Ulam–Harris notation to uniquely refer to individuals in the genealogical tree $\mathcal{T}$ of Z (see for example page 290 in [80]). For each individual $u \in \mathcal{T}$, we write b_u and d_u for its birth and death times, respectively, and $\{z_u(r)\colon r \in [b_u, d_u]\}$ for its spatial trajectory. The skeleton space carries the following processes:

(iii.1) $(\mathfrak{a};\mathbf{P}_{\mu,\nu})$ is a random measure, such that conditional on Z, $\mathfrak{a}$ is a Poisson random measure that issues, for every $u \in \mathcal{T}$, $\mathcal{M}_f(D)$-valued processes $X^{\mathfrak{a},u,r} = (X^{\mathfrak{a},u,r}_t)_{t\geq 0}$ along the space-time trajectory $\{(z_u(r), r)\colon r \in (b_u, d_u]\}$ with rate

$$dr \times \left(2\alpha(z_u(r))d\mathbb{N}^*_{z_u(r)} + \int_{(0,\infty)} \Pi(z_u(r),dy)\, ye^{-w(z_u(r))y} \times dP^*_{y\delta_{z_u(r)}}\right),$$

where P^*_μ denotes the distribution of X^* started in μ. Since at most countably many processes $X^{\mathfrak{a},u,r}$ are not equal to the constant zero measure, immigration at time t that occurred in the form of processes $X^{\mathfrak{a},u,r}$ until time t can be written as

$$I^{\mathfrak{a}}_t = \sum_{u\in\mathcal{T}} \sum_{b_u<r\leq d_u\wedge t} X^{\mathfrak{a},u,r}_{t-r}.$$

The processes $(X^{\mathfrak{a},u,r}\colon u \in \mathcal{T}, b_u < r \leq d_u)$ are independent given Z and independent of X^*.

(iii.2) $(\mathfrak{b};\mathbf{P}_{\mu,\nu})$ is a random measure, such that conditional on Z, $\mathfrak{b}$ issues, for every $u \in \mathcal{T}$, at space-time point $(z_u(d_u), d_u)$ process $X^{\mathfrak{b},u}$ with law $P^*_{Y_u\delta_{z_u(d_u)}}$. Given that u is replaced by k particles at its death time d_u, the independent random variable Y_u is distributed according to the measure

$$\frac{1}{q(x)w(x)p_k(x)}\left(\alpha(x)w(x)^2\delta_0(dy)\mathbb{1}_{\{k=2\}} + w(x)^k\frac{y^k}{k!}e^{-w(x)y}\,\Pi(x,dy)\right)\Big|_{x=z_u(d_u)}.$$

The immigration at time t that occurred in the form of processes $X^{\mathfrak{b},u}$ until time t is denoted by

$$I^{\mathfrak{b}}_t = \sum_{u\in\mathcal{T}} \mathbb{1}_{\{d_u\leq t\}} X^{\mathfrak{b},u}_{t-d_u}.$$

The processes $(X^{\mathfrak{b},u}\colon u \in \mathcal{T})$ are independent of X^* and, given Z, are mutually independent and independent of $\mathfrak{a}$.

The full immigration process is given by $I = I^{\mathfrak{a}} + I^{\mathfrak{b}}$.

Proof of Theorem 2.3. Theorem 2.3 generalises Corollary 6.2 in [98] in three ways. First, the authors choose $w(x) = -\log P_{\delta_x}(\mathcal{E}_{\text{fin}})$ but after defining Z and X^* this choice is not used anymore and their argument goes through without any changes for a general martingale function w satisfying Assumption 1. Second, the authors assume that w is locally bounded away from zero. Since w is continuous by Lemma 2.1, this condition is automatically satisfied. Finally, Kyprianou et al. enforce additional regularity conditions on the underlying motion to use a comparison principle from the literature in the proof of their Lemma 6.1 (see also their Footnote 1). The comparison principle allows them to conclude that the solution $\widetilde{u}_f^B$ to (2.2) is increasing in the domain B when the support of f is a subset of B. Lemmas A.1 (i) and B.5 below show that this monotonicity holds in the more general setup of this thesis, too. $\square$

We introduce notation to refer to the different parts of the skeleton decomposition.

Notation 2.4 (Notation for Z)**.** For $t \geq 0$, we write $Z_t = \sum_{i=1}^{N_t} \delta_{\xi_i(t)}$, where N_t denotes the number of skeleton particles at time t and $(\xi_i(t) \colon i = 1, \dots, N_t)$ their (conveniently ordered) locations. Given Z_0, $(Z^{i,0} \colon i = 1, \dots, N_0)$ denote the independent subtrees of the skeleton obtained by splitting Z according to the ancestors at time 0. The Markov property implies that $Z^{i,0}$ follows the same distribution as $(Z; \mathbf{P}_{\bullet, \delta_{\xi_i(0)}})$, $i = 1, \dots, N_0$. Under $\mathbf{P}_\mu$ with $\mu \in \mathcal{M}_c(D)$, $N_0 = \langle \mathbf{1}, Z_0 \rangle$ is a Poisson random variable with mean $\langle w, \mu \rangle$.

For $t \geq 0$, let $\boldsymbol{\mathcal{F}}_t$ denote the σ-algebra generated by the processes X^*, Z and I up to time t. Using the characterisation of the immigration process from Theorem 2.3, we obtain, for all $\mu \in \mathcal{M}_f(D)$, $\nu \in \mathcal{M}_a^{\text{loc}}(D)$, $f \in p(D)$ and $s, t \geq 0$,

$$\begin{aligned} \mathbf{P}_{\mu,\nu}\big[e^{-\langle f, X_{s+t}\rangle} \big| \boldsymbol{\mathcal{F}}_t\big] &\overset{\text{(b)}}{=} \mathbf{P}_{X_t, Z_t}\big[e^{-\langle f, X_s\rangle}\big] \overset{\text{(b)}}{=} \mathbf{P}_{X_t, Z_t}\big[e^{-\langle f, X_s^* + I_s\rangle}\big] \\ &\overset{\text{(a)}}{=} \mathbf{P}_{X_t, \bullet}\big[e^{-\langle f, X_s^*\rangle}\big] \prod_{i=1}^{N_t} \mathbf{P}_{\bullet, \delta_{\xi_i(t)}}\big[e^{-\langle f, I_s\rangle}\big] \end{aligned} \tag{2.9}$$

$\mathbf{P}_{\mu,\nu}$-almost surely. By (d), under $\mathbf{P}_\mu$ and given X_t, Z_t is a Poisson random measure with intensity $w(x) X_t(dx)$. Hence, (2.9) holds $\mathbf{P}_\mu$-almost surely when $\mathbf{P}_{\mu,\nu}$ on the left-hand side is replaced by $\mathbf{P}_\mu$. To make use of this identity, we split the immigration process according to the immigration that occurred before time t and the immigration that occurred along different branches of Z after time t.

Notation 2.5 (Notation for I)**.** For $t \geq 0$, denote by $I_s^{*,t}$ the immigration at time $s + t$ that occurred along the skeleton before time t; $I^{*,t} = (I_s^{*,t})_{s \geq 0}$. In addition, for $i \in \{1, \dots, N_t\}$, let $I_s^{i,t}$ denote the immigration at time $s + t$ that occurred along the subtree of the skeleton rooted at the i-th particle at time t with location $\xi_i(t)$;

$I^{i,t} = (I_s^{i,t})_{s\geq 0}$. We have

$$X_{s+t} = X_{s+t}^* + I_s^{*,t} + \sum_{i=1}^{N_t} I_s^{i,t} \qquad \text{for all } s,t \geq 0. \tag{2.10}$$

According to (2.9) and by the Markov property, given $\boldsymbol{\mathcal{F}}_t$, $(X_{s+t}^* + I_s^{*,t})_{s\geq 0}$ follows the same distribution as $(X^*, \mathbf{P}_{X_t})$, and $I^{i,t}$ follows the same distribution as $(I; \mathbf{P}_{\bullet,\delta_{\xi_i(t)}})$, $i = 1, \ldots, N_t$. Moreover, given $\boldsymbol{\mathcal{F}}_t$, the processes $(I^{i,t} \colon i = 1, \ldots, N_t)$ are independent and independent of $I^{*,t}$.

We end this section with a note on terminology. Several different phrases have been used in the literature to refer to the skeleton decomposition. Evans and O'Connell [70] proved the first skeleton decomposition for supercritical superprocesses in the case of a conservative motion (not necessarily a diffusion) and a quadratic, spatially independent branching mechanism with $\alpha, \beta \in (0, \infty)$, and call the result "representation theorem". Their study was motivated by the "immortal particle representation" derived by Evans [69] for critical superprocesses conditioned on non-extinction. This representation is in terms of a single "immortal particle" that throws off pieces of mass. Evans' article is part of a cluster of papers that study conditioned superprocesses. Salisbury and Verzani [116] condition the exit measure of a super-Brownian motion to hit n fixed, distinct points on the boundary of a bounded smooth domain. The authors show that the resulting process can be described as the sum of a tree with n leaves that throws off mass in a Poissonian way and of a copy of the unconditioned process, and call this decomposition "backbone representation". In a follow-up article [117] they consider different conditionings and derive an "immortal particle description" where the guiding object is a tree with possibly infinitely many branches that they call "backbone" or "branching backbone". Salisbury and Sezer [115] describe the super-Brownian motion conditioned on boundary statistics in terms of a "branching backbone" or "branching backbone system". Etheridge and Williams [67] represent a critical superprocess with infinite variance conditioned to survive until a fixed time as immigration along a Poisson number of "immortal trees". An overview of decompositions of conditioned superprocesses was offered by Etheridge [66] using the names "skeleton" and "immortal skeleton". Back in our setup of supercritical superprocesses, Engländer and Pinsky [59] speak about a "decomposition with immigration", and Fleischmann and Swart [75] construct a "trimmed tree". For the analysis of continuous-state branching processes, Duquesne and Winkel [40] find a "Galton–Watson forest". In the corresponding superprocess setup, Berestycki et al. [11] identify the "prolific backbone" and call the representation itself a "backbone decomposition". The latter phrase has been used several times since [99, 98, 104, 112].

We decided to use the term "skeleton decomposition" for the following reasons. Since the words "backbone" and "spine" are used interchangeably in spoken English, using these two words to mean different things might cause confusion. Furthermore, spine/backbone describes one key, supporting element of an object and does not branch. In contrast, a skeleton carries the entire structure and determines the main features of an object. This is the correct intuition for the spine decomposition and the skeleton decomposition of branching processes as well as the distinction between them.

2.1.2 Product L^1-criticality

The first two moments of the superprocess can be expressed in terms of the underlying motion and the branching mechanism. That is, (see, for example, Proposition 2.7 in [72]) for all $\mu \in \mathcal{M}_f(D)$ and $f \in bp(D)$,

$$P_\mu[\langle f, X_t\rangle] = \langle S_t f, \mu\rangle, \tag{2.11}$$

$$\mathrm{Var}_\mu\big(\langle f, X_t\rangle\big) = \int_0^t \langle S_s\Big[\big(2\alpha + \int_{(0,\infty)} y^2\,\Pi(\cdot, dy)\big)(S_{t-s}f)^2\Big], \mu\rangle\, ds. \tag{2.12}$$

Here $\mathrm{Var}_\mu(\langle f, X_t\rangle)$ denotes the variance of $\langle f, X_t\rangle$ under P_μ. By the monotone convergence theorem, the boundedness of f in (2.11) is unnecessary, and (2.12) holds for $f \in p(D)$ as soon as $\langle S_t f, \mu\rangle < \infty$. Similarly, under Assumption 1 and for $\mu \in \mathcal{M}_f(D)$, $f \in bp(D)$, the first two moments of $\langle f, X_t^*\rangle$ (see the discussion around (2.6) for the definitions) can be expressed as

$$\mathbf{P}_\mu[\langle f, X_t^*\rangle] = \langle S_t^* f, \mu\rangle, \tag{2.13}$$

$$\mathbf{Var}_\mu\big(\langle f, X_t^*\rangle\big) = \int_0^t \langle S_s^*\Big[\big(2\alpha + \int_{(0,\infty)} y^2\,\Pi^*(\cdot, dy)\big)(S_{t-s}^*f)^2\Big], \mu\rangle\, ds. \tag{2.14}$$

The main purpose of this section is to discuss Assumption 2, that enforces conditions on the operator $L+\beta$ and consequently on its semigroup $(S_t)_{t\geq 0}$ which is the expectation semigroup of X by (2.11). Throughout the section, we suppose that Assumptions 1 and 2 hold. Key features of the local behaviour of the superdiffusion X are determined by the generalised principal eigenvalue $\lambda_c = \lambda_c(L+\beta)$. If α and Π are sufficiently smooth and $\lambda_c \leq 0$, then the superdiffusion exhibits *weak local extinction*, i.e. the total mass assigned to a compact set by the superprocess tends to zero. For quadratic branching mechanisms this was shown by Pinsky [111, Theorem 6]; for general branching mechanisms the proof of Theorem 3 (i) in [58] gives the result. This is the reason to assume $\lambda_c > 0$.

The assumption of product L^1-criticality restricts this thesis to the situation where the expectation semigroup $(S_t)_{t\geq 0}$ scales precisely exponentially on compactly supported, continuous functions. In general, writing $S_t f(x) = e^{\lambda_c t}\omega_{f,x}(t)$, the limit $\omega_{f,x} := \lim_{t\to\infty}\omega_{f,x}(t)$ exists for all $f \in C_c^+(D)$, $x \in D$. Product L^1-criticality is equivalent to $\omega_{f,x} > 0$ for all $f \neq \mathbf{0}$. The alternative is $\omega_{f,x} = 0$ for all f and x (cf. Theorem 7 in [111] and Appendix A in [62]). Some of the relevant literature for this regime was discussed in Section 1.4. The notion of product L^1-criticality comes from the criticality theory of second order elliptic operators; see Appendix B of [59] for a good summary and Chapter 4 in [110] for a comprehensive treatment.

By Theorem 4.8.6 in [110], criticality implies that the ground state ϕ is an invariant function of $e^{-\lambda_c t}S_t$, that is $e^{-\lambda_c t}S_t\phi = \phi$, and we define a conservative diffusion $(\xi = (\xi_t)_{t\geq 0}; (\mathbb{P}^\phi_x)_{x\in D})$ by

$$\frac{d\mathbb{P}^\phi_x}{d\mathbb{P}_x}\bigg|_{\sigma(\xi_s:\, s\leq t)} = \frac{\phi(\xi_t)}{\phi(x)} e^{\int_0^t (\beta(\xi_s)-\lambda_c)\, ds} \quad \text{on } \{t < \tau_D\}, \tag{2.15}$$

$$\mathbb{P}^\phi_x[g(\xi_t)] = \phi(x)^{-1} e^{-\lambda_c t} S_t[\phi g](x), \tag{2.16}$$

for all $x \in D$, $t \geq 0$, $g \in p(D)$. Product L^1-criticality is equivalent to positive recurrence of the diffusion $(\xi = (\xi_t)_{t\geq 0}; (\mathbb{P}^\phi_x)_{x\in D})$ with stationary distribution $\phi(x)\widetilde{\phi}(x)\, dx$, c.f. [110, Theorems 4.9.5 and 4.9.6], and we call it the *ergodic motion* or the *spine* (as we explain in Section 2.2 below). In particular, see Theorems 4.3.3 and 4.8.6 in [110],

$$\langle \mathbb{P}^\phi_\cdot[g(\xi_t)], \phi\widetilde{\phi}\rangle = \langle g, \phi\widetilde{\phi}\rangle \qquad \text{for all } g \in p(D), \tag{2.17}$$

and, for every probability measure π on D and $g \in bp(D)$,

$$\langle \mathbb{P}^\phi_\cdot[g(\xi_t)], \pi\rangle \to \langle g, \phi\widetilde{\phi}\rangle \qquad \text{as } t \to \infty. \tag{2.18}$$

If, in addition, the initial distribution π is of compact support, then (2.18) holds for all $g \in p(D)$ with $\langle g, \phi\widetilde{\phi}\rangle < \infty$. Indeed, for g bounded, (2.18) follows from Theorem 4.9.9 in [110] and the dominated convergence theorem. If the support of π, $\mathrm{supp}(\pi)$, is compactly embedded in D, choose a domain $B \subset\subset D$ with $\mathrm{supp}(\pi) \subseteq B$. There exists a constant $C > 0$ such that

$$p^\phi(x, y, t) \leq C\phi(y)\widetilde{\phi}(y) \qquad \text{for all } x \in B, y \in D, t > 1, \tag{2.19}$$

where $p^\phi(x, y, t)$ denotes the transition density of $(\xi, \mathbb{P}^\phi)$ and $\lim_{t\to\infty} p^\phi(x, y, t) = \phi(y)\widetilde{\phi}(y)$ for every $x, y \in D$; cf. Pinchover [109, (2.12) and Theorem 1.3 (ii)]. Hence, (2.18), for $\pi \in \mathcal{M}_c(D)$ and $g \in p(D)$ with $\langle g, \phi\widetilde{\phi}\rangle < \infty$, follows from the dominated convergence theorem.

Lemma 2.6 (Many-to-one lemma for X and Z)**.** *For all $\mu \in \mathcal{M}_f(D)$, $\nu \in \mathcal{M}_a^{\mathrm{loc}}(D)$ and $g \in p(D)$,*

$$e^{-\lambda_c t} P_\mu[\langle \phi g, X_t\rangle] = \langle \mathbb{P}^\phi_\cdot[g(\xi_t)], \phi\mu\rangle, \tag{2.20}$$

$$e^{-\lambda_c t}\mathbf{P}_{\bullet,\nu}\Big[\Big\langle \frac{\phi}{w} g, Z_t\Big\rangle\Big] = \Big\langle \mathbb{P}^\phi_\cdot[g(\xi_t)], \frac{\phi}{w}\nu\Big\rangle \tag{2.21}$$

$$e^{-\lambda_c t}\mathbf{P}_\mu\Big[\Big\langle \frac{\phi}{w} g, Z_t\Big\rangle\Big] = \langle \mathbb{P}^\phi_\cdot[g(\xi_t)], \phi\mu\rangle. \tag{2.22}$$

Proof. Identity (2.20) follows immediately from (2.11) and (2.16). For (2.21), notice that by (1.8) the local growth rate of Z is given by

$$\beta^Z(x) := q(x)\Big(\sum_{k=2}^{\infty} kp_k(x) - 1\Big) = \partial_s G(x,s)\big|_{s=1} = \frac{\psi_0(x,w(x))}{w(x)} \qquad \text{for all } x \in D.$$

Using the definition of $\mathbb{P}^w_x$ in (1.7), we obtain for all $x \in D$,

$$\begin{aligned}
&\mathbb{P}^w_x\Big[e^{\int_0^t \beta^Z(\xi_s)\,ds}\frac{\phi(\xi_t)}{w(\xi_t)}g(\xi_t)\Big] \\
&\qquad = w(x)^{-1}\mathbb{P}_x\Big[\exp\Big(\int_0^t \Big(\beta^Z(\xi_s) - \frac{\psi_\beta(\xi_s,w(\xi_s))}{w(\xi_s)}\Big)\,ds\Big)\phi(\xi_t)g(\xi_t)\Big] \\
&\qquad = w(x)^{-1}S_t[\phi g](x) \overset{(2.16)}{=} \frac{\phi(x)}{w(x)}e^{\lambda_c t}\mathbb{P}^\phi_x[g(\xi_t)].
\end{aligned}$$

Hence, the first moment formula for branching diffusions (see, for example, Theorem 8.5 in [80]) yields

$$e^{-\lambda_c t}\mathbf{P}_{\bullet,\nu}\Big[\Big\langle\frac{\phi}{w}g, Z_t\Big\rangle\Big] = e^{-\lambda_c t}\Big\langle\mathbb{P}^w_\cdot\Big[e^{\int_0^t \beta^Z(\xi_s)\,ds}\frac{\phi(\xi_t)}{w(\xi_t)}g(\xi_t)\Big],\nu\Big\rangle = \Big\langle\mathbb{P}^\phi_\cdot[g(\xi_t)],\frac{\phi}{w}\nu\Big\rangle.$$

Since, under $\mathbf{P}_\mu$, the initial configuration of Z is given by a Poisson random measure with intensity $w(x)\mu(dx)$, (2.22) follows from (2.21). □

We record the following consequence of Lemma 2.6.

Corollary 2.7. *For all $\mu \in \mathcal{M}^\phi_f(D)$, $((W^\phi_t(X))_{t\ge 0}; P_\mu)$ and $((W^{\phi/w}_t(Z))_{t\ge 0}; \mathbf{P}_\mu)$ are martingales with*

$$P_\mu\big[W^\phi_t(X)\big] = \mathbf{P}_\mu\big[W^{\phi/w}_t(Z)\big] = \langle\phi,\mu\rangle \qquad \textit{for all } t \ge 0.$$

Proof. Since $(\xi,\mathbb{P}^\phi_x)$ is conservative, the identity for the expectations follows immediately from (2.20) and (2.22). The Markov property of X combined with (2.20) gives the claim for X. The Markov property of Z and (2.21) imply that $(W^{\phi/w}_t(Z))_{t\ge0}$ is a $\mathbf{P}_{\bullet,\nu}$-martingale for all $\nu \in \mathcal{M}^{\text{loc}}_a(D)$ with $\langle\phi/w,\nu\rangle < \infty$. Replacing ν by a Poisson random measure with intensity $w(x)\mu(dx)$ completes the proof. □

Let $\mu \in \mathcal{M}^\phi_f(D)$, $\mu \not\equiv 0$. After dividing the right-hand side of (2.20) by $\langle\phi,\mu\rangle$, the expression can be interpreted as the expectation of $g(\xi_t)$, where ξ is the ergodic motion with starting point randomised according to the probability distribution $\frac{\phi\mu}{\langle\phi,\mu\rangle}$. With this motivation, we define, for all measurable sets A,

$$\mathbb{P}^\phi_{\phi\mu}(A) := \frac{1}{\langle\phi,\mu\rangle}\big\langle\mathbb{P}^\phi_\cdot(A),\phi\mu\big\rangle. \tag{2.23}$$

We end this section with a remark for the case that the superprocess is deterministic.

Remark 2.8. If $\ell(\{x \in D\colon \alpha(x) + \Pi(x,(0,\infty)) > 0\}) = 0$, then (2.11)–(2.12) imply that $\langle f, X_t\rangle = \langle S_t f,\mu\rangle$ for all $t \ge 0$, P_μ-almost surely, for all continuous $f \in bp(D)$.

Hence, Assumption 1 cannot be satisfied. However, under Assumption 2, (2.16) and (2.18) imply that for continuous $f \in bp(D)$ with f/ϕ bounded,

$$e^{-\lambda_c t}\langle f, X_t\rangle = e^{-\lambda_c t}\langle S_t f, \mu\rangle = \langle \mathbb{P}^{\phi}_{\cdot}[f(\xi_t)/\phi(\xi_t)], \phi\mu\rangle \to \langle f/\phi, \phi\widetilde{\phi}\rangle\langle \phi, \mu\rangle = \langle f, \widetilde{\phi}\rangle W^{\phi}_{\infty}(X)$$

P_μ-almost surely, as $t \to \infty$. Now a standard approximation argument shows that the conclusion of Theorem 1.2 holds.

2.1.3 Moment conditions

In this section, we discuss Assumption 3 and compare it to the conditions used in the literature. We work under Assumptions 1 and 2. While Assumption 3 seems to be the most useful set of conditions, we prove our results under the following weaker moment assumption.

Assumption 3'. There exists $p \in (1, 2]$, $\varphi_1, \varphi_2 \in p(D)$, $\sigma_1, \sigma_2, \sigma_3 \in [p, 2]$ and $j_1, j_2 \in \{0, 1\}$ such that,

$$\sup_{x\in D} \phi(x)^{\sigma_1 - 1}\alpha(x) < \infty, \tag{2.24}$$

$$\sup_{x\in D} \phi(x)^{\sigma_2 - 1} \int_{(0,\varphi_1(x)]} y^{\sigma_2}\, \Pi(x, dy) < \infty, \tag{2.25}$$

$$\sup_{x\in D} \phi(x)^{\sigma_3 - 1} \int_{(\varphi_1(x),\infty)} y^{\sigma_3}\, \Pi(x, dy) < \infty, \tag{2.26}$$

$$\langle \phi^{p-1}, \phi\widetilde{\phi}\rangle < \infty, \tag{2.27}$$

$$\left\langle \phi^{j_1} \int_{(0,\varphi_2(\cdot)]} y^2 e^{-w(\cdot)y}\, \Pi(\cdot, dy), \phi\widetilde{\phi} \right\rangle < \infty, \tag{2.28}$$

$$\left\langle \phi^{j_2} \int_{(\varphi_2(\cdot),\infty)} y^2 e^{-w(\cdot)y}\, \Pi(\cdot, dy), \phi\widetilde{\phi} \right\rangle < \infty. \tag{2.29}$$

Assumption 3 is the special case $\varphi_1 = \varphi_2 = \mathbf{1}$, $\sigma_1 = \sigma_2 = 2$, $\sigma_3 = p$ and $j_1 = j_2 = 0$ of Assumption 3'. Notice that with this choice, Condition (2.28) trivially holds since $\langle \phi, \widetilde{\phi}\rangle < \infty$ and $x \mapsto \int_{(0,1]} y^2\, \Pi(x, dy)$ is a bounded function by the model assumptions in Section 1.1. Therefore, the following theorem generalises Proposition 1.1 and Theorems 1.2 and 1.5.

Theorem 2.9. *Suppose Assumptions 1, 2, (2.24)–(2.26) hold and $\mu \in \mathcal{M}^{\phi}_f(D)$.*

(i) *The martingales $(W^{\phi}_t(X))_{t\geq 0}$ and $(W^{\phi/w}_t(Z))_{t\geq 0}$ are bounded in $L^p(\mathbf{P}_\mu)$, and $W^{\phi}_\infty(X) = W^{\phi/w}_\infty(Z)$ $\mathbf{P}_\mu$-almost surely.*

(ii) *Suppose that, in addition, (2.27) holds. For all $f \in p(D)$ with f/ϕ bounded, we have in $L^1(P_\mu)$*

$$\lim_{t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle = \langle f, \widetilde{\phi}\rangle W^{\phi}_\infty(X). \tag{2.30}$$

(iii) *If, in addition, Assumptions 3' and 4 hold, then there exists a measurable set Ω_0 with $P_\mu(\Omega_0) = 1$ and, on Ω_0, for all ℓ-almost everywhere continuous functions $f \in p(D)$ with f/ϕ bounded, the convergence in (2.30) holds.*

The first three moment conditions, (1.10)–(1.12) or (2.24)–(2.26), are used to guarantee that the martingale $(W_t^\phi(X))_{t\geq 0}$ is bounded in L^p (see Theorem 2.15 below). To the best of our knowledge, even though these conditions may not be optimal, they are the best conditions obtained so far to guarantee L^p-boundedness, $p \in (1, 2)$, for general superprocesses. For the case of a super-Brownian motion, similar conditions were found in [97]. Condition (1.10) appeared as the main moment assumption in [61] and [62] to establish the convergence (2.30) in distribution and in probability, respectively. The two articles that study almost sure convergence in the product L^1-critical regime (i.e. under Assumption 2) are by Chen et al. [27] and Liu et al. [102]. In both papers, α and ϕ are bounded; hence, (1.10) holds.

The article [27] is restricted to quadratic branching mechanisms, i.e. $\Pi \equiv 0$, and (1.11)–(1.12) are trivially satisfied. Liu et al. [102] do not require Π to have a p-th moment. The authors show that under their assumptions (D of finite Lebesgue measure and $(S_t)_{t\geq 0}$ intrinsically ultracontractive) the martingale limit $W_\infty^\phi(X)$ is nontrivial if and only if $\langle \int_{(1,\infty)} y \log y \, \Pi(\cdot, y/\phi), \widetilde{\phi}\rangle < \infty$, and they establish their result under this condition. In the alternative case, the martingale limit is zero almost surely, and the stated convergence (1.16) holds trivially.

The fourth assumption, (1.13) or (2.27), is a technical condition. It is only used in Proposition 3.11 below to compare the immigration after a large time t, $\sum_{i=1}^{N_t} \langle f, I_s^{i,t}\rangle$, to its expectation $\sum_{i=1}^{N_t} \mathbf{P}_\mu[\langle f, I_s^{i,t}\rangle | \boldsymbol{\mathcal{F}}_t]$. In previous work on the SLLN [27, 102], Assumption (1.13) holds since ϕ is bounded.

The technical condition can be avoided using an h-transform. The h-transform for measure-valued diffusions was introduced by Engländer and Pinsky in [59]. For $h \in C^{2,\eta}(D)$, $h > 0$, let

$$L_0^h = L + a\frac{\nabla h}{h} \cdot \nabla, \quad \beta^h(x) = \frac{(L+\beta)h(x)}{h(x)}, \quad \psi_0^h(x, z) = \frac{\psi_0(x, h(x)z)}{h(x)}. \tag{2.31}$$

If β^h, αh and $x \mapsto \int_{(0,\infty)} (y \wedge h(x)y^2)\, \Pi(x, dy)$ belong to $b(D)$, then $\psi_{\beta^h}^h(x, z) := -\beta^h(x)z + \psi_0^h(x, z)$ satisfies the assumptions from Section 1.1. We denote the space of such functions h by $\mathbb{H}(\psi_\beta)$. An $(L_0^h, \psi_{\beta^h}^h; D)$-superprocess X^h started in $h(x)\mu(dx)$ can be obtained from an $(L, \psi; D)$-superprocess X started in μ by setting $X_t^h(dx) := h(x)X_t(dx)$. This result follows immediately from a comparison of the Laplace transforms using the mild equation (1.3) and Corollary 4.1.2 in [110]; see [59] for the computation in the quadratic case. In the following, we superscript all quantities derived from X^h with an h. Clearly, the $(L, \psi; D)$-superprocess can be recovered from the $(L_0^h, \psi_{\beta^h}^h; D)$-superprocess by a transform with $1/h$.

Lemma 2.10. *Let $h \in \mathbb{H}(\psi_\beta)$ and $\mu \in \mathcal{M}_f^\phi(D)$.*

(i) *The operator $L_0^h + \beta^h$ satisfies Assumption 2 with $\phi^h = \phi/h$, $\widetilde{\phi}^h = \widetilde{\phi}h$ and $\lambda_c^h = \lambda_c$, and the process $(W_t^{\phi^h}(X^h) = e^{-\lambda_c^h t}\langle \phi^h, X_t^h\rangle : t \geq 0; P_{h\mu}^h)$ is a martingale with almost sure limit $W_\infty^{\phi^h}(X^h)$.*

(ii) *Suppose (2.30) holds P_μ-almost surely for some $f \in p(D)$, then*

$$\lim_{t\to\infty} e^{-\lambda_c^h t}\langle f/h, X_t^h\rangle = \langle f/h, \widetilde{\phi}^h\rangle W_\infty^{\phi^h}(X^h) \qquad P_{h\mu}^h\text{-almost surely.} \tag{2.32}$$

If (2.30) holds in $L^1(P_\mu)$ instead, then (2.32) holds in $L^1(P_{h\mu}^h)$.

Proof. The first part of the claim was proved by Pinsky [110, Chapter 4]. Setting $X^h := hX$, we immediately obtain $W_t^{\phi^h}(X^h) = e^{-\lambda_c^h t}\langle \phi^h, X_t^h\rangle = W_t^\phi(X)$ and, using (2.30), P_μ-almost surely (in $L^1(P_\mu)$, respectively),

$$e^{-\lambda_c^h t}\langle f/h, X_t^h\rangle = e^{-\lambda_c t}\langle f, X_t\rangle \to \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X) = \langle f/h, \widetilde{\phi}^h\rangle W_\infty^{\phi^h}(X^h) \qquad \text{as } t \to \infty.$$

□

Lemma 2.10 states that Assumption 2 and our results are invariant under h-transforms. The same is true for Assumptions 1 and 4.

Lemma 2.11. *Let $h \in \mathbb{H}(\psi_\beta)$. The $(L_0^h, \psi_{\beta^h}^h; D)$-superprocess X^h satisfies Assumption 1 with martingale function $w^h = w/h$ and the distribution of the skeleton Z^h under $\mathbf{P}_{h\mu}^h$ agrees with the distribution of Z under $\mathbf{P}_\mu$ for all $\mu \in \mathcal{M}_c(D)$. In particular, if X satisfies Assumption 4, then X^h satisfies Assumption 4.*

Proof. The claim follows immediately from the definitions. □

Exploiting the invariance under h-transforms, we can prove our main results under the following moment assumption.

Assumption 3". There exists $p \in (1, 2]$ such that Conditions (1.10)–(1.12) and (2.29) for $j_2 = 1$, $\varphi_2 = \mathbf{1}$ hold and

$$\sup_{x\in D} \int_{(1/\phi(x),\infty)} y\,\Pi(x, dy) < \infty. \tag{2.33}$$

Crucially, Assumption 3" does not require $\langle \phi^p, \widetilde{\phi}\rangle < \infty$. In the case of a quadratic branching mechanism, only boundedness of $\phi\alpha$ is required. Condition (2.33) is needed to guarantee that $\phi \in \mathbb{H}(\psi_\beta)$.

Theorem 2.12. *Suppose Assumptions 1, 2, (1.10)–(1.12) and (2.33) hold, and let $\mu \in \mathcal{M}_f^\phi(D)$.*

(i) *For all $f \in p(D)$ with f/ϕ bounded, the convergence in (2.30) holds in $L^1(P_\mu)$.*

(ii) *If, in addition, Assumptions 3" and 4 hold, then there exists a measurable set* Ω_0 *with* $P_\mu(\Omega_0) = 1$ *and, on* Ω_0*, for all* ℓ*-almost everywhere continuous functions* $f \in p(D)$ *with* f/ϕ *bounded, the convergence in* (2.30) *holds.*

Proof of Theorem 2.12 assuming Theorem 2.9. Part (i): since $\beta^\phi = \lambda_c$, $\alpha^\phi = \phi\alpha$ and $\Pi^\phi(x, dy) = \frac{1}{\phi(x)}\Pi(x, dy/\phi(x))$, Conditions (1.10), (1.11), (2.33) and the model assumptions in Section 1.1 guarantee that $\phi \in \mathbb{H}(\psi_\beta)$. By Lemma 2.10 (i), X^ϕ satisfies Assumption 2 and $\phi^\phi = \mathbf{1}$. Thus, (1.10)–(1.12) imply that X^ϕ satisfies (2.24)–(2.27) with $\varphi_1 = \phi$, $\sigma_2 = 2$, $\sigma_3 = p$ and $\sigma_1 \in [p, 2]$ arbitrary. Using Lemma 2.11, we deduce that Theorem 2.9 (ii) applies to X^ϕ, and the claim follows from Lemma 2.10 (ii).

Part (ii): X^ϕ satisfies (2.28)–(2.29) with $\varphi_2 = \phi$ and arbitrary $j_1, j_2 \in \{0, 1\}$, and Assumption 4 by Lemma 2.11. Hence, Theorem 2.9 (iii) applies to X^ϕ, and Lemma 2.10 (ii) completes the proof for fixed functions f. The existence of a common set Ω_0 will be proved in Lemma 3.4 below. □

Engländer and Winter [62] proved the convergence (2.30) in probability under the assumption of a quadratic branching mechanisms and (1.10). Their argument can easily be extended to general branching mechanisms. Since the proof relies on an h-transform with $h = \phi$ and second moment estimates, the additional conditions needed for this generalisation are (1.11), (1.12) with $p = 2$, and (2.33).

The freedom to choose $p \in (1, 2]$ allows us to analyse processes where $(W_t^\phi(X))_{t\geq 0}$ is bounded in L^p for $p \in (1, 2)$ but not in L^2. Examples of such processes are given in Chapter 4. In these cases, not only our almost sure convergence result is new but also the implied convergence in probability result. The main tool to deal with non-integer moments is a spine decomposition presented in Section 2.2, and we are not aware of any other way to obtain these conditions.

The final conditions (2.28)–(2.29) simplify to (1.14) in the case $j_1 = j_2 = 0$, $\varphi_2 = \mathbf{1}$. These assumptions guarantee that the process X^* from the skeleton decomposition has finite second moments (2.14), a fact which is only used in the transition from lattice to continuous times. In particular, the SLLN along lattice times in Theorem 3.13 below holds without it. If w is bounded away from zero, for instance when the branching mechanism is spatially independent and the motion is conservative (see Section 4.1), then (1.14) holds automatically. Since Chen et al. [27] consider a quadratic branching mechanism, the conditions automatically hold in their article. In contrast, Liu et al. [102] have no conditions of this type.

In summary, our moment conditions are weaker than those used in [27], but compared to [102], we impose stricter assumptions on the Lévy measure Π, yet allow a much larger class of underlying motions ξ and domains D.

2.1.4 The strong law of large numbers for the skeleton

Throughout this section, we suppose that Assumptions 1, 2 and 3' hold. Assumption 4 may look like a strong assumption on first glance. However, we argue that this is not so. The skeleton decomposition shows that the large-time behaviour of the superprocess is guided by the skeleton. This suggests that the total mass the superprocess assigns to a compact ball, will be asymptotically well-behaved if and only if the skeleton carrying the superprocess has asymptotically a well-behaved number of particles in that ball. We write $\mathcal{B}_0(D) := \{B \in \mathcal{B}(D) \colon \ell(\partial B) = 0\}$. To show that Assumption 4 holds, it suffices to prove that, for all $\mu \in \mathcal{M}_c(D)$, $B \in \mathcal{B}_0(D)$,

$$\liminf_{n\to\infty} e^{-\lambda_c n\delta}\Big\langle \frac{\phi}{w}\mathbb{1}_B, Z_{n\delta}\Big\rangle \geq \langle \phi\mathbb{1}_B, \widetilde{\phi}\rangle W_\infty^{\phi/w}(Z) \qquad \mathbf{P}_\mu\text{-almost surely}$$

as we will see in Lemma 3.1 (ii) below. Often it is a much easier task to prove the convergence along lattice times than along continuous times.

There are good results in the literature proving SLLNs for branching diffusions. Some of the relevant literature was reviewed in Section 1.4. A nice argument to obtain almost sure asymptotics for spatial branching particle processes from related asymptotic behaviour of the spine was found recently by Harris and Roberts [82]. However, they assume a convergence for the spine which usually does not hold in our setup. The theorem we use to verify several examples in Chapter 4 is based on a result from Engländer et al. [57]. The authors prove the convergence for strictly dyadic branching diffusions along continuous times. We require only convergence along lattice times but a more general branching generator. The following theorem is a version of their result as our proof reveals.

Theorem 2.13 (Adaptation of Theorem 6 in [57])**.** *Let $\mu \in \mathcal{M}_c(D)$, and assume that for every x in the support of μ the following conditions hold:*

(i) *There is a family of sets $D_t \in \mathcal{B}(D)$, $t \geq 0$, such that for all $\delta > 0$,*

$$\mathbf{P}_{\bullet,\delta_x}(\exists n_0 \in \mathbb{N} \colon \mathrm{supp}(Z_{n\delta}) \subseteq D_{n\delta} \textit{ for all } n \geq n_0) = 1.$$

(ii) *For every $B \subset\subset D$, there exists a constant $K > 0$ such that*

$$\sup_{y\in D_t} \left|\mathbb{P}_y^\phi[\mathbb{1}_B(\xi_{Kt})] - \langle \phi\mathbb{1}_B, \widetilde{\phi}\rangle\right| \to 0 \qquad \textit{as } t\to\infty. \tag{2.34}$$

Then, for all $\delta > 0$, $f \in p(D)$ with fw/ϕ bounded,

$$\lim_{n\to\infty} e^{-\lambda_c n\delta}\langle f, Z_{n\delta}\rangle = \langle f, w\widetilde{\phi}\rangle W_\infty^{\phi/w}(Z) \qquad \mathbf{P}_\mu\textit{-almost surely.}$$

Proof. Using Notation 2.4, we have $Z = \sum_{i=1}^{N_0} Z^{i,0}$, where given $\boldsymbol{\mathcal{F}}_0$, the processes $(Z^{i,0} \colon i = 1, \ldots, N_0)$ are independent, and $(Z^{i,0}; \mathbf{P}_\mu(\cdot|\boldsymbol{\mathcal{F}}_0))$ is equal in distribution to

$(Z; \mathbf{P}_{\bullet,\delta_{\xi_i(0)}})$. In particular, $W_\infty^{\phi/w}(Z) = \sum_{i=1}^{N_0} W_\infty^{\phi/w}(Z^{i,0})$, and

$$\begin{aligned}
&\mathbf{P}_\mu\Big(\lim_{n\to\infty} e^{-\lambda_c n\delta}\langle f, Z_{n\delta}\rangle = \langle f, w\widetilde{\phi}\rangle W_\infty^{\phi/w}(Z)\Big) \\
&\quad \geq \mathbf{P}_\mu\Big(\bigcap_{i=1}^{N_0}\Big\{\lim_{n\to\infty} e^{-\lambda_c n\delta}\langle f, Z_{n\delta}^{i,0}\rangle = \langle f, w\widetilde{\phi}\rangle W_\infty^{\phi/w}(Z^{i,0})\Big\}\Big) \\
&\quad = \mathbf{P}_\mu\Big[\prod_{i=1}^{N_0}\mathbf{P}_{\bullet,\delta_{\xi_i(0)}}\Big(\lim_{n\to\infty} e^{-\lambda_c n\delta}\langle f, Z_{n\delta}\rangle = \langle f, w\widetilde{\phi}\rangle W_\infty^{\phi/w}(Z)\Big)\Big].
\end{aligned}$$

It remains to argue that under the stated assumptions, $\mathbf{P}_{\bullet,\delta_x}(\lim_{n\to\infty} e^{-\lambda_c n\delta}\langle f, Z_{n\delta}\rangle = \langle f, w\widetilde{\phi}\rangle W_\infty^{\phi/w}(Z))$ equals 1 for every x in the support of μ. Engländer et al. [57] give a proof of this result for strictly dyadic branching diffusions in two steps. The argument can be generalised as follows. The first step is to show that with $(s_n)_{n\geq 0}$ nonnegative and non-decreasing, and $U_n = e^{-\lambda_c(s_n+\delta n)}\langle f, Z_{s_n+\delta n}\rangle$, the sequence $\mathrm{D}_f(s_n, \delta n) = |U_n - \mathbf{P}_{\bullet,\delta_x}[U_n|\sigma(Z_r\colon r \leq n\delta)]|$ converges to zero. The key to this result is an upper bound on the p-th moment of $W_t^{\phi/w}(Z)$ and is obtained via a spine decomposition of the branching diffusion. This would be possible even in our more general setup but is not needed since the required bound follows easily from Theorem 2.9 (i). The second step is to show the convergence of $\mathrm{CE}_f(s_n, \delta n) = \mathbf{P}_{\bullet,\delta_x}[U_n|\sigma(Z_r\colon r \leq \delta n)]$ for $s_n = K\delta n$ to $\langle f, w\widetilde{\phi}\rangle W_\infty^{\phi/w}(Z)$. Their assumptions for this convergence are Conditions (iii) and (iv) in their Definition 4. Condition (iii) is our Condition (i) in Theorem 2.13. From the proof in [57] it is easy to see that their Condition (iv) in Definition 4 can be relaxed to our Condition (ii), a fact that has also been used in the verification of some examples in [57]. $\square$

The following lemma is useful in the verification of the conditions of Theorem 2.13 and has been proved by Engländer et al. [57]. We give the main argument for completeness. Denote by $\|\cdot\|$ the ℓ^2-norm on $\mathbb{R}^d$.

Lemma 2.14. *Suppose for $x \in D$ there are a continuous function $a\colon [0,\infty) \to [0,\infty)$ and some $\epsilon > 0$ such that*

$$\mathbb{P}_x^\phi[\mathbb{1}_{\{\|\xi_t\|\geq a(t)\}} w(\xi_t)/\phi(\xi_t)] \leq e^{-(\lambda_c+\epsilon)t} \qquad \text{for all } t \text{ sufficiently large.} \tag{2.35}$$

Then Condition (i) in Theorem 2.13 holds with $D_t = \{y \in D\colon \|y\| < a(t)\}$. If, in addition, for every $B \subset\subset D$, there is $K > 0$ such that

$$\sup_{\|y_1\|<a(t), y_2\in B}\left|\frac{p^\phi(y_1, y_2, Kt)}{\phi(y_2)\widetilde{\phi}(y_2)} - 1\right| \to 0 \qquad \text{as } t \to \infty, \tag{2.36}$$

where p^ϕ denotes the transition density of $(\xi; \mathbb{P}^\phi)$, then also Condition (ii) in Theorem 2.13 is satisfied.

Proof. Markov's inequality and (2.21) yield for all $t \geq 0$,

$$\mathbf{P}_{\bullet,\delta_x}\big(\mathrm{supp}\big(Z_t\big) \not\subseteq D_t\big) \leq \mathbf{P}_{\bullet,\delta_x}\big[\langle \mathbb{1}_{D_t^c}, Z_t\rangle\big] = e^{\lambda_c t}\frac{\phi(x)}{w(x)}\mathbb{P}_x^{\phi}\big[\mathbb{1}_{\{\|\xi_t\|\geq a(t)\}}w(\xi_t)/\phi(\xi_t)\big].$$

The Borel–Cantelli lemma yields the first part of the lemma. The second part follows immediately from the definitions and from $\langle \phi, \widetilde{\phi}\rangle < \infty$. □

We will see in Chapter 4 that for many of the main examples of superdiffusions the SLLN for the skeleton already follows from Theorem 2.13. For those processes where Assumption 4 has not been proved yet, we believe that the particle nature of the skeleton will make it easier to obtain the SLLN for the skeleton than to derive further convergence statements in the superprocess setup. This thesis will then allow us to carry results for the branching diffusion over to the superdiffusion. We emphasize that the SLLN for the skeleton is only needed along lattice times and for compactly supported starting measures.

2.2 Spine decomposition

In this section, we use a spine decomposition of X to identify $(W_t^{\phi}(X))_{t\geq 0}$ as an L^p-bounded martingale, where $p \in (1,2]$ is determined by Assumption 3'. A similar decomposition has been used for other purposes by Engländer and Kyprianou [58] on bounded subdomains for quadratic branching mechanisms and by Liu et al. [101] in the case $\alpha = \mathbf{0}$. For the one-dimensional super-Brownian motion the spine decomposition was used by Kyprianou et al. [96, 97] to establish L^p-boundedness of martingales closely related to $(W_t^{\phi}(X))_{t\geq 0}$. Similar arguments have been used in the setup of branching diffusions in [80, 57]. See [58] for an overview of the history of spine decompositions for branching processes. Throughout this section, we suppose that Assumption 2 holds. Further conditions used are stated explicitly. Recall that $\mathcal{M}_f^{\phi}(D) = \{\mu \in \mathcal{M}_f(D) \colon \langle \phi, \mu\rangle < \infty\}$.

Theorem 2.15. *Suppose Assumptions (2.24)–(2.26) hold. For all $\mu \in \mathcal{M}_f^{\phi}(D)$, process $((W_t^{\phi}(X))_{t\geq 0}; P_\mu)$ is an L^p-bounded martingale. In particular, $((W_t^{\phi}(X))_{t\geq 0}; P_\mu)$ converges in $L^p(P_\mu)$.*

Let $\mu \in \mathcal{M}_f^{\phi}(D)$, $\mu \not\equiv 0$. We already showed in Corollary 2.7 that $(W_t^{\phi}(X))_{t\geq 0}$ is a martingale. Hence, it suffices to prove L^p-boundedness, and we can define a new probability measure Q_μ by

$$\left.\frac{dQ_\mu}{dP_\mu}\right|_{\sigma(X_s \colon s\in[0,t])} = \frac{W_t^{\phi}(X)}{\langle \phi, \mu\rangle} \quad \text{for all } t \geq 0.$$

Recall from (2.23) that $(\xi = (\xi_t)_{t\geq 0}; \mathbb{P}_{\phi\mu}^{\phi})$ is the ergodic motion with randomised starting point, and use (2.15) to obtain

$$\mathbb{P}^{\phi}_{\phi\mu}(A) = \frac{e^{-\lambda_c t}}{\langle \phi, \mu\rangle}\Big\langle \mathbb{P}_{\cdot}\big[e^{\int_0^t \beta(\xi_s)\,ds}\phi(\xi_t)\mathbb{1}_A\big], \mu\Big\rangle \qquad \text{for all } A \in \sigma(\xi_s \colon 0 \le s \le t). \tag{2.37}$$

Lemma 2.16. *For all $\mu \in \mathcal{M}^{\phi}_f(D)$, $\mu \not\equiv 0$, $f, g \in bp(D)$, $t \ge 0$,*

$$\begin{aligned} Q_\mu\Big[e^{-\langle f, X_t\rangle}\frac{\langle \phi g, X_t\rangle}{\langle \phi, X_t\rangle}\Big] \\ = P_\mu\big[e^{-\langle f, X_t\rangle}\big]\mathbb{P}^{\phi}_{\phi\mu}\Big[g(\xi_t)\exp\Big(-\int_0^t \partial_z\psi_0(\xi_s, u_f(\xi_s, t-s))\,ds\Big)\Big]. \end{aligned} \tag{2.38}$$

Notice that by definition, $\langle \phi, X_t\rangle > 0$, Q_μ-almost surely.

Proof of Lemma 2.16. We prove (2.38) only for g compactly supported since the general case then follows from the monotone convergence theorem. The continuity of ϕ implies that $f + \theta\phi g \in bp(D)$ for all $\theta \ge 0$. Use the definition of Q_μ, and interchange differentiation and integration using the dominated convergence theorem to obtain

$$\begin{aligned} Q_\mu\Big[e^{-\langle f, X_t\rangle}\frac{\langle \phi g, X_t\rangle}{\langle \phi, X_t\rangle}\Big] &= -\frac{e^{-\lambda_c t}}{\langle \phi, \mu\rangle}P_\mu\big[\partial_\theta\big|_{\theta=0}e^{-\langle f+\theta\phi g, X_t\rangle}\big] \\ &= \frac{e^{-\lambda_c t}}{\langle \phi, \mu\rangle}e^{-\langle u_f(\cdot, t), \mu\rangle}\partial_\theta\big|_{\theta=0}\langle u_{f+\theta\phi g}(\cdot, t), \mu\rangle. \end{aligned}$$

By (1.2), the definition of ψ_β, and (2.37) the claim follows when we have shown that

$$\begin{aligned} h_{f,g}(x,t) &:= \partial_\theta\big|_{\theta=0}u_{f+\theta\phi g}(x,t) \\ &= \mathbb{P}_x\Big[\phi(\xi_t)g(\xi_t)\exp\Big(-\int_0^t \partial_z\psi_\beta(\xi_s, u_f(\xi_s, t-s))\,ds\Big)\Big] \end{aligned} \tag{2.39}$$

since integration with respect to μ and differentiation can be interchanged using the dominated convergence theorem. By (1.3), for any $\theta > 0$,

$$\begin{aligned} &\frac{u_{f+\theta\phi g}(x,t) - u_f(x,t)}{\theta} \\ &\qquad = S_t[\phi g](x) - \int_0^t S_s\Big[\frac{\psi_0(\cdot, u_{f+\theta\phi g}(\cdot, t-s)) - \psi_0(\cdot, u_f(\cdot, t-s))}{\theta}\Big](x)\,ds. \end{aligned}$$

The Laplace exponent $\theta \mapsto v(\theta) := u_{f+\theta\phi g}(x,t) = -\log P_{\delta_x}[e^{-\langle f+\theta\phi g, X_t\rangle}]$ is increasing, concave and nonnegative. In particular, $\frac{v(\theta)-v(0)}{\theta}$ is decreasing in θ. Moreover, $z \mapsto \psi_0(x,z)$ is increasing, convex, and nonnegative. Hence, for all $(x,t) \in D \times [0,\infty)$,

$$0 \le \frac{v(\theta) - v(0)}{\theta} = \frac{u_{f+\theta\phi g}(x,t) - u_f(x,t)}{\theta} \le S_t[\phi g](x) \le \|\phi g\|_\infty e^{\bar{\beta}t},$$

where $\bar{\beta} = \sup_{x\in D}\beta(x)$, and $\|\cdot\|_\infty$ denotes the supremum norm. A Taylor expansion of ψ_0 yields for every $(x,t) \in D \times [0,\infty)$ some $\tilde{\theta} \in (0,\theta)$ such that

$$\frac{\psi_0(x, v(\theta)) - \psi_0(x, v(0))}{\theta}$$
$$= \partial_z \psi_0(x, v(0)) \frac{v(\theta) - v(0)}{\theta} + \Big(\partial_z \psi_0(x, v(\tilde{\theta})) - \partial_z \psi_0(x, v(0))\Big) \frac{v(\theta) - v(0)}{\theta}.$$

The first term on the right-hand side is nonnegative and increases as $\theta \downarrow 0$, the second term is dominated and tends to zero. Hence,

$$h_{f,g}(x,t) = S_t[\phi g](x) - \int_0^t S_s\big[\partial_z \psi_0(\cdot, u_f(\cdot, t-s)) h_{f,g}(\cdot, t-s)\big](x)\, ds. \tag{2.40}$$

Lemma A.1 (ii) below applied to the functions $g_1(x,t) = -\partial_z \psi_\beta(x, u_f(x,t))$, $g_2(x,t) = \partial_z \psi_0(x, u_f(x,t))$, $f_1 = \phi g$ and $f_2(x,t) = -\partial_z \psi_0(x, u_f(x,t)) h_{f,g}(x,t)$ shows that the unique solution to (2.40) is given by the right-hand side of (2.39). □

Recall the definition of Dynkin and Kuznetsov's $\mathbb{N}_x$-measures from (2.8), and let $\mu \in \mathcal{M}_f^\phi(D)$, $\mu \not\equiv 0$. On a suitable probability space with measure $P_{\mu,\phi}$, we define the following processes:

(i) $(\xi = (\xi_t)_{t \geq 0}; P_{\mu,\phi})$ is equal in distribution to $(\xi_t \colon t \geq 0; \mathbb{P}^\phi_{\phi\mu})$, that is an ergodic diffusion. We refer to this process as the *spine*.

(ii) **Continuous immigration:** $(\mathbf{n}; P_{\mu,\phi})$ a random measure such that, given ξ, $\mathbf{n}$ is a Poisson random measure which issues $\mathcal{M}_f(D)$-valued processes $X^{\mathbf{n},t} = (X^{\mathbf{n},t}_s)_{s \geq 0}$ at space-time point (ξ_t, t) with rate $2\alpha(\xi_t)\, dt \times d\mathbb{N}_{\xi_t}$. The almost surely countable set of immigration times is denoted by $\mathcal{D}^{\mathbf{n}}$; $\mathcal{D}^{\mathbf{n}}_t := \mathcal{D}^{\mathbf{n}} \cap (0,t]$. Given ξ, the processes $(X^{\mathbf{n},t} \colon t \in \mathcal{D}^{\mathbf{n}})$ are independent.

(iii) **Discontinuous immigration:** $(\mathbf{m}; P_{\mu,\phi})$ a random measure such that, given ξ, $\mathbf{m}$ is a Poisson random measure which issues $\mathcal{M}_f(D)$-valued processes $X^{\mathbf{m},t}$ at space-time point (ξ_t, t) with rate $dt \times \int_{(0,\infty)} \Pi(\xi_t, dy)\, y \times dP_{y\delta_{\xi_t}}$. The almost surely countable set of immigration times is denoted by $\mathcal{D}^{\mathbf{m}}$; $\mathcal{D}^{\mathbf{m}}_t = \mathcal{D}^{\mathbf{m}} \cap (0,t]$. Given ξ, the processes $(X^{\mathbf{m},t} \colon t \in \mathcal{D}^{\mathbf{m}})$ are independent and independent of $\mathbf{n}$ and $(X^{\mathbf{n},t} \colon t \in \mathcal{D}^{\mathbf{n}})$.

(iv) $(X = (X_t)_{t \geq 0}; P_{\mu,\phi})$ is equal in distribution to $(X = (X_t)_{t \geq 0}; P_\mu)$, i.e. it is a copy of the original process. Moreover, X is independent of $\xi, \mathbf{n}, \mathbf{m}$ and all immigration processes.

We denote by

$$X^{\mathbf{n}}_t = \sum_{s \in \mathcal{D}^{\mathbf{n}}_t} X^{\mathbf{n},s}_{t-s} \qquad \text{and} \qquad X^{\mathbf{m}}_t = \sum_{s \in \mathcal{D}^{\mathbf{m}}_t} X^{\mathbf{m},s}_{t-s}$$

the continuous and discontinuous immigration processes, respectively. We write $\Gamma_t := X_t + X^{\mathbf{n}}_t + X^{\mathbf{m}}_t$ for all $t \geq 0$, and $\stackrel{d}{=}$ denotes distributional equality.

Proposition 2.17 (Spine decomposition). *For all $\mu \in \mathcal{M}_f^{\phi}(D)$, $\mu \not\equiv 0$,*

$$(X_t\colon t \geq 0; Q_\mu) \stackrel{d}{=} (\Gamma_t = X_t + X_t^{\mathbf{n}} + X_t^{\mathbf{m}}\colon t \geq 0; P_{\mu,\phi}).$$

The proof of Proposition 2.17 is very similar to the proof of Theorem 5.2 in [96], and we omit long computations.

Proof of Proposition 2.17. Using the definitions and Campbell's formula for Poisson random measures, one easily checks that the marginal distributions agree. By definition, $((\Gamma_t, \xi_t)_{t\geq 0}; P_{\mu,\phi})$ is a time-homogeneous Markov process, and when we show that

$$P_{\mu,\phi}(\xi_t \in dx \,|\, \Gamma_t) = \frac{1}{\langle \phi, \Gamma_t\rangle}\phi(x)\Gamma_t(dx) \quad \text{ for all } t \geq 0,$$

then $((\Gamma_t)_{t\geq 0}; P_{\mu,\phi})$ is a time-homogeneous Markov process (by the argument given on page 21 of [96]). Using the definition, Lemma 2.16, and $(\Gamma_t; P_{\mu,\phi}) \stackrel{d}{=} (X_t; Q_\mu)$, we find that for all $f, g \in bp(D)$,

$$P_{\mu,\phi}\big[e^{-\langle f,\Gamma_t\rangle} P_{\mu,\phi}[g(\xi_t)\,|\,\Gamma_t]\big] = P_{\mu,\phi}\Big[e^{-\langle f,\Gamma_t\rangle}\frac{\langle \phi g, \Gamma_t\rangle}{\langle \phi, \Gamma_t\rangle}\Big],$$

and the claim follows. $\square$

For all $t \geq 0$, let $\mathcal{G}_t$ be the σ-algebra generated by ξ up to time t and by $\mathbf{n}$ and $\mathbf{m}$ restricted in the time component to $[0, t]$.

Lemma 2.18. *For all $\mu \in \mathcal{M}_f^{\phi}(D)$, $\mu \not\equiv 0$, and $t \geq 0$, $P_{\mu,\phi}$-almost surely,*

$$P_{\mu,\phi}\big[e^{-\lambda_c t}\langle \phi, \Gamma_t\rangle | \mathcal{G}_t\big] = \langle \phi, \mu\rangle + \sum_{s\in\mathcal{D}_t^{\mathbf{n}}} e^{-\lambda_c s}\phi(\xi_s) + \sum_{s\in\mathcal{D}_t^{\mathbf{m}}} e^{-\lambda_c s}\mathcal{I}_s^{\mathbf{m}}\phi(\xi_s),$$

where $(\mathcal{I}_t^{\mathbf{m}} := \langle \mathbf{1}, X_0^{\mathbf{m},t}\rangle\colon t \geq 0; P_{\mu,\phi})$ is, given ξ, a Poisson point process with intensity measure $dt \times \int_{(0,\infty)} \Pi(\xi_t, dy)\, y$.

Proof. Proposition 1.1 of [45] states that, for all $f \in p(D)$ with $P_{\delta_x}[\langle f, X_t\rangle] < \infty$,

$$\mathbb{N}_x[\langle f, X_t\rangle] = P_{\delta_x}[\langle f, X_t\rangle]. \tag{2.41}$$

Using first the definition of Γ_t, and then (2.41) and the branching property of X, we obtain

$$\begin{aligned}
&P_{\mu,\phi}[e^{-\lambda_c t}\langle \phi, \Gamma_t\rangle | \mathcal{G}_t] \\
&\quad = P_\mu\big[W_t^{\phi}(X)\big] + \sum_{s\in\mathcal{D}_t^{\mathbf{n}}} e^{-\lambda_c t}\mathbb{N}_{\xi_s}[\langle \phi, X_{t-s}\rangle] + \sum_{s\in\mathcal{D}_t^{\mathbf{m}}} e^{-\lambda_c t} P_{\mathcal{I}_s^{\mathbf{m}}\delta_{\xi_s}}[\langle \phi, X_{t-s}\rangle] \\
&\quad = P_\mu\big[W_t^{\phi}(X)\big] + \sum_{s\in\mathcal{D}_t^{\mathbf{n}}} e^{-\lambda_c t} P_{\delta_{\xi_s}}[\langle \phi, X_{t-s}\rangle] + \sum_{s\in\mathcal{D}_t^{\mathbf{m}}} e^{-\lambda_c t}\mathcal{I}_s^{\mathbf{m}} P_{\delta_{\xi_s}}[\langle \phi, X_{t-s}\rangle].
\end{aligned}$$

Since $W_t^\phi(X) = e^{-\lambda_c t}\langle \phi, X_t\rangle$, $t \geq 0$, is a P_μ- and P_{δ_x}-martingale for all $x \in D$, the claim follows. □

Finally, everything is prepared for the proof of Theorem 2.15. Throughout the thesis, we use the letters c and C for generic constants in $(0,\infty)$ and their value can change from line to line. Important constants are marked by an index indicating the order in which they occur.

Proof of Theorem 2.15. The martingale property was proved in Corollary 2.7. We have to show the L^p-boundedness. If $\mu \equiv 0$, then $X_t(D) = 0$ for all $t \geq 0$, and the statement is trivially true. Let $\mu \in \mathcal{M}_f^\phi(D)$, $\mu \not\equiv 0$. We write $W_t^\phi(\Gamma) = e^{-\lambda_c t}\langle \phi, \Gamma_t\rangle$ and $\bar{p} = p - 1 \in (0,1]$. Then $x \mapsto x^{\bar{p}}$ is concave and $(x+y)^{\bar{p}} \leq x^{\bar{p}} + y^{\bar{p}}$ for $x, y \geq 0$. Hence, the definition of Q_μ, Proposition 2.17, Jensen's inequality and Lemma 2.18, yield

$$\begin{aligned}
\frac{P_\mu[W_t^\phi(X)^p]}{\langle \phi, \mu\rangle} &= Q_\mu\big[W_t^\phi(X)^{\bar{p}}\big] = P_{\mu,\phi}\big[P_{\mu,\phi}[W_t^\phi(\Gamma)^{\bar{p}}|\mathcal{G}_t]\big] \leq P_{\mu,\phi}[P_{\mu,\phi}[W_t^\phi(\Gamma)|\mathcal{G}_t]^{\bar{p}}] \\
&\leq P_{\mu,\phi}\Big[\langle \phi, \mu\rangle^{\bar{p}} + \Big(\sum_{s\in\mathcal{D}_t^{\mathbf{n}}} e^{-\lambda_c s}\phi(\xi_s)\Big)^{\bar{p}}\Big] \\
&\quad + P_{\mu,\phi}\Big[\Big(\sum_{s\in\mathcal{D}_t^{\mathbf{m}},\mathcal{I}_s^{\mathbf{m}}\leq\varphi_1(\xi_s)} e^{-\lambda_c s}\mathcal{I}_s^{\mathbf{m}}\phi(\xi_s)\Big)^{\bar{p}} + \Big(\sum_{s\in\mathcal{D}_t^{\mathbf{m}},\mathcal{I}_s^{\mathbf{m}}>\varphi_1(\xi_s)} e^{-\lambda_c s}\mathcal{I}_s^{\mathbf{m}}\phi(\xi_s)\Big)^{\bar{p}}\Big],
\end{aligned}$$

where φ_1 is determined by Assumption 3'. The first term is deterministic. For the remaining three terms we first use that $x^{\bar{p}} \leq 1+x^{\bar{\sigma}}$ for all $\bar{\sigma} \geq \bar{p}$, then $(x+y)^{\bar{\sigma}} \leq x^{\bar{\sigma}}+y^{\bar{\sigma}}$ for all $\bar{\sigma} \in [0,1]$, and finally apply Campbell's formula to obtain

$$\begin{aligned}
&P_{\mu,\phi}\Big[\Big(\sum_{s\in\mathcal{D}_t^{\mathbf{n}}} e^{-\lambda_c s}\phi(\xi_s)\Big)^{\bar{p}}\Big] \leq 1 + \int_0^t 2e^{-\lambda_c\bar{\sigma}_1 s}\mathbb{P}_{\phi\mu}^\phi[\phi(\xi_s)^{\bar{\sigma}_1}\alpha(\xi_s)]\,ds, \\
&P_{\mu,\phi}\Big[\Big(\sum_{s\in\mathcal{D}_t^{\mathbf{m}},\mathcal{I}_s^{\mathbf{m}}\leq\varphi_1(\xi_s)} e^{-\lambda_c s}\mathcal{I}_s^{\mathbf{m}}\phi(\xi_s)\Big)^{\bar{p}}\Big] \\
&\qquad\qquad \leq 1 + \int_0^t e^{-\lambda_c\bar{\sigma}_2 s}\mathbb{P}_{\phi\mu}^\phi\Big[\int_{(0,\varphi_1(\xi_s)]}\phi(\xi_s)^{\bar{\sigma}_2}y^{\bar{\sigma}_2+1}\,\Pi(\xi_s,dy)\Big]\,ds, \\
&P_{\mu,\phi}\Big[\Big(\sum_{s\in\mathcal{D}_t^{\mathbf{m}},\mathcal{I}_s^{\mathbf{m}}>\varphi_1(\xi_s)} e^{-\lambda_c s}\mathcal{I}_s^{\mathbf{m}}\phi(\xi_s)\Big)^{\bar{p}}\Big] \\
&\qquad\qquad \leq 1 + \int_0^t e^{-\lambda_c\bar{\sigma}_3 s}\mathbb{P}_{\phi\mu}^\phi\Big[\int_{(\varphi_1(\xi_s),\infty)}\phi(\xi_s)^{\bar{\sigma}_3}y^{\bar{\sigma}_3+1}\,\Pi(\xi_s,dy)\Big]\,ds,
\end{aligned}$$

where $\bar{\sigma}_i = \sigma_i - 1 \in [\bar{p}, 1]$ with σ_i defined in Assumption 3', $i \in \{1,2,3\}$. We conclude that if Assumptions (2.24)–(2.26) hold, then there exists a constant $C_1 \in (0,\infty)$ independent of μ and t such that

$$\frac{P_\mu[W_t^\phi(X)^p]}{\langle \phi, \mu\rangle} \leq \langle \phi, \mu\rangle^{\bar{p}} + C_1 \qquad \text{for all } t \geq 0, \tag{2.42}$$

and $((W_t^\phi(X))_{t\geq 0}; P_\mu)$ is an L^p-bounded martingale. Doob's inequality yields the stated L^p-convergence. $\square$

In Section 3.3 the following lemma will be used in the comparison between the immigration process and its conditional expectation.

Lemma 2.19. *Suppose Assumptions (2.24)–(2.27) hold. For every $\mu \in \mathcal{M}_c(D)$, $\mu \not\equiv 0$, there exists a time $T > 0$ and a constant $C_2 \in (0,\infty)$ such that*

$$\mathbb{P}^\phi_{\phi\mu}\big[\phi(\xi_t)^{-1} P_{\delta_{\xi_t}}[W_s^\phi(X)^p]\big] \leq C_2 \qquad \textit{for all } s \geq 0, t \geq T.$$

Proof. According to (2.42), for all $s, t \geq 0$,

$$\mathbb{P}^\phi_{\phi\mu}\big[\phi(\xi_t)^{-1} P_{\delta_{\xi_t}}[W_s^\phi(X)^p]\big] \leq \mathbb{P}^\phi_{\phi\mu}[\phi(\xi_t)^{p-1}] + C_1.$$

Since $\mu \in \mathcal{M}_c(D)$ and $\langle \phi^{p-1}, \phi\widetilde{\phi}\rangle < \infty$ by Assumption (2.27), (2.18) implies that $\mathbb{P}^\phi_{\phi\mu}[\phi(\xi_t)^{p-1}]$ converges to $\langle \phi^{p-1}, \phi\widetilde{\phi}\rangle$, and the claim follows. $\square$

CHAPTER 3

PROOFS OF THE MAIN RESULTS

In this chapter, we prove the main results stated in Section 1.3.

3.1 Reduction to a core statement

In this section, we work under Assumption 2. We first show that it suffices to consider test functions $f = \phi\mathbb{1}_B =: \phi|_B$ for Borel sets $B \in \mathcal{B}_0(D) = \{B \in \mathcal{B}(D) \colon \ell(\partial B) = 0\}$, and that we only have to prove that $\liminf_{t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle \geq \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X)$ instead of the full convergence. The proof is based on standard approximation theory combined with an idea that appeared in Lemma 9 of [5]. We denote by $C_\ell^+(D)$ the space of nonnegative, measurable, ℓ-almost everywhere continuous functions on D.

Lemma 3.1. *Let $\mu \in \mathcal{M}_f^\phi(D)$ and either $\mathbb{T} = [0,\infty)$ or $\mathbb{T} = \delta\mathbb{N}$ for some $\delta > 0$. In addition, let either $\mathcal{A} = \mathcal{B}_0(D)$ and $\mathcal{A}^\phi = \{f \in C_\ell^+(D) \colon f/\phi \in b(D)\}$, or $\mathcal{A} = \mathcal{B}(D)$ and $\mathcal{A}^\phi = \{f \in p(D) \colon f/\phi \in b(D)\}$. We define $\mathcal{A}^{\phi/w}$ like $\mathcal{A}^\phi$ where ϕ is replaced by ϕ/w.*

(i) *If for all $B \in \mathcal{A}$,*

$$\liminf_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq \langle \phi|_B, \widetilde{\phi}\rangle W_\infty^\phi(X) \qquad P_\mu\text{-almost surely,} \tag{3.1}$$

then $\lim_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle = \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X)$ P_μ-almost surely for all $f \in \mathcal{A}^\phi$.

(ii) *If for all $B \in \mathcal{A}$, $\liminf_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle \frac{\phi}{w}\mathbb{1}_B, Z_t\rangle \geq \langle \phi|_B, \widetilde{\phi}\rangle W_\infty^{\phi/w}(Z)$ $\mathbf{P}_\mu$-almost surely, then, for all $f \in \mathcal{A}^{\phi/w}$, $\lim_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle f, Z_t\rangle = \langle f, w\widetilde{\phi}\rangle W_\infty^{\phi/w}(Z)$ $\mathbf{P}_\mu$-almost surely.*

Proof. We show only Part (i); the proof of Part (ii) is similar. Let $f \in \mathcal{A}^\phi$ and write $\mathcal{S} = \{\sum_{i=1}^k c_i\phi|_{B_i} \colon k \in \mathbb{N}, c_i \in [0,\infty), B_i \in \mathcal{A}\}$. There exists a sequence of functions $f_k \in \mathcal{S}$ such that $0 \leq f_k \leq f$ and $f_k \uparrow f$ pointwise. Using (3.1) and the monotone convergence theorem, we deduce that P_μ-almost surely,

$$\liminf_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle \geq \sup_{k\in\mathbb{N}} \liminf_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle f_k, X_t\rangle \geq \sup_{k\in\mathbb{N}}\langle f_k, \widetilde{\phi}\rangle W^\phi_\infty(X) = \langle f, \widetilde{\phi}\rangle W^\phi_\infty(X).$$

Let $c = \sup_{x\in D} f(x)/\phi(x)$. Since $0 \leq c\phi - f \leq c\phi$, the same argument can be applied to $c\phi - f$, and we conclude that P_μ-almost surely

$$\begin{aligned}\limsup_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle &= \limsup_{\mathbb{T}\ni t\to\infty}\Big(cW^\phi_t(X) - e^{-\lambda_c t}\langle c\phi - f, X_t\rangle\Big)\\ &\leq cW^\phi_\infty(X) - \liminf_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle c\phi - f, X_t\rangle\\ &\leq c\langle \phi, \widetilde{\phi}\rangle W^\phi_\infty(X) - \langle c\phi - f, \widetilde{\phi}\rangle W^\phi_\infty(X) = \langle f, \widetilde{\phi}\rangle W^\phi_\infty(X). \qquad \square\end{aligned}$$

In the next step, we use the branching property of the superprocess to restrict ourselves to compactly supported starting measures.

Lemma 3.2. *Let $\mathbb{T} = [0,\infty)$ or $\mathbb{T} = \delta\mathbb{N}_0$ for some $\delta > 0$, and in addition, let $\mathcal{A}^\phi = \{f \in C^+_\ell(D) \colon f/\phi \in b(D)\}$ or $\mathcal{A}^\phi = \{f \in p(D) \colon f/\phi \in b(D)\}$.*

(i) *If for all $\mu \in \mathcal{M}_c(D)$ and $f \in \mathcal{A}^\phi$,*

$$\lim_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle = \langle f, \widetilde{\phi}\rangle W^\phi_\infty(X) \qquad P_\mu\text{-almost surely,} \tag{3.2}$$

then (3.2) holds for all $\mu \in \mathcal{M}^\phi_f(D)$.

(ii) *If convergence (3.2) holds in $L^1(P_\mu)$ for all $\mu \in \mathcal{M}_c(D)$, then it holds for all $\mu \in \mathcal{M}^\phi_f(D)$.*

Proof. Let $\mu \in \mathcal{M}^\phi_f(D)$, and take a sequence of domains $B_k \subset\subset D$, $B_k \subsetneq B_{k+1}$, with $D = \bigcup_{k=1}^\infty B_k$; $\hat{B}_k := B_k \setminus B_{k-1}$, where $B_0 := \emptyset$. On a suitable probability space, let $X^{\hat{B}_k}$, $k \in \mathbb{N}$, be independent $(L, \psi_\beta; D)$-superprocesses, where $X^{\hat{B}_k}$ is started in $\mathbb{1}_{\hat{B}_k}\mu$. By the branching property, $X^{B_k} := \sum_{l=1}^k X^{\hat{B}_l}$, $X^{D\setminus B_k} := \sum_{l=k+1}^\infty X^{\hat{B}_l}$ and $X := X^{B_k} + X^{D\setminus B_k}$ are $(L, \psi_\beta; D)$-superprocesses with starting measures $\mathbb{1}_{B_k}\mu$, $\mathbb{1}_{D\setminus B_k}\mu$ and μ, respectively. In particular,

$$W^\phi_t(X) = e^{-\lambda_c t}\langle \phi, X^{B_k}_t + X^{D\setminus B_k}_t\rangle = W^\phi_t(X^{B_k}) + W^\phi_t(X^{D\setminus B_k}),$$

and the martingale limits $W^\phi_\infty(X^{D\setminus B_k}) := \liminf_{t\to\infty} W^\phi_t(X^{D\setminus B_k})$, $k \in \mathbb{N}$, are decreasing in k. Fatou's Lemma yields

$$P_\mu\big[W^\phi_\infty(X^{D\setminus B_k})\big] = P_\mu\big[\lim_{t\to\infty} W^\phi_t(X^{D\setminus B_k})\big] \leq \liminf_{t\to\infty} P_\mu\big[W^\phi_t(X^{D\setminus B_k})\big] = \langle \phi, \mathbb{1}_{D\setminus B_k}\mu\rangle.$$

In particular, $\langle \phi, \mu\rangle < \infty$ implies that $(W^\phi_\infty(X^{D\setminus B_k}) \colon k \in \mathbb{N})$ converges to zero in $L^1(P_\mu)$ as $k \to \infty$ and, since the sequence is monotonically decreasing, almost sure convergence follows. We conclude that $\lim_{k\to\infty} W^\phi_\infty(X^{B_k}) = W^\phi_\infty(X)$ almost surely and in $L^1(P_\mu)$.

For Part (i), Lemma 3.1 (i) implies that it suffices to show

$$\liminf_{\mathbb{T}\ni t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle \geq \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X) \qquad P_\mu\text{-almost surely}$$

for all $f \in \mathcal{A}^\phi$. Since $\mathbb{1}_{B_k}\mu \in \mathcal{M}_c(D)$, the assumption implies that, for all $k \in \mathbb{N}$,

$$\liminf_{t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle \geq \liminf_{t\to\infty} e^{-\lambda_c t}\langle f, X_t^{B_k}\rangle \geq \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X^{B_k}) \qquad P_\mu\text{-almost surely},$$

and taking $k \to \infty$ yields the claim.

To show Part (ii), let $c = \sup_{x\in D} f(x)/\phi(x)$, and estimate for fixed $k \in \mathbb{N}$,

$$\begin{aligned} P_\mu\big[\big|e^{-\lambda_c t}\langle f, X_t\rangle - \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X)\big|\big] &\leq cP_\mu\big[e^{-\lambda_c t}\langle \phi, X_t^{D\setminus B_k}\rangle\big] \\ &\quad + P_\mu\big[\big|e^{-\lambda_c t}\langle f, X_t^{B_k}\rangle - \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X^{B_k})\big|\big] + \langle f, \widetilde{\phi}\rangle P_\mu\big[W_\infty^\phi(X^{D\setminus B_k})\big]. \end{aligned}$$

The second term on the right-hand side tends to zero as $t \to \infty$ by assumption. The first term is equal to $c\langle \phi, \mathbb{1}_{D\setminus D_k}\mu\rangle$ and, therefore, tends to zero as $k \to \infty$, and so does the third term. □

Let $\mathcal{M}(D)$ be the set of all σ-finite measures on D.

Remark 3.3. The superprocess X can be defined for starting measures $\mu \in \mathcal{M}(D)$ via the branching property; see also Section 1.4.4.1 in [44]. The proof of Lemma 3.2 then shows that (3.2) for all $\mu \in \mathcal{M}_c(D)$ implies (3.2) for all $\mu \in \mathcal{M}(D)$ with $\langle \phi, \mu\rangle < \infty$.

Finally, we show that it suffices to consider fixed test functions. The argument is borrowed from Chen and Shiozawa [28, Theorem 3.7].

Lemma 3.4 **(Chen and Shiozawa [28]).** *Let $\mu \in \mathcal{M}_f^\phi(D)$. If for every $B \in \mathcal{B}_0(D)$, P_μ-almost surely, $\lim_{t\to\infty} e^{-\lambda_c t}\langle \phi|_B, X_t\rangle = \langle \phi|_B, \widetilde{\phi}\rangle W_\infty^\phi(X)$, then there exists a measurable set Ω_0 such that $P_\mu(\Omega_0) = 1$ and, on Ω_0, the convergence in (1.16) holds for all $f \in C_\ell^+(D)$ with f/ϕ bounded.*

Proof. Take a countable base $(B_k)_{k\in\mathbb{N}}$ of $\mathcal{B}_0(D)$ which is closed under finite unions, and let

$$\Omega_0 = \Big\{\lim_{t\to\infty} e^{-\lambda_c t}\langle \phi|_{B_k}, X_t\rangle = \langle \phi|_{B_k}, \widetilde{\phi}\rangle W_\infty^\phi(X) \text{ for all } k \in \mathbb{N}\Big\}.$$

Then $P_\mu(\Omega_0) = 1$ by assumption. On $\{W_\infty^\phi(X) = 0\}$, convergence (1.16) trivially holds for all $f \in p(D)$ with f/ϕ bounded. On $\{W_\infty^\phi(X) > 0\} \cap \Omega_0$, we define

$$\chi_t(B) := e^{-\lambda_c t}\frac{\langle \phi|_B, X_t\rangle}{W_\infty^\phi(X)} \qquad \text{and} \qquad \chi(B) = \langle \phi|_B, \widetilde{\phi}\rangle, \qquad \text{for all } B \in \mathcal{B}(D).$$

Since $(B_k)_{k\in\mathbb{N}}$ is a base of $\mathcal{B}_0(D)$, $\liminf_{t\to\infty}\chi_t(U) \geq \chi(U)$ for all $U \in \mathcal{B}_0(D)$ open. As, in addition, $\lim_{t\to\infty}\chi_t(D) = \chi(D) = 1$ is finite, the Portmanteau theorem (cf. Theorem 13.35 in [94]) implies that χ_t converges to χ in the weak sense. For every

$f \in C_\ell^+(D)$ with f/ϕ bounded, $g := f/\phi \in bp(D)$ is ℓ-almost everywhere continuous, and since χ is absolutely continuous with respect to ℓ, $\lim_{t\to\infty}\langle g, \chi_t\rangle = \langle g, \chi\rangle$, which is equivalent to

$$\lim_{t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle = W_\infty^\phi(X)\langle f, \widetilde{\phi}\rangle \qquad \text{on } \Omega_0 \cap \{W_\infty^\phi(X) > 0\}. \qquad \square$$

3.2 Martingale limits

In this section, we prove Proposition 1.1, that is, we show that the martingale limits for the superprocess and its skeleton agree almost surely. We assume only Assumptions 1 and 2 throughout this section. Recall from (2.13) that $(S_t^*)_{t\geq 0}$ denotes the expectation semigroup of X^*.

Lemma 3.5. *Let $f \in p(D)$ with f/ϕ bounded. For all $x \in D$,*

$$\theta_t^*(x) := e^{-\lambda_c t} S_t^* f(x)/\phi(x) \to 0 \qquad \text{as } t \to \infty,$$

and for $t > 0$, the function $x \mapsto \theta_t^*(x)$ is continuous. Moreover, $\theta_t^*(x)$ is uniformly bounded in t and x, and if $f = \phi$, $\theta_t^*(x)$ is non-increasing in t.

Proof. Let $c = \sup_{x\in D} f(x)/\phi(x)$. By (2.6) and (2.15), for all $(x,t) \in D \times [0,\infty)$,

$$0 \leq \theta_t^*(x) = \mathbb{P}_x^\phi\Big[e^{\int_0^t [\beta^*(\xi_s)-\beta(\xi_s)]\,ds} f(\xi_t)/\phi(\xi_t)\Big] \tag{3.3}$$

$$\leq c\mathbb{P}_x^\phi\Big[e^{\int_0^t [\beta^*(\xi_s)-\beta(\xi_s)]\,ds}\Big] = ce^{-\lambda_c t}S_t^*\phi(x)/\phi(x). \tag{3.4}$$

Since $\beta^* - \beta \leq \mathbf{0}$, (3.4) implies that, for $f = \phi$, $\theta_t^*(x)$ is non-increasing in t. Moreover, (3.3) and Theorem 7.2.4 in [119] (see also Theorem 4.9.7 in [110]) imply that $\theta_t^*(x)$ is continuous in x for $t > 0$. The dominated convergence theorem and (3.4) yield

$$\lim_{t\to\infty} e^{-\lambda_c t} S_t^*\phi(x)/\phi(x) = \mathbb{P}_x^\phi\Big[\exp\Big(\int_0^\infty [\beta^*(\xi_s) - \beta(\xi_s)]\,ds\Big)\Big].$$

By Assumption 2, the diffusion $(\xi = (\xi_t)_{t\geq 0}; \mathbb{P}_x^\phi)$ is positive recurrent, and Theorem 4.9.5 (ii) in [110] yields

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \min\{\beta(\xi_s) - \beta^*(\xi_s), 1\}\,ds = \big\langle \min\{\beta - \beta^*, 1\}, \phi\widetilde{\phi}\big\rangle > 0 \qquad \mathbb{P}_x^\phi\text{-almost surely,}$$

where the limit is positive since $\ell(\{x \in D\colon \alpha(x) + \Pi(x,(0,\infty)) > 0\}) > 0$ by Remark 2.8. Hence, $\int_0^\infty [\beta^*(\xi_s) - \beta(\xi_s)]\,ds = -\infty$ holds $\mathbb{P}_x^\phi$-almost surely, and the claim is established. $\square$

The following lemma gives a useful bound for the $\widehat{p}$-th moment of $\langle f, I_t\rangle$. We will apply the bound to $\widehat{p} = 1$ and to $\widehat{p} = p$ with p from Assumption 3'.

Lemma 3.6. *For $\widehat{p} \geq 1$, $f \in p(D)$, $x \in D$ and $t \geq 0$,*

$$\mathbf{P}_{\bullet,\delta_x}\left[\langle f, I_t\rangle^{\widehat{p}}\right] \leq w(x)^{-1}\mathbf{P}_{\delta_x}\left[\langle f, I_t\rangle^{\widehat{p}}\right],$$

where the inequality is an equality in the case $\widehat{p} = 1$.

Proof. Using Notation 2.5, $I_t = \sum_{i=1}^{N_0} I_t^{i,0}$, where under $\mathbf{P}_{\delta_x}$, N_0 is Poisson-distributed with mean $w(x)$ and $\boldsymbol{\mathcal{F}}_0$-measurable, and $(I_t^{i,0}; \mathbf{P}_{\delta_x}(\cdot|\boldsymbol{\mathcal{F}}_0))$ is equal in distribution to $(I_t; \mathbf{P}_{\bullet,\delta_x})$. Using the monotonicity of the $\ell^{\widehat{p}}$-norm, we derive

$$\begin{aligned}\mathbf{P}_{\delta_x}\left[\langle f, I_t\rangle^{\widehat{p}}\right] &= \mathbf{P}_{\delta_x}\Big[\Big(\sum_{i=1}^{N_0}\langle f, I_t^{i,0}\rangle\Big)^{\widehat{p}}\Big] \\ &\geq \mathbf{P}_{\delta_x}\Big[\sum_{i=1}^{N_0}\langle f, I_t^{i,0}\rangle^{\widehat{p}}\Big] = \mathbf{P}_{\delta_x}\left[N_0\mathbf{P}_{\bullet,\delta_x}[\langle f, I_t\rangle^{\widehat{p}}]\right] = w(x)\mathbf{P}_{\bullet,\delta_x}\left[\langle f, I_t\rangle^{\widehat{p}}\right].\end{aligned}$$

For $\widehat{p} = 1$ the inequality is an equality. Rearranging terms completes the proof. □

We now come to the main part of this section. First, we employ the skeleton decomposition to compute the conditional expectation of $\langle f, X_{s+t}\rangle$.

Proposition 3.7. *For all $\mu \in \mathcal{M}_f^{\phi}(D)$, $f \in p(D)$ with f/ϕ bounded, and $s, t \geq 0$,*

$$\mathbf{P}_\mu\left[\langle f, X_{s+t}\rangle|\boldsymbol{\mathcal{F}}_t\right] = \langle S_s^* f, X_t\rangle + \Big\langle \frac{S_s f}{w}, Z_t\Big\rangle - \Big\langle \frac{S_s^* f}{w}, Z_t\Big\rangle \qquad \mathbf{P}_\mu\text{-almost surely.}$$

Proof. By Notation 2.5, $\langle f, X_{s+t}\rangle = \langle f, X_{s+t}^* + I_s^{*,t}\rangle + \sum_{i=1}^{N_t}\langle f, I_s^{i,t}\rangle$, where the random variable $(X_{s+t}^* + I_s^{*,t}; \mathbf{P}_\mu(\cdot|\boldsymbol{\mathcal{F}}_t))$ is equal in distribution to $(X_s^*; \mathbf{P}_{X_t})$ and $(I_s^{i,t}; \mathbf{P}_\mu(\cdot|\boldsymbol{\mathcal{F}}_t))$ to $(I_s; \mathbf{P}_{\bullet,\delta_{\xi_i(t)}})$, $i = 1, \ldots, N_t$. Hence, $\mathbf{P}_\mu$-almost surely,

$$\begin{aligned}\mathbf{P}_\mu[\langle f, X_{s+t}\rangle|\boldsymbol{\mathcal{F}}_t] &= \mathbf{P}_\mu\Big[\langle f, X_{s+t}^* + I_s^{*,t}\rangle + \sum_{i=1}^{N_t}\langle f, I_s^{i,t}\rangle\Big|\boldsymbol{\mathcal{F}}_t\Big] \\ &= \mathbf{P}_{X_t}[\langle f, X_s^*\rangle] + \sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]. \end{aligned} \tag{3.5}$$

The first term on the right can be rewritten using (2.13) to obtain $\mathbf{P}_{X_t}[\langle f, X_s^*\rangle] = \langle S_s^* f, X_t\rangle$. For the second term, we use Lemma 3.6 and Theorem 2.3 (iii.c) to derive

$$\sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle] = \sum_{i=1}^{N_t} w(\xi_i(t))^{-1}\mathbf{P}_{\delta_{\xi_i(t)}}[\langle f, X_s - X_s^*\rangle]. \tag{3.6}$$

Since f/ϕ is bounded and $\mu \in \mathcal{M}_f^{\phi}(D)$, $\langle \frac{S_s f}{w}, Z_t\rangle$ is finite $\mathbf{P}_\mu$-almost surely. Hence, (3.5), (3.6), (2.11) and (2.13) yield $\mathbf{P}_\mu$-almost surely,

$$\mathbf{P}_\mu\Big[\sum_{i=1}^{N_t}\langle f, I_s^{i,t}\rangle\Big|\boldsymbol{\mathcal{F}}_t\Big] = \sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle] = \Big\langle \frac{S_s f}{w}, Z_t\Big\rangle - \Big\langle \frac{S_s^* f}{w}, Z_t\Big\rangle \tag{3.7}$$

as required. $\square$

Proof of Proposition 1.1 and Theorem 2.9 (i). Proposition 3.7 yields,

$$\begin{aligned}&\mathbf{P}_\mu\big[W^\phi_{s+t}(X)|\boldsymbol{\mathcal{F}}_t\big]\\&\quad=e^{-\lambda_c t}\langle e^{-\lambda_c s}S^*_s\phi,X_t\rangle+e^{-\lambda_c t}\Big\langle e^{-\lambda_c s}\frac{S_s\phi}{w},Z_t\Big\rangle-e^{-\lambda_c t}\Big\langle e^{-\lambda_c s}\frac{S^*_s\phi}{w},Z_t\Big\rangle\end{aligned}\tag{3.8}$$

$\mathbf{P}_\mu$-almost surely, and we are interested in the limit as $s\to\infty$. The first and last term tend to zero $\mathbf{P}_\mu$-almost surely by Lemma 3.5 and the dominated convergence theorem. The second term is independent of s since $e^{-\lambda_c s}S_s\phi=\phi$. Hence, the right-hand side of (3.8) converges to $W^{\phi/w}_t(Z)$. According to Theorem 2.15, $((W^\phi_t(X))_{t\geq 0};\mathbf{P}_\mu)$ is an L^p-bounded martingale, and we can interchange, on the left-hand side of (3.8), the limit $s\to\infty$ with the integration to obtain

$$\mathbf{P}_\mu\big[W^\phi_\infty(X)|\boldsymbol{\mathcal{F}}_t\big]=\lim_{s\to\infty}\mathbf{P}_\mu\big[W^\phi_{s+t}(X)|\boldsymbol{\mathcal{F}}_t\big]=W^{\phi/w}_t(Z)\qquad\mathbf{P}_\mu\text{-almost surely.}\tag{3.9}$$

Since $W^\phi_\infty(X)$ is measurable with respect to $\boldsymbol{\mathcal{F}}_\infty=\sigma(\bigcup_{t\geq 0}\boldsymbol{\mathcal{F}}_t)$, (1.15) follows by taking $t\to\infty$ in (3.9). Moreover, (3.9) shows that $(W^{\phi/w}_t(Z))_{t\geq 0}$ is a uniformly integrable martingale, and since $W^\phi_\infty(X)=W^{\phi/w}_\infty(Z)$ is in $L^p(\mathbf{P}_\mu)$, L^p-boundedness of $(W^{\phi/w}_t(Z))_{t\geq 0}$ follows by Jensen's inequality. $\square$

3.3 Convergence in $L^1(P_\mu)$

In this section, we prove the WLLN in the form of Theorem 1.5 or Theorem 2.9 (ii). We suppose that Assumptions 1, 2 and (2.24)–(2.27) hold and begin with an L^p-estimate for the immigration that occurred after a large time t. Recall Notations 2.4 and 2.5.

Proposition 3.8. *For every $\mu\in\mathcal{M}_c(D)$ and $f\in p(D)$ with f/ϕ bounded, there exists a time $T>0$ and a constant $C_3\in(0,\infty)$ such that for all $s\geq 0$, $t\geq T$,*

$$e^{-\lambda_c p(s+t)}\mathbf{P}_\mu\Big[\Big|\sum_{i=1}^{N_t}\Big(\langle f,I^{i,t}_s\rangle-\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f,I_s\rangle]\Big)\Big|^p\Big]\leq C_3e^{-\lambda_c(p-1)t}.$$

Proof. For $\mu\equiv 0$ the claim is trivial. Let $\mu\not\equiv 0$. It was noted in [16, Lemma 1] that, for $p\in[1,2]$, $n\in\mathbb{N}$ and $(Y_i\colon i\in\{1,\dots,n\})$ independent, centered random variables (or martingale differences),

$$P\Big[\Big|\sum_{i=1}^n Y_i\Big|^p\Big]\leq 2^p\sum_{i=1}^n P\big[|Y_i|^p\big].$$

For $s,t\geq 0$, we apply this inequality to $\mathbf{P}_\mu[\cdot|\boldsymbol{\mathcal{F}}_t]$, $n=N_t$ and $Y_i=\langle f,I^{i,t}_s\rangle-$

$\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]$ to bound

$$\mathbf{P}_\mu\Big[\Big|\sum_{i=1}^{N_t}\Big(\langle f, I_s^{i,t}\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]\Big)\Big|^p\Big|\boldsymbol{\mathcal{F}}_t\Big] \le 2^p\sum_{i=1}^{N_t}\mathbf{P}_\mu\big[\big|\langle f, I_s^{i,t}\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]\big|^p\big|\boldsymbol{\mathcal{F}}_t\big].$$

Using $|x-y|^p \le x^p + y^p$ for $x, y \ge 0$, $(I_s^{i,t}; \mathbf{P}_\mu(\cdot|\boldsymbol{\mathcal{F}}_t)) \overset{d}{=} (I_s; \mathbf{P}_{\bullet,\delta_{\xi_i(t)}})$ and Jensen's inequality, we can bound the right-hand side by

$$2^p\sum_{i=1}^{N_t}\Big(\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle^p] + \mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]^p\Big) \le 2^{p+1}\sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle^p].$$

Now Lemma 3.6, the identity $X_s = X_s^* + I_s$ under $\mathbf{P}_{\delta_{\xi_i(t)}}$, and the monotonicity of $x \mapsto x^p$ on $[0,\infty)$ yield

$$\begin{aligned}\mathbf{P}_\mu\Big[\Big|\sum_{i=1}^{N_t}\Big(\langle f, I_s^{i,t}\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]\Big)\Big|^p\Big|\boldsymbol{\mathcal{F}}_t\Big] &\le 2^{p+1}\sum_{i=1}^{N_t}\frac{\mathbf{P}_{\delta_{\xi_i(t)}}[\langle f, X_s\rangle^p]}{w(\xi_i(t))}\\ &= 2^{p+1}\Big\langle\frac{\mathbf{P}_{\delta_\cdot}[\langle f, X_s\rangle^p]}{w}, Z_t\Big\rangle.\end{aligned}$$

Writing $c = \sup_{x\in D} f(x)/\phi(x) < \infty$, (2.22) and (2.23) yield

$$\begin{aligned}&e^{-\lambda_c p(s+t)}\mathbf{P}_\mu\Big[\Big|\sum_{i=1}^{N_t}\Big(\langle f, I_s^{i,t}\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]\Big)\Big|^p\Big]\\ &\qquad\le 2^{p+1}c^p e^{-\lambda_c pt}\mathbf{P}_\mu\Big[\Big\langle\frac{\mathbf{P}_{\delta_\cdot}[W_s^\phi(X)^p]}{w}, Z_t\Big\rangle\Big]\\ &\qquad= 2^{p+1}c^p e^{-\lambda_c(p-1)t}\langle\phi,\mu\rangle\mathbb{P}^\phi_{\phi\mu}\big[\phi(\xi_t)^{-1}\mathbf{P}_{\delta_{\xi_t}}[W_s^\phi(X)^p]\big].\end{aligned}$$

Since $\mu \in \mathcal{M}_c(D)$, Lemma 2.19 yields a time $T > 0$ and a constant $C_2 \in (0,\infty)$ such that the right-hand side is bounded by $2^{p+1}c^p e^{-\lambda_c(p-1)t}\langle\phi,\mu\rangle C_2$ for all $s \ge 0$, $t \ge T$, and the proof is complete. □

We are now in the position to prove Theorems 1.5 and 2.9 (ii).

Proof of Theorems 1.5 and 2.9 (ii). According to Lemma 3.2 (ii), it suffices to consider $\mu \in \mathcal{M}_c(D)$, and without loss of generality, we work on the skeleton space. Using the skeleton decomposition in the form of (2.10), we write for $s, t \ge 0$,

$$\begin{aligned}&e^{-\lambda_c(s+t)}\langle f, X_{s+t}\rangle - \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X)\\ &= e^{-\lambda_c(s+t)}\langle f, X_{s+t}^* + I_s^{*,t}\rangle + e^{-\lambda_c(s+t)}\sum_{i=1}^{N_t}\Big(\langle f, I_s^{i,t}\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]\Big)\\ &+ \Big(e^{-\lambda_c(s+t)}\sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle] - \langle f, \widetilde{\phi}\rangle W_t^{\phi/w}(Z)\Big) + \langle f, \widetilde{\phi}\rangle\big(W_t^{\phi/w}(Z) - W_\infty^\phi(X)\big)\\ &=: \Xi_1(s,t) + \Xi_2(s,t) + \Xi_3(s,t) + \Xi_4(s,t).\end{aligned}$$

It suffices to show that

$$\limsup_{s\to\infty}\limsup_{t\to\infty}\mathbf{P}_\mu[|\Xi_i(s,t)|]=0 \qquad \text{for all } i\in\{1,\dots,4\}. \tag{3.10}$$

Verification for $i=1$: since $(X^*_{s+t}+I^{*,t}_s;\mathbf{P}_\mu(\cdot|\boldsymbol{\mathcal{F}}_t))\overset{d}{=}(X^*_s;\mathbf{P}_{X_t})$, the first moment formulas (2.13) and (2.11), and (2.16) yield

$$\mathbf{P}_\mu[|\Xi_1(s,t)|]=e^{-\lambda_c(s+t)}\mathbf{P}_\mu[\mathbf{P}_{X_t}[\langle f,X^*_s\rangle]]=e^{-\lambda_c(s+t)}\langle S_tS^*_sf,\mu\rangle=\langle\mathbb{P}^\phi_\cdot[\theta^*_s(\xi_t)],\phi\mu\rangle,$$

where $\theta^*_s(x)=e^{-\lambda_cs}S^*_sf(x)/\phi(x)$. By Lemma 3.5, $\theta^*_s(x)$ is uniformly bounded in s and x, and converges to zero as $s\to\infty$. Using the ergodicity of $(\xi;\mathbb{P}^\phi)$, cf. (2.18), and the dominated convergence theorem, we conclude

$$\lim_{s\to\infty}\lim_{t\to\infty}\mathbf{P}_\mu[|\Xi_1(s,t)|]=\lim_{s\to\infty}\langle\theta^*_s,\phi\widetilde{\phi}\rangle\langle\phi,\mu\rangle=0.$$

Verification for $i=2$: Proposition 3.8 implies that $\Xi_2(s,t)$ converges to zero in $L^p(\mathbf{P}_\mu)$ as $t\to\infty$ for every fixed $s>0$. By monotonicity of norms, (3.10) for $i=2$ follows.

Verification for $i=3$: we use (3.7) and (2.16) to rewrite

$$\begin{aligned}e^{-\lambda_c(s+t)}\sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f,I_s\rangle]&=e^{-\lambda_c(s+t)}\Big\langle\frac{S_sf-S^*_sf}{w},Z_t\Big\rangle\\&=e^{-\lambda_ct}\Big\langle\mathbb{P}^\phi_\cdot\big[f(\xi_s)/\phi(\xi_s)\big]-\theta^*_s,\frac{\phi}{w}Z_t\Big\rangle.\end{aligned}$$

Let $\Upsilon_s(x):=\mathbb{P}^\phi_x[f(\xi_s)/\phi(\xi_s)]-\theta^*_s(x)$. Since f/ϕ is bounded, Υ is uniformly bounded in s and x, and by (2.18) and Lemma 3.5, $\lim_{s\to\infty}\Upsilon_s(x)=\langle f,\widetilde{\phi}\rangle$. Moreover, $\Xi_3(s,t)=e^{-\lambda_ct}\langle\Upsilon_s-\langle f,\widetilde{\phi}\rangle,\frac{\phi}{w}Z_t\rangle$ by the definition of $W^{\phi/w}_t(Z)$. The many-to-one lemma for Z, i.e. (2.22), yields

$$\mathbf{P}_\mu[|\Xi_3(s,t)|]\le e^{-\lambda_ct}\mathbf{P}_\mu\Big[\Big\langle\big|\Upsilon_s-\langle f,\widetilde{\phi}\rangle\big|,\frac{\phi}{w}Z_t\Big\rangle\Big]=\langle\mathbb{P}^\phi_\cdot\big[\big|\Upsilon_s(\xi_t)-\langle f,\widetilde{\phi}\rangle\big|\big],\phi\mu\rangle.$$

Since Υ_s is bounded and $\phi(x)\mu(dx)$ is a finite measure, (2.18) implies that the right-hand side converges to $\langle|\Upsilon_s-\langle f,\widetilde{\phi}\rangle|,\phi\widetilde{\phi}\rangle\langle\phi,\mu\rangle$ as $t\to\infty$, and this expression converges to zero as $s\to\infty$ by the dominated convergence theorem.

Verification for $i=4$: since $(W^{\phi/w}_t(Z))_{t\ge0}$ is an $L^p(\mathbf{P}_\mu)$-bounded martingale by Theorem 2.9 (i), it converges to $W^{\phi/w}_\infty(Z)=W^\phi_\infty(X)$ in $L^1(\mathbf{P}_\mu)$. Hence, (3.10) for $i=4$ holds. □

3.4 Asymptotics for the immigration process and the SLLN along lattice times

In this section, we analyse the asymptotic behaviour of the immigration process I. Lemma 3.1 (i) and Theorem 2.3 imply that, instead of $e^{-\lambda_c t}\langle f, X_t\rangle$, we can study $e^{-\lambda_c t}\langle f, I_t\rangle$ for the proof of Theorem 1.2 if the latter converges to the correct limit. To show this, we first study the conditional expectation of the immigration after a large time t as stated in (3.7). From now on, we work under Assumptions 1, 2, 3' and 4.

Lemma 3.9. *For all $\mu \in \mathcal{M}_c(D)$, $s, \delta > 0$, and for all $f \in p(D)$ with f/ϕ bounded,*

$$\lim_{\delta\mathbb{N}\ni t\to\infty} e^{-\lambda_c t}\sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle] = e^{\lambda_c s}\Big(\langle f,\widetilde{\phi}\rangle - \langle e^{-\lambda_c s}S_s^* f,\widetilde{\phi}\rangle\Big)W_\infty^\phi(X)$$

$\mathbf{P}_\mu$*-almost surely.*

Proof. We apply the SLLN for the skeleton (Assumption 4) to the functions f_1, f_2 given by

$$f_1(x) := \frac{S_s f(x)}{w(x)} = e^{\lambda_c s}\mathbb{P}_x^\phi\big[f(\xi_s)/\phi(\xi_s)\big]\frac{\phi(x)}{w(x)}, \qquad f_2(x) := \frac{S_s^* f(x)}{w(x)} = e^{\lambda_c s}\theta_s^*(x)\frac{\phi(x)}{w(x)}.$$

By Lemmas 2.1 and 3.5 and Theorem 4.9.7 in [110], f_1 and f_2 are continuous for any $s > 0$. Hence, (3.7) and Assumption 4 yield $\mathbf{P}_\mu$-almost surely,

$$\lim_{\delta\mathbb{N}\ni t\to\infty} e^{-\lambda_c t}\sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle] = \Big\langle\frac{S_s f}{w}, w\widetilde{\phi}\Big\rangle W_\infty^{\phi/w}(Z) - \Big\langle\frac{S_s^* f}{w}, w\widetilde{\phi}\Big\rangle W_\infty^{\phi/w}(Z).$$

By (2.16) and (2.17), $\langle S_s f,\widetilde{\phi}\rangle = e^{\lambda_c s}\langle f,\widetilde{\phi}\rangle$, and Theorem 2.9 (i) completes the proof. □

Proposition 3.10. *For all $\mu \in \mathcal{M}_c(D)$, $\delta > 0$, and for all $f \in p(D)$ with f/ϕ bounded,*

$$\lim_{\delta\mathbb{N}\ni s\to\infty}\lim_{\delta\mathbb{N}\ni t\to\infty} e^{-\lambda_c(s+t)}\sum_{i=1}^{N_t}\mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle] = \langle f,\widetilde{\phi}\rangle W_\infty^\phi(X) \qquad \mathbf{P}_\mu\text{-almost surely.}$$

Proof. The claim follows immediately from Lemmas 3.9 and 3.5 and the dominated convergence theorem. □

Recall from Notation 2.5 that, given $\boldsymbol{\mathcal{F}}_t$, $I^{i,t}$ denotes the immigration occurring along the skeleton descending from particle i at time t.

Proposition 3.11. *For all $\mu \in \mathcal{M}_c(D)$, $s, \delta > 0$, and all $f \in p(D)$ with f/ϕ bounded,*

$$\lim_{\delta\mathbb{N}\ni t\to\infty} e^{-\lambda_c(s+t)}\left|\sum_{i=1}^{N_t}\Big(\langle f, I_s^{i,t}\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle f, I_s\rangle]\Big)\right| = 0 \qquad \mathbf{P}_\mu\text{-almost surely.}$$

Proof. By the Borel–Cantelli lemma, it is sufficient to show that for all $\epsilon > 0$ there is a $n_0 \in \mathbb{N}$ such that

$$\sum_{n=n_0}^{\infty} \mathbf{P}_\mu\Big(e^{-\lambda_c(s+n\delta)}\Big|\sum_{i=1}^{N_{n\delta}}\Big(\langle f, I_s^{i,n\delta}\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(n\delta)}}[\langle f, I_s\rangle]\Big)\Big| > \epsilon\Big) < \infty. \tag{3.11}$$

To bound the left-hand side in (3.11), we first use Markov's inequality, and then Proposition 3.8 to obtain a time $T > 0$ and a constant $C_3 \in (0,\infty)$ such that, for $n_0 \geq T$, an upper bound is given by

$$\epsilon^{-p}\sum_{n=n_0}^{\infty} e^{-\lambda_c p(s+n\delta)}\mathbf{P}_\mu\Big[\Big|\sum_{i=1}^{N_{n\delta}}\Big(\langle f, I_s^{i,n\delta}\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(n\delta)}}[\langle f, I_s\rangle]\Big)\Big|^p\Big]$$
$$\leq \epsilon^{-p} C_3 \sum_{n=n_0}^{\infty} e^{-\lambda_c(p-1)n\delta} < \infty. \qquad \square$$

Combining Propositions 3.10 and 3.11, the asymptotic behaviour of the immigration process can be characterised as follows:

Corollary 3.12. *For all $\mu \in \mathcal{M}_c(D)$, $\delta > 0$, and all $f \in p(D)$ with f/ϕ bounded,*

$$\lim_{\delta\mathbb{N}\ni s\to\infty}\lim_{\delta\mathbb{N}\ni t\to\infty} e^{-\lambda_c(s+t)}\sum_{i=1}^{N_t}\langle f, I_s^{i,t}\rangle = \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X) \qquad \mathbf{P}_\mu\text{-almost surely.}$$

Now we are in the position to prove the SLLN along lattice times.

Theorem 3.13. *For all $\mu \in \mathcal{M}_f^\phi(D)$, $\delta > 0$, and all $f \in p(D)$ with f/ϕ bounded,*

$$\lim_{n\to\infty} e^{-\lambda_c n\delta}\langle f, X_{n\delta}\rangle = \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X) \qquad P_\mu\text{-almost surely.}$$

Proof. By Lemma 3.2 (i), it suffices to consider $\mu \in \mathcal{M}_c(D)$. Moreover, without loss of generality, we work on the skeleton space from Theorem 2.3. Corollary 3.12 yields $\mathbf{P}_\mu$-almost everywhere,

$$\liminf_{\delta\mathbb{N}\ni t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle = \liminf_{\delta\mathbb{N}\ni s\to\infty}\liminf_{\delta\mathbb{N}\ni t\to\infty} e^{-\lambda_c(s+t)}\langle f, X_{s+t}\rangle$$
$$\geq \liminf_{\delta\mathbb{N}\ni s\to\infty}\liminf_{\delta\mathbb{N}\ni t\to\infty} e^{-\lambda_c(s+t)}\sum_{i=1}^{N_t}\langle f, I_s^{i,t}\rangle = \langle f, \widetilde{\phi}\rangle W_\infty^\phi(X).$$

Lemma 3.1 (i) completes the proof. $\square$

3.5 Transition from lattice to continuous times

In this section, we extend the convergence along lattice times in Theorem 3.13 to convergence along continuous times and conclude our main results. We work under

Assumptions 1, 2, 3' and 4. For $\kappa > 0$, let U^κ be the resolvent operator in integral form, that is,

$$U^\kappa f(x) := \int_0^\infty e^{-\kappa t}\mathbb{P}_x^\phi[f(\xi_t)]\,dt \qquad \text{for all } f \in bp(D), x \in D.$$

The argument for the transition from lattice to continuous times proceeds in two steps. First we use the resolvent operator to bring the semigroup of $(\xi;(\mathbb{P}_x^\phi)_{x\in D})$ into the argument. The semigroup property gives us a martingale which, combined with Doob's L^p-inequality, enables us to control the behaviour between times $n\delta$ and $(n+1)\delta$. Second, we remove the resolvent operator by taking $\kappa \to \infty$ in $\kappa U^\kappa f(x)$. It is an analysis of hitting times for diffusion processes which allows us to control the convergences in this step.

The main idea for the proof is borrowed from [102], but we employ the skeleton decomposition to replace the stochastic analysis and the martingale measures used there.

Proposition 3.14. *For all $\mu \in \mathcal{M}_c(D)$, $\kappa > 0$ and $f \in bp(D)$,*

$$\lim_{t\to\infty} e^{-\lambda_c t}\langle \phi\kappa U^\kappa f, X_t\rangle = \langle \phi f, \widetilde{\phi}\rangle W_\infty^\phi(X) \qquad P_\mu\text{-almost surely.}$$

Proof. Without loss of generality, we assume that $\mu \not\equiv 0$ and work on the skeleton space. Since κU^κ is linear with $\kappa U^\kappa \mathbf{1} = \mathbf{1}$, the same argument that led to Lemma 3.1 shows that it suffices to prove that, for all $f \in bp(D)$,

$$\liminf_{t\to\infty} e^{-\lambda_c t}\langle \phi\kappa U^\kappa f, X_t\rangle \geq \langle \phi f, \widetilde{\phi}\rangle W_\infty^\phi(X) \qquad \mathbf{P}_\mu\text{-almost surely.} \tag{3.12}$$

Let $f, g \in bp(D)$ with $\kappa U^\kappa f \geq g$, $\delta, t > 0$, and let n be such that $n\delta \leq t < (n+1)\delta$. Then

$$\begin{aligned} e^{-\lambda_c t}\langle \phi\kappa U^\kappa f, X_t\rangle &\geq \Big(e^{-\lambda_c t}\langle \phi\kappa U^\kappa f, X_t\rangle - e^{-\lambda_c t}\langle \phi\mathbb{P}_\cdot^\phi\big[\kappa U^\kappa f(\xi_{(n+1)\delta-t})\big], X_t\rangle\Big) \\ &\quad + \Big(e^{-\lambda_c t}\langle \phi\mathbb{P}_\cdot^\phi[g(\xi_{(n+1)\delta-t})], X_t\rangle - e^{-\lambda_c n\delta}\langle \phi\mathbb{P}_\cdot^\phi[g(\xi_\delta)], X_{n\delta}\rangle\Big) \\ &\quad + \Big(e^{-\lambda_c n\delta}\langle \phi\mathbb{P}_\cdot^\phi[g(\xi_\delta)], X_{n\delta}\rangle - \langle \phi g, \widetilde{\phi}\rangle W_\infty^\phi(X)\Big) + \langle \phi g, \widetilde{\phi}\rangle W_\infty^\phi(X) \\ &=: \Theta_{1,\kappa U^\kappa f}(n,\delta,t) + \Theta_{2,g}(n,\delta,t) + \Theta_{3,g}(n,\delta) + \langle \phi g, \widetilde{\phi}\rangle W_\infty^\phi(X). \end{aligned} \tag{3.13}$$

If we show, for all $f, g \in bp(D)$, g of compact support, that $\mathbf{P}_\mu$-almost surely,

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\sup_{t\in[n\delta,(n+1)\delta]} |\Theta_{1,\kappa U^\kappa f}(n,\delta,t)| = 0, \tag{3.14}$$

$$\limsup_{n\to\infty}\sup_{t\in[n\delta,(n+1)\delta]} |\Theta_{2,g}(n,\delta,t)| = 0 \qquad \text{for all } \delta > 0, \tag{3.15}$$

$$\limsup_{n\to\infty} |\Theta_{3,g}(n,\delta)| = 0 \qquad \text{for all } \delta > 0, \tag{3.16}$$

then we can choose $g = \mathbb{1}_B \kappa U^\kappa f$ for $B \subset\subset D$ in (3.13) to obtain

$$\liminf_{t\to\infty} e^{-\lambda_c t}\langle \phi\kappa U^\kappa f, X_t\rangle \geq \langle \phi \mathbb{1}_B \kappa U^\kappa f, \widetilde{\phi}\rangle W^\phi_\infty(X) \qquad \mathbf{P}_\mu\text{-almost surely.}$$

Choosing a sequence $B_j \subset\subset D$, $B_j \subseteq B_{j+1}$, $D = \bigcup_{j=1}^\infty B_j$, the factor $\langle \phi \mathbb{1}_{B_j}\kappa U^\kappa f, \widetilde{\phi}\rangle$ increases, as $j \to \infty$, to

$$\langle \phi\kappa U^\kappa f, \widetilde{\phi}\rangle = \int_0^\infty \kappa e^{-\kappa t}\langle \phi \mathbb{P}^\phi_\cdot[f(\xi_t)], \widetilde{\phi}\rangle\, dt \overset{(2.17)}{=} \int_0^\infty \kappa e^{-\kappa t}\langle \phi f, \widetilde{\phi}\rangle\, dt = \langle \phi f, \widetilde{\phi}\rangle,$$

and (3.12) follows. It remains to verify (3.14)–(3.16).

Verification of (3.14): Fubini's theorem and the Markov property of $(\xi; \mathbb{P}^\phi)$ yield, for all $x \in D$ and $s > 0$,

$$\begin{aligned}\left|\kappa U^\kappa f(x) - \mathbb{P}^\phi_x[\kappa U^\kappa f(\xi_s)]\right| &= \left|\int_0^\infty \kappa e^{-\kappa t}\mathbb{P}^\phi_x[f(\xi_t)]\, dt - \int_0^\infty \kappa e^{-\kappa t}\mathbb{P}^\phi_x[f(\xi_{t+s})]\, dt\right| \\ &\leq 2(1-e^{-\kappa s})\|f\|_\infty.\end{aligned}$$

Using the linearity of integration and the definition of $W^\phi_t(X)$, we obtain

$$\sup_{t\in[n\delta,(n+1)\delta]}\left|\Theta_{1,\kappa U^\kappa f}(n,\delta,t)\right| \leq 2(1-e^{-\kappa\delta})\|f\|_\infty \sup_{t\in[n\delta,(n+1)\delta]} W^\phi_t(X).$$

Since the martingale $(W^\phi_t(X))_{t\geq 0}$ has a finite limit, (3.14) is established.

Verification of (3.15): Let $g \in bp(D)$ be compactly supported. By (2.16),

$$\Theta_{2,g}(n,\delta,t) = e^{-\lambda_c(n+1)\delta}\Big(\langle S_{(n+1)\delta-t}[\phi g], X_t\rangle - \langle S_\delta[\phi g], X_{n\delta}\rangle\Big) \qquad \text{for all } t \in [0,(n+1)\delta].$$

The Markov property of X, and (2.11), imply that $(\Theta_{2,g}(n,\delta,t)\colon t \in [n\delta,(n+1)\delta]; \mathbf{P}_\mu)$ is a martingale. Hence, (3.15) follows from the Borel–Cantelli lemma, Doob's L^p-inequality (cf. Theorem II.1.7 in [113]) and $\mathbf{P}_\mu[\langle \phi g, X_{(n+1)\delta}\rangle | \boldsymbol{\mathcal{F}}_{n\delta}] = \langle S_\delta[\phi g], X_{n\delta}\rangle$ when we prove that for sufficiently large n_0,

$$\sum_{n=n_0}^\infty e^{-\lambda_c p(n+1)\delta}\mathbf{P}_\mu\Big[\big|\langle \phi g, X_{(n+1)\delta}\rangle - \mathbf{P}_\mu[\langle \phi g, X_{(n+1)\delta}\rangle|\boldsymbol{\mathcal{F}}_{n\delta}]\big|^p\Big] < \infty. \tag{3.17}$$

By (2.10) and (3.5), we have for all $s,t>0$, $\mathbf{P}_\mu$-almost surely, given $\boldsymbol{\mathcal{F}}_t$,

$$\begin{aligned}&\langle \phi g, X_{s+t}\rangle - \mathbf{P}_\mu[\langle \phi g, X_{s+t}\rangle|\boldsymbol{\mathcal{F}}_t] \\ &\quad = \langle \phi g, X^*_{s+t} + I^{*,t}_s\rangle - \mathbf{P}_{X_t}[\langle \phi g, X^*_s\rangle] + \sum_{i=1}^{N_t}\Big(\langle \phi g, I^{i,t}_s\rangle - \mathbf{P}_{\bullet,\delta_{\xi_i(t)}}[\langle \phi g, I_s\rangle]\Big).\end{aligned} \tag{3.18}$$

The monotonicity of L^p-norms and $(X^*_{s+t} + I^{*,t}_s; \mathbf{P}_\mu(\cdot|\boldsymbol{\mathcal{F}}_t)) \overset{d}{=} (X^*_s; \mathbf{P}_{X_t})$ imply

$$\mathbf{P}_\mu\big[|\langle \phi g, X^*_{s+t} + I^{*,t}_s\rangle - \mathbf{P}_{X_t}[\langle \phi g, X^*_s\rangle]|^p\big] \leq \mathbf{P}_\mu\big[\mathbf{Var}_{X_t}(\langle \phi g, X^*_s\rangle)\big]^{p/2}. \tag{3.19}$$

Recall the definitions of X^*, β^*, π^*, and $S^*_t f$ from page 23. Denote $c^*_1(x) = 2\alpha(x)$, $c^*_2 = \int_{(0,\varphi_2(x)]} y^2\,\Pi^*(x,dy)$, $c^*_3(x) = \int_{(\varphi_2(x),\infty)} y^2\,\Pi^*(x,dy)$, $c^*(x) = \sum_{i=1}^3 c^*_i(x)$ for all $x \in D$, where φ_2 is determined by Assumption 3'. We notice that $\beta^* \le \beta$ implies that $S^*_t f \le S_t f$ for all $f \in p(D)$. Using (2.14), (2.11), and the semigroup property of S, we obtain

$$\begin{aligned}\mathbf{P}_\mu\big[\mathbf{Var}_{X_t}(\langle \phi g, X^*_s\rangle)\big] &= \int_0^s \langle S_t S^*_r\big[c^*(S^*_{s-r}[\phi g])^2\big], \mu\rangle\, dr\\ &\le \int_0^s \langle S_{t+r}\big[c^*(S_{s-r}[\phi g])^2\big], \mu\rangle\, dr.\end{aligned}$$

Recall the definition of $\mathbb{P}^\phi_{\phi\mu}$ from (2.23), and use (2.16) to deduce

$$\mathbf{P}_\mu\big[\mathbf{Var}_{X_t}(\langle \phi g, X^*_s\rangle)\big] \le \langle \phi, \mu\rangle \|g\|_\infty \int_0^s e^{\lambda_c(s+t)}\mathbb{P}^\phi_{\phi\mu}\big[c^*(\xi_{t+r})S_{s-r}[\phi g](\xi_{t+r})\big]\, dr. \quad (3.20)$$

Writing $\bar\beta = \max\{\sup_{x\in D}\beta(x), 0\}$, we notice that $S_{s-r}[\phi g](x) \le e^{\bar\beta s}\|\phi g\|_\infty$. Further, (2.16) implies $S_{s-r}[\phi g](x) \le e^{\lambda_c s}\|g\|_\infty \phi(x)$. Hence, for all $i \in \{1,2,3\}$,

$$\begin{aligned}&\mathbb{P}^\phi_{\phi\mu}\big[c^*_i(\xi_{t+r})S_{s-r}[\phi g](\xi_{t+r})\big]\\ &\quad\le \min\Big\{e^{\bar\beta s}\|\phi g\|_\infty \mathbb{P}^\phi_{\phi\mu}[c^*_i(\xi_{t+r})], e^{\lambda_c s}\|g\|_\infty \mathbb{P}^\phi_{\phi\mu}\big[c^*_i(\xi_{t+r})\phi(\xi_{t+r})\big]\Big\}.\end{aligned} \quad (3.21)$$

The right-hand side of (3.21) is bounded for large t as we now explain: for $i = 1$, boundedness of α, and therefore c^*_1, entails the assertion. For $i \in \{2,3\}$, (2.18) and Conditions (2.28) and (2.29), respectively, yield the claim. Combining (3.19)–(3.21), we obtain a time $T > 0$ and a constant $C \in (0,\infty)$, which may depend on s, g and μ, such that

$$e^{-\lambda_c p(s+t)}\mathbf{P}_\mu\big[\big|\langle \phi g, X^*_{s+t} + I^{*,t}_s\rangle - \mathbf{P}_{X_t}[\langle \phi g, X^*_s\rangle]\big|^p\big] \le Ce^{-\lambda_c pt/2} \qquad \text{for all } t \ge T. \quad (3.22)$$

Since $|x+y|^p \le 2^p(|x|^p + |y|^p)$ for all $x, y \in \mathbb{R}$, (3.18), (3.22) and Proposition 3.8 yield (3.17).

Verification of (3.16)**:** since $\langle \phi \mathbb{P}^\phi_\cdot[g(\xi_\delta)], \widetilde\phi\rangle = \langle \phi g, \widetilde\phi\rangle$ by (2.17), (3.16) follows from Theorem 3.13. □

In the second step, we remove the resolvent operator from Proposition 3.14. The proof is essentially borrowed from the lower bound in Theorem 2.2 in [102]. We present the argument here for completeness. Recall that $\mathcal{B}_0(D) = \{B \in \mathcal{B}(D) \colon \ell(\partial B) = 0\}$ and $\phi|_B = \phi \mathbb{1}_B$.

Proposition 3.15. *For all $\mu \in \mathcal{M}_c(D)$ and $B \in \mathcal{B}_0(D)$,*

$$\liminf_{t\to\infty} e^{-\lambda_c t}\langle \phi|_B, X_t\rangle \ge \langle \phi|_B, \widetilde\phi\rangle W^\phi_\infty(X) \qquad P_\mu\text{-almost surely.} \quad (3.23)$$

Proof. The claim is trivial when $\ell(B) = 0$. When (3.23) is proved for $B \in \mathcal{B}_0(D)$ with

$B \subset\subset D$, then, for arbitrary $B \in \mathcal{B}_0(D)$, we choose a sequence of sets $B_k \in \mathcal{B}_0(D)$, with $B_k \subset\subset D$, $B_k \subseteq B_{k+1}$ and $B = \bigcup_{k\in\mathbb{N}} B_k$, and the monotone convergence theorem yields, P_μ-almost surely,

$$\liminf_{t\to\infty} e^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq \sup_{k\in\mathbb{N}} \liminf_{t\to\infty} e^{-\lambda_c t}\langle \phi|_{B_k}, X_t\rangle \geq \langle \phi|_B, \widetilde{\phi}\rangle W_\infty^\phi(X).$$

Hence, let $B \in \mathcal{B}_0(D)$, $B \subset\subset D$, contain a non-empty, open ball. For small $\epsilon > 0$, let $B_\epsilon = \{x \in B\colon \operatorname{dist}(x, \partial B) \geq \epsilon\} \neq \emptyset$ and denote by $\sigma_{B_\epsilon} = \inf\{t > 0\colon \xi_t \in B_\epsilon\}$ the hitting time of B_ϵ. We write $U^\kappa(x, A) = U^\kappa \mathbb{1}_A(x)$ for all $A \in \mathcal{B}(D)$. Since $\{\xi_t \in B_\epsilon\} \subseteq \{\sigma_{B_\epsilon} \leq t\}$, for all $x \in D$,

$$\kappa U^\kappa(x, B_\epsilon) \leq \int_0^\infty \kappa e^{-\kappa t}\mathbb{P}_x^\phi(\sigma_{B_\epsilon} \leq t)\, dt = \mathbb{P}_x^\phi[e^{-\kappa\sigma_{B_\epsilon}}] \leq \mathbb{1}_B(x) + \mathbb{1}_{B^c}(x)\mathbb{P}_x^\phi[e^{-\kappa\sigma_{B_\epsilon}}],$$

where $B^c := D \setminus B$. In particular,

$$e^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq e^{-\lambda_c t}\langle \phi\kappa U^\kappa \mathbb{1}_{B_\epsilon}, X_t\rangle - e^{-\lambda_c t}\langle \phi|_{B^c}\mathbb{P}_\cdot^\phi[e^{-\kappa\sigma_{B_\epsilon}}], X_t\rangle,$$

and Proposition 3.14 yields, P_μ-almost surely,

$$\liminf_{t\to\infty} e^{-\lambda_c t}\langle \phi|_B, X_t\rangle \geq \langle \phi|_{B_\epsilon}, \widetilde{\phi}\rangle W_\infty^\phi(X) - \limsup_{t\to\infty} e^{-\lambda_c t}\langle \phi|_{B^c}\mathbb{P}_\cdot^\phi[e^{-\kappa\sigma_{B_\epsilon}}], X_t\rangle. \tag{3.24}$$

The first term on the right converges to $\langle \phi|_B, \widetilde{\phi}\rangle W_\infty^\phi(X)$ as $\epsilon \to 0$. Thus, we have to show that the second term vanishes as $\epsilon \to 0$. Heuristically, if the SLLN holds, then the lim sup is a limit with value

$$\langle \phi|_{B^c}\mathbb{P}_\cdot^\phi[e^{-\kappa\sigma_{B_\epsilon}}], \widetilde{\phi}\rangle W_\infty^\phi(X).$$

Since B_ϵ has positive distance to B^c, this value converges to zero as $\kappa \to \infty$. Hence, we first take $\kappa \to \infty$ and then $\epsilon \to 0$. Of course, we do not know the SLLN, yet. Thus, we artificially reintroduce the resolvent operator in order to apply Proposition 3.14.

Continuing rigorously, let $B'_\epsilon := \{x \in B\colon \operatorname{dist}(x, \partial B_\epsilon) \leq \epsilon/2\}$. The situation is sketched in Figure I-1.

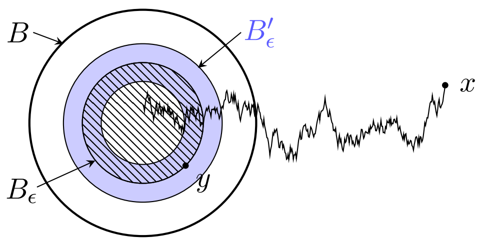

Figure I-1. The big ball with thick boundary is B, the small, hatched ball is B_ϵ and the shaded area denotes B'_ϵ. The diffusion is started in $x \in B^c$.

When ξ starts outside B, then $\xi_{\sigma_{B_\epsilon}} \in \partial B_\epsilon$, and we obtain for all $x \in B^c$,

$$\begin{aligned}\mathbb{P}_x^\phi[e^{-\kappa\sigma_{B_\epsilon}}] &= \mathbb{P}_x^\phi\Big[e^{-\kappa\sigma_{B_\epsilon}}\frac{U^\kappa(\xi_{\sigma_{B_\epsilon}}, B'_\epsilon)}{U^\kappa(\xi_{\sigma_{B_\epsilon}}, B'_\epsilon)}\Big] \\ &\le \frac{1}{\inf_{y\in\partial B_\epsilon} U^\kappa(y, B'_\epsilon)}\mathbb{P}_x^\phi\big[e^{-\kappa\sigma_{B_\epsilon}}U^\kappa(\xi_{\sigma_{B_\epsilon}}, B'_\epsilon)\big]. \end{aligned} \tag{3.25}$$

For $t \ge 0$, let $\mathcal{H}_t := \sigma(\xi_s : 0 \le s \le t)$. By the Markov property of ξ, the second factor on the right-hand side of (3.25) can be estimated by

$$\begin{aligned}\mathbb{P}_x^\phi\big[e^{-\kappa\sigma_{B_\epsilon}}U^\kappa(\xi_{\sigma_{B_\epsilon}}, B'_\epsilon)\big] &= \mathbb{P}_x^\phi\Big[e^{-\kappa\sigma_{B_\epsilon}}\mathbb{P}_x^\phi\Big[\int_0^\infty e^{-\kappa t}\mathbb{1}\{\xi_{t+\sigma_{B_\epsilon}} \in B'_\epsilon\}\,dt\Big|\mathcal{H}_{\sigma_{B_\epsilon}}\Big]\Big] \\ &= \mathbb{P}_x^\phi\Big[\int_{\sigma_{B_\epsilon}}^\infty e^{-\kappa t}\mathbb{1}\{\xi_t \in B'_\epsilon\}\,dt\Big] \le U^\kappa(x, B'_\epsilon).\end{aligned} \tag{3.26}$$

Writing $\Phi(\kappa,\epsilon) := \inf_{y\in\partial B_\epsilon}\kappa U^\kappa(y, B'_\epsilon)$, (3.25) and (3.26) yield $\mathbb{1}_{B^c}(x)\mathbb{P}_x^\phi[e^{-\kappa\sigma_{B_\epsilon}}] \le \kappa U^\kappa(x, B'_\epsilon)/\Phi(\kappa,\epsilon)$, and Proposition 3.14 entails, P_μ-almost surely,

$$\limsup_{t\to\infty} e^{-\lambda_c t}\langle \phi|_{B^c}\mathbb{P}_\cdot^\phi[e^{-\kappa\sigma_{B_\epsilon}}], X_t\rangle \le \frac{1}{\Phi(\kappa,\epsilon)}\langle \phi|_{B'_\epsilon}, \widetilde{\phi}\rangle W_\infty^\phi(X). \tag{3.27}$$

Clearly, $\langle \phi|_{B'_\epsilon}, \widetilde{\phi}\rangle W_\infty^\phi(X)$ converges to zero as $\epsilon \to 0$. Thus, it remains to bound $\kappa U^\kappa(y, B'_\epsilon)$ for $y \in \partial B_\epsilon$, and therefore $\Phi(\kappa,\epsilon)$, away from zero. We write $b_0(x)$ for the vector whose j-th component is given by $b_j(x) + \frac{1}{2}\sum_{i=1}^d \partial_{x_i}a_{i,j}(x)$, $j \in \{1,\dots,d\}$, $x \in D$. Since $B \subset\subset D$,

$$\begin{aligned} c(B,\phi) &:= \frac{\inf_{x\in B}\phi(x)}{\sup_{x\in B}\phi(x)}, & \tilde{\beta} &:= \sup_{x\in B}|\beta(x)|, \\ \tilde{b}_0 &:= \sup_{x\in B}|b_0(x)|, & \tilde{a} &:= \sup_{x\in B}\sup_{|v|=1} v^T a(x) v, \end{aligned}$$

satisfy $c(B,\phi), \tilde{a} \in (0,\infty)$ and $\tilde{\beta}, \tilde{b}_0 \in [0,\infty)$. For all $T > 0$,

$$\begin{aligned}\kappa U^\kappa(y, B'_\epsilon) &= \int_0^\infty \kappa e^{-\kappa t}\mathbb{P}_y^\phi(\xi_t \in B'_\epsilon)\,dt \\ &= \int_0^\infty e^{-t}\mathbb{P}_y^\phi(\xi_{t/\kappa} \in B'_\epsilon)\,dt \ge \int_0^T e^{-t}\mathbb{P}_y^\phi(\xi_{t/\kappa} \in B'_\epsilon)\,dt.\end{aligned} \tag{3.28}$$

For $y \in \partial B_\epsilon$, use the definition of B'_ϵ and (2.15) to estimate

$$\begin{aligned}\mathbb{P}_y^\phi(\xi_{t/\kappa} \in B'_\epsilon) &\ge \mathbb{P}_y^\phi\Big(\sup_{0\le s\le t/\kappa}|\xi_s - y| \le \epsilon/2\Big) \\ &\ge c(B,\phi)e^{-(\lambda_c+\tilde{\beta})t/\kappa}\mathbb{P}_y\Big(\sup_{0\le s\le t/\kappa}|\xi_s - y| \le \epsilon/2\Big).\end{aligned} \tag{3.29}$$

To estimate the probability on the right-hand side, we use Theorem 2.2.2 in [110]. Since this theorem is stated for a diffusion generator in non-divergence form, we introduced the function b_0. In particular, for κ so large that $t\tilde{b}_0/\kappa \leq \epsilon/4$ and for all $y \in \partial B_\epsilon$, we deduce

$$\mathbb{P}_y\Big(\sup_{0\leq s\leq t/\kappa} |\xi_s - y| \leq \epsilon/2\Big) \geq 1 - 2d\exp\Big(-\frac{\epsilon^2\kappa}{32\tilde{a}td}\Big). \tag{3.30}$$

Combining (3.28)–(3.30), we obtain for all $\epsilon > 0$,

$$\liminf_{\kappa\to\infty} \Phi(\kappa,\epsilon) \geq \int_0^T e^{-t}c(B,\phi)\,dt > 0.$$

Since the right-hand side does not depend on ϵ, taking first $\kappa \to \infty$ and then $\epsilon \to 0$ in (3.27) and (3.24) completes the proof. □

We are now in the position to conclude our main results.

Proof of Theorem 2.9 (iii). The P_μ-almost sure convergence in (2.30) for every given ℓ-almost everywhere continuous $f \in p(D)$ with f/ϕ bounded follows from Proposition 3.15 and Lemmas 3.1 (i) and 3.2 (i). The existence of a common set Ω_0 for all such test functions follows from Lemma 3.4. □

Proof of Theorem 1.2. The $L^1(P_\mu)$-convergence in (1.16) was proved in Theorem 1.5, the remainder follows from Theorem 2.9 (iii). □

Proof of Corollary 1.4. The claim follows immediately from Theorem 1.2, (2.20) and (2.18). □

CHAPTER 4

EXAMPLES

In this chapter, we explore our assumptions by verifying them for many classical examples of superdiffusions from the literature. Moreover, we give several examples to illustrate the implications and boundaries of the SLLN. For all examples considered, this thesis proves the SLLN, and for some even the WLLN, for the first time.

4.1 Spatially independent branching mechanisms

In this section, we consider superdiffusions with a conservative motion and a spatially independent branching mechanism and write $\psi(z) = \psi_\beta(x, z)$ to simplify notation. Under these conditions, the total mass process $(\langle \mathbf{1}, X_t\rangle : t \geq 0)$ is a continuous state branching process (CSBP) with branching mechanism ψ; cf. [118, 11]. We exclude the trivial case of a linear branching mechanism $\psi(z) = -\beta z$ (see Remark 2.8 for the result in this situation). Since ψ is strictly convex, $\psi(\infty) := \lim_{z\to\infty} \psi(z)$ exists in $[-\infty, 0) \cup \{\infty\}$. Writing $z^* = \sup\{z \geq 0 : \psi(z) \leq 0\}$, we have $z^* \in (0, \infty)$ if and only if $\beta > 0$ and $\psi(\infty) = \infty$, and in that case (cf. Proposition 1.1 in [118]),

$$P_\mu\big[e^{-z^*\langle \mathbf{1}, X_t\rangle}\big] = e^{-z^*\langle \mathbf{1}, \mu\rangle} \qquad \text{for all } \mu \in \mathcal{M}_f(D), t \geq 0.$$

In particular, Assumption 1 is satisfied with $w(x) = z^*$ for all $x \in D$. In this CSBP context, the skeleton decomposition was proved by Berestycki et al. in [11] a few years before [98]. The martingale function z^* is related to the event of weak extinction $\mathcal{E}_{\text{lim}} = \{\lim_{t\to\infty}\langle \mathbf{1}, X_t\rangle = 0\}$ by the identity $P_{\delta_x}(\mathcal{E}_{\text{lim}}) = e^{-z^*}$ which holds even if $\beta \leq 0$ or $\psi(\infty) < 0$. To compare the martingale function $w(x) = -\log P_{\delta_x}(\mathcal{E}_{\text{lim}})$ to the classical choice $w(x) = -\log P_{\delta_x}(\mathcal{E}_{\text{fin}})$, where $\mathcal{E}_{\text{fin}}$ denotes the event of extinction after finite time, notice that $\mathcal{E}_{\text{fin}} \subseteq \mathcal{E}_{\text{lim}}$, and for all $\mu \in \mathcal{M}_f(D)$,

$$P_\mu(\mathcal{E}_{\text{fin}}) = P_\mu(\mathcal{E}_{\text{lim}}) = e^{-z^*\langle \mathbf{1}, \mu\rangle} > 0 \qquad \text{if } \psi(\infty) = \infty \text{ and } \int^\infty \frac{1}{\psi(z)}\, dz < \infty. \quad (4.1)$$

Otherwise, $P_\mu(\mathcal{E}_{\text{fin}}) = 0$, and on $\mathcal{E}_{\text{lim}}$, the total mass of X drifts to zero while staying positive at all finite times; cf. [78, 118].

From now on, assume $\beta > 0$, $\psi(\infty) = \infty$ and $w(x) = z^*$. In this case, Assumption 3 simplifies to

$$\phi \text{ bounded} \quad \text{and} \quad \int_{(1,\infty)} y^p \, \Pi(dy) < \infty \quad \text{for some } p \in (1, 2].$$

In the following, we present two families of superprocesses for which the SLLN is proved by Theorem 1.2. As far as we know, these results are new. Apart from the intrinsic interest, the results are very useful since the analysed processes are frequently employed to obtain further examples of superprocesses with interesting properties via h-transform. For those examples the SLLN follows from Lemma 2.10.

We begin with the inward Ornstein–Uhlenbeck process (OU-process) which has attracted a wide interest in the literature. Specifically, its asymptotic behaviour is the subject of recent research articles [104, 112].

Example 4.1 (Inward OU-process)**.** Let $d \geq 1$, $D = \mathbb{R}^d$, $L = \frac{1}{2}\Delta - \gamma x \cdot \nabla$ with $\gamma > 0$, ψ spatially independent with $\beta \in (0, \infty)$, $\psi(\infty) = \infty$ and $\int_{(1,\infty)} y^p \, \Pi(dy) < \infty$ for some $p \in (1, 2]$. Then Theorem 1.2 applies with $\phi = \mathbf{1}$, $\widetilde{\phi}(x) = (\gamma/\pi)^{d/2} e^{-\gamma \|x\|^2}$ and $\lambda_c = \beta$.

The generator L corresponds to the positive recurrent inward OU-process with transition density

$$p_{\text{in-OU}}(x, y, t) = \left(\frac{\gamma}{\pi(1 - e^{-2\gamma t})} \right)^{d/2} \exp\left(- \frac{\gamma}{1 - e^{-2\gamma t}} \|y - e^{-\gamma t} x\|^2 \right) \tag{4.2}$$

for all $x, y \in \mathbb{R}^d, t > 0$. Hence, $\lambda_c = \lambda_c(L + \beta) = \beta > 0$, L is product L^1-critical, $\phi = \mathbf{1}$ and $\widetilde{\phi}(x) = (\gamma/\pi)^{d/2} e^{-\gamma \|x\|^2}$ (see Chapter 4 in [110] or Example 3 in [111]). Thus, Assumptions 1–3 are satisfied. Using the estimate for $p^\phi = p_{\text{in-OU}}$ in (2.19), we obtain that Condition (2.35) holds for $a(t) = \sqrt{(\lambda_c/\gamma + \delta)t}$ with $\delta > 0$ (see Example 10 in [57]), and using (4.2), we deduce that (2.36) holds with $K = 1$. Hence, Theorem 2.13 applies, and Assumption 4 is satisfied.

Example 4.2 (Outward OU-process)**.** Let $d \geq 1$, $D = \mathbb{R}^d$, $L = \frac{1}{2}\Delta + \gamma x \cdot \nabla$ with $\gamma > 0$, ψ spatially independent with $\beta \in (\gamma d, \infty)$, $\psi(\infty) = \infty$ and $\int_{(1,\infty)} y^p \, \Pi(dy) < \infty$ for some $p \in (1, 2]$. Then Theorem 1.2 applies with $\phi(x) = (\gamma/\pi)^{d/2} e^{-\gamma \|x\|^2}$, $\widetilde{\phi} = \mathbf{1}$ and $\lambda_c = \beta - \gamma d$.

The generator L corresponds to the conservative, transient outward OU-process. The operator $L_1 := L + \gamma d$ is the formal adjoint of the inward OU-process with parameter γ. Hence, L_1 is critical with ground states $\phi_1(x) = (\gamma/\pi)^{d/2} e^{-\gamma \|x\|^2}$ and $\widetilde{\phi}_1 = \mathbf{1}$ by Example 4.1 (see Theorem 4.3.3 in [110] or Example 2 in [111]). Writing $L_1 = L + \beta - (\beta - \gamma d)$, we deduce that Assumptions 1–3 hold and $\phi, \widetilde{\phi}$ and λ_c have been correctly identified. The corresponding ergodic motion is the inward OU-process with parameter γ. Thus, Conditions (2.35) and (2.36) can be verified using (4.2), (2.19), $a(t) = e^{\gamma(1+\delta)t}$ for some $\delta > 0$ and $K > 1 + \delta$, and Theorem 2.13 implies that Assumption 4 holds.

The SLLN describes the asymptotic behaviour of the mass in compact sets. In general, one cannot draw conclusions for the scaling of the total mass from the local behaviour [58, 60]. Example 4.2 illustrates this fact. Since the total mass process is a CSBP with branching mechanism ψ, $Y_t = e^{-\beta t}\langle \mathbf{1}, X_t\rangle$ converges to a finite random variable Y_∞ with $P_\mu(Y_\infty = 0) = P_\mu(\mathcal{E}_{\text{lim}})$ if $\beta > 0$ and $\int_{(1,\infty)} y \log y \, \Pi(dy) < \infty$; cf. [78]. In particular, in Example 4.2, the local growth rate $\lambda_c = \beta - \gamma d$ is strictly smaller than the global growth rate β. The reason is the transient nature of the underlying diffusion which allows mass to leave compact sets permanently and is reflected in the decay of ϕ at infinity. In particular, the function $\mathbf{1}$ is not an allowed test function in Theorem 1.2, but the focus is on functions of the form $\mathbb{1}_B$ for B a compact set.

We call the diffusion corresponding to the generator $L = \frac{1}{2}\nabla \cdot a\nabla + b \cdot \nabla$ *symmetric* if $b = a\nabla Q$ for some $Q \in C^{2,\eta}(D)$. The inward and outward OU-processes constitute examples of symmetric diffusions with $Q(x) = -\frac{\gamma}{2}\|x\|^2$ and $Q(x) = \frac{\gamma}{2}\|x\|^2$, respectively. Chen et al. [27] studied superdiffusions with a symmetric motion but insisted that Q is bounded. Hence, their results are not applicable to Examples 4.1 and 4.2. The result from Liu et al. [102] is not applicable since the domain is not of finite Lebesgue measure.

Engländer and Winter [62] proved convergence in probability in (1.16) for the situation of a quadratic branching mechanism. It is straightforward to extend their argument to general branching mechanisms, but the method requires second moments. Hence, if $\int_{(1,\infty)} y^p \, \Pi(dy) < \infty$ for some $p \in (1, 2)$ but not for $p = 2$, then even the convergence in probability in Examples 4.1 and 4.2 is new.

4.2 Quadratic branching mechanisms

In this section, we consider the classical situation of a quadratic branching mechanism studied by Engländer, Pinsky and Winter [59, 62] and Chen, Ren and Wang [27]. Our assumptions on the branching mechanism in this section are $\alpha, \beta \in C^\eta(D)$, $\alpha(x) > 0$ for all $x \in D$, $\lambda_c := \lambda_c(L + \beta) < \infty$ and $\Pi \equiv 0$. We write $\psi(x, z) = -\beta(x)z + \alpha(x)z^2$ and call ψ a *generalised quadratic branching mechanism* (GQBM). In Section 1.1 we insisted that α and β are bounded. This assumption can be relaxed as follows. First suppose that β is bounded from above but not necessarily from below. Engländer and Pinsky [59] showed that there is a unique $\mathcal{M}_f(D)$-valued Markov process $X = (X_t)_{t\geq 0}$ such that

$$P_\mu[e^{-\langle f, X_t\rangle}] = e^{-\langle u_f(\cdot,t), \mu\rangle} \qquad \text{for all continuous } f \in bp(D) \text{ and all } \mu \in \mathcal{M}_f(D),$$

where u_f is the minimal, nonnegative solution $u \in C(D \times [0,\infty))$, $(x,t) \mapsto u(x,t)$ twice continuously differentiable in $x \in D$ and once in $t \in (0,\infty)$, to

$$\begin{aligned} \partial_t u(x,t) &= Lu(x,t) - \psi(x,u(x,t)) && \text{for all } (x,t) \in D \times (0,\infty), \\ u(x,0) &= f(x) && \text{for all } x \in D. \end{aligned} \tag{4.3}$$

If α and β are bounded, the minimal solution of (4.3) equals the unique solution to (1.3) by Lemma A1 in [59]. Hence, the two definitions are consistent.

Now let $\beta \in C^\eta(D)$ with $\lambda_c = \lambda_c(L+\beta) < \infty$ be not necessarily bounded from above. By definition (1.9), there exists $\lambda \in \mathbb{R}$ and $h \in C^{2,\eta}(D)$, $h > 0$, such that $(L+\beta)h = \lambda h$. Recall the definition of h-transforms from Section 2.1.3. An $(L,\psi;D)$-superprocess can be defined by $X = \frac{1}{h}X^h$, where X^h is the $(L_0^h, \psi^h; D)$-superprocess with $\beta^h = \lambda$ and $\alpha^h = \alpha h$; cf. [59]. Since h is not necessarily bounded from below, the process X may take values in the space of σ-finite measures $\mathcal{M}(D)$. While we have considered mainly finite measure-valued processes in this thesis, it is natural to consider also processes with values in the space $\mathcal{M}(D)$ via the branching property, and, as noted in Remark 3.3, in our results the space of starting measures $\mathcal{M}_f^\phi(D)$ can be enlarged to the space of all $\mu \in \mathcal{M}(D)$ with $\langle \phi, \mu \rangle < \infty$.

Engländer and Pinsky [59] proved the skeleton decomposition for supercritical superdiffusions with GQBMs long before [98]. We only record the existence of a martingale function in the following lemma. Recall that $\mathcal{E}_{\text{fin}}$ denotes the event of extinction after a finite time.

Lemma 4.3 **(Engländer and Pinsky [59]).** *Let ψ be a GQBM and $\lambda_c > 0$. The function $x \mapsto w(x) := -\log P_{\delta_x}(\mathcal{E}_{\text{fin}})$ is strictly positive, belongs to $C^{2,\eta}(D)$ and satisfies (1.6).*

Proof. By Theorem 3.1 and Corollary 4.2 in [59], $w \in C^{2,\eta}(D)$, $w(x) > 0$ for all $x \in D$ and

$$P_\mu(\mathcal{E}_{\text{fin}}) = e^{-\langle w,\mu\rangle} \qquad \text{for all } \mu \in \mathcal{M}_c(D). \tag{4.4}$$

Recall the notation from the beginning of Section 2.1.1. Let $B \subset\subset D$ be a domain, and $\mu \in \mathcal{M}_f(D)$ with $\operatorname{supp}(\mu) \subseteq B$. Then the support of the exit measure $\widetilde{X}_t^B$ is P_μ-almost surely compact (see the discussion following (2.3)). Since $\mathcal{E}_{\text{fin}}$ is a tail event, the Markov property and (4.4) yield $P_\mu[e^{-\langle \widetilde{w}, \widetilde{X}_t^B\rangle}] = e^{-\langle w,\mu\rangle}$. Choose a sequence of functions $w_j \in C_c^+(D)$ with $w_j \uparrow w$ pointwise, and a sequence of domains $B_k \subset\subset D$, $B_k \subseteq B_{k+1}$, $D = \bigcup_{k=1}^\infty B_k$. By Lemma A1 in [59], $\widetilde{u}_{w_j}^{B_k}$ is increasing in j. Moreover, Lemma B.5 below shows that, for fixed j and sufficiently large k, $\widetilde{u}_{w_j}^{B_k}$ is increasing in k with $\lim_{k\to\infty} \widetilde{u}_{w_j}^{B_k} = u_{w_j}$ pointwise. It follows that, for all $\mu \in \mathcal{M}_c(D)$,

$$\begin{aligned} P_\mu[e^{-\langle w, X_t\rangle}] &= \lim_{j\to\infty}\lim_{k\to\infty} P_\mu[e^{-\langle \widetilde{w}_j, \widetilde{X}_t^{B_k}\rangle}] \\ &= \lim_{k\to\infty}\lim_{j\to\infty} P_\mu[e^{-\langle \widetilde{w}_j, \widetilde{X}_t^{B_k}\rangle}] = \lim_{k\to\infty} P_\mu[e^{-\langle \widetilde{w}, \widetilde{X}_t^{B_k}\rangle}] = e^{-\langle w,\mu\rangle}. \qquad \square \end{aligned}$$

In the remainder of this section, we choose w to be $w(x) = -\log P_{\delta_x}(\mathcal{E}_{\text{fin}})$, and let $Z = (Z_t)_{t\geq 0}$ be a strictly dyadic branching particle diffusion, where the spatial motion is defined by (1.7) and the branching rate is given by $q = \alpha w$ (in accordance with (1.8)).

One advantage of allowing unbounded α and β is that the setup is now invariant under h-transforms: for any $h \in C^{2,\eta}(D)$, $h > 0$, ψ^h is a GQBM. Moreover,

$$\begin{aligned} w^h(x) &:= -\log P^h_{\delta_x}(\exists t \geq 0\colon \langle \mathbf{1}, X^h_t\rangle = 0) \\ &= -\log P_{h(x)^{-1}\delta_x}(\exists t \geq 0\colon \langle h, X_t\rangle = 0) = w(x)/h(x), \end{aligned} \tag{4.5}$$

and Lemmas 2.10 and 2.11 remain valid for GQBMs and $\mathbb{H}(\psi) = \{h \in C^{2,\eta}(D)\colon h > 0\}$. We record the following result.

Theorem 4.4. *Let ψ be a GQBM, and suppose Assumption 2 holds and $\phi\alpha$ is bounded. Let $\mu \in \mathcal{M}^\phi_f(D)$.*

(i) *For all $f \in p(D)$ with f/ϕ bounded, the convergence in (1.16) holds in $L^1(P_\mu)$.*

(ii) *If, in addition, Assumption 4 holds, then there exists a measurable set Ω_0 with $P_\mu(\Omega_0) = 1$, and on Ω_0, the convergence in (1.16) holds for all ℓ-almost everywhere continuous $f \in p(D)$ with f/ϕ bounded.*

Proof. Let X^ϕ be an $(L^\phi_0, \psi^\phi; D)$-superprocess. Since $\beta^\phi = \lambda_c$ and $\alpha^\phi = \phi\alpha$ are bounded, X^ϕ satisfies the assumptions of Section 1.1. Moreover, X^ϕ satisfies Assumption 1 by Lemma 4.3, Assumption 2 with $\phi^\phi = \mathbf{1}$ by Lemma 2.10 (i), and Assumption 3". Hence, Theorem 2.12 (i) and Lemma 2.10 (ii) yield the first part of the claim. Lemmas 2.11, 2.10 (ii) and 3.4, and Theorem 2.12 (ii) yield the second part. $\square$

The h-transforms are one way to relate two superprocesses to each other; another is monotonicity.

Lemma 4.5. *Let ψ_β and $\hat{\psi}_{\hat{\beta}}$ be two branching mechanisms as defined in Section 1.1 with $\psi_\beta \geq \hat{\psi}_{\hat{\beta}}$. Let X and $\hat{X}$ be $(L, \psi_\beta; D)$- and $(L, \hat{\psi}_{\hat{\beta}}; D)$-superprocesses, respectively.*

(i) *For all $\mu \in \mathcal{M}_f(D)$, $f \in bp(D)$, $t \geq 0$, $P_\mu[e^{-\langle f, X_t\rangle}] \geq P_\mu[e^{-\langle f, \hat{X}_t\rangle}]$.*

(ii) *Let $w(x) = -\log P_{\delta_x}(\exists t \geq 0\colon \langle \mathbf{1}, X_t\rangle = 0)$ and $\hat{w}(x) = -\log P_{\delta_x}(\exists t \geq 0\colon \langle \mathbf{1}, \hat{X}_t\rangle = 0)$ for all $x \in D$. Then $w \leq \hat{w}$.*

Proof. Part (i) is proved in Appendix B below. Part (ii) follows from Part (i) and the identity $w(x) = \lim_{t\to\infty} \lim_{\theta\to\infty} -\log P_{\delta_x}[e^{-\theta\langle \mathbf{1}, X_t\rangle}]$. $\square$

We saw in Example 4.2 that the SLLN describes the asymptotics of the mass in compact sets, not necessarily the global growth. A second distinction between the local and global behaviour can be observed on the event $\{W^\phi_\infty(X) = 0\} \setminus \mathcal{E}_{\text{fin}}$. Engländer and Turaev [61, Problem 14] raised the question whether this event can have positive probability. Suppose Assumption 2 holds. Engländer [53] observed that if $\lim_{t\to\infty} e^{-\lambda_c t}\langle f, X_t\rangle = \langle f, \widetilde{\phi}\rangle W^{\phi}_{\infty}(X)$ in distribution for all $f \in C_c^+(D)$, $\mu \in \mathcal{M}_c(D)$, and if the support of X is transient, then

$$P_\mu(W^{\phi}_{\infty}(X) = 0) > P_\mu(\mathcal{E}_{\text{fin}}) \qquad \text{for all } \mu \in \mathcal{M}_c(D), \mu \not\equiv 0. \tag{4.6}$$

Here the support of $(X; P_\mu)$ is *recurrent* if

$$P_\mu(X_t(B) > 0 \text{ for some } t \geq 0 \,|\, \mathcal{E}^c_{\text{fin}}) = 1 \qquad \text{for every open } B \subseteq D, B \neq \emptyset,$$

and *transient* otherwise. See [59] for a detailed discussion of recurrence and transience of the support of superdiffusions.

We study three examples in this section. In the first example, α and β are bounded but w is unbounded. In the second example α is bounded, but β, ϕ and w are unbounded. Both examples are based on a recurrent motion but while the support of the superprocess is recurrent in the second, it is transient in the first example. The third example considers a large class of processes containing super-Brownian motion with compactly supported growth rate β and instances of non-symmetric underlying motions.

The domain for all these examples is $D = \mathbb{R}^d$, and therefore, none of them is covered in Liu et al.'s [102] article. Chen et al.'s [27] article is not applicable to the first two examples since they are based on the inward-OU process as underlying motion (because, as in Section 4.1, Q is unbounded) and not to the third because the motion is non-symmetric (for some processes in the considered class), and the variance parameter α is unbounded, whereas [27] requires α to be bounded.

The motivation for the first example comes from Example 5.1 in [59].

Example 4.6. Let $d \geq 1$, $D = \mathbb{R}^d$, $L = \frac{1}{2}\Delta - \gamma x \cdot \nabla$ with $\gamma > 0$, $\beta \in (0, \infty)$ constant, $\alpha(x) = e^{-\gamma\|x\|^2}$, $\Pi \equiv 0$. Then Theorem 1.2 applies with $\phi = \mathbf{1}$, $\widetilde{\phi}(x) = (\gamma/\pi)^{d/2}e^{-\gamma\|x\|^2}$ and $\lambda_c = \beta$. Moreover, $w(x) = (\beta + \gamma d)e^{\gamma\|x\|^2}$, the support of X is transient, and (4.6) holds.

There are two ways to prove (1.16) for this example. First, we perform an h-transform with $h(x) = (\gamma/\pi)^{-d/2}e^{\gamma\|x\|^2}$ to obtain

$$L_0^h = \frac{1}{2}\Delta + \gamma x \cdot \nabla, \quad \beta^h = \beta + \gamma d, \quad \alpha^h = (\pi/\gamma)^{d/2}.$$

In Example 4.2, we showed that Theorem 1.2 applies to the $(L_0^h, \psi^h; \mathbb{R}^d)$-superprocess. The $(L, \psi; \mathbb{R}^d)$-superprocess can be recovered by an h-transform with $h_2 = 1/h$, and Lemma 2.10 yields that Assumption 2 is satisfied with the stated ϕ, $\widetilde{\phi}$ and λ_c, and that (1.16) holds. Alternatively, we can deduce (1.16) by a direct application of Theorems 1.2 and 2.13. Assumption 1 holds by Lemma 4.3, and Assumption 3 holds since α and ϕ are bounded. To verify Assumption 4, we notice that $w^h(x) = \beta^h/\alpha^h = (\beta + \gamma d)(\gamma/\pi)^{d/2}$

by (4.1), and (4.5) yields

$$w(x) = w^h(x)h(x) = (\beta + \gamma d)e^{\gamma\|x\|^2} \qquad \text{for all } x \in \mathbb{R}^d.$$

Thus, w is not bounded from above. The verification of (2.35) and (2.36) is the same as in Example 4.2 since w/ϕ is of the same order and the ergodic motion is the same. Hence, the conditions hold with $a(t) = e^{\gamma(1+\delta)t}$ for some $\delta > 0$ and $K > 1 + \delta$.

To see that the support of X is transient, notice that the support is invariant under h-transforms, and the support of the $(L_0^h; \psi^h; \mathbb{R}^d)$-superprocess is transient by Theorem 4.6 in [59] and Example 2 in [111].

Example 4.6 should be compared to Example 4.2 for a quadratic branching mechanism. In both examples, the support of the superprocess is transient, and the event $\{W_\infty^\phi(X) = 0\} \setminus \mathcal{E}_{\text{fin}}$ has positive probability. Hence, in both examples, mass can escape to infinity which is reflected in the SLLN by virtue of the fact that $W_\infty^\phi(X) = 0$. However, the motion in Example 4.6 is recurrent, and the SLLN captures not only the local but also the global growth of mass.

The unbounded w in Example 4.6 can be interpreted as follows. Heuristically, since the local growth rate β is bounded away from zero, on average a large population is generated everywhere in space. Risk for the branching process comes from areas of a relatively large variance for the total mass process. In contrast, when the variance parameter α is very small, then extinction is unlikely and w becomes large.

The motivation for the next example comes from Example 10 in [57]. For $B \in \mathcal{B}(D)$, $f_1, f_2 \in p(B)$, we write $f_1 \asymp f_2$ if there are constants $0 < c \le C < \infty$ such that $cf_1(x) \le f_2(x) \le Cf_1(x)$ for all $x \in B$.

Example 4.7. Let $d \ge 1$, $D = \mathbb{R}^d$, $L = \frac{1}{2}\Delta - \gamma x \cdot \nabla$, $\beta(x) = c_1\|x\|^2 + c_2$, where $c_1, c_2 > 0$, $\gamma > \sqrt{2c_1}$. Write $\vartheta := \frac{1}{2}(\gamma - \sqrt{\gamma^2 - 2c_1})$. Then Assumption 2 holds with $\lambda_c = \vartheta d + c_2$, $\phi(x) = e^{\vartheta\|x\|^2}$ and $\widetilde{\phi}(x) = ce^{(\vartheta-\gamma)\|x\|^2}$, where $c = (\frac{\gamma - 2\vartheta}{\pi})^{d/2}$. Suppose that $\Pi \equiv 0$ and $\alpha \in C^\eta(D)$ with $\alpha \asymp 1/\phi$ on $\mathbb{R}^d$. Then Theorem 4.4 applies, $w \asymp \phi$, and the support of X is recurrent.

Let $h(x) = e^{\vartheta\|x\|^2}$. Using $-\gamma + 2\vartheta = -\sqrt{\gamma^2 - 2c_2}$ and $\vartheta^2 - \gamma\vartheta + \frac{c_1}{2} = 0$, we observe that

$$L_0^h = \frac{1}{2}\Delta - \sqrt{\gamma^2 - 2c_1}\,x \cdot \nabla, \quad \beta^h = \vartheta d + c_2, \quad \alpha^h(x) = h(x)\alpha(x) \asymp 1 \quad \text{on } \mathbb{R}^d.$$

The $(L_0^h, \psi^h; \mathbb{R}^d)$-superprocess, denoted by X^h, satisfies Assumption 2 by Example 4.1 with $\phi^h = \mathbf{1}$, $\widetilde{\phi}^h \asymp e^{(2\vartheta-\gamma)\|x\|^2}$ and $\lambda_c^h = \vartheta d + c_2$. Hence, Lemma 2.10 (i) shows that X satisfies Assumption 2, ϕ, $\widetilde{\phi}$ and λ_c have been correctly identified, and $\phi\alpha$ is bounded. When we have verified Assumption 4 for X^h, then Lemma 2.11 will yield Assumption 4 for X, and Theorem 4.4 applies.

To this end, choose constants $c_3, c_4 \in (0, \infty)$ such that $c_3/h \le \alpha \le c_4/h$. Let

$\overline{\psi}(x,z) := -\beta^h z + c_3 z^2$ and $\underline{\psi}(x,z) := -\beta^h z + c_4 z^2$, and denote by $\overline{w}$ and $\underline{w}$ the martingale functions corresponding to the event of extinction after finite time for the $(L_0^h, \overline{\psi}; \mathbb{R}^d)$- and $(L_0^h, \underline{\psi}; \mathbb{R}^d)$-superprocesses, respectively. Since $\overline{\psi} \le \psi^h \le \underline{\psi}$, Lemma 4.5 (ii) and (4.1) imply

$$\beta^h/c_3 = \overline{w}(x) \ge w^h(x) \ge \underline{w}(x) = \beta^h/c_4 \qquad \text{for all } x \in D.$$

Hence, $w^h \asymp \mathbf{1} = \phi^h$. Now the verification of (2.35) and (2.36) for X^h is the same as in Example 4.1, and we can choose $a(t) = \sqrt{(\lambda_c/(\gamma - 2\vartheta) + \delta)t}$ for some $\delta > 0$ and $K = 1$. Theorem 2.13 yields Assumption 4.

The support of X is recurrent since the support is invariant under h-transforms, and the support of the $(L_0^h, \psi^h; \mathbb{R}^d)$-superprocess is recurrent according to Theorem 4.4 (b) in [59].

The next example covers a large class of processes. The underlying motion is a Brownian motion with or without a compactly supported drift term. Depending on the choice of that drift, the motion can be symmetric or non-symmetric. For a choice of b which makes L non-symmetric see Example 13 in [57]. The article by Chen et al. [27] excludes non-symmetric motions. The example is motivated by Example 22 in [61] and Examples 12 and 13 in [57].

Example 4.8. Let $d \in \{1,2\}$, $D = \mathbb{R}^d$, $L = \frac{1}{2}\Delta + b \cdot \nabla$, where all components of b belong to $C^{1,\eta}(\mathbb{R}^d)$ for some $\eta \in (0,1]$ and are of compact support. Let $\beta_0 \in C^\eta(\mathbb{R}^d)$ be nonnegative and of compact support, $\beta_0 \ne \mathbf{0}$. There exists $\theta > 0$ such that $\lambda_c(L + \theta\beta_0) > 0$, and we let $\beta = \theta\beta_0$, $\lambda_c = \lambda_c(L+\beta)$. Write

$$\varrho(x) = \|x\|^{(1-d)/2} e^{-\sqrt{2\lambda_c}\|x\|}, \qquad \text{for all } x \in \mathbb{R}^d \setminus \{0\}.$$

Let $\alpha \in C^\eta(D)$, $\alpha(x) > 0$ for all $x \in \mathbb{R}^d$ and $\alpha \asymp 1/\varrho$ on $\mathbb{R}^d \setminus B$ for an open ball B around the origin. Then Theorem 4.4 applies with $\phi, \widetilde{\phi}, w \asymp \varrho$ on $\mathbb{R}^d \setminus B$, and the support of X is recurrent.

The existence of θ is proved in Theorems 4.6.3 and 4.6.4 of Pinsky's book [110], and $L + \beta - \lambda_c$ is critical by Theorem 4.6.7 in the same book. Denote by G the Green's function corresponding to the operator $L - \lambda_c$. Then $\phi \asymp G(\cdot, 0)$ on $\mathbb{R}^d \setminus B$ by Theorems 4.6.3 and 7.3.8 in [110]. Pinsky showed in Example 7.3.11 that the Green's function G_1 of $\frac{1}{2}\Delta - \lambda_c$ satisfies $G_1(\cdot, 0) \asymp \varrho$ on $\mathbb{R}^d \setminus B$. Since b is compactly supported, $G_1(\cdot, 0) \asymp G(\cdot, 0)$ on $\mathbb{R}^d \setminus B$, and the estimate for ϕ is established. The same argument yields the same estimate for $\widetilde{\phi}$, and Assumption 2 holds. Moreover, $\phi\alpha$ is bounded.

To check Assumption 4 we use Theorem 2.13. An h-transform of the $(L, \psi; \mathbb{R}^d)$-superprocess with $h = \phi$ gives an $(L_0^\phi, \psi^\phi; \mathbb{R}^d)$-superprocess, where L_0^ϕ corresponds to a conservative, positive recurrent motion, and $\psi^\phi(x,z) = -\lambda_c z + \phi(x)\alpha(x) z^2$. Since $\phi\alpha \asymp \mathbf{1}$, $w^\phi \asymp \mathbf{1}$ by the same argument as in Example 4.7. Hence, (4.5) implies $w/\phi \asymp \mathbf{1}$, and Conditions (i) and (ii) of Theorem 2.13 have been verified in Examples 12 and 13 of [57] with $a(t) = \sqrt{2(\|\beta\|_\infty + \delta)t}$, $K > \sqrt{(\|\beta\|_\infty + \delta)/\lambda_c}$ for $\delta > 0$. Since w^ϕ is bounded and the diffusion corresponding to L_0^ϕ is recurrent, Theorem 4.4 in [59] shows that the support of X^ϕ, and therefore X, is recurrent.

4.3 Bounded domains

In the situation that D is a bounded Lipschitz domain and L is a uniformly elliptic operator with smooth coefficients, Liu et al. [102] prove the SLLN for a general branching mechanism with arbitrary $\beta \in b(D)$, and α and Π as in Section 1.1.

However, the Wright–Fisher diffusion on domain $D = (0,1)$ is a diffusion process whose diffusion matrix $a(x) = x(1-x)$ is not uniformly elliptic. The process has attracted a wide interest in the literature (see for example [77, 74, 12]). Fleischmann and Swart [74] studied the large-time behaviour of the corresponding superprocess with spatially independent, quadratic branching mechanism on $[0,1]$. They conjecture a SLLN for the process restricted to $D = (0,1)$ (see above (23) in [74]) but prove only convergence in L^2. The Wright–Fisher diffusion is not conservative, so the arguments in Section 4.1 are not applicable. However, Theorem 1.2 applies, and the following theorem proves the conjecture for all Lebesgue-almost everywhere continuous test functions $f \in p(D)$ with f/ϕ bounded. (Fleischmann and Swart do not assume any continuity.)

Theorem 4.9 (Super-Wright–Fisher diffusion)**.** *Let $D = (0,1)$, $\beta \in (1,\infty)$, $\alpha > 0$, $\Pi \equiv 0$ and*

$$L = \frac{1}{2}x(1-x)\frac{d^2}{dx^2} = \frac{1}{2}\frac{d}{dx}x(1-x)\frac{d}{dx} + \frac{2x-1}{2}\frac{d}{dx}.$$

Then Theorem 1.2 applies with $\phi(x) = 6x(1-x)$, $\widetilde{\phi} = \mathbf{1}$ and $\lambda_c = \beta - 1$.

Proof. Let $h(x) = 6x(1-x)$. Fleischmann and Swart proved in Lemma 20 of [74] that the generator

$$L_0^h = \frac{1}{2}\frac{d}{dx}x(1-x)\frac{d}{dx} + \frac{1-2x}{2}\frac{d}{dx}$$

corresponds to an ergodic diffusion with invariant law $h(x)\ell(dx)$ on D. Using $\beta^h = \beta - 1$, we deduce that $\lambda_c(L_0^h + \beta^h) = \beta - 1$, $\phi^h = \mathbf{1}$, $\widetilde{\phi}^h = h$, and using Lemma 2.10, Assumption 2 for the $(L, \psi; D)$ superprocess as well as the stated identities for ϕ, $\widetilde{\phi}$ and λ_c are established. Assumption 1 holds by Lemma 4.3; the boundedness of α and ϕ implies that Assumption 3 is satisfied. To verify Assumption 4, we notice that Condition (i) of Theorem 2.13 is trivially satisfied for $D_t = D$, and (2.34) for $D_t = D$ and $K = 1$ has been proved in Lemma 20 of [74]. Hence, Assumption 4 follows from Theorem 2.13, and Theorem 1.2 applies. □

Appendices to Part I

APPENDIX A

FEYNMAN–KAC ARGUMENTS

In this appendix, we prove an integral identity that is used several times in the thesis. Versions of this result appeared in Lemma A.I.1.5 of [43] and Lemma 4.1.2 of [44] but the format and assumptions are different. Like in the remainder of Part I of this thesis, $(\xi = (\xi_t)_{t\geq 0}\colon (\mathbb{P}_x)_{x\in D})$ is a diffusion as described in Section 1.1.

Lemma A.1. *Let $T > 0$ and either $B = D$ or $B \subset\subset D$ open. Write $\tau = \inf\{t \geq 0\colon \xi_t \notin B\}$, and $A = D$ if $B = D$; $A = \overline{B}$ if $B \subset\subset D$.*

(i) *Let $f_1 \in b(A)$, $g_1\colon A\times[0,T] \to \mathbb{R}$ measurable and bounded from above and $f_2, g_2 \in b(A \times [0,T])$. If for all $(x,t) \in A \times [0,T]$,*

$$\begin{aligned} v(x,t) = \mathbb{P}_x\Big[e^{\int_0^{t\wedge\tau}(g_1+g_2)(\xi_r,t-r)\,dr} f_1(\xi_{t\wedge\tau})\Big] \\ + \mathbb{P}_x\Big[\int_0^{t\wedge\tau} e^{\int_0^s (g_1+g_2)(\xi_r,t-r)\,dr} f_2(\xi_s, t-s)\,ds\Big], \end{aligned} \tag{A.1}$$

then, for all $(x,t) \in A \times [0,T]$,

$$\begin{aligned} v(x,t) &= \mathbb{P}_x\Big[e^{\int_0^{t\wedge\tau} g_1(\xi_r,t-r)\,dr} f_1(\xi_{t\wedge\tau})\Big] \\ &+ \mathbb{P}_x\Big[\int_0^{t\wedge\tau} e^{\int_0^s g_1(\xi_r,t-r)\,dr}\Big(f_2(\xi_s,t-s) + g_2(\xi_s,t-s)v(\xi_s,t-s)\Big)\,ds\Big]. \end{aligned}$$

(ii) *The statement of (i) remains valid when $f_1 \in bp(A)$, $f_2, g_1, g_2\colon A\times[0,T] \to \mathbb{R}$ measurable with g_1 bounded from above, g_2 nonnegative, f_2 nonpositive and g_1+g_2 bounded from above. Notice that in this case, v might attain the value $-\infty$.*

Proof. For all $t \geq 0$, write

$$\mathcal{Y}_t = e^{\int_0^{t\wedge\tau}(g_1+g_2)(\xi_r,t-r)\,dr} f_1(\xi_{t\wedge\tau}), \quad \text{and} \quad \mathcal{Z}_t = \int_0^{t\wedge\tau} e^{\int_0^s (g_1+g_2)(\xi_r,t-r)\,dr} f_2(\xi_s,t-s)\,ds.$$

By assumption, $v(x,t) = \mathbb{P}_x[\mathcal{Y}_t + \mathcal{Z}_t]$ for all $(x,t) \in A \times [0,T]$. The Markov property implies

$$\int_0^t \mathbb{P}_x\Big[\mathbb{1}_{\{s<\tau\}} e^{\int_0^s g_1(\xi_r,t-r)\,dr} g_2(\xi_s,t-s)\mathbb{P}_{\xi_s}[\mathcal{Y}_{t-s}]\Big]\,ds$$
$$= \int_0^t \mathbb{P}_x\Big[\mathbb{1}_{\{s<\tau\}} e^{\int_0^s g_1(\xi_r,t-r)\,dr} g_2(\xi_s,t-s) e^{\int_s^{t\wedge\tau}(g_1+g_2)(\xi_r,t-r)\,dr} f_1(\xi_{t\wedge\tau})\Big]\,ds.$$

If g_2 is bounded, then Fubini's theorem and the fundamental theorem of calculus (FTC) for Lebesgue integrals imply that the right-hand side equals

$$\mathbb{P}_x\Big[f_1(\xi_{t\wedge\tau}) e^{\int_0^{t\wedge\tau} g_1(\xi_r,t-r)\,dr} \int_0^{t\wedge\tau} g_2(\xi_s,t-s) e^{\int_s^{t\wedge\tau} g_2(\xi_r,t-r)\,dr}\,ds\Big]$$
$$= \mathbb{P}_x\Big[f_1(\xi_{t\wedge\tau}) e^{\int_0^{t\wedge\tau} g_1(\xi_r,t-r)\,dr}\Big(e^{\int_0^{t\wedge\tau} g_2(\xi_r,t-r)\,dr} - 1\Big)\Big].$$

In the situation of (ii), the same identity can be obtained by truncating g_2 before the application of the FTC and using the monotone convergence theorem afterwards. The Markov property and Fubini's theorem (in case (ii) its application is justified by the nonpositivity of the integrand) yield

$$\int_0^t \mathbb{P}_x\Big[\mathbb{1}_{\{s<\tau\}} e^{\int_0^s g_1(\xi_r,t-r)\,dr} g_2(\xi_s,t-s)\mathbb{P}_{\xi_s}[\mathcal{Z}_{t-s}]\Big]\,ds$$
$$= \int_0^t \mathbb{P}_x\Big[\mathbb{1}_{\{s<\tau\}} e^{\int_0^s g_1(\xi_r,t-r)\,dr} g_2(\xi_s,t-s) \int_s^{t\wedge\tau} e^{\int_s^u (g_1+g_2)(\xi_r,t-r)\,dr} f_2(\xi_u,t-u)\,du\Big]\,ds$$
$$= \mathbb{P}_x\Big[\int_0^{t\wedge\tau} e^{\int_0^u g_1(\xi_r,t-r)\,dr} f_2(\xi_u,t-u) \int_0^u g_2(\xi_s,t-s) e^{\int_s^u g_2(\xi_r,t-r)\,dr}\,ds\,du\Big].$$

As above, the FTC implies that the right-hand side equals

$$\mathbb{P}_x\Big[\int_0^{t\wedge\tau} e^{\int_0^u g_1(\xi_r,t-r)\,dr} f_2(\xi_u,t-u)\Big(e^{\int_0^u g_2(\xi_r,t-r)\,dr} - 1\Big)\,du\Big].$$

Since $v(x,t) = \mathbb{P}_x[\mathcal{Y}_t + \mathcal{Z}_t]$ for all $(x,t) \in A \times [0,T]$, we conclude that in the situation of (i),

$$\mathbb{P}_x\Big[\int_0^{t\wedge\tau} e^{\int_0^s g_1(\xi_r,t-r)\,dr}\Big(f_2(\xi_s,t-s) + g_2(\xi_s,t-s)v(\xi_s,t-s)\Big)\,ds\Big]$$
$$= \mathbb{P}_x\Big[f_1(\xi_{t\wedge\tau}) e^{\int_0^{t\wedge\tau} g_1(\xi_r,t-r)\,dr}\Big(e^{\int_0^{t\wedge\tau} g_2(\xi_r,t-r)\,dr} - 1\Big)$$
$$+ \int_0^{t\wedge\tau} e^{\int_0^s (g_1+g_2)(\xi_r,t-r)\,dr} f_2(\xi_s,t-s)\,ds\Big].$$

In the situation of (ii), this use of linearity is justified since none of the summed integrals can take the value $+\infty$. Since f_1 is bounded and g_1, g_1+g_2 are bounded from above, the first term on the right can be written as the difference of two finite integrals and (A.1) yields the claim. □

APPENDIX B

MONOTONICITY

In this appendix, a generalised version of the mild equation (1.3) is studied. For the exit measures $\widetilde{X}_t^B$ from Section 2.1.1, we establish monotonicity in the domain, and we prove Lemma 4.5. As a by-product, we re-prove the existence and uniqueness of the solutions to (1.3) and (2.7).

To this end, we assume only that β is bounded above, not necessarily from below. More specifically, the setup is as follows. Let $\beta\colon D \to \mathbb{R}$ be measurable with $\bar{\beta} = \max\{\sup_{x\in D}\beta(x), 0\} < \infty$, $\alpha \in bp(D)$, Π a kernel from D to $(0,\infty)$ such that $x \mapsto \int_{(0,\infty)}(y \wedge y^2)\,\Pi(x,dy)$ belongs to $bp(D)$, and let $(\xi;\mathbb{P})$ be a diffusion as described in Section 1.1. We denote by $lbp(D\times[0,\infty))$ the space of all functions $f \in p(D\times[0,\infty))$ with $\|f\|_{\infty,T} := \sup_{t\in[0,T]}\|f(\cdot,t)\|_\infty < \infty$ for all $T > 0$.

For $f \in bp(D)$ and $g \in lbp(D\times[0,\infty))$, we are interested in solutions to the integral equation

$$u(x,t) + \int_0^t S_s\big[\psi_0(\cdot, u(\cdot,t-s))\big](x)\,ds = S_t f(x) + \int_0^t S_s[g(\cdot,t-s)](x)\,ds \tag{B.1}$$

for all $(x,t) \in D\times[0,\infty)$, where $S_t f(x) = \mathbb{P}_x[e^{\int_0^t \beta(\xi_s)\,ds} f(\xi_t)]$ for $f \in p(D)$. A similar analysis to ours has been carried out by Dynkin (Chapter 4, Sections 1 and 3 in [43]) for the case $\beta = \mathbf{0}$ and $g = \mathbf{0}$. The greater generality of (B.1) allows us to handle the general setup of this thesis and to prove Lemma 4.5.

Note that $z \mapsto \psi_0(x,z)$, defined in (1.1) is increasing, convex, and nonnegative. In particular, any nonnegative solution u to (B.1), satisfies

$$0 \le u(x,t) \le e^{\bar{\beta}t}\|f\|_\infty + \int_0^t e^{\bar{\beta}s}\|g(\cdot,t-s)\|_\infty\,ds \quad \text{for all } (x,t) \in D\times[0,\infty).$$

Hence, any nonnegative solution to (B.1) is an element of $lbp(D\times[0,\infty))$. Moreover, ψ_0 is locally Lipschitz continuous in the sense that for every fixed $c > 0$ there exists

$\mathcal{L}(c) \in [0,\infty)$ such that

$$|\psi_0(x,z_1) - \psi_0(x,z_2)| \le \mathcal{L}(c)|z_1 - z_2| \qquad \text{for all } z_1, z_2 \in [0,c], x \in D. \tag{B.2}$$

We use the following version of Gronwall's lemma; for a proof see Theorem A.5.1 in [68].

Lemma B.1 (**Gronwall's lemma**). *Let $T > 0$, $C, \rho \ge 0$ and $h \in b([0,T])$. If*

$$h(t) \le C + \rho \int_0^t h(s)\, ds \qquad \textit{for all } t \in [0,T],$$

then $h(t) \le Ce^{\rho t}$ for all $t \in [0,T]$.

Lemmas B.2 and B.3 in the case $\beta = \mathbf{0}$ and $g = \mathbf{0}$ are Theorems 4.1.1 and 4.3.1 in [44].

Lemma B.2 (**Uniqueness**). *Let $f, \hat{f} \in bp(D)$, $g, \hat{g} \in lbp(D \times [0,\infty))$, and suppose that u and $\hat{u}$ are nonnegative solutions to (B.1) for (f,g) and $(\hat{f},\hat{g})$, respectively. Then, for every $T > 0$, there exists a constant $C > 0$ such that*

$$\|u - \hat{u}\|_{\infty,T} \le C\big(\|f - \hat{f}\|_\infty + \|g - \hat{g}\|_{\infty,T}\big).$$

In particular, the solution to (B.1) is unique.

Proof. Fix $T > 0$, and let $c \ge \max\{\|u\|_{\infty,T}, \|\hat{u}\|_{\infty,T}\}$. Then (B.2) yields

$$|\psi_0(x,\hat{u}(x,t)) - \psi_0(x,u(x,t))| \le \mathcal{L}(c)|\hat{u}(x,t) - u(x,t)| \quad \text{for all } (x,t) \in D \times [0,T].$$

Writing $h(x,t) = |u(x,t) - \hat{u}(x,t)|$ and $M = e^{\bar{\beta}T}$, (B.1) implies that

$$h(x,t) \le M\|f - \hat{f}\|_\infty + MT\|g - \hat{g}\|_{\infty,T} + \int_0^t M\mathcal{L}(c)\|h(\cdot,s)\|_\infty\, ds$$

for all $(x,t) \in D \times [0,T]$. Lemma B.1 yields the claim. □

Lemma B.3 (**Existence**). *Let $f \in bp(D)$ and $g \in lbp(D \times [0,\infty))$. There exists a nonnegative solution $u \in lbp(D \times [0,\infty))$ to (B.1).*

Proof. Fix $T > 0$ and let $M = e^{\bar{\beta}T}$. For $k \in [0,\infty)$ and $u \in lbp(D \times [0,T])$, i.e. $u \in p(D \times [0,T])$ with $\|u\|_{\infty,T} < \infty$, we define for all $(x,t) \in D \times [0,T]$,

$$\begin{aligned} F_k u(x,t) = e^{-kt} S_t f(x) &+ \int_0^t e^{-ks} S_s[g(\cdot, t-s)](x)\, ds \\ &+ \int_0^t e^{-ks} S_s\big[ku(\cdot,t-s) - \psi_0(\cdot, u(\cdot,t-s))\big](x)\, ds. \end{aligned}$$

Let $c \ge M\|f\|_\infty + MT\|g\|_{\infty,T}$ and $k \ge \mathcal{L}(c)$. Write $v(x,t) = e^{\bar{\beta}t}\|f\|_\infty + \int_0^t e^{\bar{\beta}s}\, ds \|g\|_{\infty,T}$ for all $x \in D$, $t \in [0,T]$. We show the following:

(i) $\mathbf{0} \le F_k \mathbf{0} \le v$ on $D \times [0,T]$.

(ii) If $\mathbf{0} \le u_1 \le u_2 \le v$ on $D \times [0,T]$, then $F_k u_1 \le F_k u_2$ on $D \times [0,T]$.

(iii) $F_k v \le v$ on $D \times [0,T]$.

Indeed, $F_k\mathbf{0}(x,t) = e^{-kt}S_t f(x) + \int_0^t e^{-ks}S_s[g(\cdot,t-s)](x)\,ds \in [0, v(x,t)]$ since f and g are nonnegative and $k \ge 0$. For (ii), we combine $u_1 \le u_2 \le v \le c$ on $D \times [0,T]$ with (B.2), to obtain

$$\begin{aligned}
&F_k u_2(x,t) - F_k u_1(x,t) \\
&\quad = \int_0^t e^{-ks} S_s\big[k(u_2 - u_1)(\cdot, t-s) - \big(\psi_0(\cdot, u_2(\cdot,t-s)) - \psi_0(\cdot, u_1(\cdot,t-s))\big)\big](x)\,ds \\
&\quad \ge \int_0^t e^{-ks} S_s\big[(k - \mathcal{L}(c))(u_2 - u_1)(\cdot,t-s)\big](x)\,ds \ge 0.
\end{aligned}$$

To show (iii), we use that ψ_0 is nonnegative, the definition of v and Fubini's theorem to obtain

$$\begin{aligned}
F_k v(x,t) &\le e^{-kt}e^{\bar\beta t}\|f\|_\infty + \int_0^t e^{-ks}e^{\bar\beta s}\|g\|_{\infty,T}\,ds \\
&\qquad + \int_0^t e^{-ks}e^{\bar\beta s}k\Big(e^{\bar\beta(t-s)}\|f\|_\infty + \int_0^{t-s} e^{\bar\beta r}\,dr\|g\|_{\infty,T}\Big)\,ds \\
&= \Big(e^{-kt} + \int_0^t ke^{-ks}\,ds\Big)e^{\bar\beta t}\|f\|_\infty \\
&\qquad + \Big(\int_0^t e^{(\bar\beta - k)s}\,ds + \int_0^t ke^{-ks}\int_s^t e^{\bar\beta r}\,dr\,ds\Big)\|g\|_{\infty,T} \\
&= v(x,t).
\end{aligned}$$

In the next step, we construct a solution to (B.1) via a Picard iteration. Let $u_0 = \mathbf{0}$ and $u_n = F_k u_{n-1}$ for all $n \in \mathbb{N}$. We show by induction that $\mathbf{0} \le u_{n-1} \le u_n \le v$ on $D \times [0,T]$ for all $n \in \mathbb{N}$. For $n = 1$, this is statement (i). The induction step follows from (ii)–(iii). In particular, $(u_n)_{n\in\mathbb{N}_0}$ has a pointwise limit u, which is a fixed point of F_k by the dominated convergence theorem. Lemma A.1 (i) applied to $g_1 = \beta$, which is bounded from above, and the bounded functions $g_2 = -k$, $f_1 = f$ and $f_2(x,t) = g(x,t) + ku(x,t) - \psi_0(x, u(x,t))$, shows that u solves (B.1). □

Choosing $g = \mathbf{0}$, Lemmas B.2 and B.3 imply the existence of a unique solution to (1.3) and (2.7). The following lemma will be used in the proof of Lemma 4.5.

Lemma B.4 (Monotonicity in (f,g)). *Let $f, \hat f \in bp(D)$, $g, \hat g \in lbp(D \times [0,\infty))$ with $f \le \hat f$ and $g \le \hat g$, and denote by u and $\hat u$ the unique solutions to (B.1) corresponding to (f,g) and $(\hat f, \hat g)$, respectively. Then $u \le \hat u$.*

Proof. Since the solution is unique according to Lemma B.2, the claim follows immediately from the construction of the solution via Picard iteration in the proof of Lemma B.3. □

Proof of Lemma 4.5 (i). According to (1.3), u_f is the unique solution to (B.1) with $g = \mathbf{0}$. Moreover, (1.3) for $\hat{u}_f$, and Lemma A.1 (i) applied to $g_1 = \beta$, $g_2 = \hat{\beta} - \beta$, $f_1 = f$ and $f_2(x,t) = -\hat{\psi}_0(x, \hat{u}_f(x,t))$, imply that $\hat{u}_f$ satisfies

$$\begin{aligned}\hat{u}_f(x,t) &= S_t f(x) + \int_0^t S_s\big[-\hat{\psi}_0(\cdot, \hat{u}_f(\cdot, t-s)) + (\hat{\beta}(\cdot) - \beta(\cdot))\hat{u}_f(\cdot, t-s)\big](x)\, ds \\ &= S_t f(x) - \int_0^t S_s\big[\psi_0(\cdot, \hat{u}_f(\cdot, t-s))\big](x)\, ds \\ &\qquad + \int_0^t S_s\big[\psi_\beta(\cdot, \hat{u}_f(\cdot, t-s)) - \hat{\psi}_{\hat{\beta}}(\cdot, \hat{u}_f(\cdot, t-s))\big](x)\, ds.\end{aligned}$$

In particular, $\hat{u}_f$ solves (B.1) with $g(x,t) = \psi_\beta(x, \hat{u}_f(x,t)) - \hat{\psi}_{\hat{\beta}}(x, \hat{u}_f(x,t)) \geq 0$. Now Lemma B.4 yields the claim. □

The following lemma is used in the proofs of Theorem 2.3 and Lemma 4.3.

Lemma B.5 (Monotonicity in B). *Let $B \subset\subset D$ and $f \in bp(D)$ such that the support of f, supp(f), is compactly embedded in B. There exists a unique nonnegative solution $u_f^B \in lbp(D \times [0,\infty))$ to*

$$u(x,t) = \mathbb{P}_x\Big[e^{\int_0^{t\wedge\tau_B} \beta(\xi_s)\, ds} f(\xi_{t\wedge\tau_B})\Big] - \mathbb{P}_x\Big[\int_0^{t\wedge\tau_B} e^{\int_0^s \beta(\xi_r)\, dr} \psi_0(\xi_s, u(\xi_s, t-s))\, ds\Big]. \tag{B.3}$$

Moreover, if B_1 and B_2 are domains with supp$(f) \subseteq B_1 \subseteq B_2$, then $u_f^{B_1} \leq u_f^{B_2}$.

Proof. Let $T > 0$, $c \geq e^{\bar{\beta}T}\|f\|_\infty$, $k \geq \mathcal{L}(c)$. For $u \in lbp(D \times [0,T])$, define

$$\begin{aligned}F_k u(x,t) &= \mathbb{P}_x\Big[e^{\int_0^t [\beta(\xi_s) - k]\, ds} f(\xi_t)\, \mathbb{1}_{\{t<\tau_B\}}\Big] \\ &\qquad + \int_0^t \mathbb{P}_x\Big[e^{\int_0^s [\beta(\xi_r) - k]\, dr}\big[k u(\xi_s, t-s) - \psi_0(\xi_s, u(\xi_s, t-s))\big]\mathbb{1}_{\{s<\tau_B\}}\Big]\, ds.\end{aligned}$$

Since $kz - \psi_0(x,z) \geq kz - \mathcal{L}(c)z \geq 0$ for all $z \in [0,c]$, $F_k u$ is increasing in B for all u with $u(x,t) \leq e^{\bar{\beta}t}\|f\|_\infty =: v(x,t)$. As in Lemmas B.2 and B.3, the unique solution to (B.3) can be obtained as a pointwise limit of the increasing sequence $u_0 = \mathbf{0}$, $u_{n+1} = F_k u_n$ with $u_n \leq v$ for all n. Denote by $u_n^{(1)}$ and $u_n^{(2)}$ the iterates for the operators $F_k^{(1)}$ and $F_k^{(2)}$ corresponding to B_1 and B_2, respectively. We show by induction that $u_n^{(1)} \leq u_n^{(2)}$ for every $n \in \mathbb{N}_0$. For $n = 0$ this is trivial. For the induction step, we first use the induction hypothesis and monotonicity of $F_k^{(1)}$ (see (ii) in the proof of Lemma B.3) and then the monotonicity of F_k in B to deduce

$$u_{n+1}^{(1)} = F_k^{(1)} u_n^{(1)} \leq F_k^{(1)} u_n^{(2)} \leq F_k^{(2)} u_n^{(2)} = u_{n+1}^{(2)}.$$ □

For $B \subset\subset D$, the tuple $(L, \psi^*_{\beta^*}; B)$ satisfies the assumptions of Section 1.1, where the motion is killed at the boundary of B. Hence, Lemma B.5 implies that the $(L, \psi^*_{\beta^*}; D)$-superprocess can be obtained as a distributional limit of $(L, \psi^*_{\beta^*}; B)$-superprocesses using an increasing sequence of compactly embedded domains to approximate D; see the argument before Corollary 6.2 in [98] or Lemma A2 and Theorem A1 in [59].

Part II

Vulnerability of robust preferential attachment networks

Scale-free networks with small power law exponent are known to be robust, meaning that their qualitative topological structure cannot be altered by random removal of even a large proportion of nodes. By contrast, it has been argued in the science literature that such networks are highly vulnerable to a targeted attack, and removing a small number of key nodes in the network will dramatically change the topological structure.

Here we analyse a class of preferential attachment networks in the robust regime and prove four main results supporting this claim: after removal of an arbitrarily small proportion $\epsilon > 0$ of the oldest nodes (1) the asymptotic degree distribution has exponential instead of power law tails; (2) the largest degree in the network drops from being of the order of a power of the network size n to being just logarithmic in n; (3) the typical distances in the network increase from order $\log \log n$ to order $\log n$; and (4) the network becomes vulnerable to random removal of nodes. Importantly, all our results explicitly quantify the dependence on the proportion ϵ of removed vertices. For example, we show that the critical proportion of nodes that have to be retained for survival of the giant component undergoes a steep increase as ϵ moves away from zero, and a comparison of this result with similar ones for other networks reveals the existence of two different universality classes of robust network models.

The key techniques in our proofs are a local approximation of the network by a branching random walk with two killing boundaries, and an understanding of the particle genealogies in this process, which enters into estimates for the spectral radius of an associated operator.

CHAPTER 5

INTRODUCTION

5.1 Motivation

The problem of resilience of networks to either random or targeted attack is crucial to many instances of real world networks, ranging from social networks (like collaboration networks) via technological networks (like electrical power grids), to communication networks (like the World Wide Web). Of particular importance is whether the connectivity of a network relies on a small number of hubs and whether their loss will cause a large-scale breakdown. Albert, Albert and Nakarado [3] argue that "the power grid is robust to most perturbations, yet disturbances affecting key transmission substations greatly reduce its ability to function". Experiments of Albert, Jeong, and Barabási [4], Holme, Kim, Yoon and Han [86] and more recently of Mishkovski, Biey and Kocarev [105] find robustness under random attack but vulnerability to the removal of a small number of key nodes in several other networks. The latter paper includes a study of data related to the human brain, as well as street, collaboration and power grid networks. One should expect this qualitative behaviour across the range of real world networks and it should therefore also be present in the key mathematical models of large complex networks.

A well established feature of many real world networks is that in a suitable range of values k the proportion of nodes with degree k has a decay of order $k^{-\tau}$ for a power law exponent τ. The robustness of networks with small power law exponent under random attack has been observed heuristically by Callaway et al. [25] and Cohen et al. [32], but there seems to be controversy in these early papers about the extent of the vulnerability in the case of targeted attack, see the discussion in [39] and [33]. As Bollobás and Riordan [21, Section 10] point out, such heuristics, informative as they may be, are often quite far away from a mathematical proof that applies to a given model. In their seminal paper [21] they provide the first rigorous proof of robustness in the case of a specific preferential attachment model with power law exponent $\tau = 3$, and later Dereich and Mörters [37] proved for a class of preferential attachment models with tunable power law exponent that networks are robust under random attack if the power law exponent satisfies $\tau \leq 3$, but not when $\tau > 3$, thus revealing the precise location of the phase transition in the behaviour of preferential attachment networks. However, the question of vulnerability of robust networks when a small number of privileged nodes is removed has not been studied systematically in the mathematical literature so far.

It is our aim to give evidence for the vulnerability of robust networks by providing rigorous proof that preferential attachment networks in the robust regime $\tau \leq 3$ undergo a radical change under a targeted attack, i.e. when an arbitrarily small proportion $\epsilon > 0$ of the most influential nodes in the network is removed. Our main results, presented in Section 5.3, show how precisely this change affects the degree structure, the length of shortest paths and the connectivity in the network. The results take the form of limit theorems revealing explicitly the dependence of the relevant parameters on ϵ. Not only does this provide further insight into the topology of the network and the behaviour as ϵ tends to zero, it also allows a comparison to other network models, and thus exposes two classes of robust networks with rather different behaviour; see Section 5.5. Our mathematical analysis of the network combines probabilistic and combinatorial arguments with analytic techniques informed by new probabilistic insights. It is crucially based on the local approximation of preferential attachment networks by a branching random walk with a killing boundary recently found in [37]. In this approximation the removal of a proportion of old vertices corresponds to the introduction of a second killing boundary. On the one hand this adds an additional level of complexity to the process, as the mathematical understanding of critical phenomena in branching models on finite intervals is only just emerging; see for example [81]. On the other hand compactness of the typespace for this branching process opens up new avenues that are exploited, for example, in the form of spectral estimates based on rather subtle information on the shape of principal eigenfunctions of an operator associated with the branching process.

5.2 Mathematical framework

The established mathematical model for a large network is a sequence $(\mathsf{G}_n : n \in \mathbb{N})$ of (random or deterministic) graphs G_n with vertex set V_n and an edge set E_n consisting of (directed or undirected) edges between the vertices. We assume that the size $|\mathsf{V}_n|$ of the vertex set is increasing to infinity in probability, so that results about the limiting behaviour in the sequence of graphs may be seen as predictions for the behaviour of large networks. In all cases of interest here the average number of edges per vertex converges in probability to a finite limit and the topology of a bounded neighbourhood of a typical vertex stabilizes. An important example for this is the proportion of vertices with a given degree in G_n, which in the relevant models converges and allows us to talk about the asymptotic degree distribution. The mathematical models of power law networks therefore have an asymptotic degree distribution with the probability of degree k decaying like $k^{-\tau}$, as $k \to \infty$, for some $\tau > 1$. Our focus here is on the global properties emerging in network models with asymptotic power law degree distributions.

A crucial global feature of a network is its connectivity, and in particular the existence of a large connected component. To describe this, we denote by C_n a connected component in G_n with maximal number of nodes. The graph sequence $(\mathsf{G}_n : n \in \mathbb{N})$ has *a giant component* if there exists a constant $\zeta > 0$ such that

$$\frac{|\mathsf{C}_n|}{\mathbb{E}|\mathsf{V}_n|} \to \zeta \qquad \text{as } n \to \infty,$$

where the convergence holds in probability. We remark that for the models usually considered the issue is not the convergence itself but the positivity of the limit ζ. If a giant component exists and the length of the shortest path between any two vertices in the largest component of G_n is asymptotically bounded by a multiple of $\log n$, then the network is called *small*. If it is asymptotically bounded by a constant multiple of $\log\log n$, then the network is called *ultrasmall*; see Section 1.2 in [85].

To model a *random attack* on the network, each vertex in G_n is kept independently with probability $p \in [0,1]$ and otherwise it is removed from the vertex set together with all its adjacent edges, i.e., we run *vertex percolation* on G_n with *retention probability* p. The resulting graph is denoted by $\mathsf{G}_n(p)$. A simple coupling argument shows that there exists a critical parameter $p_c \in [0,1]$ such that the sequence $(\mathsf{G}_n(p) : n \in \mathbb{N})$ has a giant component if $p_c < p \leq 1$, and it does not have a giant component if $0 \leq p < p_c$. If $p_c = 0$, i.e. if the giant component cannot be destroyed by percolation with any positive retention parameter, then the network is called *robust*. To study the resilience of networks to a *targeted attack*, we consider models in which the construction of the network favours certain vertices in such a way that these privileged vertices have a better chance of getting a high degree than others. When G_n is a graph on n vertices, we label these by 1 to n and assume that vertices are ordered in decreasing order of privilege. This assumption allows an attacker to target the most privileged vertices without knowledge of the entire graph. The *damaged graph* G_n^ϵ, for some $\epsilon \in (0,1)$, is obtained from G_n by the removal of all vertices with label less or equal to ϵn together with all adjacent edges. In particular, the new vertex set is $\mathsf{V}_n^\epsilon = \{\lfloor \epsilon n \rfloor + 1, \ldots, n\}$, and we let C_n^ϵ be a connected component in G_n^ϵ with maximal number of nodes. Write $\mathsf{G}_n^\epsilon(p)$ for the graph obtained from G_n by first removing all vertices with label at most ϵn and then running vertex percolation on the remaining graph. Note that we would get the same graph when reversing the order in which these two attacks are performed. However, we always start with the targeted attack for definiteness.

We investigate the problem of vulnerability of random networks to targeted attack in the context of *preferential attachment networks*. This class of models has been popularised by Barabási and Albert [9] and has received considerable attention in the scientific literature. The idea is that a sequence of graphs is constructed by successively adding vertices. Together with a new vertex, new edges are introduced by that connect it to existing vertices at random with a probability depending on the degree of the existing node; the higher the degree the more likely the connection. Despite the relatively simple principle on which this model is based it shows a good match of global features with real networks. For example, the asymptotic degree distributions follow a power law, and variations in the attachment probabilities allow for tuning of the power law exponent τ; see [36]. If the power law exponent satisfies $\tau < 3$, then the network is robust and ultrasmall [37, 35].

The first mathematically rigorous study of resilience in preferential attachment networks was performed by Bollobás and Riordan [21] for the so-called LCD model. This model variant has the advantage of having an explicit static description, which makes it easier to analyse than models that have only a dynamic description. It also has a fixed power law exponent $\tau = 3$, hence, Bollobás and Riordan [21] prove only results for this specific exponent. They show that the network is robust and identify a critical proportion $\epsilon_c < 1$ such that the removal of the oldest $\lfloor \epsilon n \rfloor$ vertices leads to the destruction of the giant component if and only if $\epsilon \geq \epsilon_c$. Note that this is not in line with the notion of vulnerability that we are interested in as we only want to remove a *small* proportion of old vertices.

We consider the question of vulnerability in the following model variant introduced in [36]. Let $\mathbb{N}_0$ be the set of nonnegative integers and fix a function $f\colon \mathbb{N}_0 \to (0,\infty)$, which we call the *attachment rule*. The most important case is if f is affine, i.e. $f(k) = \gamma k + \beta$ for parameters $\gamma \in [0,1)$ and $\beta > 0$, but non-linear functions are allowed.

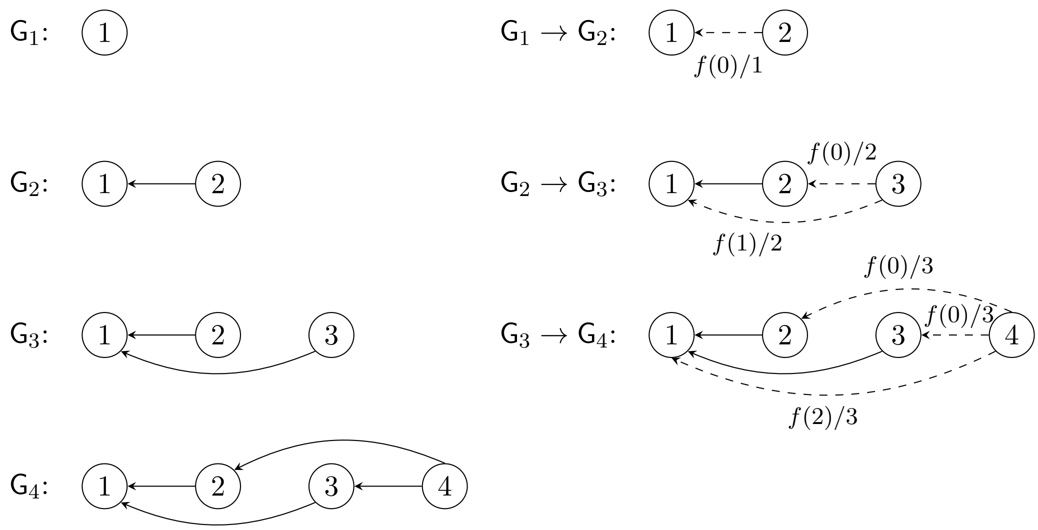

Figure II-1. One possible evolution from graph G_1 to G_4. Potential edges are displayed as dashed arrows together with their probabilities.

Given an attachment rule f, we define a growing sequence $(\mathsf{G}_n : n \in \mathbb{N})$ of random graphs by the following dynamics:

- Start with one vertex labelled 1 and no edges, i.e. G_1 is given by $\mathsf{V}_1 := \{1\}$, and $\mathsf{E}_1 := \emptyset$;

- Given the graph G_n, we construct G_{n+1} from G_n by adding a new vertex labelled $n+1$ and, for each $m \le n$ independently, inserting the directed edge $(n+1, m)$ with probability
$$\frac{f(\text{indegree of } m \text{ at time } n)}{n} \wedge 1. \tag{5.1}$$

The first few steps in one possible evolution of the graphs are displayed in Figure II-1. Formally we are dealing with a sequence of directed graphs but all edges point from the younger to the older vertex. Hence, the directions can be recreated from the undirected, labelled graph. For all structural questions, particularly regarding connectivity and the length of shortest paths, we regard $(\mathsf{G}_n : n \in \mathbb{N})$ as an undirected network. Dereich and Mörters consider in [36, 37] concave attachment rules f. Denoting the asymptotic slope of f by
$$\gamma := \lim_{k \to \infty} \frac{f(k)}{k}, \tag{5.2}$$
they show that for $\gamma \in (0,1)$ the sequence $(\mathsf{G}_n : n \in \mathbb{N})$ has an asymptotic degree distribution which follows a power law with exponent
$$\tau = \frac{\gamma + 1}{\gamma}.$$

For $\gamma \ge 1$, i.e. $\tau \le 2$, the mean of the asymptotic degree distribution is infinite and a radically different topology can be expected. Results on power law networks in this regime have been derived for example in [65, 15]; we restrict ourselves to the finite mean case $\gamma < 1$. In the case $\gamma < \frac{1}{2}$, or equivalently $\tau > 3$, there exists a critical percolation parameter $p_c > 0$ such that $(\mathsf{G}_n(p) : n \in \mathbb{N})$ has a giant component if and only if $p > p_c$.[1] If however $\gamma \ge \frac{1}{2}$, or equivalently $\tau \le 3$, the sequence $(\mathsf{G}_n(p) : n \in \mathbb{N})$ has a giant component for all $p \in (0,1]$, i.e. $(\mathsf{G}_n : n \in \mathbb{N})$ is robust. This is the regime of interest in this thesis.

5.3 Statement of the main results

In this section, we study the case of an affine attachment rule $f(k) = \gamma k + \beta$ with $\beta > 0$ and $\gamma \in [\frac{1}{2}, 1)$. Recall that for this choice the preferential attachment network is robust. We use the symbol $a(\epsilon) \asymp b(\epsilon)$ to indicate that there are constants $0 < c < C$ and some $\epsilon_0 > 0$ such that $cb(\epsilon) \le a(\epsilon) \le Cb(\epsilon)$ for all $0 < \epsilon < \epsilon_0$.

[1]The results of [37] are formulated for edge percolation, whereas we consider vertex percolation. It is not hard to see that for the existence or nonexistence of the giant component this makes no difference.

Theorem 5.1. (Loss of connectivity) *For any $\epsilon \in (0,1)$, there exists $p_c(\epsilon) \in (0,1]$ such that the damaged network*

$$(\mathsf{G}_n^\epsilon(p)\colon n \in \mathbb{N}) \textit{ has a giant component } \Leftrightarrow\ p > p_c(\epsilon). \tag{5.3}$$

If $\gamma = \frac{1}{2}$, then

$$p_c(\epsilon) \asymp \frac{1}{\log(1/\epsilon)}. \tag{5.4}$$

If $\gamma > \frac{1}{2}$, then, as $\epsilon \downarrow 0$,

$$p_c(\epsilon) = \frac{2\gamma - 1}{\sqrt{\beta(\gamma+\beta)}} \epsilon^{\gamma-1/2} \big[1 + O\big(\epsilon^{\gamma-1/2}(\log \epsilon)\big)\big].$$

Theorem 5.1 shows that the removal of an arbitrarily small proportion of old nodes makes the network vulnerable to percolation, but does not destroy the giant component. The steep increase of $p_c(\epsilon)$ as ϵ leaves zero shows that, even when a small proportion of old nodes has been removed from the network, the removal of further old nodes is much more destructive than the removal of a similar proportion of randomly chosen nodes.

Since $\epsilon^{\gamma-1/2}$ is strictly decreasing in γ, this effect is stronger the closer γ is to $\frac{1}{2}$. This result might be perceived as slightly counterintuitive since the preferential attachment becomes stronger as γ increases and therefore we might expect older nodes to be more privileged and a targeted attack to be more effective than in the small γ regime. However, the effect of the stronger preferential attachment is more than compensated by the fact that networks with a small value of γ have a (stochastically) smaller number of edges and are therefore a-priori more vulnerable. Note also that $p_c(\epsilon)$ may be equal to 1 if ϵ is not sufficiently small in which case (5.3) implies that the damaged network has no giant component. In the case $\gamma = \frac{1}{2}$, the implied constants in (5.4) can be made explicit as $c = \frac{1}{\gamma+\beta}$ and $C = \frac{1}{\beta}$; see Proposition 6.5 below.

From here onwards we additionally assume that $\beta \leq 1$. Under this condition, $f(n) < n + 1$ for all $n \in \mathbb{N}_0$, and the minimum in (5.1) is always attained by its first argument. To gain further insight into the topology of the damaged graph, we look at the asymptotic indegree distribution and at the largest indegree in the network. It was proved in Theorem 1.1 (b) of [36] that outdegrees are asymptotically Poisson-distributed, and therefore indegrees are solely responsible for the power law behaviour as well as the dynamics of maximal degrees.

For a probability measure ν on the nonnegative integers, we write $\nu_{\geq k} := \nu(\{k, k+1, \ldots\})$ and $\nu_k := \nu(\{k\})$. Let $\mathcal{Z}[m,n]$ be the indegree of vertex m in G_n at time $n \geq m$. Since for $m > \lfloor \epsilon n \rfloor$, the indegree of m in G_n and G_n^ϵ agree, writing $X^\epsilon(n)$ for the empirical indegree distribution in G_n^ϵ,

$$X_k^\epsilon(n) = \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n \rfloor + 1}^{n} \mathbb{1}_{\{k\}}(\mathcal{Z}[m,n]), \qquad \text{for } k \in \mathbb{N}_0.$$

Write $\mathsf{M}(\mathsf{G})$ for the maximal indegree in a directed graph G, and for $s, t > 0$, let

$$B(s,t) := \int_0^1 x^{s-1}(1-x)^{t-1}\, dx$$

denote the beta function at (s,t). Before we make statements about the network after the targeted attack, recall the situation in the undamaged network. In Theorem 1.1 (a) of [36], Dereich and Mörters show that the empirical indegree distribution $X^0(n)$ in G_n satisfies almost surely

$$\lim_{n\to\infty} X^0(n) = \mu$$

in total variation norm. The limit is the probability measure μ on the nonnegative integers given by

$$\mu_{\geq k} = \frac{B(k + \frac{\beta}{\gamma}, \frac{1}{\gamma})}{B(\frac{\beta}{\gamma}, \frac{1}{\gamma})} \qquad \text{for } k \in \mathbb{N}_0,$$

and satisfies $\lim_{k\to\infty} \log \mu_{\geq k} / \log k = -1/\gamma$. Moreover, Theorem 1.1 and 1.5 and Proposition 1.10 in [36] show that the maximal indegree satisfies, in probability,

$$\frac{\log \mathsf{M}(\mathsf{G}_n)}{\log(n^\gamma)} \to 1 \qquad \text{as } n \to \infty.$$

Our result shows that in the damaged network the asymptotic degree distribution is no longer a power law but has exponential tails. The maximal degree grows only logarithmically, not polynomially.

Theorem 5.2. (Collapse of large degrees) *Let $\epsilon \in (0,1)$. Almost surely,*

$$\lim_{n\to\infty} X^\epsilon(n) = \mu^\epsilon$$

in total variation norm. The limit is the probability measure μ^ϵ on the nonnegative integers given by

$$\mu^\epsilon_{\geq k} = \int_\epsilon^1 \frac{1}{1-\epsilon} B\big(k, \tfrac{\beta}{\gamma}\big)^{-1} \int_{y^\gamma}^1 x^{\frac{\beta}{\gamma}-1}(1-x)^{k-1}\, dx\, dy \qquad \textit{for } k \in \mathbb{N}. \tag{5.5}$$

It satisfies $\lim_{k\to\infty} \log \mu^\epsilon_{\geq k}/k = \log(1-\epsilon^\gamma)$. Moreover, the maximal indegree satisfies, in probability,

$$\frac{\mathsf{M}(\mathsf{G}_n^\epsilon)}{\log n} \to -\frac{1}{\log(1-\epsilon^\gamma)} \qquad \textit{as } n \to \infty. \tag{5.6}$$

It is worth mentioning that $\mu = \mu^0$, so Theorem 5.2 remains valid for $\epsilon = 0$. Moreover, the result holds also for $\gamma \in (0, \frac{1}{2})$ by the same proof. Theorem 5.2 shows in particular that, by removing the $\lfloor \epsilon n \rfloor$ oldest vertices, we have removed all vertices with a degree bigger than a given constant multiple of $\log n$. This justifies the comparison of our vulnerability results with empirical studies of real world networks such as [32], in which all nodes whose degree exceeds a given threshold are removed. Note also that, as $\epsilon \downarrow 0$, the right-hand side in (5.6) is asymptotically equivalent to $\epsilon^{-\gamma}$ and the growth of the maximal degree is the faster the larger γ.

Denote the graph distance in a graph G by d_G. Preferential attachment networks are ultrasmall for sufficiently small power law exponents. For our model, Mönch [107], see also [35, 38], has shown that for independent random vertices V_n, W_n chosen uniformly from C_n,

$$\text{if } \gamma = \frac{1}{2}, \text{ then} \qquad \mathsf{d}_{\mathsf{G}_n}(V_n, W_n) \sim \frac{\log n}{\log\log n},$$
$$\text{if } \gamma > \frac{1}{2}, \text{ then} \qquad \mathsf{d}_{\mathsf{G}_n}(V_n, W_n) \sim \frac{4\log\log n}{\log(\gamma/(1-\gamma))},$$

meaning that the ratio of the left- and right-hand side converges to one in probability as $n \to \infty$. Removing an arbitrarily small proportion of old vertices however leads to a massive increase in the typical distances, as our third main theorem reveals. We say that a sequence of events $(\mathcal{E}_n : n \in \mathbb{N})$ holds *with high probability* if $\mathbb{P}(\mathcal{E}_n) \to 1$ as $n \to \infty$.

Theorem 5.3. (Increase of typical distances) *Let $\epsilon > 0$ be sufficiently small so that $(\mathsf{G}_n^\epsilon : n \in \mathbb{N})$ has a giant component, and let V_n, W_n be chosen independently and uniformly from C_n^ϵ. Then, for all $\delta > 0$,*

$$\mathsf{d}_{\mathsf{G}_n^\epsilon}(V_n, W_n) \geq \frac{1-\delta}{\log(1/p_c(\epsilon))} \log n \qquad \textit{with high probability.}$$

Our proof gives the result for all values $\gamma \in [0, 1)$, $\epsilon > 0$, with $p_c(\epsilon) < 1$, but if $\gamma < \frac{1}{2}$, even without removal of old vertices the typical distances in the network are known to be of order $\log n$, so that this is not surprising. We believe that there is an upper bound matching the lower bound above, but the proof would be technical and the result much less interesting.

In the next two sections we discuss some further ramifications of our main results.

5.4 Non-linear attachment rules

So far we have presented results for the case of affine attachment rules f, given by $f(k) = \gamma k + \beta$. While the fine details of the network behaviour often depend on the exact model definition, we expect the principal scaling and macroscopic features to be independent of these details. To investigate this *universality*, we now discuss to what extent Theorem 5.1 remains true when we look at more general non-linear attachment rules f.

We consider two classes of attachment rules.

(1) A function $f\colon \mathbb{N}_0 \to (0,\infty)$ is called a *L-class* attachment rule if there exists $\gamma \in [0,1)$ and $0 < \beta_l \le \beta_u$ such that $\gamma k + \beta_l \le f(k) \le \gamma k + \beta_u$ for all k. Note that the parameter γ for a L-class rule is uniquely defined by (5.2).

(2) A concave function $f\colon \mathbb{N}_0 \to (0,\infty)$ with $\gamma := \lim_{k\to\infty} f(k)/k \in [0,1)$ is called a *C-class* attachment rule. Note that concavity of f implies that the limit above exists and that f is non-decreasing.

The asymptotic slope of the attachment rule determines the key features of the model. For example, Dereich and Mörters [36] show that, for certain C-class attachment rules with $\gamma > 0$, the asymptotic degree distribution is a power law with exponent $\tau = 1 + 1/\gamma$. The following theorem shows that γ also determines the scaling of the critical percolation parameter for the damaged network.

Theorem 5.4. (Loss of connectivity, non-linear case) *Let f be a L-class or C-class attachment rule. For all $\epsilon \in (0,1)$,*

$$p_{\mathrm{c}}(\epsilon) := \inf\big\{p\colon (\mathsf{G}_n^{\epsilon}(p)\colon n \in \mathbb{N}) \textit{ has a giant component}\big\} > 0.$$

Moreover, if f is in the L-class and

$$\text{if } \gamma = \frac{1}{2}, \text{ then} \qquad \lim_{\epsilon\downarrow 0} \frac{\log p_{\mathrm{c}}(\epsilon)}{\log\log(1/\epsilon)} = -1,$$

$$\text{if } \gamma > \frac{1}{2}, \text{ then} \qquad \lim_{\epsilon\downarrow 0} \frac{\log p_{\mathrm{c}}(\epsilon)}{\log \epsilon} = \gamma - \frac{1}{2}.$$

If f is in the C-class, the statement remains true in the case $\gamma > \frac{1}{2}$, and in the case $\gamma = \frac{1}{2}$ if the limit is replaced by a $\limsup_{\epsilon\downarrow 0}$ and the equality by '$\le$'.

Theorem 5.4 implies that the damaged network $(\mathsf{G}_n^{\epsilon}\colon n \in \mathbb{N})$ is not robust. But as $\lim_{\epsilon\downarrow 0} p_{\mathrm{c}}(\epsilon) = 0$ it is still 'asymptotically robust' for $\epsilon \downarrow 0$ in the sense that when less than order n old vertices are destroyed, then the critical percolation parameter remains zero. We formulate this as a corollary. For two graphs $\mathsf{G} = (\mathsf{V},\mathsf{E})$ and $\tilde{\mathsf{G}} = (\tilde{\mathsf{V}},\tilde{\mathsf{E}})$, we write $\mathsf{G} \ge \tilde{\mathsf{G}}$ if there is a coupling such that $\mathsf{V} \supseteq \tilde{\mathsf{V}}$ and $\mathsf{E} \supseteq \tilde{\mathsf{E}}$.

Corollary 5.5. *Let f be a L-class or C-class attachment rule with $\gamma \ge \frac{1}{2}$, and let $(m_n\colon n \in \mathbb{N})$ be a sequence of natural numbers with $\lim_{n\to\infty} m_n/n = 0$. The network $(\mathsf{G}_n^{(m_n)}\colon n \in \mathbb{N})$, consisting of the graphs G_n damaged by removal of the oldest m_n vertices along with all adjacent edges, is robust.*

Proof. Let $p \in (0,1)$. By Theorem 5.4, there exists $\epsilon > 0$ such that $p_{\mathrm{c}}(\epsilon) < p$. Choose $n_0 \in \mathbb{N}$ such that $m_n/n < \epsilon$ for all $n \ge n_0$. Then $\mathsf{G}_n^{(m_n)} \ge \mathsf{G}_n^{\epsilon}$ for all $n \ge n_0$, implying $\mathsf{G}_n^{(mn)}(p) \geq \mathsf{G}_n^{\epsilon}(p)$. Since $(\mathsf{G}_n^{\epsilon}(p) \colon n \in \mathbb{N})$ has a giant component, so does $(\mathsf{G}_n^{(mn)}(p) \colon n \in \mathbb{N})$. $\square$

Theorem 5.4 is derived from Theorem 5.1 using the monotonicity of the network in the attachment rule. Its appeal lies in the large class of functions to which it applies. The L-class attachment rules are all positive, bounded perturbations of linear functions. In Figure II-2 we see several examples: on the left a concave function which is also in the C-class, then a convex function and a function which is convex in one and concave

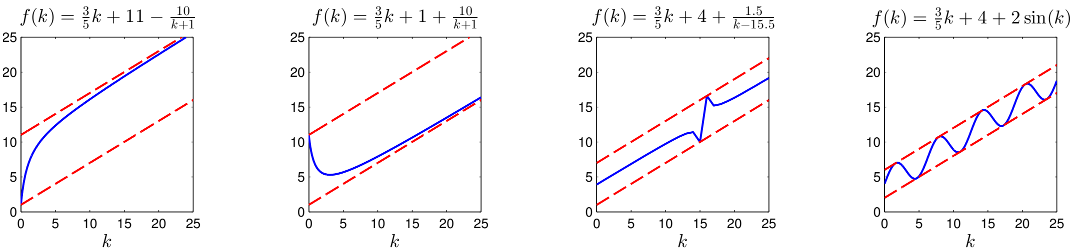

Figure II-2. Examples for L-class attachment rules. The blue curve is the attachment rule, the red, dashed lines are linear lower and upper bounds.

in another part of its domain. The latter examples are not monotone, and all three are asymptotically vanishing perturbations of an affine attachment rule. The example of an L-class attachment rule on the right shows that this may also fail.

The C-class attachment rules are always non-decreasing as positive concave functions and always have a linear lower bound with the same asymptotic slope γ as the function itself. However, when the perturbation $k \mapsto f(k) - \gamma k$ is not bounded, then there exists no linear function with slope γ which is an upper bound to the attachment rule; any linear upper bound will be steeper. Two examples are displayed in Figure II-3.

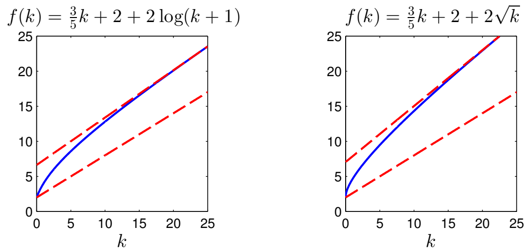

Figure II-3. Examples for C-class attachment rules. The blue curve is the attachment rule, the red, dashed lines are linear lower and upper bounds. The slope of the upper bound is strictly larger than γ.

5.5 Vulnerability of other network models

We would like to investigate to what extent our results are common to robust random network models rather than specific to preferential attachment networks. Again our focus is on Theorem 5.1 and we look at two types of networks, the configuration model and the inhomogeneous random graphs. Both types have an explicit static description, and are therefore much easier to analyse than the preferential attachment networks studied in our main theorems.

5.5.1 Configuration model

A targeted attack can be planned particularly well when the degree sequence of the network is known. A random graph model with fixed degree sequence is given by the configuration model. For $n \in \mathbb{N}$, denote by $\mathbf{d_n} = (d_i^n)_{i=1}^n \in \mathbb{N}^n$, with $\sum_{i=1}^n d_i^n$ even, the desired degrees. To simplify notation, we write d_i instead of d_i^n for the degree of vertex i at time n. The multigraph $\mathsf{G}_n^{(\mathrm{CM})}$ on vertex set $\{1, \dots, n\}$ is constructed as follows: to every vertex i attach d_i half-edges. Combine the half-edges into pairs by a uniformly random matching of the set of all half-edges. Each pair of half-edges is then joined to form an edge of $\mathsf{G}_n^{(\mathrm{CM})}$. The configuration model has received a lot of attention in the literature; see [85] and the references therein. A good targeted attack in the configuration model is the removal of the vertices with the highest degree, and we denote by $\mathsf{G}_n^{(\mathrm{CM}),\epsilon}$ the network after removal of the $\lfloor \epsilon n \rfloor$ vertices with the largest degree.

Let $n_k = |\{i \le n \colon d_i = k\}|$ be the number of vertices with degree k, and assume that there exists a $\mathbb{N}$-valued random variable D with $0 < \mathbb{E}D < \infty$ and $\mathbb{P}(D = 2) < 1$, such that, as $n \to \infty$,

$$n_k/n \to \mathbb{P}(D = k) \quad \text{for all } k \in \mathbb{N} \qquad \text{and} \qquad \frac{1}{n}\sum_{k=1}^{\infty} k n_k \to \mathbb{E}D. \tag{5.7}$$

In particular, the law of D is the weak limit of the empirical degree distribution in $(\mathsf{G}_n^{(\mathrm{CM})} \colon n \in \mathbb{N})$, and the network is robust if $\mathbb{E}[D^2] = \infty$; see [87, Theorem 3.5]. Our focus is on the case that the distribution of D is a power law with exponent $\tau = 1 + 1/\gamma$, $\gamma \ge \frac{1}{2}$.

Theorem 5.6. *Let $\epsilon \in (0, 1)$, $\gamma \in [\frac{1}{2}, 1)$ and suppose there is a constant $C > 0$ such $\mathbb{P}(D > k) \sim Ck^{-1/\gamma}$ as $k \to \infty$. Then there exists $p_{\mathrm{c}}^{(\mathrm{CM})}(\epsilon) > 0$ such that*

$$(\mathsf{G}_n^{(\mathrm{CM}),\epsilon}(p) \colon n \in \mathbb{N}) \textit{ has a giant component } \Leftrightarrow p > p_{\mathrm{c}}^{(CM)}(\epsilon).$$

Moreover,

$$p_{\mathrm{c}}^{(\mathrm{CM})}(\epsilon) \asymp \begin{cases} \frac{1}{\log(1/\epsilon)} & \text{if } \gamma = \frac{1}{2}, \\ \epsilon^{2\gamma - 1} & \text{if } \gamma > \frac{1}{2}. \end{cases}$$

We observe the same basic phenomenon as in the corresponding preferential attachment models: while the undamaged network is robust, after removal of an arbitrarily small proportion of privileged nodes the network becomes vulnerable to random removal of vertices. However, when $\gamma > \frac{1}{2}$, then the increase of the critical percolation parameter $p_{\mathrm{c}}(\epsilon)$ as ϵ leaves zero is less steep than in the corresponding preferential attachment model.

Note that our assumptions imply that $0 < \mathbb{E}D < \infty$. In the case $\mathbb{E}D = \infty$, Bhamidi et al. [15, Theorem 3.3] show a more extreme form of vulnerability, where the network can be disconnected with high probability by deleting a bounded number of vertices.

5.5.2 Inhomogeneous random graphs

Inhomogeneous random graphs are a generalisation of the classical Erdős–Rényi random graph. Let $\kappa\colon (0,1] \times (0,1] \to (0,\infty)$ be a symmetric kernel. The inhomogeneous random graph $\mathsf{G}_n^{(\kappa)}$ corresponding to kernel κ has the vertex set $\mathsf{V}_n = \{1,\dots,n\}$, and any pair of distinct vertices i and j is connected by an edge independently with probability

$$\mathbb{P}(\{i,j\} \text{ present in } \mathsf{G}_n^{(\kappa)}) = \tfrac{1}{n}\kappa\big(\tfrac{i}{n}, \tfrac{j}{n}\big) \wedge 1. \tag{5.8}$$

Many features of this model class are discussed by Bollobás, Janson and Riordan [20], and van der Hofstad [85]. The first inhomogeneous random graph model we consider is a version of the *Chung–Lu model*; see for example [29, 30, 31]. The relevant kernel is

$$\kappa^{(\mathrm{CL})}(x,y) = x^{-\gamma} y^{-\gamma} \qquad \text{for } x, y \in (0,1].$$

This is an example of a kernel of the form $\kappa(x,y) = \chi(x)\chi(y)$, for some χ, which are called kernels of *rank one*; see [20]. Note that a similar factorisation occurs in the configuration model since the probability that vertices i and j are directly connected is roughly proportional to $d_i d_j$. Therefore, the configuration model can be classified as a *rank one model*, too. By Theorem 3.13 and Corollary 13.1 in [20], the network corresponding to $\kappa^{(\mathrm{CL})}$ has an asymptotic degree distribution which is a power law with exponent $\tau = 1 + 1/\gamma$.

The second inhomogeneous random graph model we consider is chosen such that the edge probabilities agree (at least asymptotically, cf. Lemma 7.9 below) with those in a preferential attachment network, and the asymptotic degree distribution is a power law with exponent $\tau = 1 + 1/\gamma$. The relevant kernel is

$$\kappa^{(\mathrm{PA})}(x,y) = \frac{1}{(x \wedge y)^{\gamma}(x \vee y)^{1-\gamma}} \qquad \text{for } x, y \in (0,1].$$

Note that, if $\gamma \neq \frac{1}{2}$, this kernel is not of rank one but *strongly inhomogeneous*. The two kernels $\kappa^{(\mathrm{CL})}$ and $\kappa^{(\mathrm{PA})}$ allow us to demonstrate the difference between rank one models and preferential attachment models within one model class.

We denote by $\mathsf{G}_n^{(\mathrm{CL})}$ and $\mathsf{G}_n^{(\mathrm{PA})}$ the inhomogeneous random graphs with kernel $\kappa^{(\mathrm{CL})}$ and kernel $\kappa^{(\mathrm{PA})}$, respectively. If $\gamma \geq \frac{1}{2}$, then $(\mathsf{G}_n^{(\mathrm{CL})}\colon n \in \mathbb{N})$ and $(\mathsf{G}_n^{(\mathrm{PA})}\colon n \in \mathbb{N})$ are robust by Theorem 3.1 and Example 4.11 in [20]. Since the kernels $\kappa^{(\mathrm{CL})}$ and $\kappa^{(\mathrm{PA})}$ are decreasing in both components, vertices with small labels are favoured in the corresponding models. We denote by $\mathsf{G}_n^{(\mathrm{CL}),\epsilon}$ and $\mathsf{G}_n^{(\mathrm{PA}),\epsilon}$ what remains of the graphs $\mathsf{G}_n^{(\mathrm{CL})}$ and $\mathsf{G}_n^{(\mathrm{PA})}$, respectively, after removal of all vertices with label at most ϵn along with their adjacent edges.

The following theorem confirms that, like in the preferential attachment and in the configuration model, the removal of a positive fraction of key vertices makes the networks vulnerable to random removal of nodes. Notice that $\kappa^{(\mathrm{CL})}$ and $\kappa^{(\mathrm{PA})}$ agree for $\gamma = \frac{1}{2}$ so that we only have to state a result for $\mathsf{G}_n^{(\mathrm{CL}),\epsilon}$ in this regime.

Theorem 5.7. *Let $\gamma \in [\frac{1}{2}, 1)$, $\star \in \{\mathrm{CL}, \mathrm{PA}\}$, and $\epsilon \in (0,1)$. There exists $p_{\mathrm{c}}^{(\star)}(\epsilon) > 0$ such that*

$$(\mathsf{G}_n^{(\star),\epsilon}(p) \colon n \in \mathbb{N}) \text{ has a giant component } \Leftrightarrow\ p > p_{\mathrm{c}}^{(\star)}(\epsilon).$$

Moreover,

$$p_{\mathrm{c}}^{(\mathrm{CL})}(\epsilon) = \begin{cases} \frac{1}{\log(1/\epsilon)} & \text{if } \gamma = \frac{1}{2}, \\ (2\gamma - 1)\epsilon^{2\gamma-1}[1 + O(\epsilon^{2\gamma-1})] & \text{if } \gamma > \frac{1}{2}, \end{cases}$$

and

$$p_{\mathrm{c}}^{(\mathrm{PA})}(\epsilon) \asymp \epsilon^{\gamma-1/2} \qquad \text{if } \gamma > \tfrac{1}{2}.$$

The fact that the Chung–Lu model is vulnerable to targeted attacks has also been remarked by van der Hofstad in Section 9.1 of [85].

Summarising, we note that vulnerability to a targeted attack is a universal feature of robust networks, holding not only for preferential attachment networks but also for configuration models and various classes of inhomogeneous random graphs. In the case $2 < \tau < 3$, studying the asymptotic behaviour of the critical percolation parameter $p_{\mathrm{c}}(\epsilon)$ as a function of the proportion ϵ of removed vertices reveals *two universality classes* of networks, that are distinguished by the critical exponent measuring the polynomial rate of decay of $p_{\mathrm{c}}(\epsilon)$ as $\epsilon \downarrow 0$. In terms of the power law exponent τ, this critical exponent equals $\frac{3-\tau}{\tau-1}$ in the case of the configuration model and the Chung–Lu model, but is only half this value in the case of preferential attachment networks and inhomogeneous random graphs with a strongly inhomogeneous kernel. The same classification of networks has emerged in a different context in [35], where it was noted that the typical distances in networks of the two classes differ by a factor of two. The key feature of the configuration model and the rank one inhomogeneous random graphs seems to be that the connection probability of two vertices factorises. By contrast, the connection probabilities in preferential attachment networks have a more complex structure giving privileged nodes a stronger advantage.

5.6 The local neighbourhood in the network

Dereich and Mörters [37] have shown that the (not too large) graph neighbourhood of a uniformly chosen vertex in G_n can be coupled to a branching random walk on the negative half-line. Although we cannot make direct use of this coupling result in our proofs, it is helpful to formulate our ideas in this framework. Therefore, we now explain heuristically that a suitable exploration of the local neighbourhood of a given vertex $v_0 \in \mathsf{G}_n^\epsilon$ reveals a graph that can be approximated by the genealogical tree of a two-type branching random walk with two killing boundaries. A complete definition of the branching process used in our analysis is given in Section 6.1, and the coupling is proved rigorously in Chapter 8 below.

Firstly, we associate to every vertex in G_n a location on the negative half-line such that the youngest vertex is located at the origin, and the distance between vertex j and vertex $j+1$ is represented as $1/j$. In particular, the vertex labelled v is located at $s_n(v) := -\sum_{j=v}^{n-1} \frac{1}{j}$, the location of the oldest vertex scales like $-\log n$, and vertices with label at most $\lfloor \epsilon n \rfloor$, which we remove when damaging the network, are asymptotically located to the left of $\log \epsilon$. The location of a vertex is determined by its age in the network with old vertices being located further left than young vertices; Figure II-4 has a sketch. As the graph size increases, the location of any fixed vertex moves to the left and the vertex locations $(s_n(v) \colon v \in \{1, \dots, n\})$ become dense on the negative half-line.

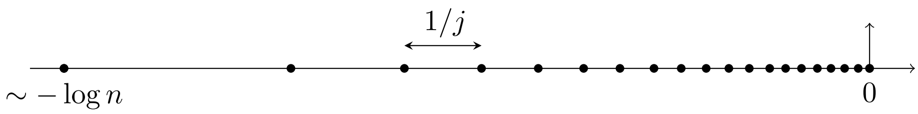

Figure II-4. Vertex locations if $n = 20$. Vertices are ordered from the oldest on the left to the youngest on the right.

We run an exploration from vertex $v_0 \in \mathsf{G}_n^\epsilon$ and successively create particles in the branching random walk that approximate the discovered vertices. We stop as soon as there is no longer a one-to-one correspondence between the nodes in the two processes. For example, this could happen if in the network a vertex is rediscovered and the explored subgraph is no longer a tree. A careful analysis, carried out in Section 8.1 below, shows that when the order in which vertices are explored is chosen in a suitable way, then we do not stop until either the whole component is discovered or at least c_n vertices have been found, where $\lim_{n\to\infty} c_n^2/n = 0$.

To start, we place a particle at the location of vertex v_0 and declare it to be the root of the branching random walk. Then we explore all direct neighbours of v_0 in G_n^ϵ. The locations of the particles in the first generation of the branching random walk are chosen to approximate the locations of these direct neighbours. To this end, we distinguish offspring located to the left and right of v_0. For a given interval $[a, b]$ on the left of $s_n(v_0)$, i.e. $[a, b] \subseteq [\log \epsilon, s_n(v_0)]$, the number of vertices located in $[a, b]$ is a sum of independent Bernoulli random variables by the definition of the model. Write $\mathcal{Z}[u, v]$ for the indegree of vertex u at time v, $\triangle\mathcal{Z}[u, v_0 - 1] = \mathcal{Z}[u, v_0] - \mathcal{Z}[u, v_0 - 1]$. The probability that v_0 has a direct neighbour labelled $u < v_0$, is given by

$$\mathbb{P}\big(\triangle\mathcal{Z}[u, v_0 - 1] = 1\big) = \frac{1}{v_0 - 1}\mathbb{E}\big[f(\mathcal{Z}[u, v_0 - 1])\big] = \frac{\beta}{v_0 - 1}\prod_{j=u}^{v_0-2}(1 + \gamma/j), \tag{5.9}$$

where the last equality follows from the fact that $(f(\mathcal{Z}[u,n])\prod_{j=u}^{n-1}\frac{1}{1+\gamma/j}: n \geq u)$ is a martingale by the definition of the network. The right-hand side of (5.9) can be approximated by

$$\frac{\beta}{u-1}e^{-(\log(v_0-1)-\log(u-1))}\prod_{j=u}^{v_0-2}e^{\gamma/j} \approx \frac{\beta}{1-\gamma}\big[1-e^{-(1-\gamma)/(u-1)}\big]e^{-(1-\gamma)\sum_{j=u}^{v_0-1}\frac{1}{j}}$$
$$= \beta\int_{-\sum_{j=u-1}^{v_0-1}\frac{1}{j}}^{-\sum_{j=u}^{v_0-1}\frac{1}{j}} e^{(1-\gamma)t}\,dt.$$

Hence,

$$\mathbb{P}\big(\triangle\mathcal{Z}[u,v_0-1]=1\big) \approx \int_{s_n(u-1)-s_n(v_0)}^{s_n(u)-s_n(v_0)} \beta e^{(1-\gamma)t}\,dt. \tag{5.10}$$

Since the location of u can be written as $s_n(v_0)$ plus the displacement $s_n(u)-s_n(v_0)$, asymptotically, we can approximate the displacements of the direct neighbours of v_0 on its left by the points of a Poisson point process Π with intensity measure $\beta e^{(1-\gamma)t}$ on $(-\infty,0]$ that lie in $[\log\epsilon - s_n(v_0), 0]$. We emphasise that Π describes the displacements, not the particle locations. Hence, in the branching random walk, the relative positions of the offspring to the left of a particle with location λ are given by the points of Π that lie in $[\log\epsilon-\lambda, 0]$.

In the next step, we motivate the point process that describes the relative positions of the offspring on the right in the branching random walk. Note that in the network every direct neighbour u of v_0 with $u > v_0$ increases the indegree of v_0 and therefore the probability that v_0 has further offspring on its right. The distance between the i-th and $(i+1)$-st right neighbour of v_0 in the network is given by

$$T_{v_0}[i] := \sup\Big\{\sum_{j=k}^{l}\frac{1}{j}: \mathcal{Z}[v_0,k]=i=\mathcal{Z}[v_0,l]\Big\}.$$

Suppose the i-th neighbour of v_0 is born at time k. For given $t>0$, we have

$$\mathbb{P}\big(T_{v_0}[i] > t \,|\, \mathcal{Z}[v_0,k-1] < \mathcal{Z}[v_0,k]=i\big) = \mathbb{P}\big(\mathcal{Z}[v_0,l]=i \,|\, \mathcal{Z}[v_0,k]=i\big), \tag{5.11}$$

where l is the smallest integer with $\sum_{j=k}^{l}\frac{1}{j} > t$. Plugging in the connection probabilities given in the model definition, we deduce that (5.11) is equal to

$$\prod_{j=k}^{l-1}\Big(1-\frac{f(i)}{j}\Big) \approx \exp\Big(-f(i)\sum_{j=k}^{l-1}\frac{1}{j}\Big) \approx \exp\big(-f(i)t\big).$$

Hence, the distance between the i-th and $(i+1)$-st right neighbour of v_0 is approximately exponentially distributed with rate $f(i)$. For a precise statement, see Lemma 7.2 below. Consequently, the displacements of the direct neighbours on the right of v_0 are well approximated by the jump times in $[0,-s_n(v_0)]$ of the pure jump Markov process

$Z = (Z_t : t \geq 0)$ that starts in $Z_0 = 0$ and jumps from i to $i+1$ after an exponential waiting time of rate $f(i)$, independently of the previous jumps. Therefore, in the branching random walk, the relative positions of the offspring to the right of the root, that is located in λ, are given by the jump times of Z in $[0, -\lambda]$.

When the exploration is continued, the information gathered from the already explored neighbourhood leads to a size-biasing effect. Indeed, in the network the edges between a vertex v and its direct neighbours on the right $u > v$ are not independent. If v was discovered as a direct neighbour of a vertex w on its right, i.e. $w > v$, then we already know that v has indegree at least one. Consequently, we expect v to have more direct neighbours on its right than without this information. Mathematically, this leads to a size biasing effect, and the displacements of particles on the right of v are given by the jump times in $[0, -s_n(v)]$ of Z started in *one* instead of zero. In contrast, if v was discovered as a direct neighbour of a vertex w with smaller label, i.e. $w < v$, then we do not have that information and the displacements are again the jump times in $[0, -s_n(v)]$ of Z started in zero. Similarly, for the direct neighbours on the left of v, there is no size-biasing effect as a consequence of the independence between the edges on the left. Of course, there are several further dependencies coming from the previously explored subgraph. However, we show in Chapter 8 that the error accrued by adjusting only for the immediate parent is asymptotically negligible when we discover not more than c_n vertices, where $\lim_{n\to\infty} c_n^3/n = 0$.

To be able to use different offspring distributions depending on the relative location of the parent, each vertex is equipped with a mark α in $\{\ell, \mathrm{r}\}$ to indicate the relative location of the parent, where the non-numerical symbols ℓ and r stand for 'left' and 'right', respectively. The relative positions of the offspring can be generated as the points of Π on $(-\infty, 0]$ and the jump times of Z (with initial state depending on the mark) on $[0, \infty)$. All offspring particles located on the left of $\log \epsilon$ or on the right of 0 are immediately removed. In other words, the approximating tree is the genealogical tree of a two-type branching random walk with two killing boundaries.

An equivalent description is as a multitype branching process with type space $\Phi := [\log \epsilon, 0] \times \{\ell, \mathrm{r}\}$, where the first component indicates the location of a particle and the second indicates its mark. Whilst the branching random walk interpretation offers more intuition, the two killing boundaries make the mathematical analysis difficult. Hence in our analysis, we will use the interpretation of the process as multitype branching process with the larger typespace Φ.

5.7 Main ideas of the proofs

Understanding the local neighbourhood of vertices in the network is the key to many of its properties. As in [37], the survival probability of the approximating killed branching random walk is equal to the asymptotic relative size ζ of the largest component. This result allows us to determine, for example, the critical parameter of percolation from knowledge about the survival probability of the percolated branching process. To form the percolated branching process with retention probability p from the original process every particle is kept with probability p and removed together with its line of descent with probability $1-p$ independently of all other particles.

It is instructive to continue the comparison of the damaged and undamaged networks in the setup of this branching process. In [37], where the undamaged network is analysed, the branching random walk has only *one* killing boundary on the right. It turns out that on the set of survival, the leftmost particle drifts away from the killing boundary, such that it does not feel the boundary anymore. As a consequence, the unkilled process carries all information needed to determine whether the killed branching random walk survives with positive probability and, therefore, whether the network has a giant component. The *two* killing boundaries in the branching random walk describing the damaged network prevent us from using this analogy; every particle is exposed to the threat of absorption.

To survive indefinitely, a genealogical line of descent has to move within the (space-time) strip $[\log \epsilon, 0] \times \mathbb{N}_0$. To understand the optimal strategy for survival, observe that, in the network with strong preferential attachment, old vertices typically have a large degree and therefore are connected to many young vertices while young vertices themselves have only a few connections. This means that in the branching random walk without killing, particles produce many offspring to the right but only a few to the left. Hence, if a particle is located near the left killing boundary, it represents an old vertex in the graph and is very fertile, but its offspring are mostly located further to the right and are therefore less fertile. A particle near the right killing boundary, however, represents a young vertex and has itself a small number of offspring, which then however have a good chance of being fertile since they are necessarily located further left in the strip. As a result, the optimal survival strategy for a particle is to have an ancestral line of particles whose locations are alternating between positions near the left and the right killing boundary. This intuition is the basis for our proofs.

Continuing more formally, in the proof of Theorem 5.1 we show that positivity of the survival probability can be characterised in terms of the largest eigenvalue ρ_ϵ of an operator that describes the spatial distribution of offspring of a given particle. The branching random walk survives percolation with retention parameter p if its growth rate $p\rho_\epsilon$ exceeds the value one, so that $p_c(\epsilon) = 1/\rho_\epsilon$. Our intuition allows us to guess the form of the eigenfunction corresponding to ρ_ϵ, which, relatively to the particle density, has its mass concentrated in two bumps near the left and right killing boundary. From this guess we obtain sufficiently accurate estimates for the largest eigenvalue, and therefore for the critical percolation parameter, as long as the preferential attachment effect is strong enough. This is the case if $\gamma \geq \frac{1}{2}$, allowing us to prove Theorem 5.1.

By contrast, for $\gamma < \frac{1}{2}$ we know that the network is not robust, i.e. we have $p_c(0) > 0$. It would be of interest to understand the behaviour of $p_c(\epsilon) - p_c(0)$ as $\epsilon \downarrow 0$.

Our methods can be applied to this case, but the resulting bounds are very rough. The reason is that in this regime the preferential attachment is much weaker, and the intuitive idea underlying our estimates gives a less accurate picture.

The idea for the proof of Theorem 5.3 is based on the branching process comparison, too. To bound the probability that two typical vertices V and W are connected by a path of length at most h, we study the expected number of such paths. That is given by the number of paths of length at most $h-1$ starting from V multiplied by the probability that the terminal vertex in the path connects to W. By our branching process heuristics, the number of such paths can be approximated by the number of particles in the first $h-1$ generations of the branching random walk, which is of order ρ_ϵ^h where $\rho_\epsilon = 1/p_c(\epsilon)$ as before. The probability of connecting any vertex with label at least ϵn to W is bounded from above by $f(m)/\epsilon n$, where m is the maximal degree in the network. Since $m = o(n)$ by Theorem 5.2, this implies that the probability of a connection between V and W is bounded from above by $\exp(h \log(1/p_c(\epsilon)) - \log n + o(\log n))$ and therefore goes to zero if $h \leq (1-\delta) \log n / \log(1/p_c(\epsilon))$, $\delta > 0$, which yields the result.

Theorem 5.2 is relatively soft by comparison. The independence of the indegrees of distinct vertices allows us to study them separately, and we again use the continuous approximation to describe the expected empirical indegree evolution. The limit theorem for the empirical distribution itself follows from a standard concentration argument. The asymptotic result for the maximal degrees is only slightly more involved and is based on fairly standard extreme value arguments.

5.8 Overview

The outline of Part II of this thesis is as follows. We start with the main steps of the proofs in Chapter 6. The multitype branching process, which locally approximates a connected component in the network, is defined in Section 6.1, and its key properties are stated. The main part of the proof of Theorem 5.1 then follows in Section 6.2. The analysis of the multitype branching process is conducted in Section 6.3. Chapter 7 is devoted to the study of the topology of the damaged graph. In Section 7.1 the typical and maximal degree of vertices is analysed; in Section 7.2 typical distances are studied. The couplings between the network and the approximating branching process that underlie our proofs are provided in Chapter 8. We then look at model variations in Chapter 9. The derivation of Theorem 5.4 from Theorem 5.1 is presented in Section 9.1. This is the only section which requires consideration of non-linear attachment rules. We finish in Section 9.2 by studying the question of vulnerability in other network models.

CHAPTER 6

CONNECTIVITY AND BRANCHING PROCESSES

From here until the end of Chapter 8, we restrict our attention to linear attachment rules $f(k) = \gamma k + \beta$, for $\gamma \in [0, 1)$ and $\beta > 0$. Unless stated otherwise, ϵ is a fixed value in $(0, 1)$. The goal of this chapter is to prove Theorem 5.1. To this end, we couple the local neighbourhood of a vertex in G_n^ϵ to a multitype branching process. The branching process is introduced in Section 6.1, and Theorem 5.1 is deduced in Section 6.2. Properties of the branching processes which are needed in the analysis are proved in Section 6.3. The proof of the coupling between network and branching process is deferred to Chapter 8.

6.1 The approximating branching process

As explained in Section 5.6, the local neighbourhood of a vertex in G_n^ϵ can be approximated by a multitype branching process with type space $\Phi = [\log \epsilon, 0] \times \{\ell, \mathrm{r}\}$. A typical element of Φ is denoted by $\phi = (\lambda, \alpha)$. The intuitive picture is that λ encodes the spatial position of the particle which we call *location*. The second coordinate α indicates on which side of the particle its parent is located, and we refer to α as the *mark*. In view of (5.10), a particle of type $(\lambda, \alpha) \in \Phi$ produces offspring to its left with displacements having the same distribution as those points of the Poisson point process Π on $(-\infty, 0]$ with intensity measure

$$\beta e^{(1-\gamma)t} \mathbb{1}_{(-\infty,0]}(t)\, dt \tag{6.1}$$

that lie in $[\log \epsilon - \lambda, 0]$. Since these offspring have their parent on the right, they are of mark r.

We denote by Z an increasing, integer-valued process, which jumps from i to $i + 1$ after an exponential waiting time of rate $f(i)$, independently of the previous jumps. We write P for the distribution of Z started in zero and E for the corresponding expectation. By $(\hat{Z}_t \colon t \geq 0)$ we denote a version of the process started in $\hat{Z}_0 = 1$ under the measure P.

The distribution of the offspring to the right depends on the mark of the parent. As motivated in Section 5.6, when the particle is of type (λ, ℓ), then the displacements of the offspring follow the same distribution as the jump times of $(Z_t \colon t \in [0, -\lambda])$, but when the particle is of type (λ, r), then the displacements follow the same distribution as the jump times of $(\hat{Z}_t \colon t \in [0, -\lambda])$. All offspring on the right have their parent on the left, so their mark is ℓ. Observe that the chosen offspring distributions ensure that new particles have again a location in $[\log \epsilon, 0]$. The offspring distribution to the right is *not* a Poisson point process: the more particles are born, the higher the rate at which new particles arrive.

We call the branching process thus constructed the *idealized branching process* (IBP). It can be interpreted as a labelled tree, where every node represents a particle and is connected to its children and (apart from the root) to its parent. We equip node x with label $\phi(x) = (\lambda(x), \alpha(x))$, where $\lambda(x)$ denotes its location and $\alpha(x)$ its mark, and write $|x|$ for the generation of x. To obtain a branching process approximation to the percolated graph $\mathsf{G}_n^\epsilon(p)$, we define the *percolated* IBP by associating to every offspring in the IBP an independent Bernoulli(p) random variable. If the random variable is zero, then we delete the offspring together with its line of descent. If it equals one, then the offspring is retained in the percolated IBP.

Let S^ϵ be a random variable such that e^{S^ϵ} is uniformly distributed on $[\epsilon, 1]$, that is,

$$\mathbb{P}(-S^\epsilon \leq t) = \frac{1}{1-\epsilon}\big(1 - e^{-t}\big) \qquad \text{for } t \in [0, -\log \epsilon]. \tag{6.2}$$

Recalling the definitions from the beginning of Section 5.6, the location of a uniformly chosen vertex in V_n^ϵ converges weakly to S^ϵ. Denote by $\zeta^\epsilon(p)$ the survival probability of the tree which is with probability p equal to the genealogical tree of the percolated IBP started with one particle of mark ℓ in location S^ϵ, and equals the empty tree otherwise. Let $\mathsf{C}_n^\epsilon(p)$ be a connected component in $\mathsf{G}_n^\epsilon(p)$ of maximal size.

Theorem 6.1. *For all $\epsilon \in (0,1)$ and $p \in (0,1]$, in probability,*

$$\frac{|\mathsf{C}_n^\epsilon(p)|}{\mathbb{E}|\mathsf{V}_n^\epsilon(p)|} \to \zeta^\epsilon(p)/p \qquad \text{as } n \to \infty.$$

The proof of Theorem 6.1 is postponed to Chapter 8. The theorem describes the asymptotic size of the largest component in the network in terms of the survival probability of the percolated IBP. To make use of this connection, we have to understand the branching process.

For any measurable, complex-valued, bounded function g on Φ, and $\phi \in \Phi$, let

$$A_p g(\phi) := E_{\phi,p}\Big[\sum_{|x|=1} g(\lambda(x), \alpha(x))\Big],$$

where the expectation $E_{\phi,p}$ refers to the percolated IBP starting with a single particle of type ϕ, percolated with retention probability p. We write $A = A_1$ for the operator corresponding to the unpercolated branching process and $E_\phi := E_{\phi,1}$. Recall that all quantities associated with the IBP, and in particular A_p, depend on the fixed value of ϵ. We denote by $C(\Phi)$ the complex Banach space of continuous functions on Φ equipped with the supremum norm. The following proposition, which summarizes properties of A_p, is proved in Section 6.3.1.

Proposition 6.2. *For all $\epsilon \in (0,1)$ and $p \in (0,1]$, the operator $A_p\colon C(\Phi) \to C(\Phi)$ is linear, strictly positive and compact with spectral radius $\rho_\epsilon(A_p) \in (0,\infty)$. Moreover, $A_p = pA$ and $\rho_\epsilon(A_p) = p\rho_\epsilon(A)$.*

The survival probability of the percolated IBP has the following property.

Theorem 6.3. *For all $\epsilon \in (0,1)$ and $p \in (0,1]$,*

$$\zeta^\epsilon(p) > 0 \qquad \Leftrightarrow \qquad \rho_\epsilon(A_p) > 1.$$

Theorem 6.3 is proved in Section 6.3.2. Combined with Theorem 6.1 and Proposition 6.2, it gives a characterisation of the critical percolation parameter for the network $(\mathsf{G}_n^\epsilon(p)\colon n \in \mathbb{N})$.

Corollary 6.4. *The network $(\mathsf{G}_n^\epsilon(p)\colon n \in \mathbb{N})$ has a giant component if and only if $p > \rho_\epsilon(A)^{-1}$.*

Notice that the corollary implies that $(\mathsf{G}_n^\epsilon\colon n \in \mathbb{N})$ has no giant component when $\rho_\epsilon(A) \le 1$. Moreover, the first statement of Theorem 5.1 follows from the corollary by taking $p_c(\epsilon) = \rho_\epsilon(A)^{-1} \wedge 1$.

To complete the proof of Theorem 5.1, it remains to estimate the spectral radius $\rho_\epsilon(A)$. This estimation is performed in Section 6.2 below using that (see, e.g., Theorem 45.1 in [84]) for a linear and bounded operator A on a complex Banach space, the spectral radius is given by

$$\rho(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} = \inf\{\|A^n\|^{\frac{1}{n}} : n \in \mathbb{N}\}. \tag{6.3}$$

By the definition of the Poisson point process Π in (6.1), the intensity measure of Π equals

$$\mathbb{1}_{(-\infty,0]}(t)\, M(dt), \qquad \text{for } M(dt) := \beta e^{(1-\gamma)t}\, dt.$$

We denote by Π^ℓ the point process given by the jump times of $(Z_t\colon t \ge 0)$ and by Π^{r} the point process given by the jump times of $(\hat{Z}_t\colon t \ge 0)$. A simple computation (cf. Lemma 1.12 in [37]) shows that with $M^\alpha(dt) := a_\alpha e^{\gamma t}\, dt$, where $a_\ell = \beta$ and $a_{\mathrm{r}} = \gamma + \beta$, the intensity measure of Π^α is given by $\mathbb{1}_{[0,\infty)}(t) M^\alpha(dt)$ for $\alpha \in \{\ell, \mathrm{r}\}$. Hence, for any bounded, measurable function g on Φ and $(\lambda,\alpha)\in\Phi$,

$$\begin{aligned} Ag(\lambda,\alpha) &= E_{(\lambda,\alpha)}\Big[\sum_{|x|=1} g(\lambda(x),\alpha(x))\Big] \\ &= \int_{\log\epsilon-\lambda}^{0} g(\lambda+t,\mathrm{r})\beta e^{(1-\gamma)t}\,dt + \int_0^{-\lambda} g(\lambda+t,\ell)a_\alpha e^{\gamma t}\,dt. \end{aligned} \tag{6.4}$$

6.2 Proof of Theorem 5.1

Subject to the considerations of the previous section, Theorem 5.1 follows from the following proposition.

Proposition 6.5. *(a) If $\gamma=\frac12$, then* $\dfrac{1}{\gamma+\beta}\dfrac{1}{\log(1/\epsilon)} \le \rho_\epsilon(A)^{-1} \le \dfrac{1}{\beta}\dfrac{1}{\log(1/\epsilon)}$.

(b) If $\gamma>\frac12$, then

$$\begin{aligned} &\Big(1+\log(\epsilon^{1-2\gamma})\epsilon^{\gamma-1/2}+\big[\log(\epsilon^{1-2\gamma})\epsilon^{\gamma-1/2}\big]^2\Big)^{-1/2} \\ &\qquad\qquad \le \frac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1}\epsilon^{-\gamma+1/2}\rho_\epsilon(A)^{-1} \le \big(1-\epsilon^{\gamma-1/2}\big)^{-1}. \end{aligned}$$

Proof of Proposition 6.5 (a). Denote by $C([\log\epsilon,0])$ the complex Banach space of continuous functions on $[\log\epsilon,0]$. For all $h_0\in C([\log\epsilon,0])$, $\lambda\in[\log\epsilon,0]$, let

$$\bar{A}h_0(\lambda) := \int_{\log\epsilon-\lambda}^{0} h_0(\lambda+t)e^{t/2}\,dt + \int_0^{-\lambda} h_0(\lambda+t)e^{t/2}\,dt.$$

Note that A and $\bar{A}$ map real-valued functions to real-valued functions and nonnegative functions to nonnegative functions, and they are monotone. For A this observation implies

$$\|A^n\| = \sup\big\{\|A^n g\| \colon g\in C(\Phi), g \text{ is } [0,1]\text{-valued}\big\} = \|A^n\mathbf{1}\|, \tag{6.5}$$

where $\|A^n g\| = \sup\{|A^n g(\phi)| \colon \phi\in\Phi\}$ and $\mathbf{1}$ denotes the constant function with value 1. Combining (6.3) and (6.5), we deduce that $\rho_\epsilon(A)=\lim_{n\to\infty}\|A^n\mathbf{1}\|^{\frac1n}$. The same argument shows that

$$\rho_\epsilon(\bar{A}) = \lim_{n\to\infty}\|\bar{A}^n\mathbf{1}\|^{\frac1n}. \tag{6.6}$$

Defining $h(\lambda,\alpha):=h_0(\lambda)$ for all $(\lambda,\alpha)\in\Phi$, (6.4) yields for $h_0\in C([\log\epsilon,0])$ with $h_0\ge 0$,

$$\beta\bar{A}h_0(\lambda)\le Ah(\lambda,\alpha)\le(\gamma+\beta)\bar{A}h_0(\lambda) \qquad \text{for all } (\lambda,\alpha)\in\Phi.$$

In particular, by the monotonicity and linearity of A and $\bar{A}$,

$$\beta^n\bar{A}^n\mathbf{1}(\lambda)\le A^n\mathbf{1}(\lambda,\alpha)\le(\gamma+\beta)^n\bar{A}^n\mathbf{1}(\lambda) \qquad \text{for } (\lambda,\alpha)\in\Phi, n\in\mathbb{N},$$

implying $\rho_\epsilon(A) \in [\beta\rho_\epsilon(\bar{A}), (\gamma+\beta)\rho_\epsilon(\bar{A})]$. To complete the proof it suffices to show that $\rho_\epsilon(\bar{A}) = \log(1/\epsilon)$, which we can achieve by 'guessing' the principal eigenfunction of $\bar{A}$. Indeed, the result follows from (6.6) and

$$\bar{A}^{n+1}\mathbf{1}(\lambda) = 2(1-\epsilon^{1/2})(\log(1/\epsilon))^n e^{-\lambda/2} \qquad \text{for } \lambda \in [\log\epsilon, 0], n \in \mathbb{N}_0. \tag{6.7}$$

We show (6.7) by induction over n. For $n = 0$,

$$\bar{A}\mathbf{1}(\lambda) = \int_{\log\epsilon-\lambda}^{0} e^{t/2}\,dt + \int_0^{-\lambda} e^{t/2}\,dt = 2(1-e^{-\lambda/2}\epsilon^{1/2}+e^{-\lambda/2}-1) = 2(1-\epsilon^{1/2})e^{-\lambda/2}.$$

Moreover, with $h_0(\lambda) := e^{-\lambda/2}$ we have

$$\bar{A}h_0(\lambda) = \int_{\log\epsilon-\lambda}^{0} e^{-(\lambda+t)/2}e^{t/2}\,dt + \int_0^{-\lambda} e^{-(\lambda+t)/2}e^{t/2}\,dt = e^{-\lambda/2}\log(1/\epsilon).$$

Thus, $\lambda \mapsto e^{-\lambda/2}$ is an eigenfunction of $\bar{A}$ with eigenvalue $\log(1/\epsilon)$, and (6.7) follows. $\square$

Proof of the lower bound in Proposition 6.5 (b). We analyse the ancestral lines of particles in the branching process at a fixed time $n \geq 2$. Going back two steps in the ancestral line of every particle alive, we can divide the population at time n in four groups depending on the relative positions of parent and child in the transitions from generation $n-2$ to $n-1$ and from $n-1$ to n: (1) in both steps the child is to the left of its parent, (2) in the first step the child is to the left and in the second it is to the right of its parent, (3) first right, then left, (4) in both steps the child is to the right of its parent. The cases are depicted in Figure II-5.

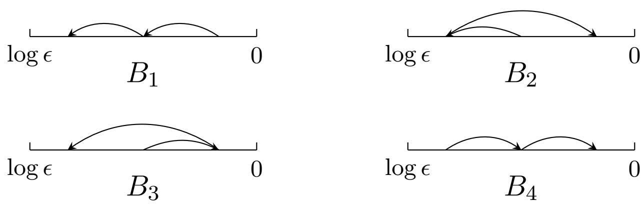

Figure II-5. Possible genealogy of a particle contributing to the respective operators.

We denote by B_i, $i \in \{1,\dots,4\}$, the mean operators corresponding to the four scenarios. Using the point processes Π, Π^ℓ and Π^{r}, for any bounded, measurable function g on Φ and any type $(\lambda,\alpha) \in \Phi$,

$$B_1 g(\lambda,\alpha) := \mathbb{E}\Big[\sum_{\substack{\mathfrak{p}\in\Pi \\ \log\epsilon-\lambda\leq\mathfrak{p}}} \sum_{\substack{\mathfrak{q}\in\Pi \\ \log\epsilon-\lambda-\mathfrak{p}\leq\mathfrak{q}}} g(\lambda+\mathfrak{p}+\mathfrak{q},\mathrm{r})\Big],$$

$$B_2 g(\lambda,\alpha) := \mathbb{E}\Big[\sum_{\substack{\mathfrak{p}\in\Pi \\ \log\epsilon-\lambda\leq\mathfrak{p}}} \sum_{\substack{\mathfrak{q}\in\Pi^{\mathrm{r}} \\ \mathfrak{q}\leq-(\lambda+\mathfrak{p})}} g(\lambda+\mathfrak{p}+\mathfrak{q},\ell)\Big],$$

$$B_3 g(\lambda,\alpha) := \mathbb{E}\Big[\sum_{\substack{\mathfrak{p}\in\Pi^\alpha \\ \mathfrak{p}\leq-\lambda}} \sum_{\substack{\mathfrak{q}\in\Pi \\ \log\epsilon-\lambda-\mathfrak{p}\leq\mathfrak{q}}} g(\lambda+\mathfrak{p}+\mathfrak{q},\mathrm{r})\Big],$$

$$B_4 g(\lambda, \alpha) := \mathbb{E}\Big[\sum_{\substack{\mathfrak{p} \in \Pi^\alpha \\ \mathfrak{p} \le -\lambda}} \sum_{\substack{\mathfrak{q} \in \Pi^\ell \\ \mathfrak{q} \le -(\lambda+\mathfrak{p})}} g(\lambda + \mathfrak{p} + \mathfrak{q}, \ell) \Big].$$

Intuitively, going back the ancestral line of a typical particle in the population at a late time, for a few generations the ancestral particles may be in group (4), because of the high fertility of particles positioned near the left boundary of $[\log \epsilon, 0]$. But this behaviour is not sustainable, as after a few generations in this group the offspring particle will typically be near the right end of the interval and will therefore be pushed into the right killing boundary so that it is likely to die out. Over a longer period the ancestral particles are much more likely to be in groups (2) and (3), as this behaviour is sustainable over long periods when the ancestral line is hopping more or less regularly between positions near the left and the right boundary of the interval $[\log \epsilon, 0]$. A similar pattern can also be observed when studying typical paths in the random graph model; see our discussion in Section 5.7. The aim is now to turn this heuristics into useful bounds on high iterates of the operator A.

It is helpful to understand how the operators B_i act on the constant function $\mathbf{1}$ as well as on the functions $g_1(\lambda, \alpha) := e^{-\gamma\lambda}$ and $g_2(\lambda, \alpha) := e^{-(1-\gamma)\lambda}$. We can write

$$B_3 g(\lambda, \alpha) = \int_0^{-\lambda} \int_{\log \epsilon - \lambda - t}^{0} g(\lambda + t + s, \mathrm{r})\, M(ds)\, M^\alpha(dt),$$

where $M(dt) = \beta e^{(1-\gamma)t}\, dt$ and $M^\alpha(dt) = a_\alpha e^{\gamma t}\, dt$ with $a_\alpha \le \gamma + \beta$ are the intensity measures of the point processes Π and Π^α. From this we obtain, for $(\lambda, \alpha) \in \Phi$,

$$\begin{aligned}
B_3 \mathbf{1}(\lambda, \alpha) &\le \int_{-\infty}^{-\lambda} \int_{-\infty}^{0} M(ds)\, M^{\mathrm{r}}(dt) = \frac{\beta(\gamma+\beta)}{\gamma(1-\gamma)} e^{-\gamma\lambda}, \\
B_3 g_1(\lambda, \alpha) &\le \int_{-\infty}^{-\lambda} \int_{\log \epsilon - t - \lambda}^{\infty} e^{-\gamma(\lambda+t+s)}\, M(ds)\, M^{\mathrm{r}}(dt) = \frac{\beta(\gamma+\beta)}{(2\gamma-1)^2} \epsilon^{1-2\gamma} e^{-\gamma\lambda}, \\
B_3 g_2(\lambda, \alpha) &\le \int_{-\infty}^{-\lambda} \int_{\log \epsilon}^{0} e^{-(1-\gamma)(\lambda+t+s)}\, M(ds)\, M^{\mathrm{r}}(dt) = \frac{\beta(\gamma+\beta)}{2\gamma-1} \log(1/\epsilon) e^{-\gamma\lambda}.
\end{aligned}$$

Moreover, similarly elementary calculations for B_1, B_2 and B_4 imply

$$\begin{aligned}
B_1 \mathbf{1}(\lambda, \alpha) &\le \tfrac{\beta^2}{(1-\gamma)^2}, & B_2 \mathbf{1}(\lambda, \alpha) &\le \tfrac{\beta(\gamma+\beta)}{\gamma(2\gamma-1)} \epsilon^{1-2\gamma} e^{-(1-\gamma)\lambda}, \\
B_4 \mathbf{1}(\lambda, \alpha) &\le \tfrac{\beta(\gamma+\beta)}{\gamma} \log(1/\epsilon) e^{-\gamma\lambda},
\end{aligned}$$

and

$$\begin{aligned}
B_1 g_1(\lambda, \alpha) &\le \tfrac{\beta^2}{2\gamma-1} \log(1/\epsilon) \epsilon^{1-2\gamma} e^{-(1-\gamma)\lambda}, & B_1 g_2(\lambda, \alpha) &\le \beta^2 (\log \epsilon)^2 e^{-(1-\gamma)\lambda}, \\
B_2 g_1(\lambda, \alpha) &\le \tfrac{\beta(\gamma+\beta)}{2\gamma-1} \log(1/\epsilon) \epsilon^{1-2\gamma} e^{-(1-\gamma)\lambda}, & B_2 g_2(\lambda, \alpha) &\le \tfrac{\beta(\gamma+\beta)}{(2\gamma-1)^2} \epsilon^{1-2\gamma} e^{-(1-\gamma)\lambda}, \\
B_4 g_1(\lambda, \alpha) &\le \beta(\gamma+\beta) (\log \epsilon)^2 e^{-\gamma\lambda}, & B_4 g_2(\lambda, \alpha) &\le \tfrac{\beta(\gamma+\beta)}{2\gamma-1} \log(1/\epsilon) e^{-\gamma\lambda}.
\end{aligned}$$

Summarising, there exists $C_\epsilon > 0$ such that $B_i\mathbf{1}(\phi) \le C_\epsilon g_1(\phi)$ for all $i \in \{1,\dots,4\}$, $\phi \in \Phi$, and denoting

$$b_{\mathrm{sm}} := b_1 := b_4 := \beta(\gamma+\beta)(\log\epsilon)^2, \qquad b_{\mathrm{bg}} := b_2 := b_3 := \tfrac{\beta(\gamma+\beta)}{(2\gamma-1)^2}\epsilon^{1-2\gamma},$$

where sm stands for 'small' and bg for 'big', we have

$$\begin{aligned} &B_i g_1(\phi) \le b_{\mathrm{bg}} \log(\epsilon^{1-2\gamma}) g_2(\phi), && B_i g_2(\phi) \le b_i g_2(\phi) && \text{for } i \in \{1,2\},\\ &B_i g_1(\phi) \le b_i g_1(\phi), && B_i g_2(\phi) \le b_{\mathrm{bg}} \log(\epsilon^{1-2\gamma})\epsilon^{2\gamma-1} g_1(\phi) && \text{for } i \in \{3,4\}. \end{aligned}$$

Using that by definition $A^2 = \sum_{i=1}^4 B_i$, our estimate for $B_i\mathbf{1}$ and monotonicity of B_i yield

$$\begin{aligned} A^{2(n+1)}\mathbf{1}(\phi) &= \sum_{i_0,\dots,i_n\in\{1,\dots,4\}} B_{i_n}\circ\cdots\circ B_{i_0}\mathbf{1}(\phi)\\ &\le 4C_\epsilon \sum_{i_1,\dots,i_n\in\{1,\dots,4\}} B_{i_n}\circ\cdots\circ B_{i_1} g_1(\phi). \end{aligned} \tag{6.8}$$

Up to constants, the estimates for B_3 and B_4 preserve g_1 but change g_2 into g_1, whereas the estimates for B_1 and B_2 preserve g_2 and change g_1 into g_2. Hence, we split the sequence of indices into blocks containing only 1 or 2 and blocks containing only 3 or 4. We write m for the number of blocks, k_j for the length of block j and $\bar{k}_j := \sum_{i=1}^{j-1} k_i + 1$ for the first index in block j. Then

$$\begin{aligned} &\sum_{i_1,\dots,i_n\in\{1,\dots,4\}} B_{i_n}\circ\cdots\circ B_{i_1} g_1(\phi)\\ &\qquad\qquad = \sum_{m=1}^{n+1} \sum_{\substack{k_1+\ldots+k_m=n\\ k_1\in\mathbb{N}_0,k_2,\ldots,k_m\in\mathbb{N}}} \sum_{(i_1,\ldots,i_n)} B_{i_n}\circ\cdots\circ B_{i_1} g_1(\phi), \end{aligned} \tag{6.9}$$

where the last sum is over all sequences of indices $(i_1,\dots,i_n)$ with $i_{\bar{k}_j},\dots,i_{\bar{k}_{j+1}-1} \in \{3,4\}$ for j odd and $i_{\bar{k}_j},\dots,i_{\bar{k}_{j+1}-1} \in \{1,2\}$ for j even. We insist that formally the first block contains the indices 3 or 4 — the case that this does not hold is covered by $k_1 = 0$. Hence, in the first block, operators B_3 and B_4 encounter g_1, which is preserved. To determine the constants, we only have to keep track of how often B_4 is used; we call this number l_1. The first operator belonging to a new block j causes a factor $b_{\mathrm{bg}}\log(\epsilon^{1-2\gamma})$ and if the change is from a $\{1,2\}$ to a $\{3,4\}$ block, then an additional $\epsilon^{2\gamma-1}$ is obtained. For the subsequent steps within block j, we again have to track how often the operator causing the smaller constant b_{sm}, B_1 or B_4, is used. This number is called l_j. After applying all n operators, the function $g_1(\phi)\mathbb{1}_{\mathrm{odd}}(m) + g_2(\phi)\mathbb{1}_{\mathrm{even}}(m)$ remains and is bounded it by $\epsilon^{-\gamma}$. This sequence of estimates allows us to upper bound the right-hand side of (6.9) by

$$\begin{aligned}
&\sum_{m=1}^{n+1} \sum_{\substack{k_1+\ldots+k_m=n\\ k_1\in\mathbb{N}_0,k_2,\ldots,k_m\in\mathbb{N}}} b_{\mathrm{bg}}^{m-1}(\log(\epsilon^{1-2\gamma}))^{m-1}\epsilon^{(2\gamma-1)(\lceil\frac{m}{2}\rceil-1)}\epsilon^{-\gamma}\\
&\qquad\qquad \cdot \sum_{l_1=0}^{k_1}\Big[\binom{k_1}{l_1} b_{\mathrm{sm}}^{l_1} b_{\mathrm{bg}}^{k_1-l_1}\Big] \prod_{j=2}^{m}\sum_{l_j=0}^{k_j-1}\Big[\binom{k_j-1}{l_j} b_{\mathrm{sm}}^{l_j} b_{\mathrm{bg}}^{k_j-1-l_j}\Big]\\
&= \epsilon^{-\gamma}\sum_{m=1}^{n+1} \sum_{\substack{k_1+\ldots+k_m=n\\ k_1\in\mathbb{N}_0,k_2,\ldots,k_m\in\mathbb{N}}} b_{\mathrm{bg}}^{m-1}(\log(\epsilon^{1-2\gamma}))^{m-1}\epsilon^{(2\gamma-1)(\lceil\frac{m}{2}\rceil-1)}(b_{\mathrm{sm}}+b_{\mathrm{bg}})^{k_1}\\
&\qquad\qquad \cdot \prod_{j=2}^{m}(b_{\mathrm{sm}}+b_{\mathrm{bg}})^{k_j-1}\\
&= \epsilon^{-\gamma}\sum_{m=1}^{n+1} \sum_{\substack{k_1+\ldots+k_m=n\\ k_1\in\mathbb{N}_0,k_2,\ldots,k_m\in\mathbb{N}}} b_{\mathrm{bg}}^{m-1}(\log(\epsilon^{1-2\gamma}))^{m-1}\epsilon^{(2\gamma-1)(\lceil\frac{m}{2}\rceil-1)}(b_{\mathrm{sm}}+b_{\mathrm{bg}})^{n-(m-1)}.
\end{aligned}$$

Given m, the number of configurations $k_1\in\mathbb{N}_0, k_2,\ldots,k_m\in\mathbb{N}$ with $k_1+\ldots+k_m=n$ is the number of arrangements of $m-1$ dividers and $n-(m-1)$ balls, which equals $\binom{n}{m-1}$. Since $\lceil\frac{m}{2}\rceil-1 \geq \frac{m-1}{2}-\frac{1}{2}$, an application of the binomial theorem yields

$$\sum_{i_1,\ldots,i_n\in\{1,\ldots,4\}} B_{i_n}\circ\cdots\circ B_{i_1}g_1(\phi) \leq \epsilon^{-2\gamma+1/2}\big(b_{\mathrm{bg}}\log(\epsilon^{1-2\gamma})\epsilon^{\gamma-1/2}+b_{\mathrm{bg}}+b_{\mathrm{sm}}\big)^n. \quad (6.10)$$

Combining (6.8) and (6.10), we conclude that, for all $\phi\in\Phi$,

$$A^{2(n+1)}\mathbf{1}(\phi) \leq 4C_\epsilon\epsilon^{-2\gamma+1/2}b_{\mathrm{bg}}^n\big(\log(\epsilon^{1-2\gamma})\epsilon^{\gamma-1/2}+1+\tfrac{b_{\mathrm{sm}}}{b_{\mathrm{bg}}}\big)^n.$$

Now (6.3) yields, for all $\epsilon\in(0,1)$,

$$\rho_\epsilon(A)^{-1} \geq \tfrac{2\gamma-1}{\sqrt{\beta(\gamma+\beta)}}\epsilon^{\gamma-1/2}\Big(1+\log(\epsilon^{1-2\gamma})\epsilon^{\gamma-1/2}+\big[\log(\epsilon^{1-2\gamma})\epsilon^{\gamma-1/2}\big]^2\Big)^{-1/2}. \qquad \square$$

The insight gained in the proof of the lower bound, enables us to 'guess' an approximating eigenfunction, which is the main ingredient in the proof of the upper bound.

Proof of the upper bound in Proposition 6.5 (b). Let $c_{\mathrm{r}} := 1$ and $c_\ell := \beta/(\gamma+\beta)$ and, for $(\lambda,\alpha)\in\Phi$, let

$$g_{\mathrm{e}}(\lambda,\alpha) := c_\alpha\epsilon^\gamma e^{-\gamma\lambda}\mathbb{1}_{[\log\epsilon,\frac{\log\epsilon}{2}]}(\lambda)+\sqrt{\beta/(\gamma+\beta)}\epsilon^{1/2}e^{-(1-\gamma)\lambda}\mathbb{1}_{(\frac{\log\epsilon}{2},0]}(\lambda).$$

Notice that $a_\alpha/c_\alpha = \gamma+\beta$ for $\alpha\in\{\ell,\mathrm{r}\}$.

Recall that we write $|x|$ for the generation of a particle x in the IBP, and $\lambda(x)$ for its location. If $(\lambda,\alpha)\in[\log\epsilon,\frac{\log\epsilon}{2}]\times\{\ell,\mathrm{r}\}$, then

$$\begin{aligned}Ag_{\mathrm{e}}(\lambda,\alpha) &\geq E_{(\lambda,\alpha)}\Big[\sum_{\substack{|x|=1\\ \lambda(x)>\frac{\log\epsilon}{2}}} g_{\mathrm{e}}(\lambda(x),\ell)\Big]\\ &= a_\alpha\sqrt{\beta/(\gamma+\beta)}\epsilon^{1/2}e^{-(1-\gamma)\lambda}\int_{-\lambda+\frac{\log\epsilon}{2}}^{-\lambda} e^{(2\gamma-1)t}\,dt\\ &= c_\alpha\tfrac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1}\epsilon^{1/2}e^{-\gamma\lambda}\big[1-\epsilon^{\gamma-1/2}\big] = \tfrac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1}\epsilon^{-\gamma+1/2}\big[1-\epsilon^{\gamma-1/2}\big]g_{\mathrm{e}}(\lambda,\alpha).\end{aligned}$$

If $(\lambda,\alpha)\in(\frac{\log\epsilon}{2},0]\times\{\ell,\mathrm{r}\}$, then

$$\begin{aligned}Ag_{\mathrm{e}}(\lambda,\alpha) \geq E_{(\lambda,\alpha)}\Big[\sum_{\substack{|x|=1\\ \lambda(x)\leq\frac{\log\epsilon}{2}}} g_{\mathrm{e}}(\lambda(x),\mathrm{r})\Big] &= \beta c_{\mathrm{r}}\epsilon^\gamma e^{-\gamma\lambda}\int_{-\lambda+\log\epsilon}^{-\lambda+\frac{\log\epsilon}{2}} e^{(1-2\gamma)t}\,dt\\ &= \tfrac{\beta}{2\gamma-1}\epsilon^\gamma e^{-\gamma\lambda}e^{(1-2\gamma)(\log\epsilon-\lambda)}\big[1-e^{-(1-2\gamma)\frac{\log\epsilon}{2}}\big]\\ &= \tfrac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1}\epsilon^{-\gamma+1/2}[1-\epsilon^{\gamma-1/2}]g_{\mathrm{e}}(\lambda,\alpha).\end{aligned}$$

By monotonicity of A, this implies

$$\|A^n\|\geq\big(\tfrac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1}\epsilon^{-\gamma+1/2}[1-\epsilon^{\gamma-1/2}]\big)^n.$$

Taking the n-th root on both sides, an application of (6.3) yields the required bound for $\rho_\epsilon(A)$. □

Proof of Theorem 5.1. The result follows immediately from Corollary 6.4, and Propositions 6.2 and 6.5. □

6.3 A multitype branching process

In this section, we analyse the IBP and its relation to the associated operator A. We begin by collecting properties of A in Section 6.3.1, and then use these properties to prove necessary and sufficient conditions for the multitype branching process to survive with positive probability in Section 6.3.2. Throughout, we use the notation introduced in Section 6.1, and write $P_{\phi,p}$ for the distribution of the percolated IBP with retention probability p started with one particle of type $\phi\in\Phi$, abbreviating $P_\phi:=P_{\phi,1}$.

6.3.1 Proof of Proposition 6.2

Lemma 6.6. *For all nonnegative $g\in C(\Phi)$ with $g\not\equiv 0$, we have $\min_{\phi\in\Phi}A^2g(\phi)>0$.*

Proof. If $g\in C(\Phi)$, $g\geq 0$, $g\not\equiv 0$, then there exist $\log\epsilon\leq\lambda_1<\lambda_2\leq 0$ and $\alpha_0\in\{\ell,\mathrm{r}\}$

such that g is strictly positive on $[\lambda_1, \lambda_2] \times \{\alpha_0\}$. Hence, it suffices to show that

$$\min_{\phi \in \Phi} P_\phi\big(\exists x\colon |x| = 2, \phi(x) \in [\lambda_1, \lambda_2] \times \{\alpha_0\}\big) > 0.$$

By the definition of the process, any particle produces offspring in any given interval of positive length with, uniformly in the start type, strictly positive probability. The two steps allow the time needed to ensure that the relative position of the parent gives $\alpha(x) = \alpha_0$. □

Lemma 6.7. *The operator $A\colon C(\Phi) \to C(\Phi)$ is compact.*

Proof. According to (6.4), we can write for $g \in C(\Phi)$ and $(\lambda, \alpha) \in \Phi$,

$$Ag(\lambda, \alpha) = \int_{\log \epsilon}^{0} g(t, \mathrm{r})\kappa_\ell(\lambda, t)\, dt + \int_{\log \epsilon}^{0} g(t, \ell)\kappa_{\mathrm{r}}(\lambda, \alpha, t)\, dt,$$

with $\kappa_\ell(\lambda, t) = \mathbb{1}_{[\log \epsilon, \lambda]}(t)\beta e^{(1-\gamma)(t-\lambda)}$ and $\kappa_{\mathrm{r}}(\lambda, \alpha, t) = \mathbb{1}_{[\lambda, 0]}(t) a_\alpha e^{\gamma(t-\lambda)}$. Thus A can be written as the sum of two operators, which are both compact by the Arzelà–Ascoli theorem. □

We summarize some standard properties of compact, positive operators in the following proposition.

Proposition 6.8. *Let X be a complex Banach space and $A\colon X \to X$ be a linear, compact and strictly positive operator.*

(i) The spectral radius of A, $\rho = \rho(A)$, is a strictly positive eigenvalue of A with one dimensional eigenspace, generated by a strictly positive eigenvector φ. The eigenvalue ρ is also the spectral radius of the adjoint A^* and the corresponding eigenspace is generated by a strictly positive eigenvector ν_0. We rescale φ and ν_0 such that $\|\varphi\| = 1$ and $\nu_0(\varphi) = 1$ to make the choice unique.

(ii) There exists $\theta_0 \in [0, \rho)$ such that $|\theta| \le \theta_0$ for all $\theta \in \sigma(A) \setminus \{\rho\}$, where $\sigma(A)$ is the spectrum of A.

(iii) For any $\theta > \theta_0$ and $g \in X$, we have $A^n g = \rho^n \nu_0(g)\varphi + O(\theta^n)$ for all $n \in \mathbb{N}$.

Proof. Statements (i) and (ii) are immediate from the Krein–Rutman theorem, see Theorem 3.1.3 (ii) in [110], and the general form of the spectrum of compact operators. Statement (iii) then follows from the spectral decomposition of a compact operator on a complex Banach space. See for example [84] and there in particular Theorem 49.1 and Proposition 50.1. □

Now all results are collected to establish Proposition 6.2.

Proof of Proposition 6.2. Identity $A_p = pA$ holds by definition and implies $\rho_\epsilon(A_p) = p\rho_\epsilon(A)$. Moreover, it is clear that it suffices to prove the first sentence of the statement for $p = 1$. Linearity is immediate from the definition, positivity was shown in Lemma 6.6, and compactness is the content of Lemma 6.7. The positive spectral radius follows immediately from Proposition 6.8 (i). □

6.3.2 Proof of Theorem 6.3

We start with a moment estimate for the total number of offspring of a particle. In the sequel, we write $|\mathrm{IBP}_n|$ for the number of particles in generation n of the IBP.

Lemma 6.9. *We have* $\sup_{\phi\in\Phi} E_\phi\big[|\mathrm{IBP}_1|^2\big] < \infty$.

Proof. Let Π, Z and $\hat{Z}$ be independent realisations of the Poisson point process and the pure jump processes defined in Section 6.1. Let $\phi = (\lambda, \alpha) \in \Phi$. By the definition of the IBP, under $P_{(\lambda,\alpha)}$,

$$|\mathrm{IBP}_1| \overset{d}{=} \begin{cases} \Pi([\log\epsilon - \lambda, 0]) + Z_{-\lambda} & \text{if } \alpha = \ell, \\ \Pi([\log\epsilon - \lambda, 0]) + \hat{Z}_{-\lambda} & \text{if } \alpha = \mathrm{r}, \end{cases}$$

where $\overset{d}{=}$ denotes distributional equality. Since f is non-decreasing, $\hat{Z}$ stochastically dominates Z. This implies that for all $\phi \in \Phi$,

$$E_\phi\big[|\mathrm{IBP}_1|^2\big] \le 2\Big(\mathbb{E}\big[\Pi([\log\epsilon, 0])^2\big] + E\big[(\hat{Z}_{-\log\epsilon})^2\big]\Big).$$

The first term on the right is finite because Π is a Poisson point process with finite intensity measure. The second summand was computed in Lemma 1.12 of [37] and found to be finite. □

The next result is a classical fact about branching processes. We give a proof since we could not find a reference for the result in sufficient generality; see Theorem III.11.2 in [83] for a special case.

Lemma 6.10. *For all* $p \in [0, 1]$, $N \in \mathbb{N}$ *and* $\phi \in \Phi$,

$$P_{\phi,p}\big(1 \le |\mathrm{IBP}_n| \le N \textit{ infinitely often}\big) = 0.$$

Proof. We split the proof in two parts. First we show that $\delta := \inf_{\phi\in\Phi} P_{\phi,p}(|\mathrm{IBP}_1| = 0) > 0$, then we conclude the statement from this result. By definition of the percolated

IBP, for all $(\lambda, \alpha) \in \Phi$,

$$\begin{aligned} P_{(\lambda,\alpha),p}\big(|\mathrm{IBP}_1| = 0\big) &\geq P_{(\lambda,\alpha),1}\big(|\mathrm{IBP}_1| = 0\big) \\ &= \begin{cases} P\big(\{\Pi([\log\epsilon - \lambda, 0]) = 0\} \cap \{Z_{-\lambda} = 0\}\big) & \text{if } \alpha = \ell \\ P\big(\{\Pi([\log\epsilon - \lambda, 0]) = 0\} \cap \{\hat{Z}_{-\lambda} = 0\}\big) & \text{if } \alpha = \mathrm{r} \end{cases} \\ &\geq P\big(\Pi([\log\epsilon, 0]) = 0\big) P\big(\hat{Z}_{-\log\epsilon} = 0\big) > 0. \end{aligned}$$

Since the lower bound is independent of (λ, α), the claim that $\delta > 0$ is established.

For the second step of the proof, we set $p = 1$ to simplify notation. The proof for general p is identical. Fix $N \in \mathbb{N}$, set $\tau_0 := 0$ and, for $k \geq 1$, let $\tau_k := \inf\{n > \tau_{k-1} \colon |\mathrm{IBP}_n| \in [1, N]\}$, where $\inf \emptyset := \infty$. The strong Markov property implies, for all $\phi \in \Phi$ and $k \in \mathbb{N}$,

$$P_\phi(\tau_k < \infty) \leq P_\phi(\tau_1 < \infty) \sup_\nu P_\nu(\tau_1 < \infty)^{k-1},$$

where the supremum is over all counting measure ν on Φ such that $\nu(\Phi) \in [1, N]$. Under P_ν, $\nu = \sum_{i=1}^n \delta_{\phi_i}$, the branching process is started with n particles of types $\phi_1, \dots, \phi_n$. When all original ancestors have no offspring in the first generation, then the branching process suffers immediate extinction and $\tau_1 = \infty$. Hence, for all such ν,

$$P_\nu(\tau_1 < \infty) = 1 - P_\nu(\tau_1 = \infty) \leq 1 - P_\nu\big(|\mathrm{IBP}_1| = 0\big) \leq 1 - \delta^{\nu(\Phi)} \leq 1 - \delta^N.$$

We conclude, for all $\phi \in \Phi$,

$$\begin{aligned} P_\phi\big(1 \leq |\mathrm{IBP}_n| \leq N \text{ infinitely often}\big) &= \lim_{k\to\infty} P_\phi(\tau_k < \infty) \leq \lim_{k\to\infty} \sup_\nu P_\nu(\tau_1 < \infty)^{k-1} \\ &\leq \lim_{k\to\infty} (1 - \delta^N)^{k-1} = 0. \end{aligned}$$

□

Proof of Theorem 6.3. Throughout the proof, we write $\rho := \rho_\epsilon(A_p)$ and φ for the corresponding strictly positive eigenfunction with $\|\varphi\| = 1$ from Proposition 6.8 (i). First suppose $\rho \leq 1$. By Lemma 6.10, $P_{\phi,p}(\lim_{n\to\infty} |\mathrm{IBP}_n| \in \{0, \infty\}) = 1$. By Proposition 6.8 (iii), the assumption $\rho \leq 1$ implies that

$$E_{\phi,p}\big[|\mathrm{IBP}_n|\big] = A_p^n \mathbf{1}(\phi) = \rho^n \nu_0(\mathbf{1}) \varphi(\phi) + o(1).$$

Hence, $\sup_{n\in\mathbb{N}} E_{\phi,p}[|\mathrm{IBP}_n|] < \infty$, and we conclude that $\lim_{n\to\infty} |\mathrm{IBP}_n| = 0$ $P_{\phi,p}$-almost surely for all $\phi \in \Phi$ and, therefore, $\zeta^\epsilon(p) = 0$.

Now suppose that $\rho > 1$, and denote $W_n = \frac{1}{\rho^n} \sum_{|x|=n} \varphi(\phi(x))$ for $n \in \mathbb{N}$. Then $(W_n \colon n \in \mathbb{N})$ is under $P_{\phi,p}$ a nonnegative martingale with respect to the filtration generated by the branching process. Hence, $W := \lim_{n\to\infty} W_n$ exists almost surely. Given Lemma 6.9, Biggins and Kyprianou show in Theorem 1.1 of [17] that $E_{\phi,p}[W] = \varphi(\phi)$ and therefore, $P_{\phi,p}(W > 0) > 0$. This implies in particular that the branching process survives with positive probability irrespective of the start type. □

We now investigate continuity of the survival probability as a function of the attachment rule. For this purpose we emphasise dependence on f by adding it as an additional argument to several quantities. The result is used in the proof of Theorem 6.1 in Chapter 8 below.

Lemma 6.11. *Let* $p \in (0,1]$. *Then* $\lim_{\delta \downarrow 0} \zeta^\epsilon(p, f-\delta) = \zeta^\epsilon(p,f)$.

Proof. Observe that there exists a natural coupling of the IBP(f) with the IBP($f-\delta$) such that every particle in the IBP($f-\delta$) is also present in the IBP(f), and hence, $\zeta^\epsilon(p, f-\delta)$ is increasing as $\delta \downarrow 0$. We can therefore assume that $\zeta^\epsilon(p,f) > 0$, that is $\rho(f) := \rho_\epsilon(A_p, f) > 1$. By the continuity of A_p in the attachment rule (see (6.4)), there exists $\delta_0 > 0$ such that $\rho_\epsilon(A_p, f-\delta_0) > 1$. In the proof of Theorem 6.3 we have seen that this implies that the IBP($f-\delta_0$) survives with positive probability, irrespective of the start type, and similar to Lemma 6.6, we conclude

$$\inf_{\phi \in \Phi} P_{\phi,p}\big(\mathrm{IBP}(f-\delta_0) \text{ survives}\big) > 0. \tag{6.11}$$

Recall the definition of the martingale $(W_n \colon n \in \mathbb{N})$ and its almost sure limit W from the proof of Theorem 6.3. We have $E_{\phi,p}[W] = \varphi(\phi)$, and

$$W = \frac{1}{\rho(f)} \sum_{|x|=1} W(\phi(x)) \qquad P_{\phi,p}\text{-almost surely,}$$

where, conditionally on the first generation, $(W(\phi(x)) \colon |x|=1)$ are independent copies of the random variable W under $P_{\phi(x),p}$. In particular, $\phi \mapsto P_{\phi,p}(W=0)$ is a fixed point of the operator $Hg(\phi) = E_{\phi,p}[\prod_{|x|=1} g(\phi(x))]$ on the set of $[0,1]$-valued, measurable functions. As the only $[0,1]$-valued fixed points of H are the constant function $\mathbf{1}$ and the extinction probability $\phi \mapsto P_{\phi,p}(\mathrm{IBP}(f) \text{ dies out})$, we deduce that $W > 0$ almost surely on survival. Let $c > 0$ and $N \in \mathbb{N}$. On the space of the coupling between IBP(f) and IBP($f-\delta$),

$$\begin{aligned}
\zeta^\epsilon(p,f) - \zeta^\epsilon(p,f-\delta) &= P(\mathrm{IBP}(f) \text{ survives}, \mathrm{IBP}(f-\delta) \text{ dies out}) \\
&\le P\big(W \le c, \mathrm{IBP}(f) \text{ survives}\big) + P\Big(W > c, \exists n \ge N \colon |\mathrm{IBP}_n(f)| < \frac{c\rho(f)^n}{2\max\varphi}\Big) \\
&\quad + P\Big(|\mathrm{IBP}_n(f)| \ge \frac{c\rho(f)^n}{2\max\varphi}\ \forall n \ge N, \mathrm{IBP}(f-\delta) \text{ dies out}\Big) \\
&=: \Theta_1(c) + \Theta_2(c,N) + \Theta_3(c,N,\delta).
\end{aligned}$$

Since the offspring distribution of an individual particle is continuous in δ uniformly on the type space, the probability that IBP(f) and IBP($f-\delta$) agree until generation N tends to one as $\delta \downarrow 0$. On this event, when $|\mathrm{IBP}_N(f)| \ge C\rho(f)^N$ for some $C > 0$, then the probability that the IBP($f-\delta$) subsequently dies out is bounded from above by

$$\sup_{\phi\in\Phi} P_{\phi,p}\big(\mathrm{IBP}(f-\delta) \text{ dies out}\big)^{\lceil C\rho(f)^N\rceil}.$$

By (6.11), this expression tends to zero as $N \to \infty$ when $\delta \le \delta_0$. Hence, for all $c > 0$,

$$0 \le \limsup_{\delta\downarrow 0} \big(\zeta^\epsilon(p,f) - \zeta^\epsilon(p,f-\delta)\big) \le \Theta_1(c) + \limsup_{N\to\infty} \Theta_2(c,N).$$

On the event $\{W > c\} \cap \{W_n \to W\}$, there is a finite stopping time N_0 such that $W_n \ge W/2$ for all $n \ge N_0$ and we deduce that $\rho(f)^{-n}|\mathrm{IBP}_n(f)| \ge W_n/\max\varphi \ge c/(2\max\varphi)$. Since W_n converges to W almost surely, we conclude that $\lim_{N\to\infty}\Theta_2(c,N) = 0$. Finally, $\Theta_1(c)$ tends to zero as $c \downarrow 0$ because W is positive on the event of survival. $\square$

CHAPTER 7

THE TOPOLOGY OF THE DAMAGED GRAPH

We investigate the empirical indegree distribution and the maximal indegree of the damaged network in Section 7.1, and typical distances in Section 7.2.

7.1 Degrees: proof of Theorem 5.2

The following lemma formalises basic facts about the indegrees $\mathcal{Z}[m,n]$.

Lemma 7.1. *For given $n \in \mathbb{N}$, the random variables $(\mathcal{Z}[m,n] \colon m \leq n)$ are independent. Fix $\hat{m} \in \{1,\dots,n\}$ and let $(\mathcal{Z}^m[\hat{m},n] \colon m \leq n)$, be independent copies of the random variable $\mathcal{Z}[\hat{m},n]$.*

(i) There is a coupling between $(\mathcal{Z}[m,n] \colon 1 \leq m \leq \hat{m})$ and $(\mathcal{Z}^m[\hat{m},n] \colon 1 \leq m \leq \hat{m})$ such that

$$\mathcal{Z}[m,n] \geq \mathcal{Z}^m[\hat{m},n] \qquad \text{for all } 1 \leq m \leq \hat{m}.$$

(ii) There is a coupling between $(\mathcal{Z}[m,n] \colon \hat{m} \leq m \leq n)$ and $(\mathcal{Z}^m[\hat{m},n] \colon \hat{m} \leq m \leq n)$ such that

$$\mathcal{Z}[m,n] \leq \mathcal{Z}^m[\hat{m},n] \qquad \text{for all } \hat{m} \leq m \leq n.$$

Proof. The independence of $(\mathcal{Z}[m,n] \colon m \leq n)$ is immediate from the network construction. Consequently, to prove (i) and (ii) it suffices to couple $\mathcal{Z}[m,n]$ and $\mathcal{Z}^m[\hat{m},n]$ for fixed $1 \leq m \leq \hat{m} \leq n$ in such a way that $\mathcal{Z}[m,n] \geq \mathcal{Z}^m[\hat{m},n]$. Equivalently, we show that $\mathcal{Z}[m,n]$ stochastically dominates $\mathcal{Z}[\hat{m},n]$. Let $\mathcal{Y}^m = (\mathcal{Y}^m_l \colon l \in \mathbb{N}_0)$ be the Markov process given by $\mathcal{Y}^m_l = \mathcal{Z}[m,m+l]$. Then $\mathcal{Y}^m_0 = 0$ and, for all $l,k \in \mathbb{N}_0$ with $k \leq l$,

$$\mathbb{P}(\mathcal{Y}^m_{l+1} = k+1 | \mathcal{Y}^m_l = k) = 1 - \mathbb{P}(\mathcal{Y}^m_{l+1} = k | \mathcal{Y}^m_l = k) = \frac{f(k)}{m+l}$$

is decreasing in m. Hence, in every step, the probability that $\mathcal{Y}^m$ jumps is at least the probability that $\mathcal{Y}^{\hat{m}}$ jumps. Since $\mathcal{Z}[m,n] = \mathcal{Y}^m_{n-m}$, $\mathcal{Z}[\hat{m},n] = \mathcal{Y}^{\hat{m}}_{n-\hat{m}}$ and $n-m \geq n-\hat{m}$, the claim is established. □

Our goal is to determine the asymptotic behaviour of $\max_{m\in \mathsf{V}_n^\epsilon} \mathcal{Z}[m,n]$ and

$$X^\epsilon_{\geq k}(n) = \frac{1}{n-\lfloor \epsilon n\rfloor} \sum_{m=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{1}_{\{k,k+1,\dots\}}(\mathcal{Z}[m,n]).$$

Lemma 7.1 allows us to replace the independent random variables in these sequences by groups of independent and identically distributed random variables.

Dereich and Mörters observe in [36], see for example Corollary 4.3, that the indegrees in network $(\mathsf{G}_n\colon n\in\mathbb{N})$ are closely related to the pure jump process $(Z_t)_{t\geq 0}$. Since the indegrees are not altered by the targeted attack, the same holds in the damaged network $(\mathsf{G}_n^\epsilon\colon n\in\mathbb{N})$. We now explain this connection. Let $\psi(k) := \sum_{j=1}^{k-1}\frac{1}{j}$ for all $k\in\mathbb{N}$, which we consider as a time change, mapping 'real time' epochs k to an 'artificial time' $\psi(k)$. The artificial time spent by the process $\mathcal{Z}[m,\cdot]$ in state i is

$$T_m[i] := \sup\Big\{\sum_{j=k}^{l}\frac{1}{j}\colon \mathcal{Z}[m,k]=i=\mathcal{Z}[m,l]\Big\}.$$

Let $k\in\mathbb{N}_0$ and n_k the last real time that $\mathcal{Z}[m,\cdot]$ spends in k, that is, $\mathcal{Z}[m,n_k]=k$, $\mathcal{Z}[m,n_k+1]=k+1$. Then $\sum_{i=0}^{k} T_m[i] = \sum_{j=m}^{n_k}\frac{1}{j} = \psi(n_k+1)-\psi(m)$. In particular,

$$\mathcal{Z}[m,n]\leq k \quad\Leftrightarrow\quad \sum_{i=0}^{k} T_m[i] \geq \psi(n+1)-\psi(m). \tag{7.1}$$

By definition, there exists a sequence of independent random variables $(T[i]\colon i\in\mathbb{N}_0)$ such that $T[i]$ is exponentially distributed with mean $1/f(i)$ and

$$Z_t\leq k \quad\Leftrightarrow\quad \sum_{i=0}^{k} T[i] > t. \tag{7.2}$$

The next lemma provides a coupling between the artificial times $T_m[i]$ and the exponential times $T[i]$. In combination with (7.1) and (7.2), this allows us to study the jump process $(Z_t\colon t\geq 0)$ instead of the involved dynamics of the indegree process $\mathcal{Z}[m,\cdot]$. The proof of the lemma is identical to the proof of Lemma 4.1 in [36], and is omitted. However, the argument behind the result was sketched in the paragraph of (5.11). We denote by $\tau_{m,i} = \inf\{\psi(k)\colon \mathcal{Z}[m,k]=i\}$ the artificial first entrance time of $\mathcal{Z}[m,\cdot]$ into state i. If $\tau_{m,i}=\psi(k)$, we write $\triangle\tau_{m,i}=k^{-1}$.

Lemma 7.2. *There exists a constant $\eta>0$ such that for all $m\in\mathbb{N}$ there is a coupling such that, for all $i\in\mathbb{N}_0$ with $f(i)\triangle\tau_{m,i}\leq\frac{1}{2}$,*

$$T[i]-\eta f(i)\triangle\tau_{m,i}\leq T_m[i]\leq T[i]+\triangle\tau_{m,i} \qquad \textit{almost surely,}$$

and the random variables $((T[i], T_m[i]) \colon i \in \mathbb{N}_0)$ *are independent.*

By definition $\triangle\tau_{m,i} \leq m^{-1}$. Hence, Lemma 7.2 yields a coupling such that when $f(k)/m \leq \frac{1}{2}$, then

$$\sum_{i=0}^{k} T[i] - (k+1)\eta f(k)/m \leq \sum_{i=0}^{k} T_m[i] \leq \sum_{i=0}^{k} T[i] + (k+1)/m.$$

In particular, for $f(k)/m \leq \frac{1}{2}$, the equivalence (7.1) implies

$$P\Big(\sum_{i=0}^{k} T[i] \geq \psi(n+1) - \psi(m) + \eta(k+1)f(k)/m\Big) \leq \mathbb{P}(\mathcal{Z}[m,n] \leq k)$$
$$\leq P\Big(\sum_{i=0}^{k} T[i] \geq \psi(n+1) - \psi(m) + (k+1)/m\Big).$$

If $m \leq n$ and $m - \vartheta n = O(1)$ for some $\vartheta \in (0,1]$, then $\psi(n+1) - \psi(m) = \sum_{j=m}^{n} \frac{1}{j} = -\log\vartheta + o(1)$. Hence, for $m \leq n$, $m - \vartheta n = O(1)$ and $k = O(\log n)$,

$$\mathbb{P}(\mathcal{Z}[m,n] \leq k) = P\Big(\sum_{i=0}^{k} T[i] \geq -\log\vartheta + o(1)\Big), \tag{7.3}$$

where the random null sequence $o(1)$ is bounded by a deterministic null sequence of order $O((\log n)^2/n)$.

We proceed by estimating the distribution function of $\sum_{i=0}^{k} T[i]$. The following identity for the incomplete beta function will be of use.

Lemma 7.3. *Let* $a \in (0,1]$, $c > 0$ *and* $k \in \mathbb{N}_0$. *Then*

$$\sum_{i=0}^{k} \binom{k}{i} \frac{(-a)^i}{i+c} = a^{-c} \int_0^a x^{c-1}(1-x)^k \, dx.$$

Proof. Denote the left-hand side by $\theta(k,a,c)$. For $x > 0$, we have

$$\frac{\partial}{\partial x}\big[x^c \theta(k,x,c)\big] = \frac{\partial}{\partial x}\Big[\sum_{i=0}^{k} \binom{k}{i} (-1)^i \frac{x^{i+c}}{i+c}\Big] = \sum_{i=0}^{k} \binom{k}{i} (-1)^i x^{i+c-1} = x^{c-1}(1-x)^k.$$

Integrating both sides between 0 and a, and dividing by a^c, we obtain the claim. $\square$

Lemma 7.4. *For* $k \in \mathbb{N}_0$ *and* $t \geq 0$,

$$P(Z_t \geq k+1) = P\Big(\sum_{i=0}^{k} T[i] \leq t\Big) = B\big(k+1, \tfrac{\beta}{\gamma}\big)^{-1} \int_{e^{-\gamma t}}^{1} x^{\frac{\beta}{\gamma}-1}(1-x)^k \, dx. \tag{7.4}$$

Proof. Let $k \in \mathbb{N}_0$. The probability density for $\sum_{i=0}^{k} T[i]$ is given by (see for example Problem 12, Chapter 1 in [71])

$$t \mapsto \sum_{i=0}^{k} \frac{\prod_{j=0,j\neq i}^{k} f(j)}{\prod_{j=0,j\neq i}^{k}[f(j)-f(i)]} f(i)e^{-f(i)t}\mathbb{1}_{[0,\infty)}(t).$$

Using $f(j) = \gamma j + \beta$, we can rewrite for all $i \in \{0, \dots, k\}$,

$$\frac{\prod_{j=0,j\neq i}^{k} f(j)}{\prod_{j=0,j\neq i}^{k}[f(j)-f(i)]} = \frac{\frac{\beta}{f(i)}\frac{k!}{k!}\prod_{j=1}^{k} f(j)}{\gamma^k(-1)^i i!(k-i)!} = \frac{\beta}{\gamma}\Big(\prod_{j=1}^{k}\frac{f(j)}{\gamma j}\Big)\binom{k}{i}\frac{(-1)^i}{i+\frac{\beta}{\gamma}}.$$

We obtain

$$P\Big(\sum_{i=0}^{k} T[i] \leq t\Big) = \frac{\beta}{\gamma}\Big(\prod_{j=1}^{k}\frac{f(j)}{\gamma j}\Big)\sum_{i=0}^{k}\binom{k}{i}\frac{(-1)^i}{i+\frac{\beta}{\gamma}}\big(1-e^{-\beta t}e^{-\gamma i t}\big).$$

The factor in front of the sum equals $B(k+1,\frac{\beta}{\gamma})^{-1}$. Thus, it remains to show that the sum agrees with the integral in (7.4). Using Lemma 7.3, we derive

$$\begin{aligned}
&\sum_{i=0}^{k}\binom{k}{i}\frac{(-1)^i}{i+\frac{\beta}{\gamma}}\big(1-e^{-\beta t}e^{-\gamma i t}\big)\\
&\qquad = \int_0^1 x^{\frac{\beta}{\gamma}-1}(1-x)^k\,dx - e^{-\beta t}(e^{-\gamma t})^{-\frac{\beta}{\gamma}}\int_0^{e^{-\gamma t}} x^{\frac{\beta}{\gamma}-1}(1-x)^k\,dx\\
&\qquad = \int_{e^{-\gamma t}}^1 x^{\frac{\beta}{\gamma}-1}(1-x)^k\,dx.
\end{aligned}$$

□

Proposition 7.5. *For all $k \in \mathbb{N}_0$,*

$$\mathbb{E}\big[X^{\epsilon}_{\geq k+1}(n)\big] \to \int_{\epsilon}^{1}\frac{1}{1-\epsilon}B\big(k+1,\tfrac{\beta}{\gamma}\big)^{-1}\int_{y^{\gamma}}^{1} x^{\frac{\beta}{\gamma}-1}(1-x)^k\,dx\,dy \qquad \text{as } n\to\infty.$$

Proof. Let $\delta > 0$, $\Delta = \lfloor \delta\epsilon n\rfloor$, $N = 1+\lfloor\frac{n-\lfloor\epsilon n\rfloor}{\Delta}\rfloor$, $m_j = \lfloor\epsilon n\rfloor+1+j\Delta$ for $j = 0,\dots,N-1$, $m_N = n+1$, $\Delta_j = \Delta$ for $j = 1,\dots,N-1$ and $\Delta_N = m_N - m_{N-1} \in [0,\Delta)$. Combining Lemma 7.1 with (7.3), we obtain

$$\begin{aligned}
\frac{1}{n-\lfloor\epsilon n\rfloor}\sum_{m=\lfloor\epsilon n\rfloor+1}^{n}\mathbb{P}(\mathcal{Z}[m,n]>k) &\leq \frac{1}{n-\lfloor\epsilon n\rfloor}\sum_{j=0}^{N-1}\Delta_{j+1}\mathbb{P}(\mathcal{Z}[m_j,n]>k)\\
&\leq \frac{1}{n-\lfloor\epsilon n\rfloor}\sum_{j=0}^{N-1}\delta\epsilon n\,P\Big(\sum_{i=0}^{k}T[i]\leq -\log(\epsilon+j\delta\epsilon)+o(1)\Big).
\end{aligned}$$

Hence,

$$\limsup_{n\to\infty}\frac{1}{n-\lfloor\epsilon n\rfloor}\sum_{m=\lfloor\epsilon n\rfloor+1}^{n}\mathbb{P}(\mathcal{Z}[m,n]>k)$$

is bounded from above by

$$\sum_{j=0}^{\lceil \frac{1-\epsilon}{\delta\epsilon} \rceil} \frac{\delta\epsilon}{1-\epsilon} P\Big(\sum_{i=0}^{k} T[i] \leq -\log\big(\epsilon + j\frac{\delta\epsilon}{1-\epsilon}(1-\epsilon)\big)\Big).$$

Taking $\delta \to 0$, we conclude

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n-\lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{P}(\mathcal{Z}[m,n] > k) \\ \leq \int_0^1 P\Big(\sum_{i=0}^{k} T[i] \leq -\log\big(\epsilon + y(1-\epsilon)\big)\Big)\, dy. \end{aligned} \tag{7.5}$$

Similarly,

$$\begin{aligned} \frac{1}{n-\lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{P}(\mathcal{Z}[m,n] > k) &\geq \frac{1}{n-\lfloor \epsilon n \rfloor} \sum_{j=1}^{N} \Delta_j \mathbb{P}(\mathcal{Z}[m_j-1,n] > k) \\ &\geq \frac{1}{n-\lfloor \epsilon n \rfloor} \sum_{j=1}^{N-1} \Delta_j P\Big(\sum_{i=0}^{k} T[i] \leq -\log(\epsilon + j\delta\epsilon) + o(1)\Big), \end{aligned}$$

and as above we see that lim inf satisfies the reverse inequality in (7.5). Lemma 7.4 yields the claim. □

The following proposition proves the first part of Theorem 5.2.

Proposition 7.6. *Let μ^ϵ be the probability measure on $\mathbb{N}_0$ that satisfies (5.5). Then, almost surely, $\lim_{n\to\infty} X^\epsilon(n) = \mu^\epsilon$ in total variation norm, and*

$$\lim_{k\to\infty} \log \mu^\epsilon_{\geq k}/k = \log(1-\epsilon^\gamma). \tag{7.6}$$

Proof. Dereich and Mörters (pp 1238–1239 in [36]) give a simple argument based on Chernoff's inequality to upgrade the convergence of the expected empirical degree distribution to convergence of the empirical degree distribution in total variation norm. Given Proposition 7.5, the proof remains valid for the damaged network and is therefore omitted.

To establish (7.6), we write $a(k) \asymp b(k)$ if there exist constants $0 < c \leq C < \infty$ such that $ca(k) \leq b(k) \leq Ca(k)$ for all large k. By Stirling's formula, $B(k+1, \beta/\gamma)^{-1} \asymp k^{\beta/\gamma}$. Moreover,

$$\begin{aligned} \frac{1}{1-\epsilon}\int_\epsilon^1 \int_{y^\gamma}^1 x^{\frac{\beta}{\gamma}-1}(1-x)^k \, dx\, dy &\asymp \int_\epsilon^1 \int_{y^\gamma}^1 (1-x)^k\, dx\, dy \\ &= \frac{1}{k+1}\int_\epsilon^1 (1-y^\gamma)^{k+1}\, dy \asymp \frac{(1-\epsilon^\gamma)^{k+1}}{(k+1)^2}. \end{aligned}$$

In the first estimate we used that $x^{\frac{\beta}{\gamma}-1}$ is bounded from zero and infinity; in the second we employed Laplace's method (see for example Section 3.5 of [103]). In particular, (7.6) holds. $\square$

To complete the proof of Theorem 5.2, it remains to derive the asymptotic behaviour of the maximal indegree. The statement follows from the next two lemmas.

Lemma 7.7 (Upper bound)**.** *Let* $c > -\frac{1}{\log(1-\epsilon^\gamma)}$. *Then,*

$$\mathbb{P}\big(\max_{m\in\mathsf{V}_n^\epsilon} \mathcal{Z}[m,n] \le c\log n\big) \to 1 \qquad \text{as } n\to\infty.$$

Proof. Write $k_n = \lfloor c\log n\rfloor$, $\underline{m} = \lfloor \epsilon n\rfloor + 1$ and $\Delta = n - \lfloor \epsilon n\rfloor$. Moreover, let $\mathcal{Z}^m[\underline{m},n]$, $m\le n$, be independent copies of $\mathcal{Z}[\underline{m},n]$. Lemma 7.1 and (7.3) yield

$$\begin{aligned}\mathbb{P}\big(\max_{m\in\mathsf{V}_n^\epsilon} \mathcal{Z}[m,n] \le c\log n\big) &\ge \mathbb{P}\big(\max_{m\in\mathsf{V}_n^\epsilon} \mathcal{Z}^m[\underline{m},n] \le k_n\big) = \mathbb{P}\big(\mathcal{Z}[\underline{m},n]\le k_n\big)^\Delta \\ &= P\Big(\sum_{i=0}^{k_n} T[i] \ge -\log\epsilon + o(1)\Big)^\Delta \qquad (7.7)\\ &= \exp\Big(-\Delta P\Big(\sum_{i=0}^{k_n} T[i] \le -\log\epsilon + o(1)\Big)(1+o(1))\Big),\end{aligned}$$

using a Taylor expansion in the last equality. As above, uniformly for t in compact subintervals of $(0,\infty)$,

$$\int_{e^{-\gamma t}}^1 x^{\frac{\beta}{\gamma}-1}(1-x)^k\,dx \asymp \int_{e^{-\gamma t}}^1 (1-x)^k\,dx \asymp \exp\Big(k\log(1-e^{-\gamma t}) - \log k\Big).$$

Thus, Lemma 7.4 and Stirling's formula yield for $\vartheta\in(0,1)$ and $t = -\log\vartheta + o(1)$,

$$\begin{aligned}P\Big(\sum_{i=0}^{k} T[i]\le t\Big) &\asymp \exp\Big(k\log(1-e^{-\gamma t}) + (\tfrac{\beta}{\gamma}-1)\log k\Big)\\ &= \exp\Big(k\log(1-\vartheta^\gamma)(1+o(1))\Big)\end{aligned} \qquad (7.8)$$

as $k\to\infty$. Using this estimate for $k=k_n$ and $\vartheta=\epsilon$, the exponent on the right-hand side of (7.7) tends to zero as $n\to\infty$ if $c > -1/\log(1-\epsilon^\gamma)$. $\square$

Lemma 7.8 (Lower bound)**.** *Let* $c < -\frac{1}{\log(1-\epsilon^\gamma)}$. *Then,*

$$\mathbb{P}\big(\max_{m\in\mathsf{V}_n^\epsilon} \mathcal{Z}[m,n] \le c\log n\big) \to 0 \qquad \text{as } n\to\infty.$$

Proof. The idea of the proof is to restrict the maximum to an arbitrarily small proportion of the oldest vertices. Let $\delta>0$, and write $k_n := \lfloor c\log n\rfloor$, $\Delta = \lfloor \delta\epsilon n\rfloor$ and $\overline{m} = \lfloor\epsilon n\rfloor + \Delta$. Moreover, let $\mathcal{Z}^m[\overline{m},n]$, $m\le n$, be independent copies of $\mathcal{Z}[\overline{m},n]$.

According to Lemma 7.1, there is a coupling such that

$$\max_{m\in \mathsf{V}_n^\epsilon} \mathcal{Z}[m,n] \geq \max_{m=\lfloor \epsilon n\rfloor+1,\dots,\lfloor \epsilon n\rfloor+\Delta} \mathcal{Z}[m,n] \geq \max_{m=\lfloor \epsilon n\rfloor+1,\dots,\lfloor \epsilon n\rfloor+\Delta} \mathcal{Z}^m[\overline{m},n].$$

Arguing as in (7.7), (7.3) yields that

$$\mathbb{P}\big(\max_{m\in \mathsf{V}_n^\epsilon} \mathcal{Z}[m,n] \leq c\log n\big)$$

is bounded from above by

$$\exp\Big(-\Delta P\Big(\sum_{i=0}^{k_n} T[i] \leq -\log(\epsilon(1+\delta)) + o(1)\Big)\big(1+o(1)\big)\Big).$$

Now (7.8) with $\vartheta = \epsilon(1+\delta)$ implies that the exponent on the right-hand side tends to $-\infty$ if $c < -1/\log(1-(\epsilon(1+\delta))^\gamma)$. Since δ was arbitrary, the claim is established. $\square$

Proof of Theorem 5.2. The result follows immediately from Proposition 7.6 and Lemmas 7.7 and 7.8. $\square$

7.2 Distances: proof of Theorem 5.3

In this section, we study the typical distance between two uniformly chosen vertices in C_n^ϵ, and prove Theorem 5.3. We write $\triangle\mathcal{Z}[m,n] = \mathcal{Z}[m,n+1] - \mathcal{Z}[m,n]$ and, for $m \geq n$, $\mathcal{Z}[m,n] = 0$. In the graph, the indegree of vertex m at time m is zero by definition, but we will also use the distribution of the process $(\mathcal{Z}[m,n]\colon n \geq m)$ for different initial values. Formally, the evolution of $\mathcal{Z}[m,\cdot]$ with initial value k is obtained by using the attachment rule $g(l) := f(k+l)$, and we denote its distribution by $\mathbb{P}^k$, using $\mathbb{E}^k$ for the corresponding expectation; we abbreviate $\mathbb{P} := \mathbb{P}^0$, $\mathbb{E} := \mathbb{E}^0$. We further write $\hat{n} := \inf\{n \in \mathbb{N}\colon f(n)/n \leq 1\} \vee 2$. Note that $\gamma < 1$ implies $\hat{n} \in \mathbb{N}$. We observe some facts about the indegree distribution. These are adaptations of results in [37].

Lemma 7.9 (Lemma 2.7 in [37])**.** *For all $k \in \mathbb{N}_0$ and $m,n \in \mathbb{N}$ with $k \leq m$, $\hat{n} \leq m \leq n$,*

$$\mathbb{P}^k(\triangle\mathcal{Z}[m,n] = 1) \leq \frac{f(k)}{(m-1)^\gamma n^{1-\gamma}}. \tag{7.9}$$

Proof. Observe that $(f(\mathcal{Z}[m,n])\prod_{j=m}^{n-1}\frac{1}{1+\gamma/j}\colon n \geq m)$ is a martingale, and therefore

$$\mathbb{P}^k(\triangle\mathcal{Z}[m,n] = 1) = \mathbb{E}^k\Big[\frac{f(\mathcal{Z}[m,n])}{n}\Big] = \frac{f(k)}{n}\prod_{j=m}^{n-1}(1+\gamma/j) \leq \frac{f(k)}{(m-1)^\gamma n^{1-\gamma}}. \qquad \square$$

Lemma 7.10 (Lemma 2.10 in [37])**.** *For all $k \in \mathbb{N}_0$, $m, m' \in \mathbb{N}$, $\hat{n} \leq m \leq m'$, $k \leq m$, there exists a coupling of the process $(\mathcal{Z}[m,n]\colon n \geq m)$ under the conditional probability*

$\mathbb{P}^k(\cdot|\triangle\mathcal{Z}[m,m'] = 1)$ *and the process* $(\mathcal{Z}[m,n]\colon n \geq m)$ *under* $\mathbb{P}^{k+1}$ *such that, apart from time* m'*, the jump times of the first process are a subset of the jump times of the latter.*

The proof of the lemma is similar to the proof of Lemma 2.10 in [37], and we omit it. After these preliminary results, we now begin our analysis of typical distances in the network $(\mathsf{G}_n^\epsilon\colon n \in \mathbb{N})$. Recall, that for this type of questions, we consider G_n^ϵ to be an undirected graph. For $v, w \in \mathsf{V}_n^\epsilon$ and $h \in \mathbb{N}_0$, let

$$\mathcal{S}_h(v,w) := \{(v_0,\dots,v_h)\colon v_i \in \mathsf{V}_n^\epsilon, v_i \neq v_j \text{ for } i \neq j, v_0 = v, v_h = w\}$$

be the set of all self-avoiding paths of length h between v and w, and let $\mathcal{S}_h(v) = \{\mathsf{p}\colon \mathsf{p} \in \mathcal{S}_h(v,w) \text{ for some } w \in \mathsf{V}_n^\epsilon\}$ the set of all self-avoiding paths of length h starting in v.

Definition 7.11. Let $\theta \in (0,\infty)$ and $\mathsf{G} = (\mathsf{V},\mathsf{E})$ be an undirected graph with $\mathsf{V} \subseteq \mathbb{N}$. A self-avoiding path $\mathsf{p} = (v_0,\dots,v_h)$ in G is θ-admissible (or admissible) if, for all $i \in \{1,\dots,h\}$, we have $\{v_{i-1},v_i\} \in \mathsf{E}$ and

$$\big|\{w \in \mathsf{V}\colon v_{i-1} < w \leq v_i, \{v_{i-1},w\} \in \mathsf{E}\}\big| \leq \theta. \tag{7.10}$$

Note that (7.10) is automatically satisfied if $v_i < v_{i-1}$. In the graph G_n^ϵ, condition (7.10) can be written as $\mathcal{Z}[v_{i-1},v_i] \leq \theta$. We further denote, for $v, w \in \mathsf{V}_n^\epsilon$, $h \in \mathbb{N}_0$ and $\theta \in (0,\infty)$,

$$N_h^\theta(v,w) := \big|\{\mathsf{p} \in \mathcal{S}_h(v,w)\colon \mathsf{p} \text{ is } \theta\text{-admissible in } \mathsf{G}_n^\epsilon\}\big|,$$
$$N_h^\theta(v) := \big|\{\mathsf{p} \in \mathcal{S}_h(v)\colon \mathsf{p} \text{ is } \theta\text{-admissible in } \mathsf{G}_n^\epsilon\}\big|,$$

and, for $h > 0$, $N_{\leq h}^\theta(v) = \sum_{k=0}^{\lfloor h\rfloor} N_k^\theta(v)$, $N_{\leq h}^\theta(v,w) = \sum_{k=0}^{\lfloor h\rfloor} N_k^\theta(v,w)$. The dependence of $\mathcal{S}_h(v,w)$, $\mathcal{S}_h(v)$, $N_h^\theta(v,w)$ etc. on n is suppressed in the notation, but it will always be clear from the context which graph is considered. We write $\mathrm{IBP}^\epsilon(f)$ for the idealized branching process with type space $[\log\epsilon, 0] \times \{\ell, \mathrm{r}\}$ generated with attachment rule f if we want to emphasize f and ϵ. The proof of the following lemma is deferred to Section 8.3.

Lemma 7.12. *Let* $\delta > 0$ *such that* $\gamma(1+\delta) < 1$, $\underline{\epsilon} \in (0,\epsilon)$, *and* $(\theta_n\colon n \in \mathbb{N})$ *a positive sequence with* $\theta_n = o(n)$. *For all sufficiently large* n, $v_0 \in \mathsf{V}_n^\epsilon$, *and* $h \in \mathbb{N}_0$,

$$\mathbb{E}\big[N_h^{\theta_n}(v_0)\big] \leq E_{(s_n(v_0),\ell)}\big[\big|\mathrm{IBP}_h^{\underline{\epsilon}}((1+\delta)f)\big|\big], \qquad \textit{where } s_n(v_0) := -\sum_{j=v_0}^{n-1}\frac{1}{j}.$$

We are now in the position to prove Theorem 5.3. For two vertices $v, w \in \mathsf{V}_n^\epsilon$ in different components of G_n^ϵ, the distance $\mathsf{d}_{\mathsf{G}_n^\epsilon}(v,w)$ between them is defined to be infinite.

Proof of Theorem 5.3. Let $v, w \in \mathsf{V}_n^\epsilon$, $h \in \mathbb{N}$. With $\theta_n := (\log n)^2$, (5.6) yields

$$\begin{aligned}\mathbb{P}\big(\mathsf{d}_{\mathsf{G}_n^\epsilon}(v,w) \le h\big) &\le \mathbb{P}\Big(\mathsf{d}_{\mathsf{G}_n^\epsilon}(v,w) \le h, \max_{m \in \mathsf{V}_n^\epsilon} \mathcal{Z}[m,n] \le \theta_n\Big) + \mathbb{P}\Big(\max_{m \in \mathsf{V}_n^\epsilon} \mathcal{Z}[m,n] > \theta_n\Big) \\ &\le \mathbb{P}\big(N_{\le h}^{\theta_n}(v,w) \ge 1\big) + o(1), \end{aligned} \tag{7.11}$$

where the error bound is uniform in v, w and h. Markov's inequality yields, for every $v, w \in \mathsf{V}_n^\epsilon$ with $v \neq w$ and for every $h \in \mathbb{N}$,

$$\mathbb{P}\big(N_{\le h}^{\theta_n}(v,w) \ge 1\big) \le \mathbb{E}\big[N_{\le h}^{\theta_n}(v,w)\big] = \sum_{k=1}^{h} \sum_{\mathsf{p} \in \mathcal{S}_k(v,w)} \mathbb{P}\big(\mathsf{p} \text{ is } \theta_n\text{-admissible in } \mathsf{G}_n^\epsilon\big). \tag{7.12}$$

Let $\mathsf{p} = (v_0, \dots, v_k) \in \mathcal{S}_k(v,w)$. We write $w_i^+ := v_{i-1} \vee v_i$, $w_i^- := v_{i-1} \wedge v_i$ and

$$\mathcal{E}_i := \big\{\triangle\mathcal{Z}[w_i^-, w_i^+ - 1] = 1, \mathcal{Z}[v_{i-1}, v_i] \le \theta_n\big\} \qquad \text{for every } i \in \{1, \dots, k\}.$$

We have

$$\begin{aligned}\mathbb{P}(\mathsf{p} \text{ is } \theta_n\text{-admissible in } \mathsf{G}_n^\epsilon) &= \mathbb{P}\Big(\bigcap_{i=1}^{k} \big\{\triangle\mathcal{Z}[w_i^-, w_i^+ - 1] = 1, \mathcal{Z}[v_{i-1}, v_i] \le \theta_n\big\}\Big) \\ &= \mathbb{P}\Big(\mathcal{E}_k \Big| \bigcap_{i=1}^{k-1} \mathcal{E}_i\Big) \mathbb{P}((v_0, \dots, v_{k-1}) \text{ is } \theta_n\text{-admissible in } \mathsf{G}_n^\epsilon). \end{aligned} \tag{7.13}$$

To estimate the probability $\mathbb{P}(\mathcal{E}_k | \bigcap_{i=1}^{k-1} \mathcal{E}_i)$, we first note that the only edge in the self-avoiding path p on whose presence the event $\{v_{k-1}, v_k\} \in \mathsf{E}_n^\epsilon$ can depend is $\{v_{k-2}, v_{k-1}\}$. The possible arrangements of these two edges are sketched in Figure II-6. When $v_{k-2} < v_{k-1}$ (cases A, B and C in Figure II-6), then in addition, we have knowledge of edges whose left vertex is v_{k-2} because $\mathcal{Z}[v_{k-2}, v_{k-1}] \le \theta_n$. However, these are always independent of $\{v_{k-1}, v_k\}$. If $v_{k-1} < v_k$ (cases A, D and E in Figure II-6), then event $\mathcal{E}_k$ requires that $\mathcal{Z}[v_{k-1}, v_k] \le \theta_n$. Since edges with left vertex v_{k-1} depend only on edges whose left vertex is also v_{k-1}, the only relevant conditioning occurs in cases D and E in Figure II-6 by requiring $\{v_{k-1}, v_{k-2}\}$ to be present.

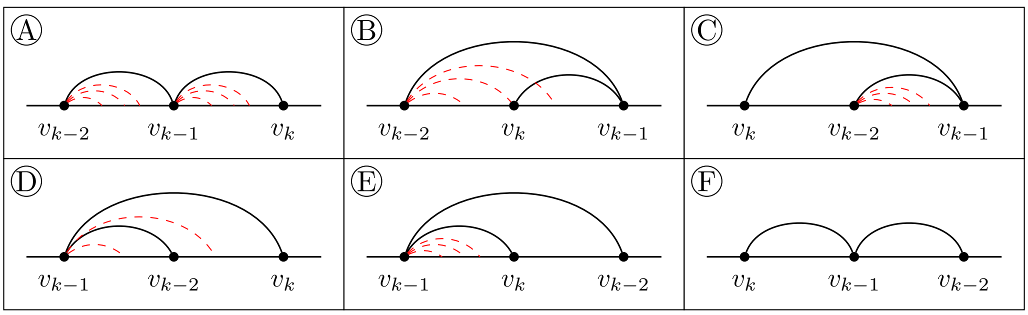

Figure II-6. Possible interactions of two edges on a self-avoiding path. The red, dashed edges have to be considered to decide if the number of right-neighbours is small enough to declare the path admissible.

We deduce

$$\mathbb{P}\Big(\mathcal{E}_k\Big|\bigcap_{i=1}^{k-1}\mathcal{E}_i\Big)$$
$$=\begin{cases}\mathbb{P}(\triangle\mathcal{Z}[w_k^-,w_k^+-1]=1,\mathcal{Z}[v_{k-1},v_k]\le\theta_n) & \text{in A,B,C,F,}\\ \mathbb{P}(\triangle\mathcal{Z}[v_{k-1},v_k-1]=1,\mathcal{Z}[v_{k-1},v_k]\le\theta_n|\triangle\mathcal{Z}[v_{k-1},v_{k-2}-1]=1) & \text{in D,E.}\end{cases}$$

Using Lemma 7.10 and (7.9), we can bound the probability in both cases by $f(1)/(\epsilon n)$. Combining this estimate with (7.12) and (7.13), we obtain

$$\begin{aligned}\mathbb{P}\big(N^{\theta_n}_{\le h}(v,w)\ge 1\big)&\le\sum_{k=1}^{h}\sum_{\mathsf{p}\in\mathcal{S}_{k-1}(v)}\frac{f(1)}{\epsilon n}\mathbb{P}(\mathsf{p}\text{ is }\theta_n\text{-admissible in }\mathsf{G}_n^\epsilon)\\&=\frac{f(1)}{\epsilon n}\sum_{k=0}^{h-1}\mathbb{E}\big[N_k^{\theta_n}(v)\big].\end{aligned}$$

Lemma 7.12 yields for small $\bar\delta>0$ and $\underline{\epsilon}:=\epsilon-\bar\delta>0$,

$$\mathbb{P}\big(N^{\theta_n}_{\le h}(v,w)\ge 1\big)\le\frac{f(1)}{\epsilon n}\sum_{k=0}^{h-1}E_{(s_n(v),\ell)}\Big[\big|\mathrm{IBP}_k^{\underline{\epsilon}}((1+\bar\delta)f)\big|\Big].$$

We denote by $\bar\rho$ the spectral radius of the operator A associated to $\mathrm{IBP}^{\underline{\epsilon}}((1+\bar\delta)f)$, and by $\bar\varphi$ the corresponding eigenfunction. Choose a constant C such that $C\ge\max_\phi\bar\varphi(\phi)/\min_\phi\bar\varphi(\phi)$ for all sufficiently small $\bar\delta$. That is possible since the eigenfunctions are continuous in $\bar\delta$ (this can be seen along the lines of Note 3 to Chapter II on pages 568-569 of [91]). By Theorem 6.3, and by the assumption that G_n^ϵ has a giant component, $\rho_\epsilon(A)>1$. Combining this fact with the continuity of the spectral radius with respect to the operator (see Chapter II.5 in [91]), we obtain $\bar\rho>1$ for all small $\bar\delta$. Hence, for all $v,w\in\mathsf{V}_n^\epsilon$, $v\ne w$,

$$\begin{aligned}\mathbb{P}\big(N^{\theta_n}_{\le h}(v,w)\ge 1\big)&\le\frac{f(1)}{\epsilon n}\sum_{k=0}^{h-1}\frac{1}{\min_\phi\bar\varphi(\phi)}E_{(s_n(v),\ell)}\Big[\sum_{\substack{x\in\mathrm{IBP}^{\underline{\epsilon}}((1+\bar\delta)f)\\|x|=k}}\bar\varphi(\phi(x))\Big]\\&=\frac{f(1)}{\epsilon n}\sum_{k=0}^{h-1}\bar\rho^k\frac{\bar\varphi(s_n(v),\ell)}{\min_\phi\bar\varphi(\phi)}\\&\le\frac{f(1)C}{\epsilon n}\frac{\bar\rho^h}{\bar\rho-1}=\frac{f(1)C}{\epsilon(\bar\rho-1)}\exp\big(h\log\bar\rho-\log n\big).\end{aligned}$$

In particular, for $\delta>0$ and $h_n:=(1-\delta^2)\frac{\log n}{\log\bar\rho}$, we showed that

$$\sup_{v,w\in\mathsf{V}_n^\epsilon,v\ne w}\mathbb{P}(N^{\theta_n}_{\le h_n}(v,w)\ge 1)=o(1).$$

For independent, uniformly chosen vertices V_n,W_n in C_n^ϵ, we have $V_n\ne W_n$ with high probability. According to (7.11), this implies $\mathbb{P}(\mathsf{d}_{\mathsf{G}_n^\epsilon}(V_n, W_n) \leq h_n) = o(1)$. Choosing $\bar{\delta}$ so small that $\log \bar{\rho} \leq (1+\delta) \log \rho_\epsilon(A)$, it follows that, with high probability,

$$\mathsf{d}_{\mathsf{G}_n^\epsilon}(V_n, W_n) \geq (1-\delta^2)\frac{\log n}{\log \bar{\rho}} \geq (1-\delta)\frac{\log n}{\log \rho_\epsilon(A)}.$$

Since $\rho_\epsilon(A) = 1/p_{\mathrm{c}}(\epsilon)$ by Corollary 6.4, the proof is complete. $\square$

CHAPTER 8

APPROXIMATION BY A BRANCHING PROCESS

In this chapter, we compare the connected components in the network to the multitype branching process (the IBP) defined in Section 6.1. We begin by coupling the local neighbourhood of a uniformly chosen vertex to the IBP in Sections 8.1 and 8.2. These local considerations allow us to draw conclusions about the existence or nonexistence of the giant component from knowledge of the branching process; see Section 8.4. For the analysis of the typical distances in the network, knowing the local neighbourhood is insufficient. We show in Section 8.3 that a slightly larger IBP dominates the network globally in a suitable way.

8.1 Coupling the network to a tree

The proof of the coupling follows the lines of [37] for the undamaged network, but unfortunately we cannot use their results directly as the coupling in [37] makes extensive use of vertices which are removed in the damaged network. Note however that the removal of the old vertices significantly reduces the risk of cycles in the local neighbourhood of a vertex, and therefore, the coupling here will be successful for much longer than the coupling in [37].

In the first step, we couple the local neighbourhood of a vertex v_0 in G_n^ϵ to a labelled tree $\mathbb{T}_n^\epsilon(v_0)$, thus ruling out cycles in that subgraph. In Section 8.2, we then study the large n-asymptotics of the offspring distributions to arrive at the IBP.

Every vertex v in the labelled tree $\mathbb{T}_n^\epsilon(v_0)$ is equipped with a V_n^ϵ-valued 'tag' and a 'mark' $\alpha \in \mathsf{V}_n^\epsilon \cup \{\ell\}$. The tag indicates which vertex in the network is approximated by v. We use the same notation for vertex and tag to emphasize the similarity between the tree and the network. The mark α carries information about the tag of the parent w of v in the tree. In the spirit of Section 5.6, v has mark $\alpha = \ell$ if its parent has a smaller tag, i.e. $w < v$, and we say that the parent of v is on its left. In contrast, if $w > v$ we say that the parent is on its right. It turns out that here it is beneficial to record the exact tag of w instead of only the relative position and we choose $\alpha = w$. Hence, a typical label is of the form (v, α).

To construct the coupling, we run an exploration process on the connected component of v_0. The offspring distribution of a vertex v in the tree is chosen to be the same as the distribution of direct neighbours of v in G_n^ϵ when only the vertex w is known as whose direct neighbour v is found in the exploration. That vertex w determines the mark of v. The need of this information to identify the offspring distribution is the reason why vertices in $\mathbb{T}_n^\epsilon(v_0)$ are equipped with marks, whereas vertices in $\mathsf{G}_n^\epsilon(v_0)$ are not. Note the similarity to the comparison between network and IBP sketched in Section 5.6.

Formally, for $v_0 \in \mathsf{V}_n^\epsilon$, let $\mathbb{T}_n^\epsilon(v_0)$ be the random tree with root v_0 of label (v_0, ℓ) constructed as follows: every vertex v produces independently offspring to the left, i.e. with tag $u \in \{\lfloor \epsilon n \rfloor + 1, \dots, v-1\}$, with probability

$$\mathbb{P}(v \text{ has a descendant with tag } u) = \mathbb{P}(\triangle \mathcal{Z}[u, v-1] = 1).$$

All offspring on the left are of mark v. Moreover, independently, v produces descendants to its right, i.e. with tag in $\{v+1, \dots, n\}$. Since the parent of these descendant is on the left, they are of mark ℓ. The distribution of the cumulative sum[1] of the sequence of relative positions of the right descendants depends on the mark of v. When v is of mark $\alpha = \ell$, then the cumulative sum is distributed according to the law of $(\mathcal{Z}[v, u] \colon v+1 \le u \le n)$. When v is of mark $\alpha = w \in \mathsf{V}_n^\epsilon$, $w > v$, then the cumulative sum follows the same distribution as $(\mathcal{Z}[v, u] - \mathbb{1}_{[w,\infty)}(u) \colon v+1 \le u \le n)$ conditioned on $\triangle \mathcal{Z}[v, w-1] = 1$. The percolated version $\mathbb{T}_{n,p}^\epsilon(v_0)$ is obtained from $\mathbb{T}_n^\epsilon(v_0)$ by deleting every particle in $\mathbb{T}_n^\epsilon(v_0)$ together with its line of descent with probability $1-p$, independently for all particles. In particular, with probability $1-p$, the root v_0 is deleted and $\mathbb{T}_{n,p}^\epsilon(v_0)$ is empty.

We write $\mathsf{C}_{n,p}^\epsilon(v_0)$ for the connected component in $\mathsf{G}_n^\epsilon(p)$ containing vertex v_0.

Proposition 8.1. *Suppose $(c_n \colon n \in \mathbb{N})$ is a sequence of positive integers that satisfies $\lim_{n\to\infty} c_n^2/n = 0$. Then there exists a coupling of a uniformly chosen vertex V_n in V_n^ϵ, graph $\mathsf{G}_n^\epsilon(p)$ and tree $\mathbb{T}_{n,p}^\epsilon(V_n)$ such that*

$$|\mathsf{C}_{n,p}^\epsilon(V_n)| \wedge c_n = |\mathbb{T}_{n,p}^\epsilon(V_n)| \wedge c_n \qquad \textit{with high probability.}$$

To prove Proposition 8.1, we define an exploration process which we then use to inductively collect information about the tree and the network on the same probability space. We show that the two discovered graphs agree until a stopping time, which is with high probability larger than c_n. After that time, the undiscovered part of the tree and the network can be generated independently of each other.

We begin by specifying the exploration process that is used to explore the connected

[1] For a sequence $(x_j \colon j = 1, \dots, n)$ the cumulative sum is given by $(\sum_{i=1}^j x_i \colon j = 1, \dots, n)$

component of a vertex v_0 in a labelled graph G, like $\mathsf{C}^{\epsilon}_{n,p}(v_0)$ or $\mathbb{T}^{\epsilon}_{n,p}(v_0)$. We distinguish three categories of vertices:

- *veiled vertices:* vertices for which we have not yet found a connection to the cluster of v_0,
- *active vertices:* vertices for which we already know that they belong to the cluster of v_0 but for which we have not yet explored all its immediate neighbours,
- *dead vertices:* vertices which belong to the cluster of v_0 and for which all immediate neighbours have been explored.

At the beginning of the exploration only v_0 is active and all other vertices are veiled. In the first exploration step we explore all immediate neighbours of v_0, declare v_0 as dead and all its immediate neighbours as active. The other vertices remain veiled. We now continue from the active vertex v with the smallest tag and explore all its immediate neighbours apart from v_0 from where we just came. The exploration is continued until there are no active vertices left.

We couple the exploration processes of the network and the tree started with $v_0 \in \mathsf{V}^{\epsilon}_n$ up to a stopping time T, such that up to time T both explored subgraphs (without the marks) coincide. In particular, the explored part of $\mathsf{C}^{\epsilon}_{n,p}(v_0)$ is a tree, and every tag has been used at most once by the active or dead vertices in $\mathbb{T}^{\epsilon}_{n,p}(v_0)$. If at least one of these properties fails, then we say that the *coupling fails*. We also stop the exploration, when either the number of dead and active vertices exceeds c_n or when there are no active vertices left. In this case we say that the *coupling is successful*.

Lemma 8.2. *Suppose that $p \in (0,1]$ and $(c_n : n \in \mathbb{N})$ satisfies $\lim_{n\to\infty} c_n^2/n = 0$. Then*

$$\lim_{n\to\infty} \sup_{v_0 \in \mathsf{V}^{\epsilon}_n} \mathbb{P}\big(\textit{the coupling of } \mathsf{C}^{\epsilon}_{n,p}(v_0) \textit{ and } \mathbb{T}^{\epsilon}_{n,p}(v_0) \textit{ fails}\big) = 0.$$

In the sequel, we will label some key constants by the lemma in which they appear first. The following result will be used in the proof of Lemma 8.2.

Lemma 8.3 (Adaptation of Lemma 2.12 in [37])**.** *Let $(c_n : n \in \mathbb{N})$ be such that $\lim_{n\to\infty} c_n/n = 0$. Then there exists a constant $C_{8.3} > 0$ such that for all sufficiently large n, for all disjoint sets $\mathcal{I}_0, \mathcal{I}_1 \subseteq \mathsf{V}^{\epsilon}_n$ with $|\mathcal{I}_0| \le c_n$ and $|\mathcal{I}_1| \le 1$, and for all $u, v \in \mathsf{V}^{\epsilon}_n$,*

$$\mathbb{P}\big(\triangle\mathcal{Z}[v,u] = 1 \,\big|\, \triangle\mathcal{Z}[v,i] = 1 \textit{ for } i \in \mathcal{I}_1, \triangle\mathcal{Z}[v,i] = 0 \textit{ for } i \in \mathcal{I}_0\big) \\ \le C_{8.3}\mathbb{P}\big(\triangle\mathcal{Z}[v,u] = 1 \,\big|\, \triangle\mathcal{Z}[v,i] = 1 \textit{ for } i \in \mathcal{I}_1\big).$$

Proof. We have

$$\mathbb{P}\big(\triangle\mathcal{Z}[v,u] = 1 \,\big|\, \triangle\mathcal{Z}[v,i] = 1 \text{ for } i \in \mathcal{I}_1, \triangle\mathcal{Z}[v,i] = 0 \text{ for } i \in \mathcal{I}_0\big) \\ \le \frac{\mathbb{P}(\triangle\mathcal{Z}[v,u] = 1 \,|\, \triangle\mathcal{Z}[v,i] = 1 \text{ for } i \in \mathcal{I}_1)}{\mathbb{P}(\triangle\mathcal{Z}[v,i] = 0 \text{ for } i \in \mathcal{I}_0 \,|\, \triangle\mathcal{Z}[v,i] = 1 \text{ for } i \in \mathcal{I}_1)}.$$

With n so large that $\lfloor \epsilon n \rfloor \geq \hat{n}$, Lemma 7.10 and (7.9) imply that

$$\begin{aligned}\mathbb{P}(\triangle\mathcal{Z}[v,i]=0 \text{ for } i\in\mathcal{I}_0 \,|\, \triangle\mathcal{Z}[v,i]=1 \text{ for } i\in\mathcal{I}_1) &\geq \mathbb{P}^1(\triangle\mathcal{Z}[v,i]=0 \text{ for } i\in\mathcal{I}_0)\\ \geq \prod_{i\in\mathcal{I}_0}\mathbb{P}^1(\triangle\mathcal{Z}[v,i]=0) \geq \prod_{i\in\mathcal{I}_0}\Big(1-\frac{f(1)}{(v-1)^{\gamma}i^{1-\gamma}}\Big) &\geq \Big(1-\frac{f(1)}{\epsilon n}\Big)^{c_n}.\end{aligned}$$

Since c_n/n tends to zero as $n\to\infty$, the right-hand side converges to one. $\square$

Proof of Lemma 8.2. We assume that n is so large that $\lfloor \epsilon n \rfloor \geq \hat{n}$. To distinguish the exploration processes, we use the term *descendant* for a child in the labelled tree and the term *neighbour* in the context of $\mathsf{G}_n^\epsilon(p)$. The σ-algebra generated by the exploration until the completion of step k is denoted $\mathcal{F}_k$.

Since the probability of removing v_0 is the same in $\mathsf{C}^\epsilon_{n,p}(v_0)$ and $\mathbb{T}^\epsilon_{n,p}(v_0)$, this event can be perfectly coupled. If v_0 is not removed, then we explore the immediate neighbours of v_0 in $\mathsf{G}_n^\epsilon(p)$ and the children of the root v_0 in the tree. Again these families are identically distributed and can be perfectly coupled.

Now suppose that we successfully completed exploration step k and are about to start the next step from vertex v. At this stage, every vertex in the tree can be uniquely referred to by its tag, and the subgraphs coincide. Denoting by $\mathfrak{a}$ and $\mathfrak{d}$ the set of active and dead vertices, respectively, we have $\mathfrak{a}\neq\emptyset$ and $|\mathfrak{a}\cup\mathfrak{d}|<c_n$. We continue by exploring the left descendants and neighbours of v. Since we always explore the leftmost active vertex, we cannot encounter any dead or active neighbours in this step. However, in the tree $\mathbb{T}^\epsilon_{n,p}(v_0)$ we may find a dead left descendant (i.e. an offspring whose tag agrees with the tag of a dead particle); we call this event Ia. On Ia, the vertices in the explored part of $\mathbb{T}^\epsilon_{n,p}(v_0)$ are no longer uniquely identifiable by their tag and we stop. We have

$$\begin{aligned}\mathbb{P}\big(\mathrm{Ia}\,|\,\mathcal{F}_k\big) &= \mathbb{P}\big(\exists d\in\mathfrak{d}\colon d \text{ is a left descendant of } v\,|\,\mathcal{F}_k\big)\\ &\leq \sum_{d\in\mathfrak{d}}\mathbb{P}(\triangle\mathcal{Z}[d,v-1]=1)\leq c_n\frac{f(0)}{\epsilon n}.\end{aligned}$$

In the first inequality, we used subadditivity, the definition of $\mathbb{T}^\epsilon_{n,p}(v_0)$, and omitted the event that offspring of v may be removed by percolation. Hence, $\mathbb{P}(\mathrm{Ia})=O(c_n/n)$. In the exploration to the left in the tree, we immediately check if a found left descendant has a right descendant which is dead. We denote this event by Ib and stop the exploration as soon as it occurs. The reason is that in the network this event could not happen since we always explore the leftmost active vertex. The distribution of left neighbours agrees with the distribution of the left descendants conditioned on having no dead right descendants, and we can couple both explorations such that they agree in this case. The probability of the adverse event Ib, $\mathbb{P}\big(\mathrm{Ib}\,|\,\mathcal{F}_k\big)$, is given by

$$\mathbb{P}\big(\exists u\in\mathfrak{d}^c, d\in\mathfrak{d}\colon u \text{ is a left descendant of } v,\, d \text{ is a right descendant of } u\,|\,\mathcal{F}_k\big).$$

Using the definition of $\mathbb{T}^{\epsilon}_{n,p}(v)$, this probability can be bounded from above by

$$\sum_{u\in\mathfrak{d}^c}\sum_{d\in\mathfrak{d}}\mathbb{P}\big(\triangle\mathcal{Z}[u,v-1]=1\big)\mathbb{P}\big(\triangle\mathcal{Z}[u,d-1]=1\,|\,\triangle\mathcal{Z}[u,v-1]=1\big).$$

By the definition of the exploration process, there are at most c_n dead vertices. Therefore, Lemma 7.10 and (7.9) yield

$$\mathbb{P}\big(\mathrm{Ib}|\mathcal{F}_k\big)\le c_n\sum_{u\in\mathfrak{d}^c,u\le v-1}\frac{f(0)}{(u-1)^{\gamma}(v-1)^{1-\gamma}}\frac{f(1)}{\epsilon n}\le c_n\frac{f(0)f(1)}{\epsilon n}\frac{1}{(v-1)^{1-\gamma}}\sum_{u=1}^{v-1}u^{-\gamma},$$

which implies in particular that $\mathbb{P}(\mathrm{Ib})=O(c_n/n)$.

We turn to the exploration of right descendants, resp. neighbours. When vertex v is of mark $\alpha\neq\ell$, then we already know that v has no right descendants, resp. neighbours, in $\mathfrak{d}$ since we checked this when v was discovered. We denote the event that a right descendant, resp. neighbour, is active by IIr and stop the exploration as soon as this event occurs because the tags in $\mathbb{T}^{\epsilon}_{n,p}(v_0)$ are no longer unique, resp. we found a cycle in $\mathsf{C}^{\epsilon}_{n,p}(v_0)$. According to Lemma 8.3 and (7.9), the probability $\mathbb{P}\big(\mathrm{IIr}\,\big|\,\mathcal{F}_k\big)$ can be bounded from above by

$$\begin{aligned}\mathbb{P}\Big(\exists a\in\mathfrak{a}\colon\triangle\mathcal{Z}[v,a-1]=1\,\Big|\,\triangle\mathcal{Z}[v,\alpha-1]=1,\triangle\mathcal{Z}[v,d-1]=0\,\forall\,d\in\mathfrak{d}\setminus\{\alpha\},\mathcal{F}_k\Big)&\\ \le C_{8.3}\sum_{a\in\mathfrak{a}}\mathbb{P}^1(\triangle\mathcal{Z}[v,a-1]=1)\le C_{8.3}c_n\frac{f(1)}{\epsilon n}.&\end{aligned}$$

Thus, $\mathbb{P}(\mathrm{IIr})=O(c_n/n)$. Conditional on the event that there are no active vertices in the set of right descendants, resp. neighbours, the offspring distributions in tree and network agree, and can, therefore, be perfectly coupled. When the vertex v is of mark $\alpha=\ell$, then we have not gained any information about its right descendants, yet. The event that there is a dead or active vertex in the right descendants is denoted by IIℓa. We stop when this event occurs, and use (7.9) to estimate

$$\begin{aligned}\mathbb{P}\big(\mathrm{II}\ell\mathrm{a}|\mathcal{F}_k\big)&=\mathbb{P}\big(\exists a\in\mathfrak{a}\cup\mathfrak{d}\colon a\text{ is a right descendant of }v\big|\mathcal{F}_k\big)\\&\le\sum_{a\in\mathfrak{a}\cup\mathfrak{d}}\mathbb{P}(\triangle\mathcal{Z}[v,a-1]=1)\le c_n\frac{f(0)}{\epsilon n}.\end{aligned}$$

Thus, $\mathbb{P}(\mathrm{II}\ell\mathrm{a})=O(c_n/n)$. In $\mathsf{C}^{\epsilon}_{n,p}(v_0)$ we know that v has no dead right neighbours as this would have stopped the exploration in the moment when v became active. The event that there are active vertices in the set of right neighbours is denoted by IIℓb, and we stop as soon as it occurs since a cycle is created. Using again (7.9), we find

$$\begin{aligned}\mathbb{P}\big(\mathrm{II}\ell\mathrm{b}|\mathcal{F}_k\big)&=\mathbb{P}\big(\exists a\in\mathfrak{a}\colon\triangle\mathcal{Z}[v,a-1]=1\big|\triangle\mathcal{Z}[v,d-1]=0\text{ for }d\in\mathfrak{d},\mathcal{F}_k\big)\\&\le\sum_{a\in\mathfrak{a}}\mathbb{P}(\triangle\mathcal{Z}[v,a-1]=1)\le c_n\frac{f(0)}{\epsilon n}.\end{aligned}$$

As in the case $\alpha \neq \ell$, the explorations can be perfectly coupled when the adverse events do not occur. We showed that in every step the coupling fails with a probability bounded by $O(c_n/n)$. As there are at most c_n exploration steps until we end the coupling successfully, the probability of failure is $O(c_n^2/n) = o(1)$. In other words, the coupling succeeds with high probability. $\square$

Proof of Proposition 8.1. First, consider the statement for a fixed vertex v_0. When the coupling is successful and ends because at least c_n vertices were explored, then $|\mathsf{C}^\epsilon_{n,p}(v_0)| \geq c_n$ and $|\mathbb{T}^\epsilon_{n,p}(v_0)| \geq c_n$. If the coupling is successful and ends because there are no active vertices left, then $|\mathsf{C}^\epsilon_{n,p}(v_0)| = |\mathbb{T}^\epsilon_{n,p}(v_0)|$ since the subgraphs coincide. As the coupling is successful with high probability by Lemma 8.2, $|\mathsf{C}^\epsilon_{n,p}(v_0)| \wedge c_n = |\mathbb{T}^\epsilon_{n,p}(v_0)| \wedge c_n$ with high probability. Because Lemma 8.2 shows the success of the coupling uniformly in the start vertex, the randomization of the vertex v_0 to a uniformly chosen vertex $V_n \in \mathsf{V}^\epsilon_n$ is now straightforward. $\square$

8.2 Coupling the tree to the IBP

Coupling the neighbourhood of a vertex to a labelled tree provides a great simplification of the problem since many dependencies are eliminated. However, the offspring distribution in the tree $\mathbb{T}^\epsilon_{n,p}(V_n)$ is still complicated and depends on n. Since we are mainly interested in the asymptotic size of the giant component, we now couple the tree to the IBP, which does not depend on n and is much easier to analyse. We denote by $|\mathcal{X}^\epsilon(p)|$ the total progeny of the IBP. Recall the definition of S^ϵ from (6.2).

Proposition 8.4. *Let $p \in (0,1]$, and let $(c_n \colon n \in \mathbb{N})$ be a sequence of positive integers with $\lim_{n\to\infty} c_n^3/n = 0$. Then there exists a coupling of a uniformly chosen vertex V_n in V^ϵ_n, the graph $\mathsf{G}^\epsilon_n(p)$ and the percolated IBP started with a particle of mark ℓ and location S^ϵ such that, with high probability,*

$$|\mathsf{C}^\epsilon_{n,p}(V_n)| \wedge c_n = |\mathcal{X}^\epsilon(p)| \wedge c_n.$$

Proof of Proposition 8.4. Throughout the proof, we suppose that n is so large that $\lfloor \epsilon n \rfloor \geq \hat{n}$. Instead of coupling the IBP directly to the network, we couple a projected version of the IBP to the tree $\mathbb{T}^\epsilon_{n,p}(V_n)$. As long as the number of particles is preserved under the projection, this is sufficient according to Proposition 8.1. To describe the projection, we define $\pi^\epsilon_n \colon [\log \epsilon, 0] \to \mathsf{V}^\epsilon_n$ by

$$\pi^\epsilon_n(\lambda) = v \qquad \Leftrightarrow \qquad s_n(v-1) < \lambda \leq s_n(v), \tag{8.1}$$

where $s_n(v) = -\sum_{j=v}^{n-1} \frac{1}{j}$. Since $s_n(\lfloor \epsilon n \rfloor) < \log(\lfloor \epsilon n \rfloor / n) \leq \log \epsilon$, every location in $[\log \epsilon, 0]$ can be uniquely identified with a tag in V^ϵ_n by the map π^ϵ_n. The projected IBP is again a labelled tree: the genealogical tree of the IBP with its marks is preserved, the location of a particle x is replaced by the tag $\pi^\epsilon_n(\lambda(x))$. If $s_n(\lfloor \epsilon n \rfloor + 1) < \log \epsilon$, then no particles of the IBP are projected onto $\lfloor \epsilon n \rfloor + 1$. Moreover, while for $v \geq \lfloor \epsilon n \rfloor + 3$ an interval of length $1/(v-1)$ is projected onto v, for $\lfloor \epsilon n \rfloor + 2$ only an interval of length at most $s_n(\lfloor \epsilon n \rfloor + 2) - \log \epsilon$ is used. This length is positive but may be smaller than $1/(\lfloor \epsilon n \rfloor + 1)$. As a consequence, the projected IBP can have unusually few particles at $\lfloor \epsilon n \rfloor + 1$ and $\lfloor \epsilon n \rfloor + 2$, and we treat these two tags separately.

The exploration of the two trees follows the same procedure as the exploration described in Section 8.1, and we declare the coupling successful and stop as soon as either there are no active vertices left or the number of active and dead vertices exceeds c_n. Since both objects are trees, as long as the labels for the starting vertices agree, any failure of the coupling comes from a failure in the coupling of the offspring distributions. For simplicity, we consider only the case $p = 1$. The generalisation to $p \in (0, 1]$ is straightforward.

We first show that the labels of the starting vertices can be coupled with high probability. To this end, note that the distribution of S^ϵ is chosen such that $\exp(S^\epsilon)$ is uniformly distributed on $[\epsilon, 1]$. Since $\log \epsilon \leq s_n(\lfloor \epsilon n \rfloor + 2) \leq s_n(v-1) \leq s_n(v) \leq 0$, for $v \geq \lfloor \epsilon n \rfloor + 3$, we obtain

$$\begin{aligned}\mathbb{P}\big(\pi_n^\epsilon(S^\epsilon) = v\big) &= \mathbb{P}\big(e^{s_n(v-1)} < e^{S^\epsilon} \leq e^{s_n(v)}\big) \\ &= \tfrac{1}{1-\epsilon}\big(e^{s_n(v)} - e^{s_n(v-1)}\big) = \tfrac{1}{1-\epsilon}\, e^{s_n(v-1)}\big(e^{1/(v-1)} - 1\big).\end{aligned}$$

The right-hand side is in the interval $[\frac{1}{1-\epsilon}(\frac{1}{n} - \frac{2}{vn}), \frac{1}{1-\epsilon}(\frac{1}{n} + \frac{2}{vn})]$. Moreover, the probability that V_n or $\pi_n^\epsilon(S^\epsilon)$ is in $\{\lfloor \epsilon n \rfloor + 1, \lfloor \epsilon n \rfloor + 2\}$ is of order $O(1/n)$. Hence, V_n and S^ϵ can be coupled such that

$$\mathbb{P}\big(V_n \neq \pi_n^\epsilon(S^\epsilon)\big) \leq \sum_{v = \lfloor \epsilon n \rfloor + 3}^{n} \big|\mathbb{P}\big(\pi_n^\epsilon(S^\epsilon) = v\big) - \tfrac{1}{n - \lfloor \epsilon n \rfloor}\big| + O\big(\tfrac{1}{n}\big) = O\big(\tfrac{\log n}{n}\big).$$

In the next step, we study the offspring distributions of a particle x in the IBP with label (λ, α) and $\pi_n^\epsilon(\lambda) = v$. We start with the offspring to the left. Let $u \in \{\lfloor \epsilon n \rfloor + 1, \ldots, v\}$. By the definition of the IBP, the number of projected offspring of x that have tag u is Poisson-distributed with parameter

$$\int_{(s_n(u-1)-\lambda) \vee (\log \epsilon - \lambda)}^{(s_n(u)-\lambda) \wedge 0} \beta e^{(1-\gamma)t}\, dt.$$

A vertex with tag v in $\mathbb{T}_n^\epsilon(V_n)$ produces a Bernoulli-distributed number of descendants with tag u with success probability $\mathbb{P}(\triangle \mathcal{Z}[u, v-1] = 1)$ when $u < v$, and with success probability zero when $u = v$. It is proved in Lemma 6.3 of [37] that for $u \geq \lfloor \epsilon n \rfloor + 3$ the Poisson distributions can be coupled to the Bernoulli distribution such that they disagree with a probability bounded by a constant multiple of $v^{\gamma-1} u^{-(\gamma+1)}$ for $u < v$ and $1/v$ for $u = v$. For $u \in \{\lfloor \epsilon n \rfloor + 1, \lfloor \epsilon n \rfloor + 2\}$, a similar estimate shows that the probability can be bounded by a constant multiple of $1/(\epsilon n)$. Since the number of descendants with tag in $\{\lfloor \epsilon n \rfloor + 1, \dots, v\}$ form an independent sequence of random variables, we can apply the coupling sequentially for each location and obtain a coupling of the π_n^ϵ-projected left descendants in the IBP and the left descendants in $\mathbb{T}_n^\epsilon(V_n)$. The failure probability of this coupling can be estimated by

$$\begin{aligned}\mathbb{P}(\text{left descendants of } v \text{ disagree}) &\leq \frac{3C}{\epsilon n} + \frac{C'}{v^{1-\gamma}} \sum_{u=\lfloor \epsilon n \rfloor + 3}^{v-1} \frac{1}{u^{\gamma+1}} \\ &\leq \frac{3C}{\epsilon n} + \frac{C'}{\epsilon n} \log\left(\frac{v-1}{\lfloor \epsilon n \rfloor}\right) \leq \frac{C''}{n},\end{aligned}$$

where C, C', C'' are suitable positive constants whose value can change from line to line in the sequel. We turn to the offspring on the right. Suppose that particle x in the IBP has mark $\alpha = \ell$. The cumulative sum of π_n^ϵ-projected right descendants of x follows the same distribution as $(Z_{s_n(u)-\lambda} \colon v \leq u \leq n)$. The cumulative sum of right descendants of v in $\mathbb{T}_n^\epsilon(V_n)$ is distributed according to the law of $(\mathcal{Z}[v,u] \colon v \leq u \leq n)$. The following lemma is taken from [37], and we omit its proof.

Lemma 8.5 (Lemma 6.2 in [37])**.** *Fix a level $H \in \mathbb{N}$. We can couple the processes $(Z_{s_n(u)-\lambda} \colon v \leq u \leq n)$ and $(\mathcal{Z}[v,u] \colon v \leq u \leq n)$ such that for the coupled processes $(\mathcal{Y}_u^{(1)} \colon v \leq u \leq n)$ and $(\mathcal{Y}_u^{(2)} \colon v \leq u \leq n)$,*

$$\mathbb{P}(\mathcal{Y}_u^{(1)} \neq \mathcal{Y}_u^{(2)} \text{ for some } u \leq \sigma_H) \leq C_{8.5} \frac{f(H)^2}{v-1}$$

for some constant $C_{8.5} > 0$, and where σ_H denotes the first time that one of the processes reaches or exceeds H.

In the coupling between the tree $\mathbb{T}_n^\epsilon(V_n)$ and the projected IBP we consider at most c_n right descendants. Hence, Lemma 8.5 implies that the distributions can be coupled such that

$$\mathbb{P}(\text{right descendants of } v \text{ disagree}) \leq C_{8.5} \frac{f(c_n)^2}{v-1} \leq C \frac{c_n^2}{\epsilon n},$$

for some $C > 0$. When $\alpha = \mathrm{r}$, then the cumulative sum of π_n^ϵ-projected right descendants of x follows the same distribution as $(\hat{Z}_{s_n(u)-\lambda} - 1 \colon v \leq u \leq n)$. The cumulative sum of right descendants of a vertex v with mark $w \in \mathsf{V}_n^\epsilon$, $w > v$, in $\mathbb{T}_n^\epsilon(V_n)$ is distributed according to $(\mathcal{Z}[v,u] - \mathbb{1}_{[w,\infty)}(u) \colon v \leq u \leq n)$ conditioned on $\triangle\mathcal{Z}[v, w-1] = 1$. We can couple these two distributions. Again the proof of the following lemma is given in [37] up to minor changes and therefore omitted.

Lemma 8.6 (Lemma 6.6 in [37])**.** *Fix a level $H \in \mathbb{N}$. We can couple the processes $(\hat{Z}_{s_n(u)-\lambda} - 1 \colon v \leq u \leq n)$ and $(\mathcal{Z}[v,u] - \mathbb{1}_{[w,\infty)}(u) \colon v \leq u \leq n)$ conditioned on $\triangle\mathcal{Z}[v, w-1] = 1$ such that for the coupled processes $(\mathcal{Y}_u^{(1)} \colon v \leq u \leq n)$ and $(\mathcal{Y}_u^{(2)} \colon v \leq u \leq n)$,*

$$\mathbb{P}(\mathcal{Y}_u^{(1)} \neq \mathcal{Y}_u^{(2)} \text{ for some } u \leq \sigma_H) \leq C_{8.6} \frac{f(H)^2}{v-1},$$

for some constant $C_{8.6} > 0$, and where σ_H denotes the first time that one of the processes reaches or exceeds H.

As we explore at most c_n vertices during the exploration, Lemma 8.6 implies that we can couple the offspring distribution to the right such that there is a constant $C > 0$ with

$$\mathbb{P}(\text{right descendants of } v \text{ disagree}) \leq C_{8.6} \frac{f(c_n)^2}{v-1} \leq C \frac{c_n^2}{\epsilon n}.$$

Since we explore at most c_n vertices in total, the probability that the coupling fails can be bounded by a constant multiple of $c_n/n + c_n^3/n$, which converges to zero. Thus, the two explorations can be successfully coupled with high probability and, as in the proof of Proposition 8.1, the claim follows. □

8.3 Dominating the network by a branching process

Like in the coupling, we begin with a comparison to a tree: for $\theta \in \mathbb{N}$ and $v_0 \in \mathsf{V}_n^\epsilon$, let $\mathbb{T}_n^{\epsilon,\theta}(v_0)$ be the subtree of $\mathbb{T}_n^\epsilon(v_0)$, where every particle can have at most θ offspring to the right. That is, for a particle with tag v and mark $\alpha = \ell$, the cumulative sum of the offspring to the right is distributed according to the law of $(\mathcal{Z}[v,u] \wedge \theta \colon v+1 \leq u \leq n)$. When v is of mark $\alpha = w \in \mathsf{V}_n^\epsilon$, $w > v$, then the cumulative sum follows the same distribution as $((\mathcal{Z}[v,u] - \mathbb{1}_{[w,\infty)}(u)) \wedge \theta \colon v+1 \leq u \leq n)$ conditioned on $\triangle\mathcal{Z}[v, w-1] = 1$. We refer to the particles at graph distance h from the root in $\mathbb{T}_n^{\epsilon,\theta}(v_0)$ as the h-th generation. Recall from Section 7.2 that $N_h^\theta(v_0)$ denotes the number of θ-admissible paths of length h in G_n^ϵ with initial vertex v_0.

Lemma 8.7. *For all $\theta, h, n \in \mathbb{N}$, $v_0 \in \mathsf{V}_n^\epsilon$,*

$$\mathbb{E}\big[N_h^\theta(v_0)\big] \leq \mathbb{E}\Big[\big|\big\{\textit{particles in generation } h \textit{ of } \mathbb{T}_n^{\epsilon,\theta}(v_0)\big\}\big|\Big].$$

Proof. Let $\mathsf{p} = (v_0, \ldots, v_h) \in \mathcal{S}_h(v_0)$. Using the notation and terminology from the proof of Theorem 5.3, and the definition of the tree $\mathbb{T}_n^{\epsilon,\theta}(v_0)$, one easily checks that in cases A, B, C, E and F of Figure II-6 on page 118, $\mathbb{P}(\mathcal{E}_h | \cap_{i=1}^{h-1} \mathcal{E}_i)$ agrees with the probability that in tree $\mathbb{T}_n^{\epsilon,\theta}(v_0)$ a particle with tag v_{h-1} gives birth to a particle of tag v_h given that its parent has tag v_{h-2}. In case D of Figure II-6, the tree $\mathbb{T}_n^{\epsilon,\theta}$ is allowed to have one more offspring on its right because the edge $\{v_{h-2}, v_{h-1}\}$ is not accounted for. Hence, $\mathbb{P}(\mathcal{E}_h | \cap_{i=1}^{h-1} \mathcal{E}_i)$ is bounded from above by the probability for the event in the tree. We obtain

$$\mathbb{E}[N_h^\theta(v_0)] = \sum_{\mathsf{p} \in \mathcal{S}_h(v_0)} \mathbb{P}(\mathsf{p} \text{ is } \theta\text{-admissible in } \mathsf{G}_n^\epsilon) \leq \sum_{\mathsf{p} \in \mathcal{S}_h(v_0)} \mathbb{P}(\mathsf{p} \text{ present in } \mathbb{T}_n^{\epsilon,\theta}(v_0)).$$

Particles in generation h of $\mathbb{T}_n^{\epsilon,\theta}(v_0)$, who have two ancestors with the same tag, are not represented in the sum on the right-hand side. Adding these, we obtain the result. □

Proof of Lemma 7.12. By Lemma 8.7, it suffices to show that, for every h, the number of particles in the h-th generation of $\mathbb{T}_n^{\epsilon,\theta_n}(v_0)$ is stochastically dominated by the number of particles in $\mathrm{IBP}_h^{\epsilon}((1+\delta)f)$ started in $s_n(v_0)$, or, as in the proof of Proposition 8.4, by the number of particles in the h-th generation of the π_n^{ϵ}-projected $\mathrm{IBP}_h^{\epsilon}((1+\delta)f)$ defined in (8.1). Since both processes are trees starting with the same type of particle, it suffices to compare the offspring distributions. All particles in $\mathbb{T}_n^{\epsilon,\theta_n}(v_0)$ have a tag $v > \lfloor \epsilon n \rfloor$, but the projected IBP can have offspring with tag $v \in \{\lfloor \underline{\epsilon} n \rfloor + 1, \dots, \lfloor \epsilon n \rfloor\}$. Hence, these offspring are ignored in the following, giving us a lower bound on the projected IBP. We assume that n is so large, that $n \geq \hat{n}$ and $s_n(\lfloor \epsilon n \rfloor + 1) \geq \log \underline{\epsilon}$.

Let x be a particle in the IBP of type (λ, α) with $\pi_n^{\epsilon}(\lambda) = v$. We begin with the offspring to the left, i.e. tag $u \in \{\lfloor \epsilon n \rfloor + 1, \dots, v\}$. A particle in $\mathbb{T}_n^{\epsilon,\theta_n}(v_0)$ with tag v cannot produce particles in $u = v$, therefore, the IBP clearly dominates. For $u < v$, using (7.9), the probability that a particle with tag u is a child of x, is

$$\mathbb{P}(\triangle \mathcal{Z}[u, v-1] = 1)) \leq \beta (u-1)^{-\gamma} (v-1)^{-(1-\gamma)}.$$

Writing $\bar{f}(k) = (1+\delta) f(k) = \bar{\gamma} k + \bar{\beta}$, for $k \in \mathbb{N}_0$, the number of particles with tag u produced by x in the projected IBP follows a Poisson distribution with parameter

$$\int_{s_n(u-1)-\lambda}^{s_n(u)-\lambda} \bar{\beta} e^{(1-\bar{\gamma})t}\, dt \leq \frac{\bar{\beta}}{u-1} e^{-(1-\gamma)\sum_{k=u-1}^{v-1} \frac{1}{k}} \leq \frac{\bar{\beta}}{u-1} \Big(\frac{u-2}{v-1}\Big)^{1-\gamma},$$

where we used that $\lambda \leq s_n(v)$ and $e^y - 1 \geq y$. For $\varrho > 0$, $\eta \in [0,1]$, the Poisson distribution with parameter ϱ is dominating the Bernoulli distribution with parameter η if and only if $e^{-\varrho} \leq 1 - \eta$. Since $e^{-y} \leq 1 - y + y^2/2$ for $y \geq 0$, it suffices to show that $\varrho(1 - \varrho/2) \geq \eta$. In our case, $\eta = \beta(u-1)^{-\gamma}(v-1)^{-(1-\gamma)}$, $\varrho = \eta(1+\delta)(1 - 1/(u-1))^{1-\gamma}$ and the inequality holds for all large n and $u \in \mathsf{V}_n^{\epsilon}$, $u < v$, since η is a null sequence.

We turn to the right descendants. The pure jump process corresponding to the attachment rule $\bar{f}$ is denoted by $\bar{Z}$, and we write P^l for the distribution of $\bar{Z}$ when started in l, that is, $P^l(\bar{Z}_0 = l) = 1$. First suppose that $\alpha = \ell$. The cumulative sum of π_n^{ϵ}-projected right descendants of x have the distribution of $(\bar{Z}_{s_n(u)-\lambda} \colon v \leq u \leq n)$, where $\bar{Z}_0 = 0$. The cumulative sum of right descendants of v in $\mathbb{T}_n^{\epsilon,\theta_n}(v_0)$ is distributed according to the law of $(\mathcal{Z}[v,u] \wedge \theta_n \colon v \leq u \leq n)$. We couple these distributions by defining $((\mathcal{Y}_u^{(1)}, \mathcal{Y}_u^{(2)}) \colon v \leq u \leq n)$ to be the time-inhomogeneous Markov chain which starts in $P^0(\bar{Z}_{s_n(v)-\lambda} \in \cdot) \otimes \delta_0$, has the desired marginals, and evolves from state (l,k) at time j according to a coupling of $\bar{Z}_{1/j}$ and $\mathcal{Z}[j, j+1]$ which guarantees that $\bar{Z}_{1/j} \geq \mathcal{Z}[j, j+1]$ until $\mathcal{Y}^{(2)}$ reaches state θ_n, where $\mathcal{Y}^{(2)}$ is absorbed. To prove that this coupling exists, it suffices to show that

$$e^{-\bar{f}(l)/j} = P^l(\bar{Z}_{1/j} = l) \leq \mathbb{P}^k(\mathcal{Z}[j, j+1] = k) = 1 - f(k)/j \qquad \text{for } j \in \mathsf{V}_n^{\epsilon}, k \leq \theta_n, k \leq l.$$

Since $\bar{f}$ is non decreasing, this inequality follows as above once we show that $\varrho(1-\varrho/2) \geq$

η with $\eta = f(k)/j$, $\varrho = \bar{f}(k)/j = \eta(1+\delta)$. Since $k \le \theta_n = o(n)$ and $j \ge \lfloor \epsilon n \rfloor$, η is a null sequence and the claim follows. Hence, $\mathcal{Y}_j^{(1)} \ge \mathcal{Y}_j^{(2)}$ for all j, and the domination is established.

Now suppose that $\alpha = \mathrm{r}$ and that the location of x's parent is projected onto tag w. The cumulative sum of π_n^ϵ-projected right descendants of x has the distribution of $(Y_{s_n(u)-\lambda} \colon v \le u \le n)$, where Y is a version of $\bar{Z}$ under measure P^1. The cumulative sum of right descendants of v in $\mathbb{T}_n^{\epsilon,\theta_n}(v_0)$ is distributed according to the law of $((\mathcal{Z}[v,u] - \mathbb{1}_{[w,\infty)}(u)) \wedge \theta_n \colon v \le u \le n)$ conditioned on $\triangle\mathcal{Z}[v, w-1] = 1$. We couple these distributions as in the $\alpha = \ell$ case, but, for times $j \le w-2$, the Markov chain evolves from state (l,k) according to a coupling of $Y_{1/j}$ and $\mathcal{Z}[j,j+1]$ conditioned on $\triangle\mathcal{Z}[j,w-1] = 1$ which guarantees that $Y_{1/j} \ge \mathcal{Z}[j,j+1]$ until either $j = w-2$ or $\mathcal{Y}^{(2)}$ reaches θ_n and is absorbed. To show that this coupling exists, it suffices to show that for all $j \in \mathsf{V}_n^\epsilon$, $k \le \theta_n$, $k \le l$,

$$P^{l+1}(\bar{Z}_{1/j} - 1 = l) \le \mathbb{P}^k(\mathcal{Z}[j,j+1] = k | \mathcal{Z}[j,w-1] = 1). \tag{8.2}$$

We compute

$$\begin{aligned}
\mathbb{P}^k(\mathcal{Z}[j,j+1] = k | \triangle\mathcal{Z}[j,w-1] = 1) &= 1 - \frac{\mathbb{P}^k(\triangle\mathcal{Z}[j,j] = 1, \triangle\mathcal{Z}[j,w-1] = 1)}{\mathbb{P}^k(\triangle\mathcal{Z}[j,w-1] = 1)} \\
&= 1 - \frac{\frac{f(k)}{j}\mathbb{P}^{k+1}(\triangle\mathcal{Z}[j+1,w-1] = 1)}{\mathbb{P}^k(\triangle\mathcal{Z}[j,w-1] = 1)} \\
&= 1 - \frac{f(k+1)}{j+\gamma}.
\end{aligned}$$

Since $\bar{f}$ is non-decreasing, (8.2) follows when we show that $\varrho(1-\varrho/2) \ge \eta$ with $\eta = f(k+1)/(j+\gamma)$ and $\varrho = \bar{f}(k+1)/j = \eta(1+\delta)(1+\gamma/j)$. Since $k \le \theta_n = o(n)$ and $j \ge \lfloor \epsilon n \rfloor$, η is a null sequence, and (8.2) is proved. In the transition from generation $j = w-1$ to $j = w$, $\mathcal{Y}^{(2)}$ cannot change its state while $\mathcal{Y}^{(1)}$ can increase. From generation $j = w$ onwards, the coupling explained in case $\alpha = \ell$ is used. Thus, the Markov chain can be constructed such that $\mathcal{Y}_j^{(1)} \ge \mathcal{Y}_j^{(2)}$ for all j and the domination is proved. $\square$

8.4 Proof of Theorem 6.1

Proposition 8.4 implies the following result.

Corollary 8.8. *Let $p \in (0,1]$ and $(c_n \colon n \in \mathbb{N})$ a sequence with $\lim_{n\to\infty} c_n^3/n = 0$ and $\lim_{n\to\infty} c_n = \infty$. Then, as $n \to \infty$,*

$$\begin{aligned}
\mathbb{E}\Big[\frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{v=\lfloor \epsilon n \rfloor+1}^{n} \mathbb{1}\{|\mathsf{C}_{n,p}^\epsilon(v)| \ge c_n\}\Big] &= \mathbb{P}\big(|\mathsf{C}_{n,p}^\epsilon(V_n)| \ge c_n\big) \\
&\to P(|\mathcal{X}^\epsilon(p)| = \infty) = \zeta^\epsilon(p).
\end{aligned}$$

This convergence can be strengthened to convergence in probability.

Lemma 8.9. *Let $p \in (0,1]$, and let $(c_n : n \in \mathbb{N})$ be a sequence with $\lim_{n\to\infty} c_n^3/n = 0$ and $\lim_{n\to\infty} c_n = \infty$. Then*

$$M^{\epsilon}_{n,p}(c_n) := \frac{1}{(1-\epsilon)n} \sum_{v=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{1}\{|\mathsf{C}^{\epsilon}_{n,p}(v)| \geq c_n\} \to \zeta^{\epsilon}(p) \quad \textit{in probability, as } n \to \infty.$$

To prove Lemma 8.9, we use a variance estimate for $M^{\epsilon}_{n,p}(c_n)$.

Lemma 8.10. *Let $p \in (0,1]$, and let $(c_n : n \in \mathbb{N})$ be a positive sequence. There exists a constant $C > 0$ such that*

$$\mathrm{Var}(M^{\epsilon}_{n,p}(c_n)) \leq \frac{C}{n}\Big(c_n + \frac{c_n^2}{\epsilon n}\Big).$$

The proof is almost identical to the proof of Proposition 7.1 in [37]. The necessary changes are similar to the changes made for the proofs of Proposition 8.1 and 8.4. We sketch only the main steps.

Proof sketch. Write

$$\mathrm{Var}\Big(\frac{1}{(1-\epsilon)n} \sum_{v=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{1}\{|\mathsf{C}^{\epsilon}_{n,p}(v)| \geq c_n\}\Big) = \frac{1}{(1-\epsilon)^2 n^2} \tag{8.3}$$
$$\cdot \sum_{v,w=\lfloor \epsilon n\rfloor+1}^{n} \Big(\mathbb{P}(|\mathsf{C}^{\epsilon}_{n,p}(v)| \geq c_n, |\mathsf{C}^{\epsilon}_{n,p}(w)| \geq c_n) - \mathbb{P}(|\mathsf{C}^{\epsilon}_{n,p}(v)| \geq c_n)\mathbb{P}(|\mathsf{C}^{\epsilon}_{n,p}(w)| \geq c_n)\Big).$$

To estimate the probability $\mathbb{P}(|\mathsf{C}^{\epsilon}_{n,p}(v)| \geq c_n, |\mathsf{C}^{\epsilon}_{n,p}(w)| \geq c_n)$, we run two successive explorations in the graph $\mathsf{G}^{\epsilon}_n(p)$, the first starting from v, and the second starting from w. For these explorations, we use the exploration process described below Proposition 8.1 but in every step only neighbours in the set of veiled vertices are explored. The first exploration is terminated as soon as either the number of dead and active vertices exceeds c_n or there are no active vertices left. The second exploration, additionally, stops when a vertex is found which was already unveiled in the first exploration. Let

$$\Theta_v := \{\text{the first exploration started in vertex } v \text{ stops because } c_n \text{ vertices are found}\}.$$

Then, for any $v \in \mathsf{V}^{\epsilon}_n$, $\mathbb{P}(|\mathsf{C}^{\epsilon}_{n,p}(v)| \geq c_n) = \mathbb{P}(\Theta_v)$, and in the proof of Proposition 7.1 of [37] it was shown that there exists a constant $C' > 0$, independent of v and n, such that

$$\sum_{w=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{P}\big(|\mathsf{C}^{\epsilon}_{n,p}(v)| \geq c_n, |\mathsf{C}^{\epsilon}_{n,p}(w)| \geq c_n\big) \tag{8.4}$$
$$\leq \mathbb{P}(\Theta_v)\Big(c_n + \sum_{w=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{P}(\Theta_w) + C' c_n \mathbb{P}^{c_n}(\triangle\mathcal{Z}[\lfloor \epsilon n\rfloor+1, \lfloor \epsilon n\rfloor+1] = 1)\Big).$$

Combining (8.3) and (8.4), and using (7.9), there exists a constant $C > 0$ with

$$\mathrm{Var}(M_n^\epsilon(c_n)) \le \frac{1}{(1-\epsilon)^2 n^2} \sum_{v=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{P}(\Theta_v)\Big(c_n + C' \frac{c_n f(c_n)}{\epsilon n}\Big) \le \frac{C}{n}\Big(c_n + \frac{c_n^2}{\epsilon n}\Big). \qquad \square$$

Proof of Lemma 8.9. Using Chebyshev's inequality, Corollary 8.8 and Lemma 8.10 yield the claim. $\square$

Lemma 8.9 already implies that the asymptotic relative size of a largest component in the network is bounded from above by $\zeta^\epsilon(p)$. To show that the survival probability also constitutes a lower bound, we use the following sprinkling argument.

Lemma 8.11. *Let $\epsilon \in [0,1)$, $p \in (0,1]$, $\delta \in (0, f(0))$, and define $\underline{f}(k) := f(k) - \delta$ for all $k \in \mathbb{N}_0$. Denote by $\underline{\mathsf{C}}^\epsilon_{n,p}(v)$ the connected component containing v in the network $\underline{\mathsf{G}}^\epsilon_n(p)$ constructed with the attachment rule $\underline{f}$. Let $\kappa > 0$ and $(c_n \colon n \in \mathbb{N})$ be a sequence with*

$$\lim_{n\to\infty} \big[\tfrac{1}{2}\kappa(1-\epsilon)p\delta c_n - \log n\big] = \infty \quad \text{and} \quad \lim_{n\to\infty} c_n^2/n = 0.$$

Suppose that

$$\frac{1}{n - \lfloor \epsilon n\rfloor} \sum_{v=\lfloor \epsilon n\rfloor+1}^{n} \mathbb{1}\{|\underline{\mathsf{C}}^\epsilon_{n,p}(v)| \ge 2c_n\} \ge \kappa \qquad \text{with high probability.}$$

Then there exists a coupling of the networks $(\mathsf{G}^\epsilon_n(p) \colon n \in \mathbb{N})$ and $(\underline{\mathsf{G}}^\epsilon_n(p) \colon n \in \mathbb{N})$ such that $\underline{\mathsf{G}}^\epsilon_n(p) \le \mathsf{G}^\epsilon_n(p)$ for all $n \in \mathbb{N}$, and with high probability all connected components in $\underline{\mathsf{G}}^\epsilon_n(p)$ with at least $2c_n$ vertices belong to one connected component in $\mathsf{G}^\epsilon_n(p)$.

Lemma 8.11 in the case $\epsilon = 0$ and $p = 1$ is Proposition 4.1 in [37]. The proof remains valid for $\epsilon \in [0,1)$, $p \in (0,1]$, up to obvious changes and is therefore omitted.

Proof of Theorem 6.1. Choose $c_n = (\log n)^2$. By Lemma 8.9, we have in probability

$$\limsup_{n\to\infty} \frac{|\mathsf{C}^\epsilon_{n,p}|}{(1-\epsilon)n} \le \limsup_{n\to\infty} \max\Big\{\frac{c_n}{(1-\epsilon)n}, M^\epsilon_{n,p}(c_n)\Big\} \le \zeta^\epsilon(p).$$

Moreover, for $\delta \in (0, f(0))$, Lemma 8.9 implies that $M^\epsilon_{n,p}(2c_n, f-\delta)$ converges to $\zeta^\epsilon(p, f-\delta)$ in probability. Hence, Lemmas 6.11 and 8.11 yield that for all $\delta' > 0$, $|\mathsf{C}^\epsilon_{n,p}| \ge (n - \lfloor \epsilon n\rfloor)(\zeta^\epsilon(p) - \delta')$ with high probability, as required. $\square$

CHAPTER 9

VARIATIONS AND OTHER MODELS

We study preferential attachment networks with non-linear attachment rules in Section 9.1, and inhomogeneous random graphs and the configuration model in Sections 9.2.2 and 9.2.1, respectively.

9.1 Non-linear attachment rules: proof of Theorem 5.4

Theorem 5.4 is an immediate consequence of Theorem 5.1 and a stochastic domination result on the level of the networks. We make use of the notation and terminology introduced in Section 5.4.

Proof of Theorem 5.4. First suppose that f is a L-class attachment rule with $\overline{f} \geq f \geq \underline{f}$, where $\overline{f}, \underline{f}$ are for two affine attachment rules given by $\overline{f}(k) = \gamma k + \beta_u$ and $\underline{f}(k) = \gamma k + \beta_l$. There exists a natural coupling of the networks generated by these attachment rules such that

$$\overline{\mathsf{G}}_n \geq \mathsf{G}_n \geq \underline{\mathsf{G}}_n \quad \text{for all } n \in \mathbb{N}.$$

This ordering is retained after a targeted attack and percolation, and implies the ordering $\overline{p}_{\mathrm{c}}(\epsilon) \leq p_{\mathrm{c}}(\epsilon) \leq \underline{p}_{\mathrm{c}}(\epsilon)$ of the critical percolation parameters. Applying Theorem 5.1 to $\overline{f}$ and $\underline{f}$, we obtain positive constants $C_1, \ldots, C_4$ such that, for small $\epsilon \in (0,1)$,

$$\begin{aligned} \frac{C_1}{\log(1/\epsilon)} \leq p_{\mathrm{c}}(\epsilon) &\leq \frac{C_2}{\log(1/\epsilon)} && \text{if } \gamma = \tfrac{1}{2}, \\ C_3 \epsilon^{\gamma - 1/2} \leq p_{\mathrm{c}}(\epsilon) &\leq C_4 \epsilon^{\gamma - 1/2} && \text{if } \gamma > \tfrac{1}{2}, \end{aligned}$$

and the result follows.

Now let f be a C-class attachment rule. Concavity of f implies that the increments $\gamma_k := f(k+1) - f(k)$ form a non-increasing sequence converging to γ. In particular, with $\underline{f}(k) := \gamma k + f(0)$, we get $f(k) = \sum_{l=0}^{k-1} \gamma_l + f(0) \geq \gamma k + f(0) = \underline{f}(k)$, for all $k \in \mathbb{N}_0$. To obtain a corresponding upper bound, let $\beta_j := f(j) - \gamma_j j$. Then $\beta_j > 0$ and for all $k \in \mathbb{N}_0$,

$$\begin{aligned} f(k) - \beta_j &= f(k) - f(j) + \gamma_j j \\ &= \left\{ \begin{array}{rl} -\sum_{l=k}^{j-1} \gamma_l + \gamma_j j \leq -(j-k)\gamma_j + \gamma_j j & \text{if } k \leq j \\ \sum_{l=j}^{k-1} \gamma_l + \gamma_j j \leq (k-j)\gamma_j + \gamma_j j & \text{if } k \geq j \end{array} \right\} = \gamma_j k. \end{aligned}$$

Hence, the attachment rule given by $\overline{f}_j(k) := \gamma_j k + \beta_j$, for $k \in \mathbb{N}_0$, satisfies $\overline{f}_j \geq f$, and we can use the same coupling as in the first part of the proof to obtain

$$\overline{p}_{\mathrm{c}}^{(j)}(\epsilon) \leq p_{\mathrm{c}}(\epsilon) \leq \underline{p}_{\mathrm{c}}(\epsilon),$$

where $\overline{p}_{\mathrm{c}}^{(j)}(\epsilon)$ corresponds to the network with attachment rule $\overline{f_j}$. Since $\gamma_j \downarrow \gamma$, we have $\gamma_j \in [\frac{1}{2}, 1)$ for large j. Theorem 5.1 yields, for $\gamma > \frac{1}{2}$, constants $C, C_j > 0$ such that

$$\log C_j + (\gamma_j - 1/2) \log \epsilon \leq \log p_{\mathrm{c}}(\epsilon) \leq \log C + (\gamma - 1/2) \log \epsilon.$$

Dividing by $\log \epsilon$, and then taking first $\epsilon \downarrow 0$ and then $j \to \infty$ yields the claim for $\gamma > \frac{1}{2}$. In the case $\gamma = \frac{1}{2}$ it could happen that $\gamma_j > \frac{1}{2}$ for all $j \in \mathbb{N}$. In this situation, Theorem 5.1 does not give a bound on the right scale. Therefore, we can use only the upper bound on $p_{\mathrm{c}}(\epsilon)$ which gives the stated result. □

9.2 Other models

In this section, we study vulnerability of two other classes of robust network models.

9.2.1 Configuration model: proof of Theorem 5.6

The configuration model is a natural way to construct a network with given degree sequence. It is closely related to the uniformly chosen simple graph with given degree sequence as is explained in Section 7.5 of [85]. Existence of a giant component in the configuration model has been studied by Molloy and Reed [106] and Janson and Luczak [88, 87]. Recall from Section 5.5.1 that we write D for the weak limit of the degree of a uniformly chosen vertex. Janson and Luczak [88] showed that if (5.7) holds and $\mathbb{P}(D = 2) < 1$, then

$$(\mathsf{G}_n^{(\mathrm{CM})} \colon n \in \mathbb{N}) \text{ has a giant component } \Leftrightarrow \mathbb{E}[D(D-1)] > \mathbb{E}D.$$

Janson [87] found a simple construction that allows to obtain a corresponding result for the network after random or deterministic removal of vertices (or edges), where the retention probability of a vertex can depend on its degree. Let $\boldsymbol{\pi} = (\pi_k)_{k \in \mathbb{N}}$ be a sequence of retention probabilities with $\pi_k \mathbb{P}(D = k) > 0$ for some k. Every vertex i is removed with probability $1-\pi_{d_i}$ and kept with probability π_{d_i}, independently of all other vertices. Janson describes the network after percolation as follows [87, page 90]: for each vertex i, replace it with probability $1-\pi_{d_i}$ by d_i new vertices of degree 1. Then construct the configuration model $\mathsf{G}_n^{(\mathrm{CM}),\boldsymbol{\pi}}$ corresponding to the new degree sequence and larger number of vertices, and remove from this graph uniformly at random vertices of degree 1 until the correct number of vertices for $\mathsf{G}_n^{(\mathrm{CM})}$ after percolation is reached. The removal of these surplus vertices cannot destroy or split the giant component since the vertices are of degree 1. Hence, it suffices to study the existence or nonexistence of a giant component in $\mathsf{G}_n^{(\mathrm{CM}),\boldsymbol{\pi}}$.

To construct $\mathsf{G}_n^{(\mathrm{CM}),\epsilon}(p)$, we remove the $\lfloor \epsilon n\rfloor$ vertices with the largest degree from $\mathsf{G}_n^{(\mathrm{CM})}$, and then run vertex percolation with retention probability p on the remaining graph. In general, this does not fit exactly into the setup of Janson. To emulate the behaviour, we denote by n_j the number of vertices with degree j in the graph, and let $K_n = \inf\{k\in\mathbb{N}_0\colon \sum_{j=k+1}^{\infty} n_j \le \lfloor \epsilon n\rfloor\}$. Then all vertices with degree larger than K_n are deterministically removed in $\mathsf{G}_n^{(\mathrm{CM}),\epsilon}(p)$, i.e. $\pi_j = 0$ for $j\ge K_n+1$. In addition, we deterministically remove $\lfloor \epsilon n\rfloor - \sum_{j=K_n+1}^{\infty} n_j$ vertices of degree K_n, while all other vertices are subject to vertex percolation with retention probability p. In particular, $\pi_j = p$ for $j\le K_n-1$.

Write $F(x) := \mathbb{P}(D\le x)$ for $x\ge 0$, and $[1-F]^{-1}$ for the generalised inverse of $[1-F]$, that is

$$[1-F]^{-1}(u) = \inf\{k\in\mathbb{N}_0\colon [1-F](k)\le u\}\qquad \text{for all } u\in(0,1).$$

One easily checks that $K_n\in\{m,m+1\}$ for all sufficiently large n, where $m := [1-F]^{-1}(\epsilon)$. Using this observation, it is not difficult to adapt Janson's proof (c.f. Theorem 3.5 in [87]) to show that

$$(\mathsf{G}_n^{(\mathrm{CM}),\epsilon}(p)\colon n\in\mathbb{N}) \text{ has a giant component } \Leftrightarrow\ p > p_{\mathrm{c}}(\epsilon),$$

where

$$p_{\mathrm{c}}(\epsilon) := \frac{\mathbb{E}D}{\mathbb{E}[D(D-1)\mathbb{1}_{D\le m}] - m(m-1)(\epsilon - [1-F](m))}.$$

Proof of Theorem 5.6. Let U be a uniformly distributed random variable on $(0,1)$. Then $[1-F]^{-1}(U)$ has the same distribution as D and

$$\begin{aligned}\mathbb{E}[D(D-1)\mathbb{1}_{\{D\le m\}}] &- m(m-1)(\epsilon - [1-F](m))\\ &= \mathbb{E}\big[[1-F]^{-1}(U)\big([1-F]^{-1}(U)-1\big)\mathbb{1}_{\{U\ge\epsilon\}}\big]\\ &\asymp \mathbb{E}\big[[1-F]^{-1}(U)^2\mathbb{1}_{\{U\ge\epsilon\}}\big].\end{aligned}\tag{9.1}$$

The assumption $[1-F](k)\sim Ck^{-1/\gamma}$ as $k\to\infty$ implies that $[1-F]^{-1}(u)\sim C^{\gamma}u^{-\gamma}$ as

$u \downarrow 0$. Let $u_0 > 0$ such that

$$\frac{1}{2} \le \frac{[1-F]^{-1}(u)}{C^\gamma u^{-\gamma}} \le \frac{3}{2} \qquad \text{for all } u \le u_0.$$

Since $[1-F]^{-1}(u)^2$ is not integrable around zero but bounded on $[u_0, 1)$, we deduce that the right-hand side of (9.1) is equal to

$$\begin{aligned}\mathbb{E}\big[[1-F]^{-1}(U)^2 \mathbb{1}_{\{U \ge \epsilon\}}\big] &= \int_\epsilon^1 [1-F]^{-1}(u)^2\, du \asymp \int_\epsilon^{u_0} [1-F]^{-1}(u)^2\, du \\ &\asymp \int_\epsilon^{u_0} u^{-2\gamma}\, du \asymp \begin{cases} \log(1/\epsilon) & \text{if } \gamma = \frac{1}{2}, \\ \epsilon^{1-2\gamma} & \text{if } \gamma > \frac{1}{2}. \end{cases}\end{aligned}$$ □

9.2.2 Inhomogeneous random graphs: proof of Theorem 5.7

The classical Erdős–Rényi random graph can be generalised by giving each vertex a weight, and choosing the probability for an edge between two vertices as an increasing function of their weights. Suppose that $\kappa\colon (0,1] \times (0,1] \to (0,\infty)$ is a symmetric, continuous kernel with

$$\int_0^1 \int_0^1 \kappa(x,y)\, dx\, dy < \infty \tag{9.2}$$

and recall from (5.8) that in the inhomogeneous random graph $\mathsf{G}_n^{(\kappa)}$, the edge $\{i,j\}$ is present with probability $\frac{1}{n}\kappa(\frac{i}{n}, \frac{j}{n}) \wedge 1$, independently of all other edges. We assume that vertices are ordered in decreasing order of privilege, i.e. κ is non-increasing in both components. Bollobás et al. showed in Theorem 3.1 and Example 4.11 of [20] that, for all $\epsilon \in [0,1)$,

$$(\mathsf{G}_n^{(\kappa),\epsilon}(p)\colon n \in \mathbb{N}) \text{ has a giant component } \Leftrightarrow\ p > p_{\mathrm{c}}(\epsilon) := \|T_\kappa\|_{L^2(\epsilon,1)}^{-1}, \tag{9.3}$$

where

$$T_\kappa g(x) = \int_\epsilon^1 \kappa(x,y) g(y)\, dy, \qquad \text{for all } x \in (\epsilon, 1)$$

and all measurable functions g such that the integral is well-defined, and $\|\cdot\|_{L^2(\epsilon,1)}$ denotes the operator norm on the L^2-space with respect to the Lebesgue measure on $(\epsilon, 1)$. The same result holds for a version of the Norros–Reittu model in which edges between different vertex pairs are independent, and edge $\{i,j\}$ is present with probability $1 - e^{-\kappa(i/n, j/n)/n}$ for all $i,j \in \{1,\dots,n\}$. Consequently, the estimates given in Theorem 5.7 hold for this model, too.

Proof of Theorem 5.7. Since $\kappa^{(\mathrm{CL})}$ and $\kappa^{(\mathrm{PA})}$ are positive, symmetric, continuous kernels satisfying (9.2), the first part of the theorem follows immediately from (9.3). By definition,

$$\|T_\kappa\|_{L^2(\epsilon,1)} = \sup\big\{\|T_\kappa g\|_{L^2(\epsilon,1)}\colon \|g\|_{L^2(\epsilon,1)} \le 1\big\}.$$

For a rank one kernel $\kappa(x,y) = \chi(x)\chi(y)$, the operator norm of T_κ is attained at $\chi/\|\chi\|_{L^2(\epsilon,1)}$ with $\|T_\kappa\|_{L^2(\epsilon,1)} = \|\chi\|^2_{L^2(\epsilon,1)}$. Hence,

$$\|T_{\kappa^{(\mathrm{CL})}}\|_{L^2(\epsilon,1)} = \int_\epsilon^1 x^{-2\gamma}\,dx = \begin{cases} \log(1/\epsilon) & \text{if } \gamma = 1/2, \\ \frac{1}{1-2\gamma}\big[1-\epsilon^{1-2\gamma}\big] & \text{if } \gamma \neq \frac{1}{2}, \end{cases}$$

and

$$p_{\mathrm{c}}^{(\mathrm{CL})}(\epsilon) = \begin{cases} (1-2\gamma)\frac{1}{1-\epsilon^{1-2\gamma}} & \text{if } \gamma < \frac{1}{2}, \\ \frac{1}{\log(1/\epsilon)} & \text{if } \gamma = \frac{1}{2}, \\ (2\gamma-1)\epsilon^{2\gamma-1}\frac{1}{1-\epsilon^{2\gamma-1}} & \text{if } \gamma > \frac{1}{2}. \end{cases}$$

Now suppose that $\gamma > 1/2$. By Cauchy–Schwarz's inequality and the symmetry of $\kappa^{(\mathrm{PA})}$,

$$\begin{aligned} \|T_{\kappa^{(\mathrm{PA})}}\|^2_{L^2(\epsilon,1)} &\le \int_\epsilon^1\int_\epsilon^1 \kappa^{(\mathrm{PA})}(x,y)^2\,dy\,dx = 2\int_\epsilon^1\int_\epsilon^x x^{2(\gamma-1)}y^{-2\gamma}\,dy\,dx \\ &\le 2\int_0^1 x^{2(\gamma-1)}\,dx\int_\epsilon^\infty y^{-2\gamma}\,dy = \frac{2}{(2\gamma-1)^2}\epsilon^{1-2\gamma}. \end{aligned}$$

For the lower bound, let $c_\epsilon = \sqrt{2\gamma-1}\epsilon^{\gamma-1/2}$ and $g(x) = c_\epsilon x^{-\gamma}$. Then $\|g\|_{L^2(\epsilon,1)} \le 1$, and

$$\begin{aligned} \|T_{\kappa^{(\mathrm{PA})}}\|^2_{L^2(\epsilon,1)} \ge \|T_{\kappa^{(\mathrm{PA})}}g\|^2_{L^2(\epsilon,1)} &\ge \int_\epsilon^1\Big(\int_\epsilon^x \kappa^{(\mathrm{PA})}(x,y)g(y)\,dy\Big)^2\,dx \\ &= \frac{c_\epsilon^2}{(2\gamma-1)^2}\int_\epsilon^1 x^{2(\gamma-1)}[\epsilon^{1-2\gamma}-x^{1-2\gamma}]^2\,dx \\ &\ge \frac{c_\epsilon^2}{(2\gamma-1)^2}\int_\epsilon^1 x^{2(\gamma-1)}\big[\epsilon^{2(1-2\gamma)}-2\epsilon^{1-2\gamma}x^{1-2\gamma}\big]\,dx \\ &= \frac{\epsilon^{1-2\gamma}}{(2\gamma-1)^2}\Big[1-\epsilon^{2\gamma-1}+2(2\gamma-1)\epsilon^{2\gamma-1}\log\epsilon\Big]. \end{aligned}$$

The claim follows. $\square$

BIBLIOGRAPHY

[1] L. Addario-Berry, N. Berestycki, and N. Gantert, eds. (2013). *Extremes in branching random walk and branching Brownian motion.* No. 20 in Oberwolfach Reports.

[2] W. Aiello, F. Chung, and L. Lu (2001). *A random graph model for power law graphs*. Exp. Math., **10**(1):53–66.

[3] R. Albert, I. Albert, and G. L. Nakarado (2004). *Structural vulnerability of the North American power grid*. Phys. Rev. E, **69**:025103.

[4] R. Albert, H. Jeong, and A.-L. Barabási (2000). *Error and attack tolerance of complex networks*. Nature, **406**:378–382.

[5] S. Asmussen and H. Hering (1976). *Strong limit theorems for general supercritical branching processes with applications to branching diffusions*. Z. Wahrsch. Verw. Gebiete, **36**(3):195–212.

[6] K. B. Athreya (1968). *Some results on multitype continuous time Markov branching processes*. Ann. Math. Statist., **39**:347–357.

[7] K. B. Athreya and P. E. Ney (1972). *Branching processes.* Grundlehren Math. Wiss., Springer.

[8] F. Ball and P. Donnelly (1995). *Strong approximations for epidemic models*. Stoch. Proc. Appl., **55**(1):1–21.

[9] A.-L. Barabási and R. Albert (1999). *Emergence of scaling in random networks*. Science, **286**(5439):509–512.

[10] J. Berestycki, N. Berestycki, and J. Schweinsberg (2013). *The genealogy of branching Brownian motion with absorption*. Ann. Probab., **41**(2):527–618.

[11] J. Berestycki, A. E. Kyprianou, and A. Murillo-Salas (2011). *The prolific backbone for supercritical superprocesses*. Stoch. Proc. Appl., **121**(6):1315–1331.

[12] N. Berestycki (2009). *Recent progress in coalescent theory.* Vol. 16 of Ensaios Mat. Sociedade Brasileira de Matemática, Rio de Janeiro.

[13] J. Bertoin, J. Fontbona, and S. Martínez (2008). *On prolific individuals in a supercritical continuous-state branching process*. J. Appl. Probab., **45**(3):714–726.

[14] S. Bhamidi and R. v. d. Hofstad (2012). *Weak disorder asymptotics in the stochastic mean-field model of distance*. Ann. Appl. Probab., **22**(1):29–69.

[15] S. Bhamidi, R. v. d. Hofstad, and G. Hooghiemstra (2010). *Extreme value theory, Poisson-Dirichlet distributions, and first passage percolation on random networks*. Adv. Appl. Prob., **42**(3):706–738.

[16] J. D. Biggins (1992). *Uniform convergence of martingales in the branching random walk*. Ann. Probab., **20**(1):137–151.

[17] J. D. Biggins and A. E. Kyprianou (2004). *Measure change in multitype branching*. Adv. Appl. Prob., **36**(2):544–581.

[18] B. Bollobás (1985). *Random graphs.* Academic Press.

[19] B. Bollobás, S. Janson, and O. Riordan (2005). *The phase transition in the uniformly grown random graph has infinite order*. Random Structures Algorithms, **26**(1-2):1–36.

[20] B. Bollobás, S. Janson, and O. Riordan (2007). *The phase transition in inhomogeneous random graphs*. Random Structures Algorithms, **31**(1):3–122.

[21] B. Bollobás and O. Riordan (2003). *Robustness and vulnerability of scale-free random graphs*. Internet Math., **1**(1):1–35.

[22] B. Bollobás and O. Riordan (2004). *The diameter of a scale-free random graph*. Combinatorica, **24**(1):5–34.

[23] B. Bollobás and O. Riordan (2005). *Slow emergence of the giant component in the growing m-out graph*. Random Structures Algorithms, **27**(1):1–24.

[24] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády (2001). *The degree sequence of a scale-free random graph process*. Random Structures Algorithms, **18**(3):279–290.

[25] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts (2000). *Network robustness and fragility: percolation on random graphs*. Phys. Rev. Lett., **85**(25):5468–5471.

[26] B. Chauvin, P. Flajolet, D. Gardy, and A. Mokkadem, eds. (2002). *Mathematics and computer science II. Algorithms, trees, combinatorics and probabilities*. Trends in Mathematics, Birkhäuser Verlag, Basel.

[27] Z.-Q. Chen, Y.-X. Ren, and H. Wang (2008). *An almost sure scaling limit theorem for Dawson–Watanabe superprocesses*. J. Funct. Anal., **254**(7):1988–2019.

[28] Z.-Q. Chen and Y. Shiozawa (2007). *Limit theorems for branching Markov processes*. J. Funct. Anal., **250**(2):374–399.

[29] F. Chung and L. Lu (2002). *The average distances in random graphs with given expected degrees*. Proc. Nat. Acad. Sci. USA, **99**(25):15879–15882.

[30] F. Chung and L. Lu (2002). *Connected components in random graphs with given expected degree sequences*. Ann. Comb., **6**(2):125–145.

[31] F. Chung and L. Lu (2006). *Complex graphs and networks*. Vol. 107 of Regional Conference Series in Mathematics. Co-publication of the Amer. Math. Soc. and CBMS.

[32] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin (2001). *Breakdown of the internet under intentional attack*. Phys. Rev. Lett., **86**:3682–3685.

[33] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin (2001). *Reply*. Phys. Rev. Lett., **87**:219802.

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

[35] S. Dereich, C. Mönch, and P. Mörters (2012). *Typical distances in ultrasmall random networks*. Adv. Appl. Prob., **44**:583–601.

[36] S. Dereich and P. Mörters (2009). *Random networks with sublinear preferential attachment: degree evolutions*. Electron. J. Probab., **14**:1222–1267.

[37] S. Dereich and P. Mörters (2013). *Random networks with sublinear preferential attachment: the giant component*. Ann. Probab., **41**(1):329–384.

[38] S. Dommers, R. v. d. Hofstad, and G. Hooghiemstra (2010). *Diameters in Preferential Attachment Models*. J. Stat. Phys., **139**(1):72–107.

[39] S. N. Dorogovtsev and J. F. F. Mendes (2001). *Comment on "Breakdown of the internet under intentional attack"*. Phys. Rev. Lett., **87**:219801.

[40] T. Duquesne and M. Winkel (2007). *Growth of Lévy trees*. Probab. Theory Relat. Fields, **139**(3-4):313–371.

[41] E. B. Dynkin (1992). *Superdiffusions and parabolic nonlinear differential equations*. Ann. Probab., **20**(2):942–962.

[42] E. B. Dynkin (1993). *On regularity of superprocesses*. Probab. Theory Relat. Fields, **95**(2):263–281.

[43] E. B. Dynkin (1993). *Superprocesses and partial differential equations*. Ann. Probab., **21**(3):1185–1262.

[44] E. B. Dynkin (2002). *Diffusions, superdiffusions and partial differential equations.* Vol. 50 of Amer. Math. Soc. Colloq. Publ., Amer. Math. Soc.

[45] E. B. Dynkin and S. E. Kuznetsov (2004). $\mathbb{N}$*-measures for branching exit Markov systems and their applications to differential equations*. Probab. Theory Relat. Fields, **130**(1):135–150.

[46] M. Eckhoff (2013). *The size of the giant component in preferential attachment networks*. In Extremes in branching random walk and branching Brownian motion. No. 20 in Oberwolfach Reports.

[47] M. Eckhoff, J. Goodman, R. v. d. Hofstad, and F. R. Nardi (2013). *Short paths for first passage percolation on the complete graph*. J. Stat. Phys., **151**(6):1056–1088.

[48] M. Eckhoff, J. Goodman, R. v. d. Hofstad, and F. R. Nardi (2014). *Long paths in first passage percolation on the complete graph I. Local IP dynamics*. In preparation.

[49] M. Eckhoff, J. Goodman, R. v. d. Hofstad, and F. R. Nardi (2014). *Long paths in first passage percolation on the complete graph II. Global branching dynamics*. In preparation.

[50] M. Eckhoff, A. E. Kyprianou, and M. Winkel (2013). *Spines, skeletons and the strong law of large numbers for superdiffusions*. To appear in Ann. Probab. Preprint available at arXiv:1309.6196 [math.PR].

[51] M. Eckhoff and P. Mörters (2014). *Vulnerability of robust preferential attachment networks*. Electron. J. Probab., **19**(57):1–47.

[52] M. Eckhoff and P. Mörters (2014). *The size of the giant component in preferential attachment networks*. In preparation.

[53] J. Engländer (2004). *An example and a conjecture concerning scaling limits of superdiffusions*. Statist. Probab. Lett., **66**(3):363–368.

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

[55] J. Engländer (2009). *Law of large numbers for superdiffusions: the non-ergodic case*. Ann. Inst. H. Poincaré Probab. Stat., **45**(1):1–6.

[56] J. Engländer (2014). *Spatial branching in random environments and with interaction.* Vol. 20 of Adv. Ser. Stat. Sci. Appl. Probab., World Scientific/Imperial College Press.

[57] J. Engländer, S. C. Harris, and A. E. Kyprianou (2010). *Strong law of large numbers for branching diffusions*. Ann. Inst. H. Poincaré Probab. Stat., **46**(1):279–298.

[58] J. Engländer and A. E. Kyprianou (2004). *Local extinction versus local exponential growth for spatial branching processes*. Ann. Probab., **32**(1A):78–99.

[59] J. Engländer and R. G. Pinsky (1999). *On the construction and support properties of measure-valued diffusions on $D \subseteq \mathbb{R}^d$ with spatially dependent branching*. Ann. Probab., **27**(2):684–730.

[60] J. Engländer, Y.-X. Ren, and R. Song (2013). *Weak extinction versus global exponential growth of total mass for superdiffusions*. Preprint available at arXiv:1301.6842 [math.PR].

[61] J. Engländer and D. Turaev (2002). *A scaling limit theorem for a class of superdiffusions*. Ann. Probab., **30**(2):683–722.

[62] J. Engländer and A. Winter (2006). *Law of large numbers for a class of superdiffusions*. Ann. Inst. H. Poincaré Probab. Stat., **42**(2):171–185.

[63] P. Erdős and A. Rényi (1959). *On random graphs. I*. Publ. Math. Debrecen, **6**:290–297.

[64] P. Erdős and A. Rényi (1960). *On the evolution of random graphs*. Magyar Tud. Akad. Mat. Kutató Int. Közl., **5**:17–61.

[65] H. v. d. Esker, R. v. d. Hofstad, G. Hooghiemstra, and D. Znamenski (2006). *Distances in random graphs with infinite mean degrees*. Extremes, **8**(3):111–141.

[66] A. M. Etheridge (2000). *An introduction to superprocesses.* Vol. 20 of Univ. Lecture Ser., Amer. Math. Soc.

[67] A. M. Etheridge and D. R. E. Williams (2003). *A decomposition of the $(1+\beta)$-superprocess conditioned on survival*. Proc. Roy. Soc. Edinburgh Sect. A, **133**(4):829–847.

[68] S. N. Ethier and T. G. Kurtz (2005). *Markov processes: characterization and convergence*. Wiley, New York, 2nd ed.

[69] S. N. Evans (1993). *Two representations of a conditioned superprocess*. Proc. Roy. Soc. Edinburgh Sect. A, **123**(5):959–971.

[70] S. N. Evans and N. O'Connell (1994). *Weighted occupation time for branching particle systems and a representation for the supercritical superprocess*. Canad. Math. Bull., **37**(2):187–196.

[71] W. Feller (1971). *An introduction to probability theory and its applications. Vol. II.* Wiley, New York.

[72] P. J. Fitzsimmons (1988). *Construction and regularity of measure-valued Markov branching processes*. Israel J. Math., **64**(3):337–361.

[73] K. Fleischmann and C. Mueller (2004). *Super-Brownian motion with extra birth at one point*. SIAM J. Math. Anal., **36**(3):740–772.

[74] K. Fleischmann and J. M. Swart (2003). *Extinction versus exponential growth in a supercritical super-Wright–Fisher diffusion*. Stoch. Proc. Appl., **106**(1):141–165.

[75] K. Fleischmann and J. M. Swart (2004). *Trimmed trees and embedded particle systems*. Ann. Probab., **32**(3A):2179–2221.

[76] E. N. Gilbert (1959). *Random graphs*. Ann. Math. Statist., **30**(4):1141–1144.

[77] A. Greven, A. Klenke, and A. Wakolbinger (2001). *Interacting Fisher–Wright diffusions in a catalytic medium*. Probab. Theory Relat. Fields, **120**(1):85–117.

[78] D. R. Grey (1974). *Asymptotic behaviour of continuous time, continuous state-space branching processes*. J. Appl. Probab., **11**:669–677.

[79] R. Grummt and M. Kolb (2013). *Law of large numbers for super-Brownian motions with a single point source*. Stoch. Proc. Appl., **123**(4):1183–1212.

[80] R. Hardy and S. C. Harris (2009). *A spine approach to branching diffusions with applications to L^p-convergence of martingales*. In Séminaire de Probabilités XLII. Vol. 1979 of Lecture Notes in Math., 281–330, Springer, Berlin.

[81] S. C. Harris, M. Hesse, and A. E. Kyprianou (2012). *Branching Brownian motion in a strip: survival near criticality*. Preprint available at arXiv:1212.1444 [math.PR].

[82] S. C. Harris and M. I. Roberts (2014). *A strong law of large numbers for branching processes: almost sure spine events*. Electron. Commun. Probab., **19**(28):1–6.

[83] T. E. Harris (1963). *The theory of branching processes*. Grundlehren Math. Wiss., Springer, Berlin.

[84] H. Heuser (1982). *Functional Analysis*. Wiley, Chichester.

[85] R. v. d. Hofstad (2014). *Random graphs and complex networks*. Draft book, available online at http://www.win.tue.nl/~rhofstad/NotesRGCN.html.

[86] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han (2002). *Attack vulnerability of complex networks*. Phys. Rev. E, **65**:056109.

[87] S. Janson (2009). *On percolation in random graphs with given vertex degrees*. Electron. J. Probab., **14**(5):87–118.

[88] S. Janson and M. J. Luczak (2008). *A new approach to the giant component problem*. Random Structures Algorithms, **34**(2):197–216.

[89] S. Janson, T. Łuczak, and A. Rucinski (2000). *Random graphs*. Wiley-Intersci. Ser. Discrete Math. Optim., Wiley.

[90] N. Kaplan and S. Asmussen (1976). *Branching random walks II*. Stoch. Proc. Appl., **4**(1):15–31.

[91] T. Kato (1995). *Perturbation theory for linear operators*. Classics Math., Springer, Berlin. Reprint of the 1980 edition.

[92] D. G. Kendall (1951). *Some problems in the theory of queues*. R. Stat. Soc. Ser. B Stat. Methodol., **13**(2):151–173.

[93] H. Kesten and B. P. Stigum (1966). *A limit theorem for multidimensional Galton–Watson processes*. Ann. Math. Statist., **37**(5):1211–1223.

[94] A. Klenke (2008). *Probability theory. A comprehensive course*. Universitext, Springer, London. Translated from the 2006 German original.

[95] M. A. Kouritzin and Y.-X. Ren (2014). *A strong law of large numbers for super-stable processes*. Stoch. Proc. Appl., **124**(1):505–521.

[96] A. E. Kyprianou, R.-L. Liu, A. Murillo-Salas, and Y.-X. Ren (2012). *Supercritical super-Brownian motion with a general branching mechanism and travelling waves*. Ann. Inst. H. Poincaré Probab. Stat., **48**(3):661–687.

[97] A. E. Kyprianou and A. Murillo-Salas (2013). *Super-Brownian motion: L^p-convergence of martingales through the pathwise spine decomposition*. In Advances in superprocesses and nonlinear PDEs. Vol. 38 of Springer Proc. Math. Stat., 113–121, Springer.

[98] A. E. Kyprianou, J.-L. Perez, and Y.-X. Ren (2013). *The backbone decomposition for spatially dependent supercritical superprocesses*. Preprint available at arXiv:1304.2019 [math.PR].

[99] A. E. Kyprianou and Y.-X. Ren (2012). *Backbone decomposition for continuous-state branching processes with immigration*. Statist. Probab. Lett., **82**(1):139–144.

[100] J.-F. Le Gall (1999). *Spatial branching processes, random snakes and partial differential equations*. Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel.

[101] R.-L. Liu, Y.-X. Ren, and R. Song (2009). *$L \log L$ criterion for a class of superdiffusions*. J. Appl. Probab., **46**(2):479–496.

[102] R.-L. Liu, Y.-X. Ren, and R. Song (2013). *Strong law of large numbers for a class of superdiffusions*. Acta Appl. Math., **123**(1):73–97.

[103] P. D. Miller (2006). *Applied asymptotic analysis*. Vol. 75 of Grad. Stud. Math., Amer. Math. Soc.

[104] P. Miłoś (2012). *Spatial CLT for the supercritical Ornstein–Uhlenbeck superprocess*. Preprint available at arXiv:1203.6661 [math.PR].

[105] I. Mishkovski, M. Biey, and L. Kocarev (2011). *Vulnerability of complex networks*. Commun. Nonlinear Sci. Numer. Simul., **16**(1):341–349.

[106] M. Molloy and B. Reed (1995). *A critical point for random graphs with a given degree sequence*. Random Structures Algorithms, **6**(2-3):161–179.

[107] C. Mönch (2013). *Distances in preferential attachment networks*. Ph.D. thesis, University of Bath.

[108] E. Perkins (2002). *Dawson–Watanabe superprocesses and measure-valued diffusions*. In Lectures on probability theory and statistics, Saint-Flour, 1999. Vol. 1781 of Lecture Notes in Math., 125–324, Springer, Berlin.

[109] Y. Pinchover (2013). *Some aspects of large time behavior of the heat kernel: an overview with perspectives*. In Mathematical physics, spectral theory and stochastic analysis. Vol. 232 of Oper. Theory Adv. Appl., 299–339, Springer, Basel.

[110] R. G. Pinsky (1995). *Positive harmonic functions and diffusion.* Vol. 45 of Cambridge Stud. Adv. Math., Cambridge University Press, Cambridge.

[111] R. G. Pinsky (1996). *Transience, recurrence and local extinction properties of the support for supercritical finite measure-valued diffusions*. Ann. Probab., **24**(1):237–267.

[112] Y.-X. Ren, R. Song, and R. Zhang (2014). *Central limit theorems for super Ornstein–Uhlenbeck processes*. Acta Appl. Math., **130**:9–49.

[113] D. Revuz and M. Yor (1999). *Continuous martingales and Brownian motion.* Vol. 293 of Grundlehren Math. Wiss., Springer, Berlin, 3rd ed.

[114] O. Riordan (2005). *The small giant component in scale-free random graphs*. Combin. Probab. Comput., **14**(5-6):897–938.

[115] T. S. Salisbury and A. D. Sezer (2013). *Conditioning super-Brownian motion on its boundary statistics, and fragmentation*. Ann. Probab., **41**(5):3617–3657.

[116] T. S. Salisbury and J. Verzani (1999). *On the conditioned exit measures of super Brownian motion*. Probab. Theory Relat. Fields, **115**(2):237–285.

[117] T. S. Salisbury and J. Verzani (2000). *Non-degenerate conditionings of the exit measures of super Brownian motion*. Stoch. Proc. Appl., **87**(1):25–52.

[118] Y.-C. Sheu (1997). *Lifetime and compactness of range for super-Brownian motion with a general branching mechanism*. Stoch. Proc. Appl., **70**(1):129–141.

[119] D. W. Stroock and S. R. S. Varadhan (1979). *Multidimensional diffusion processes.* Vol. 233 of Grundlehren Math. Wiss., Springer, Berlin-New York.

[120] L. Wang (2010). *An almost sure limit theorem for super-Brownian motion*. J. Theoret. Probab., **23**(2):401–416.

[121] S. Watanabe (1967). *Limit theorem for a class of branching processes*. In Markov processes and potential theory, 205–232, Wiley, New York.

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