Electronic Journal of Probability

Vol. 9 (2004), Paper no. 12, pages 316-381.

Journal URL
http://www.math.washington.edu/~ejpecp/

Hierarchical Equilibria of Branching Populations

D. A. DAWSON[1]
Carleton University
Ottawa, Canada K1S 5B6 (e.mail: ddawson@math.carleton.ca)

L. G. GOROSTIZA[2]
Centro de Investigación y de Estudios Avanzados
07000 México D.F., Mexico (e.mail: lgorosti@math.cinvestav.mx)

A. WAKOLBINGER[3]
Goethe-Universität
Frankfurt am Main, Germany (e.mail: wakolbin@math.uni-frankfurt.de)

Abstract

The objective of this paper is the study of the equilibrium behavior of a population on the hierarchical group Ω_N consisting of families of individuals undergoing critical branching random walk and in addition these families also develop according to a critical branching process. Strong transience of the random walk guarantees existence of an equilibrium for this two-level branching system. In the limit $N \to \infty$ (called the *hierarchical mean field limit*), the equilibrium aggregated populations in a nested sequence of balls $B_\ell^{(N)}$ of hierarchical radius ℓ converge to a backward Markov chain on $\mathbb{R}_+$. This limiting Markov chain can be explicitly represented in terms of a cascade of subordinators which in turn makes possible a description of the genealogy of the population.

[1]Research supported by NSERC (Canada) and a Max Planck Award for International Cooperation.
[2] Research supported by CONACYT grant 37130-E (Mexico).
[3] Research supported by DFG (SPP 1033) (Germany).

Keywords and phrases: Multilevel branching, hierarchical mean-field limit, strong transience, genealogy.

Mathematics Subject Classifications (2000): Primary 60J80; Secondary 60J60, 60G60.

Submitted October 11, 2003. Final version accepted April 2, 2004.

Contents

1 Introduction

Spatial branching processes involve two basic mechanisms, spatial migration and branching. These two mechanisms act in opposite directions: the branching causes fluctuations of the local population densities which are counteracted by the smoothing effect of the migration, and a *transient* migration is needed to sustain an equilibrium of a geographically extended population where each individual has an offspring of mean one. Multilevel branching systems (see, e.g. [DH], [GHW], [Wu]) involve branching, that is death and replication, at a collective level. For example, in two-level branching systems both individuals and families (that is, collections of individuals that trace back to a common ancestor in the individual branching) reproduce independently. In such systems, the fluctuations of the population densities are substantially enhanced compared to systems with branching on the individual level only, and in two-level branching systems a *strongly transient* migration is needed to sustain an equilibrium. It is well known that Euclidean random walks are transient if and only if the dimension is bigger than 2, and strongly transient if and only if the dimension is bigger than 4. In this sense, 2 is the critical dimension for one-level branching systems, and 4 is the critical dimension for two-level branching systems.

In the present paper, we will focus on spatial models with a *hierarchical* (or *ultrametric*) geographical structure (N islands (blocks of radius one) per archipelago (blocks of radius 2), N blocks of radius ℓ per block of radius $\ell+1$, $\ell > 1$, cf. [SF]). The migration process then is modelled by so called *hierarchical random walks*: at a certain rate depending on ℓ, an individual jumps to a randomly chosen island in distance ℓ. This ultrametric structure leads to a separation of time scales as $N \to \infty$, and makes the models particularly suitable for a thorough analysis of equilibrium states and cluster formation. It turns out that in the *hierarchichal mean field limit* (with order of N individuals per island and $N \to \infty$) there is a separation of time scales in which the population densities in the blocks of different radii evolve. For a block of radius ℓ, the natural time scale turns out to be N^{ℓ} in the case of one-level branching (see [DG2]) and $N^{\ell/2}$ in the case of two-level branching. On this time scale, the population density in a block of radius ℓ performs, as $N \to \infty$, a process whose fluctuations are governed by the branching and whose drift is given by a flow of emigration and immigration from the surrounding block. For a sequence of nested blocks, this leads to a hierarchy of branching equilibria whose structure we describe in the next subsection. For the case of two-level branching, the convergence of the population densities in nested blocks towards this hierarchy as $N \to \infty$ is stated in Theorem 2.4.1 and proved in section 5.

Generically, in our hierarchical model the migration process that sustains an equilibrium is at the border to recurrence in the case of one-level branching, and at the border to weak transience in the case of two-level branching, as $N \to \infty$. In this sense, the hierarchical one-level branching equilibria studied in [DG2] correspond to a situation "near dimension 2", and the hierarchical two-level branching equilibria studied in our paper correspond to a situation "near dimension 4". Dimension 4 is of considerable interest because it serves as a critical dimension not only for the two-level branching systems studied in this paper but also for a number of phenomena in statistical physics including the large scale fluctuations of ferromagnetic models at the critical temperature.

The structures of the family clusters in equilibrium can be best understood in terms of the genealogy of the branching system, see [DP1, DP2, StW]). We will describe the genealogy of the equilibrium population in the mean-field limit using a cascade of subordinators.

2 Overview

2.1 Hierarchies of one–and two–level Feller branching diffusions

Consider a large population whose size is fluctuating because of critical reproduction, and which is subject to emigration of individuals and immigration from a surrounding (still larger) reservoir of individuals. The immigration rate is given by the population density in the environment, which fluctuates on a slower time scale. Now consider an infinite hierarchy of such populations $\mathcal{P}_\ell$, $\ell = 1, 2, ..$, where $\mathcal{P}_{\ell+1}$ acts as an environment for $\mathcal{P}_\ell$, and think of an equilibrium situation. We will study two situations where there is a sequence of time scales such that, in the limit of large local population sizes, on the time scale associated with $\mathcal{P}_\ell$ the population density $\zeta_{\ell+1}$ of $\mathcal{P}_{\ell+1}$ remains constant, and given $\zeta_{\ell+1} = a$, the dynamics of the population density ζ_ℓ of $\mathcal{P}_\ell$ is of the form

$$d\zeta_\ell(t) = dM_\ell(t) - c_\ell(\zeta_\ell(t) - a)\, dt. \tag{2.1.1}$$

Here c_ℓ is a positive constant which describes the migration intensity into and out of $\mathcal{P}_\ell$, and M_ℓ is a martingale describing the fluctuations of ζ_ℓ.

In subsection 2.2.3 we will describe a situation in which the martingale M_ℓ has quadratic variation

$$d\langle M_\ell\rangle(t) = \zeta_\ell(t)dt, \tag{2.1.2}$$

hence in this case (2.1.1) specializes to

$$d\zeta_\ell(t) = \sqrt{\zeta_\ell(t)}dW_\ell(t) - c_\ell(\zeta_\ell(t) - a)\, dt, \tag{2.1.3}$$

where $W_1, W_2, \ldots$ are independent Wiener processes. For each ℓ (2.1.3) is the stochastic differential equation of a subcritical Feller branching diffusion with immigration ([AN, EK]).

Later on, we will consider a dynamics of the population density ζ_ℓ which is again of the form (2.1.1) but where the fluctuations are governed by a "family structure" of the population. More precisely, the martingale M_ℓ has quadratic variation

$$d\langle M_\ell\rangle(t) = \left(\int_{(0,\infty)} x^2\xi_\ell(t,x)dx\right) dt. \tag{2.1.4}$$

where $\xi_\ell(t,x)$ measures the rescaled number of families of size x within $\mathcal{P}_\ell$. The link between the population density $\zeta_\ell(t)$ and the density $\xi_\ell(t,x)$ of family sizes is given by

$$\zeta_\ell(t) = \int_{(0,\infty)} x\,\xi_\ell(t,x)dx. \tag{2.1.5}$$

The form (2.1.4) of the fluctuations of ζ_ℓ indicates that we are dealing with Feller branching diffusions of *families* rather than individuals. This *family branching* shows up in the dynamics which is described by an absolutely continuous measure-valued process with density $\xi_\ell(t,x)$ satisfying the stochastic partial differential equation (SPDE)

$$\frac{\partial}{\partial t}\xi_\ell(t,x) = \sqrt{\xi_\ell(t,x)}\,\dot{W}_\ell(t)(x) + G^*\xi_\ell(t,x) - c\left(-\frac{\partial}{\partial x}x\xi_\ell(t,x) + a\delta_0'(x)\right) \tag{2.1.6}$$

with $c = c_\ell$ where G^* is the adjoint of the generator G of a critical Feller branching diffusion given by

$$Gf = \frac{1}{2}x\frac{\partial^2}{\partial x^2}f, \tag{2.1.7}$$

$\dot{W}$ is space-time white noise and δ_0' is the derivative (in the sense of Schwartz distributions) of the δ-function at 0. We will show existence and (weak) uniqueness of the solution to (2.1.6) in Proposition 3.2.1 and paragraph 3.2.3. An equivalent formulation of (2.1.6) is

$$\frac{\partial}{\partial t}\langle \xi_\ell(t), f\rangle = \langle \sqrt{\xi_\ell(t)}\dot{W}_\ell(t) f\rangle + \frac{1}{2}\langle \xi_\ell(t), xf''\rangle - c\left(\langle \xi_\ell(t), xf'\rangle - a\lim_{\varepsilon\downarrow 0}\langle \frac{1}{\varepsilon}\delta_\varepsilon, f\rangle\right), \qquad (2.1.8)$$

where δ_ε is the Dirac measure in ε, and $f:(0,\infty)\to\mathbb{R}$ has bounded first and second derivatives with $f(x)\le \text{const}\ x$. Note that the first term in (2.1.6) (and (2.1.8)) comes from the familiy branching, the second comes from the individual branching, and the "mean reversion" term comes from the individual emigration and the immigration of infinitesimally small families.

In addition, (2.1.6) shows that the sizes of the single families develop independently according to *individual* subcritical Feller branching diffusions. We will therefore call ξ_ℓ, $\ell = 1, 2, \ldots$, a hierarchy of *two-level branching diffusions.*

Two-level branching diffusions have been introduced by Dawson and Hochberg [DH] as superprocesses over Feller branching diffusions, where all mass accumulating in 0 is removed. Therefore, these processes take their values in the measures on $(0,\infty)$. In fact, it turns out that for $t>0$ they have absolutely continuous states. In our context, in addition to the set-up of [DH], there is a "continuous immigration of small families". We will see how this fits into the framework of immigration processes from the boundary studied in Li and Shiga [LS]. For general background on superprocesses and related stochastic partial differential equations, see [D] and [DP2].

The hierarchies of branching equilibria considered in our paper are motivated through a spatial picture which we describe for the case of (2.1.2) ("one-level branching") in subsection 2.2.3 and for the case of (2.1.4) ("two-level branching") in subsection 2.4. The case of a hierarchy of one-level branching systems was studied by Dawson and Greven [DG1, DG2] in the context of super-random walks (or interacting Feller branching diffusions).

For any given $\theta > 0$ (which in the geographical model will play the role of a "global population density") we will construct the hierarchy

$$(\ldots, \zeta^\theta_{\ell+1}, \zeta^\theta_\ell, \ldots, \zeta^\theta_2, \zeta^\theta_1), \qquad (2.1.9)$$

in terms of an entrance law for a backward Markov chain where the conditional law of ζ_ℓ given $\zeta_{\ell+1} = a$ is an equilibrium state of (2.1.1). More precisely, in subsection 4.1 we will show the following result.

Proposition 2.1.1 *Let $(c_\ell)_{\ell\ge 1}$ be a sequence of positive numbers, and let us distinguish two cases:*

a) Assume $\sum_\ell c_\ell^{-1} < \infty$. For $a>0$, $\ell\in\mathbb{N}$, let $\Pi^{(1)}_\ell(a,.)$ be the equilibrium distribution of (2.1.3).

b) Assume $\sum_\ell c_\ell^{-2} < \infty$. For $a>0$, $\ell\in\mathbb{N}$ and $c = c_\ell$, let $\xi_{\ell,a}$ be an equilibrium state of (2.1.6), and $\Pi^{(2)}_\ell(a,.)$ be the distribution of $\int_{(0,\infty)} x\xi_{\ell,a}(x)dx$.

In both cases, for every $\theta > 0$ the backward Markov chain with transition probability

$$P(\zeta_\ell \in A | \zeta_{\ell+1} = a) = \Pi_\ell(a, A),$$

where Π is either $\Pi^{(1)}$ or $\Pi^{(2)}$, has a unique entrance law $\{\zeta_\ell^\theta\}_{\ell=\ldots,2,1}$ satisfying

$$\mathbb{E}\zeta_j^\theta \equiv \theta \tag{2.1.10}$$

and

$$\lim_{j\to\infty} \zeta_j^\theta = \theta \quad a.s. \tag{2.1.11}$$

2.1.1 A cascade of subordinators

To work out parallels between the one- and two-level branching situations described in subsections 2.2.3 and 2.4, and to discuss aspects relevant for the genealogy of the hierarchical branching equilibria, we write $\Pi_\ell(a, .)$ for the equilibrium distribution of (2.1.1) in the two cases (2.1.2) and (2.1.4) (which correspond to cases a) and b) in Proposition 2.1.1).

In both cases the parameter a enters as a factor into the immigration rate of a continuous state branching process, hence $\Pi_\ell(a_1 + a_2, .) = \Pi_\ell(a_1, .) * \Pi_\ell(a_2, .)$. Therefore the $\Pi_\ell(a, .)$ are infinitely divisible distributions on $(0, \infty)$ and there exist subordinators (that is, processes with stationary independent non-negative increments) $S_\ell(a)$, $a \geq 0$, such that

$$\mathcal{L}(S_\ell(a)) = \Pi_\ell(a, .). \tag{2.1.12}$$

We will see in subsection 6.2 that in case a) the S_ℓ are Gamma processes.

In both cases, the Lévy-Khinchin decomposition (see [K2], Chapt. 15) $S_\ell(a) = \sum_i H_i$ describes (asymptotically as $N \to \infty$) the partitioning of the equilibrium population in $B_\ell^{(N)}$ into subpopulations stemming from one and the same immigrant into $B_\ell^{(N)}$, given that the population density of the surrounding block $B_{\ell+1}^{(N)}$ is a. Recall from (2.1.1) and Proposition 2.1.1 that

$$\int x \Pi_\ell(a, dx) = a. \tag{2.1.13}$$

Therefore,

$$\mathbb{E}S_\ell(a) = a. \tag{2.1.14}$$

Let us denote (in either of the two cases) the Lévy measure of S_ℓ by μ_ℓ, and the second moment of μ_ℓ by m_ℓ. Assume that $S_1, S_2, \ldots$ are independent. An iteration of the Lévy-Khinchin representation (which can be interpreted in terms of the genealogy of the branching hierarchy, see subsection 4.3) will show that the Lévy measure of the iterated subordinator $S_\ell(S_{\ell+1}(\ldots S_{j-1}))$ has second moment $m_\ell + \ldots + m_{j-1}$ (see subsection 4.5). We will show that $m_\ell = 1/2c_\ell$ in the one-level case, and $m_\ell = 1/4c_\ell^2$ in the two-level case (see Remarks 6.2.3 and

3.2.7). Hence the transience condition (2.2.4) in the one-level case, and the strong transience condition (2.3.4) in the two-level case, is equivalent to the summability condition

$$\sum_{\ell=1}^{\infty} m_\ell < \infty. \tag{2.1.15}$$

Under this summability condition we will prove in subsection 4.2 that for each $\theta > 0$ the limit in distribution

$$\zeta_\ell^\theta = d\text{-}\lim_{j\to\infty} S_\ell(S_{\ell+1}\dots S_{j-1}(\theta))), \tag{2.1.16}$$

exists, has expectation θ and defines an entrance law with respect to $(\Pi_\ell(a,.))$. In particular one has

$$\zeta_\ell^\theta =^d S_\ell(S_{\ell+1}\dots S_{j-1}(\zeta_j^\theta))), \quad j > \ell. \tag{2.1.17}$$

For each $j > \ell$ this gives a decomposition of ζ_ℓ^θ, which asymptotically as $N \to \infty$, stands for the partitioning of the equilibrium population $\zeta_\ell^{(N,\theta)}$ in $B_\ell^{(N)}$ into subpopulations stemming from one and the same immigrant into $B_{j-1}^{(N)}$.

2.1.2 Genealogy

In Section 4.3 we develop a genealogy of the jumps occurring in the cascade of subordinators. The idea is that given a jump of $S_{\ell+1}(\cdot)$ at time t_i there will be a set of jumps of $S_\ell(\cdot)$ that occur in the time interval $(S_{\ell+1}(t_i-), S_{\ell+1}(t_i))$ and these level ℓ jumps will be said to be *descendants* of the level $(\ell+1)$ jump. In [BLG], continuum flows of subordinators are constructed and the corresponding genealogy of jumps is considered. In subsections 4.3, 4.4 and 4.6 we use this idea to develop the full genealogical structure of the population associated with the entrance law. This leads to a decomposition of the population into a countable collection of subpopulations of individuals having a common ancestor. For the case of critical individual branching this was done in [DG2], for the two-level branching case this is new. We will work out the parallels between the two cases in a general framework, which also sheds some new light on the results of [DG2].

Intuitively this genealogy describes the limiting genealogical structure of the spatial branching equilibria with hierarchical geographic structure described in the Introduction as the parameter $N \to \infty$ and the analogue of the "clan decomposition" of the equilibrium of super-Brownian motion (e.g. [DP1]).

2.2 Hierarchical geography, random walks and branching equilibria

2.2.1 A class of random walks

In order to give a precise formulation for the spatial system we now describe the set Ω_N of sites on which the spatial population lives. For fixed $N \in \mathbb{N}$, let Ω_N be the countably many leaves of a tree all of whose inner nodes have degree $N+1$. In other words, each node in depth $\ell+1$, $\ell = 0, 1, \dots$, has N successors in depth ℓ and one predecessor in depth $\ell+2$. For two sites $y, z \in \Omega_N$, their hierarchical distance $d_N(y,z)$ is defined as the depth of their closest common ancestor in the tree. Note that d_N is an ultrametric, that is, $d_N(y,z) \leq \max\{d_N(y,x), d_N(x,z)\}$. We define the individual migration on Ω_N as follows. Let $q_1^{(N)}, q_2^{(N)}, \dots$ be positive numbers with $\sum_\ell q_\ell^{(N)} < \infty$. At rate $q_\ell^{(N)}$ the individual makes a jump of distance ℓ (i.e. , it waits for an exponentially distributed time with parameter $\sum_\ell q_\ell^{(N)}$ and then jumps a distance j with probability $q_j^{(N)} / \sum_\ell q_\ell^{(N)}$), choosing its arrival site uniformly among all sites at a distance ℓ from its previous site.

The set Ω_N can be identified with the set of sequences in $\{0, 1, \dots, N-1\}$ almost all of whose entries are equal to zero. With component-wise addition mod N, Ω_N is a countable Abelian group (the so called *hierarchical group of order* N). Note that $d_N(y,z)$ is translation invariant; it will be written as $|y-x|$. The migration specified above is a (continuous time) random walk on Ω_N called *hierarchical random walk*. Hierarchical random walks, in particular their transience and recurrence properties, are studied in [DGW2], [DGW3].

2.2.2 A system of branching random walks

We now introduce a system of branching random walks on Ω_N. This is given by a system of particles undergoing symmetric random walks with migration rates $q_\ell^{(N)}$ together with branching. Branching and migration are assumed independent. We specify the branching mechanism as simply as possible: after a unit exponential time, an individual, independently of all the others, either is removed or splits into two, each case with probability $1/2$.

Remark 2.2.1 *[LMW], [Gr] Assume that the migration rates $q_\ell^{(N)}$ are such that the individual random walk is transient. Then, for each $\theta > 0$, there exists a unique branching equilibrium with a mean number of θ individuals per site. This equilibrium is of Poisson type, i.e. the equilibrium population Φ is the superposition $\Phi = \sum_i \Phi_i$ of a Poisson system $\sum_i \delta_{\Phi_i}$ of families (each family consists of a collection of individuals which in a historical picture can be traced back to a common ancestor).*

2.2.3 The hierarchical mean-field limit of a branching equilibrium "near dimension two"

Now consider, for a large N, the total number $Z_\ell^{(N)}(t)$ of individuals in a closed ball $B_\ell^{(N)}$ of radius ℓ at time t. Assume that the individual rate of immigration into and emigration from $B_\ell^{(N)}$ is given by

$$q_{\ell+1}^{(N)} = c_\ell N^{-\ell}, \tag{2.2.1}$$

where the c_ℓ do not depend on N and satisfy

$$\limsup_{\ell} \frac{c_{\ell+1}}{c_\ell} < \infty. \tag{2.2.2}$$

Note that $B_\ell^{(N)}$ has N^ℓ elements, and look at the time evolution of the *population density* (or *block mean*) $A_\ell^{(N)}(t) = Z_\ell^{(N)}(t)N^{-\ell}$ in $B_\ell^{(N)}$ at time scale tN^ℓ. This corresponds to the classical Feller branching diffusion approximation ([EK], Thm. 1.3 on p. 388).

We will call a random walk on Ω_N with jump rates $q_\ell^{(N)}$ of the form (2.2.1) a $(1,(c_\ell),N)$-*random walk*. The following proposition results from combining (3.2.11) and (3.2.15) in [DGW2].

Proposition 2.2.2 *Consider a $(1,(c_\ell),N)$-random walk on Ω_N with rates given by (2.2.1). Assume that*

$$\limsup_{\ell} \frac{c_{\ell+1}}{c_\ell} < N. \tag{2.2.3}$$

Then the random walk is transient if and only if

$$\sum_\ell c_\ell^{-1} < \infty. \tag{2.2.4}$$

Now assume that (c_j) satisfies (2.2.2). Then for each $N > \limsup_\ell \frac{c_{\ell+1}}{c_\ell}$, because of Proposition 2.2.2 and Remark 2.2.1 there is an equilibrium for the system of $(1,(c_\ell),N)$-branching random walks with mean θ for each $\theta > 0$. We now consider the corresponding equilibrium population densities $A_\ell^{(N)}$ in $B_\ell^{(N)}$, $\ell = 1,2,\dots$ (where here and below $(B_\ell^{(N)})$ denotes a sequence of nested balls of radii ℓ in Ω_N). In order to identify the limit as $N \to \infty$ of this sequence of population densities we must consider the dynamics of $A_\ell^{(N)}$ in its natural time scale N^ℓ. Let us first discuss on a heuristic level why, in the limit $N \to \infty$, the drift term on the r.h.s. of (2.1.1) arises on the time scale N^ℓ for the population density in $B_\ell^{(N)}$. Because of the ultrametric structure of Ω_N, an individual in $B_\ell^{(N)}$ has to make a jump of size $\geq \ell+1$ in order to leave $B_\ell^{(N)}$. Because of (2.2.1) and (2.2.2), for N large, jumps of size $> \ell+1$ happen rarely compared to jumps of size $\ell+1$ (since $q_{l+k} = o(q_{l+1})$ as $N \to \infty$ for $k > 1$). Hence the individual emigration rate from $B_\ell^{(N)}$ on time scale N^ℓ is c_ℓ (asymptotically as $N \to \infty$). Concerning immigration into $B_\ell^{(N)}$, again because of (2.2.1), it is only the environment in $B_{\ell+1}^{(N)}$ that is relevant for large N. An individual in $B_{\ell+1}^{(N)} \setminus B_\ell^{(N)}$ necessarily has to jump a distance $\ell+1$ in order to make it into $B_\ell^{(N)}$, and on average every $(N-1)$-st of these jumps will take the individual into $B_\ell^{(N)}$ (note that $B_{\ell+1}^{(N)}$ is N times as large as $B_\ell^{(N)}$). Since the block mean $A_{\ell+1}^{(N)}$ does not change on the time scale N^ℓ as $N \to \infty$, the total immigration rate into $B_\ell^{(N)}$ on this time scale is (asymptotically as $N \to \infty$) of the order

$$Z_{\ell+1}^{(N)} N^\ell q_{\ell+1}^{(N)}/N = (N^{\ell+1}A_{\ell+1}^{(N)})c_\ell/N = c_\ell A_{\ell+1}^{(N)} N^\ell. \tag{2.2.5}$$

This suggests that the limiting dynamics of the population densities $A_\ell^{(N)}$, in their natural time scales, as $N \to \infty$ is given by (2.1.1) with $a = A_{\ell+1}^{(N)}$. The separation of time scales on balls of different radii that underlies the previous discussion is a feature of the hierarchical random walks, which is due to the ultrametric structure of Ω_N (see [DGW2]). This is also explained in more detail in Remark 2.3.4 below.

Instead of branching particle systems, Dawson and Greven [DG2] consider super-random walks (or so-called interacting Feller diffusions) on Ω_N. Note also that the definition of the random walk in [DG2] is slightly different but asymptotically equivalent as $N \to \infty$ to the one used in this paper. In [DG2] the sites to which a jump is made are chosen with uniform distribution on a ball rather than on a sphere. However, the "interior" of the ball is asymptotically negligible as compared to the sphere as N goes to infinity.

A particle system analogue of Theorem 4(b) of [DG2] is the following, which we state without proof.

Proposition 2.2.3 *Consider a sequence (c_ℓ) satisfiying conditions (2.2.4) and (2.2.2) for transience. For N large enough such that (2.2.3) is met, let the individual migration process be a $(1,(c_\ell),N)$-random walk. For $\theta > 0$ let $Z_\ell^{(N,\theta)}(t)$ denote the total number of individuals in $B_\ell^{(N)}$ at time t and $A_\ell^{(N,\theta)}(t) := Z_\ell^{(N,\theta)}(t)N^{-\ell}$ be the population density at time t in $B_\ell^{(N)}$ in the Poisson type branching equilibrium population on Ω_N with mean number θ of individuals per site (see Remark 2.2.1). Let $\{\zeta_\ell^\theta\}$ be the entrance law provided by Proposition 2.1.1, case a). Then*

$$\{A_\ell^{(N,\theta)}(0)\}_{\ell\in\mathbb{N}} \Longrightarrow \{\zeta_\ell^\theta\}_{\ell\in\mathbb{N}} \quad \text{as } N \to \infty,$$

where $\Longrightarrow$ denotes weak convergence of finite dimensional distributions.

Let us now explain in which sense transient $(1,(c_\ell),N)$-random walks can be interpreted as random walks "near dimension 2".

Definition 2.2.4 *[DGW2] Let Z be an irreducible transient random walk on a countable Abelian group Γ. Its degree of transience is defined by*

$$\gamma := \sup\{\eta \geq 0 : \mathbb{E}_0 L^\eta < \infty\}, \tag{2.2.6}$$

where L is the last exit time of Z from $0 \in \Gamma$.

Expressed in more analytic terms, the degree of transience of Z is

$$\gamma = \sup\{\eta \geq 0 : \int_1^\infty t^\eta p_t(0,0)\, dt < \infty\},$$

where p_t is the transition probability of Z [DGW2, SaW].

For simple symmetric random walk on the d-dimensional lattice $\mathbb{Z}^d$, it is well-known that dimension 2 is the borderline for transience. For $d > 2$, the degree of transience is $d/2 - 1$, since the rate of decay of the transition probability is $p_t(0,0) \sim \text{const}.t^{-d/2}$.

Remark 2.2.5 *[DGW2] a) Let* $0 < c < N$. *Then the* $(1,(c^\ell),N)$*-random walk on* Ω_N *is transient if and only if* $c > 1$. *In this case its degree of transience is* $\log c/(\log N - \log c)$. *Thus for fixed c the transient* $(1,(c^\ell),N)$*-random walks on* Ω_N *have degrees of transience* $O(1/\log N)$ *and therefore asymptotically as* $N \to \infty$, *can be viewed as analogues of random walks "near (Euclidean) dimension 2".*

b) Assume that (c_ℓ) *satisfies conditions* (2.2.4) *and* (2.2.2), *and put* $c := \limsup c_{\ell+1}/c_\ell \geq 1$. *Then, for all* $N > c$ *the* $(1,(c_\ell),N)$*-random walk on* Ω_N *is transient with degree of transience in the interval* $[0, \log c/(\log N - \log c)]$.

Since certain properties of systems of branching random walks such as persistence and structure of occupation time fluctuations depend only on the degree of transience of the random walks, branching populations whose migration process is a hierarchical random walk can give insight into the behavior of a larger class of branching populations whose random walks have the same degree of transience.

2.3 Two-level branching systems in a hierarchical geography

2.3.1 Strongly transient migration

Whereas the situation described in subsection 2.2.3 gives an analogue to a situation "near dimension 2", our main focus later on will be on the analogue to a situation "near dimension 4". In this context we will consider the (stronger) mass fluctuations induced by a critical reproduction of whole *families* (of mutually related individuals), together with a (stronger) smoothing caused by a strongly transient migration.

Definition 2.3.1 *An irreducible random walk Z on a countable Abelian group* Γ *is called strongly transient if*

$$\mathbb{E}_0 L < \infty \tag{2.3.1}$$

where L denotes the last exit time of Z from 0. A transient random walk with $\mathbb{E}_0 L = \infty$ *is called weakly transient.*

Note that strong transience is equivalent to

$$\int_1^\infty t p_t(0,0) dt < \infty,$$

and a necessary condition is that the degree of transience be equal to or greater than 1 [DGW2]. Moreover, as mentioned above simple symmetric d-dimensional random walk has degree of transience $d/2 - 1$, and it is strongly transient iff $d > 4$.

In order to introduce a family of strongly transient random walks on Ω_N we replace (2.2.1) by

$$q_{\ell+1}^{(N)} = c_\ell N^{-\ell/2}, \tag{2.3.2}$$

where the c_ℓ do not depend on N and satisfy (2.2.2). We will call a random walk on Ω_N with jump rates $q_\ell^{(N)}$ of the form (2.3.2) a $(2,(c_\ell),N)$*-random walk*.

Proposition 2.3.2 *[DGW2] Consider a $(2,(c_\ell),N)$-random walk on Ω_N with rates given by (2.3.2). Assume that*

$$\limsup_\ell \frac{c_{\ell+1}}{c_\ell} < N^{1/2}. \tag{2.3.3}$$

Then the random walk is strongly transient if and only if

$$\sum_\ell c_\ell^{-2} < \infty. \tag{2.3.4}$$

Remark 2.3.3 *[DGW2] a) Let $0 < c < N^{1/2}$. Then the $(2,(c^\ell),N)$-random walk on Ω_N is strongly transient if and only if $c > 1$. In this case its degree of transience is*

$$\frac{\log N + 2\log c}{\log N - 2\log c}.$$

Thus for fixed c the strongly transient $(2,(c^\ell),N)$-random walks have degree of transience $1+O(1/\log N)$ and therefore asymptotically as $N \to \infty$ can be viewed as analogues of random walks "near (Euclidean) dimension 4".

b) Assume that (c_ℓ) satisfies conditions (2.3.4) and (2.2.2), and put $c := \limsup c_{\ell+1}/c_\ell \geq 1$. Then, for all $N > c^2$ the $(2,(c_\ell),N)$-random walk on Ω_N is strongly transient with degree of transience in the interval $[1,(\log N + 2\log c)/(\log N - 2\log c)]$.

Remark 2.3.4 *The natural time scale for the strongly transient $(2,(c^\ell),N)$-random walk in $B_\ell^{(N)}$ is $N^{\ell/2}$. More precisely, as shown in [DGW2], asymptotically as $N \to \infty$ for this random walk on the time scale $N^{\ell/2}$ only the migrations within the ball $B_\ell^{(N)}$ and the surrounding ball $B_{\ell+1}^{(N)}$ are relevant. Similarly, this occurs for the transient $(1,(c^\ell),N)$-random walk in time scale N^ℓ (see [DGW2] for details). This effect is basic for the limiting hierarchy of branching equilibria obtained in this paper.*

2.3.2 Two-level branching equilibria

A main objective of this paper is to study two-level branching systems for a migration which is on the border between strong and weak transience – recall that a strongly transient migration is required for the existence of a branching equilibrium. Thus, for Euclidean random walks, $d = 4$ is the critical dimension for a two-level branching system in the same way as dimension $d = 2$ is the critical dimension for a one- level branching system.

We are going to study two-level branching systems on Ω_N. Consider a system of $(2,(c_\ell),N)$-random walks on Ω_N such that (c_ℓ) satisfies the conditions (2.3.3) and (2.3.4) for strong transience, and recall from Remark 2.2.5 that these random walks are close to the border to weak transience at least for large N if (c_ℓ) does not grow superexponentially. Introduce, in addition to the individual branching and migration, a *family branching*: independently of everything else, after each family Φ_i at rate one either vanishes or reproduces resulting in two identical copies Φ_i',Φ_i'', each case with probability 1/2. After a reproduction event Φ_i' and Φ_i'' evolve as independent one level branching systems. This creates the basic *two-level branching system* $\Psi^{(N)}(t)$ which is started with the family system at time t_0 given by $\Psi^{(N)}(t_0) = \sum_i \delta_{\Phi_i}$ described in Remark 2.2.1.

The following result is the analogue for two-level branching systems on Ω_N of the propositon in [GHW], p. 537. The latter is formulated for $\mathbb{R}^d$ as space of positions; its proof carries over to Ω_N in an obvious way.

Proposition 2.3.5 *Assume that*
(i) the random walk on Ω_N is strongly transient and
(ii) $\Psi^{(N)}(t_0) = \sum_i \delta_{\Phi_i}$ where $\{\Phi_i\}$ corresponds to the family decomposition of an equilibrium state for the one-level branching random walk with mean number θ of individuals per site.
Then as $t_0 \to -\infty$, the two-level branching system $\Psi^{(N)}(0)$ converges in distribution to a translation invariant equilibrium $\Psi^{(N,\theta)}(0)$ with a mean number θ of individuals per site.

Remark 2.3.6 *The notation $\Psi^{(N)}(t)$ and $\Psi^{(N,\theta)}(t)$ will be used throughout to denote the two-level branching system and the equilibrium process with mean θ, respectively.*

Remark 2.3.7 *Greven and Hochberg [GH] have obtained more general conditions under which the convergence to equilibrium as in Proposition 2.3.5 occurs as well as conditions under which it fails.*

Now we consider a system of $(2,(c_\ell),N)$ random walks with (c_ℓ) satisfying conditions (2.3.4) and (2.2.2) for strong transience. Then for each $N > \left(\limsup_\ell \frac{c_{\ell+1}}{c_\ell}\right)^2$ because of Propositions 2.3.2 and 2.3.5 there is a two-level branching equilibrium $\Psi^{(N,\theta)}(t)$ with a mean number of θ individuals per site for each $\theta > 0$. A main objective of this paper is to study the equilibrium structure that arises in the limit as $N \to \infty$ of the corresponding sequence of family structures in the blocks $B_\ell^{(N)}$.

2.4 The hierarchical mean-field limit of a two-level branching equilibrium "near dimension four"

2.4.1 Local normalized family-size process

Let, for fixed N, $\Psi^{(N,\theta)}(t), t \in \mathbb{R}$, be the equilibrium process of the two-level branching system as provided by Proposition 2.3.5. Denote the number of families in $\Psi^{(N,\theta)}(t)$ having j individuals in $B_\ell^{(N)}$ by $n_\ell^{(N,\theta)}(t,j)$ and write $H_\ell^{(N,\theta)}(t) := \sum_j n_\ell^{(N,\theta)}(t,j)\delta_j$ for the *local family-size process*. For each fixed ℓ, $H_\ell^{(N,\theta)}(0)$ is a random measure on $\mathbb{N}$ which describes the population of equilibrium family sizes in the block $B_\ell^{(N)}$. We note that the process $\{H_\ell^{(N,\theta)}(t)\}_{t\in\mathbb{R}}$ can be viewed as a branching Markov chain on $\mathbb{Z}_+$ with instantaneous killing at 0, the Markov chain on $\mathbb{Z}_+$ being a standard subcritical binary branching process with immigration. Note that $\sum_j jn_\ell^{(N,\theta)}(t,j) = Z_\ell^{(N,\theta)}(t)$, the number of individuals in $B_\ell^{(N)}$ at time t . Now consider the *equilibrium normalized family size process* defined by

$$\eta_\ell^{(N,\theta)}(t) = \sum_j N^{-\ell/2} n_\ell^{(N,\theta)}(tN^{\ell/2}, j)\delta_{jN^{-\ell/2}}, \tag{2.4.1}$$

In other words, for $0 < a < b < \infty$,

$$\eta_\ell^{(N,\theta)}(t)(a,b) = N^{-\ell/2}\eta_\ell(N^{\ell/2}t, N^{\ell/2}(a,b)). \tag{2.4.2}$$

Note that the natural time scale in which to observe the subpopulation in $B_\ell^{(N)}$ in this case is $N^{\ell/2}$ and not N^ℓ as was the case for one-level branching (see Remark 2.3.4).

For each ℓ, N and t, $\eta_\ell^{(N,\theta)}(t)$ is a random measure on $(0,\infty)$. More precisely, we take as state space the set $M^1(0,\infty)$ of Radon measures μ on $(0,\infty)$ that satisfy the condition $\int x\mu(dx) < \infty$. (Note that we do not keep track of families of size 0.)

The corresponding normalized population mass in $B_\ell^{(N)}$ (the "radius ℓ block average") is given by

$$\zeta_\ell^{(N,\theta)}(t) = \int x\eta_\ell^{(N,\theta)}(t,dx) = \sum_j N^{-\ell} jn_\ell^{(N,\theta)}(N^{\ell/2}t, j) = N^{-\ell} Z_\ell^{(N,\theta)}(N^{\ell/2}t), \tag{2.4.3}$$

and in terms of $\Psi^{(N,\theta)}(t)$:

$$\zeta_\ell^{(N,\theta)}(t) = \frac{1}{N^\ell}\int \sum_{x\in\Omega_N, |x|\le\ell} \mu(x)\Psi^{(N,\theta)}(N^{\ell/2}t, d\mu). \tag{2.4.4}$$

2.4.2 Convergence theorem

We now state our main result that makes precise the sense in which the entrance law described in section 2.1 approximates the two-level spatial equilibrium in Ω_N obtained in Proposition 2.3.5 when the parameter $N \to \infty$ and the random walk satisfies (2.3.4).

Theorem 2.4.1 *(Hierarchical mean-field limit) Consider a sequence (c_ℓ) satisfying conditions (2.2.2) and (2.3.4) for strong transience of the $(2,(c_\ell),N)$-random walk. For fixed $N \in \mathbb{N}$ obeying (2.3.3), and a sequence of nested blocks $B_\ell^{(N)}$ in Ω_N, let $\{\zeta_\ell^{(N,\theta)}(0)\}_{\ell\in\mathbb{N}}$ be the* radius

ℓ block averages *(defined in (2.4.3)) of an equilibrium two-level branching system with an expected number θ of particles per site. Let $\{\zeta_\ell^\theta\}_{\ell\in\mathbb{N}}$ be the entrance law provided by Proposition 2.1.1, case b). Then*

$$\{\zeta_\ell^{(N,\theta)}(0)\}_{\ell\in\mathbb{N}} \Longrightarrow \{\zeta_\ell^\theta\}_{\ell\in\mathbb{N}} \quad \text{as } N\to\infty.$$

The proof of this result is based on the spatial ergodic theorem for the equilibrium random field on Ω_N obtained in section 5.1, a separation of time scales property derived in section 5.2 and a diffusion limit theorem for the family size processes $\{\eta_\ell^{(N,\theta)}(t)\}$ as $N\to\infty$ obtained in section 5.3. Using these results the proof of Theorem 2.4.1 is given in section 5.4.

3 Super subcritical Feller branching

In this section we continue the investigation of diffusion limits of two-level branching populations *without geographical structure*, which were introduced in [DH]. In our case, these are superprocesses whose basic process is a subcritical Feller branching diffusion killed at 0 (this killing corresponds to the removal of void families). With a view towards the application to the hierarchically structured geographical model, we will concentrate in subsection 3.1 on an initial condition of *many small families*, which in the diffusion limit corresponds (on a heuristic level) to an intial condition $\infty\delta_0$. In subsection 3.2.1 we investigate time stationary super subcritical Feller branching processes which arise as diffusion limits of two-level branching populations with a high-rate immigration of individuals. The simplest situation is to think of each immigrant individual founding a new family; in the diffusion limit this leads to *super subcritical Feller branching diffusions with immigration* of $\infty\delta_0$ at a constant rate (abbreviated by SSFBI). Again with a view towards the geographical model, we will consider the situation where (only) every once in a while a newly immigrated individual belongs to an already existing family. If this happens relatively rarely, then the diffusion limit remains to be SSFBI, see Proposition 3.2.1 and Corollary 3.2.2.

3.1 Diffusion limit of two-level branching particle systems

For $c>0$ and $\varepsilon>0$, consider the $M^1(0,\infty)$-valued family-size process $\{\tilde H^\varepsilon(t,dx)\}$ of a two-level branching particle system (without geographical structure) with branching rates equal to $1/\varepsilon$ at both levels and subcritical at the individual level with subcriticality parameter εc. (An example is the local family size process $H_\ell^{(N,\theta)}$ (defined in subsection 2.4.1) run at time scale $N^{\ell/2}$ and *with immigration suppressed*; here, $c=c_\ell$ and $\varepsilon=N^{-\ell/2}$.)

Consider the rescaled family-size process

$$\widetilde{\eta}^\varepsilon(t,(x_1,x_2)) := \varepsilon \widetilde{H}^\varepsilon(t,(x_1/\varepsilon,x_2/\varepsilon)),\ t>0. \tag{3.1.1}$$

Proposition 3.1.1 *Let $\widetilde{\eta}^{\varepsilon}(t)$ be as in (3.1.1). Assume that $\widetilde{\eta}^{\varepsilon}(0) = \varepsilon\lfloor\frac{a}{\varepsilon}\rfloor\delta_{\varepsilon\lfloor\frac{x}{\varepsilon}\rfloor}$ where $a > 0$ and $x > 0$ are fixed. Then*
(a) as $\varepsilon \to 0$,

$$\{\widetilde{\eta}^{\varepsilon}(t)\}_{t\geq 0} \Longrightarrow \{\xi(t)\}_{t\geq 0} \tag{3.1.2}$$

in the sense of weak convergence of $M_f([0,\infty))$-valued càdlàg processes, and $\xi(t)$ is the $M_f([0,\infty))$-valued superprocess starting in $\xi(0) = a\delta_x$, whose motion is the subcritical Feller branching process with generator G_c given by

$$G_c f = \left(\frac{1}{2}x\frac{\partial^2}{\partial x^2} - cx\frac{\partial}{\partial x}\right) f, \tag{3.1.3}$$

acting on functions

$$f \in C_0^2([0,\infty)) := \{f \in C^2([0,\infty)) : \lim_{x\to\infty} f(x) = 0,\ f(0) = 0\}.$$

(Here $M_f([0,\infty))$ denotes the space of finite measures on $[0,\infty)$.)
(b) The law of the process $\xi(t)$ is uniquely determined by the Laplace functional as follows:

$$\begin{aligned}\mathbb{E}_{a\delta_x}\left(\exp(-\int_{\mathbb{R}+} f(y)\xi(t,dy))\right) &= \exp(-\int V_t f(y)\,\xi(0,dy)) \\ &= \exp(-au(t,x)),\end{aligned} \tag{3.1.4}$$

where $V_t f(x) = u(t,x)$ is the unique solution of the non-linear p.d.e.

$$\begin{aligned}\frac{\partial u(t,x)}{\partial t} &= G_{c_\ell} u(t,x) - \frac{1}{2}u(t,x)^2, \\ u(0,x) &= f(x).\end{aligned} \tag{3.1.5}$$

Proof This is essentially Theorem 4.1 of [DH].

Remark As an application of (3.1.4) and (3.1.5) we obtain the compact support property of ξ appearing in Proposition 3.1.1. Assume that $\xi(0)$ has compact support and let $\bar{R}_t$ denote the range of $\xi(s)$ up to time t. Then following the method of Iscoe as in Theorem 1.8 of [LS] or Theorem A of [DLM] one can show that $\bar{R}_t$ is bounded almost surely. This involves showing (as in [DLM]) that for $c > 0$ the equation

$$\begin{aligned}\frac{1}{2}x\frac{\partial^2 u}{\partial x^2} &= xc\frac{\partial u}{\partial x} + \frac{1}{2}u^2, \\ u(0) = 0, \quad \left.\frac{\partial u}{\partial x}\right|_{x=0} &= \alpha\end{aligned}$$

has for any $\alpha > 0$ a blow-up at some finite x.

3.1.1 Evolution equation and entrance law

In Proposition 3.2.1 we will prove an extension of Proposition 3.1.1 which includes immigration. In this subsection we obtain some properties of the solution of the evolution equation that will be used there.

Let

$$C_1(0,\infty) := \{f \in C(0,\infty),\ \lim_{x\to\infty} f(x) = 0,\ |f(x)| \le \text{const} \cdot x\} \tag{3.1.6}$$

and (T_t) be the semigroup of the Feller branching diffusion with subcriticality parameter c, absorbed at zero.

Let $V_t f$ be the solution of

$$\frac{\partial}{\partial t} V_t f = \frac{1}{2} x \frac{\partial^2}{\partial x^2} V_t f - cx \frac{\partial}{\partial x} V_t f - \frac{1}{2}(V_t f)^2, \quad V_0 f = f. \tag{3.1.7}$$

Lemma 3.1.2 *Let $f \in C_1(0,\infty)$, $f \ge 0$. Then for $t > 0$, $V_t f(x)$ is differentiable at zero and*

$$(V_t f)'(0) = \int f(y)\kappa_t(dy) - \frac{1}{2}\int_0^t \int (V_s f)^2(y)\kappa_{t-s}(dy)\, ds \tag{3.1.8}$$

where κ_t is the (T_t)-entrance law given by (6.2.16) in the Appendix.

Proof. First note if $f \in C_1^+(0,\infty)$ then $V_t f(x) \le T_t f(x)$ so that V_t maps $C_1^+(0,\infty)$ into itself. Using the evolution form of (3.1.7),

$$\lim_{\varepsilon\downarrow 0} \frac{V_t f(\varepsilon)}{\varepsilon} = \lim_{\varepsilon\downarrow 0} \frac{T_t f(\varepsilon)}{\varepsilon} - \frac{1}{2}\int_0^t \lim_{\varepsilon\downarrow 0} \frac{T_{t-s}(V_s f)^2(\varepsilon)}{\varepsilon} ds. \tag{3.1.9}$$

The result then follows from (6.2.17). ■

Proposition 3.1.3 *Let $\xi^\varepsilon(t), t \ge 0$ denote the super subcritical Feller branching diffusion (without immigration) process starting in $\varepsilon^{-1}\delta_\varepsilon$ at time 0. Then as $\varepsilon \to 0$, ξ^ε converges in the sense of weak convergence of $M^1(0,\infty)$-valued continuous processes on the time interval $[t_0,\infty)$ for all $t_0 > 0$ to a measure-valued diffusion ξ^0 where, for all $t > 0$, $\xi^0(t)$ is an infinitely divisible random measure with Laplace functional given by*

$$\mathbb{E}\exp\left(-\langle \xi^0(t), f\rangle\right) = \exp\left(-(V_t f)'(0)\right). \tag{3.1.10}$$

Proof. Because $V_t f(0) = 0$ we have

$$\begin{aligned} \mathbb{E}\exp\left(-\langle \xi^0(t), f\rangle\right) &= \lim_{\varepsilon\downarrow 0} \mathbb{E}\exp\left(-\langle \xi^\varepsilon(t), f\rangle\right) && (3.1.11)\\ &= \lim_{\varepsilon\downarrow 0} \mathbb{E}_{\varepsilon^{-1}\delta_\varepsilon} \exp\left(-\langle \xi(t), f\rangle\right) \\ &= \lim_{\varepsilon\downarrow 0} \exp\left(-\varepsilon^{-1}(V_t f)(\varepsilon)\right) \\ &= \lim_{\varepsilon\downarrow 0} \exp\left(-\varepsilon^{-1}((V_t f)(\varepsilon) - (V_t f)(0)\right) \\ &= \exp\left(-(V_t f)'(0)\right). \end{aligned}$$

■

Remark 3.1.4 *Since $\xi^0(t)$ is infinitely divisible, its Laplace transform must be of the form*

$$\mathbb{E}\exp\left(-\langle\xi^0(t), f\rangle\right) = \exp\left(-\int(1 - e^{-\langle m,f\rangle})K_t(dm)\right), \quad f \in C_1((0,\infty)). \tag{3.1.12}$$

for some uniquely determined measure K_t on $M(0,\infty)$, the space of Radon measures on $(0,\infty)$. The measure K_t is the canonical measure of $\xi^0(t)$.

We can now put these results into the framework of [LS].

A crucial property of the entrance law (κ_t) given by (6.2.15), which follows immediately by partial integration from the density (6.2.16), is given by the following lemma.

Lemma 3.1.5 *For all bounded continuously differentiable functions g on $[0,\infty)$ with $g(0) = 0$,*

$$\lim_{t\to 0}\int_0^\infty g(x)\kappa_t(x)dx = g'(0). \tag{3.1.13}$$

We fix a strictly positive function $\rho \in D(G_c)$ with

$$\rho(x) = x \text{ for } x \in (0, \frac{1}{2}], \quad \rho(x) = 1 \text{ for } x \geq 1. \tag{3.1.14}$$

Note that such a ρ meets condition [A] in [LS]. We take as state space $M_\rho := \{\mu \in M(0,\infty) : \int \rho(x)\mu(dx) < \infty\}$.

Following [LS] we put

$$C_\rho(0,\infty) := \{f \in C(0,\infty) : |f| \leq \text{const}\rho, \lim_{x\to\infty} f(x) = 0\}, \tag{3.1.15}$$

$$D_\rho(G_c) := \{f \in D(G_c) : f, G_c f \in C_\rho(0,\infty)\} \tag{3.1.16}$$

and

$$\kappa_{0^+}(g) := \lim_{t\to 0}\int_0^\infty g(x)\kappa_t(x)dx, \quad g \in D_\rho(G_c). \tag{3.1.17}$$

Combining (3.1.10), (3.1.12), (3.1.13) and (3.1.17) we obtain

$$\int(1 - e^{-\langle m,f\rangle})K_t(dm) = (V_t f)'(0) = \kappa_{0^+}(V_t f), \quad f \in C_1((0,\infty)), \tag{3.1.18}$$

where $V_t f$ is the solution of (3.1.7).

3.2 Super subcritical Feller branching diffusion with individual immigration

3.2.1 Diffusion limit with immigration

We now extend Proposition 3.1.1 to include immigration, taking a fixed t_0 as origin of time. Since in our application the population from which the immigrants come is structured into families that undergo family branching we incorporate multitype immigration and label the set of possible families of immigrants by $I := [0, 1]$.

Let $M^1(I\times(0,\infty))$ denote the set of Radon measures μ on $I\times(0,\infty)$ satisfying $\int_{I\times(0,\infty)} x\mu(dy,dx) < \infty$. We denote the single atom measure corresponding to one individual of type $y_k^\varepsilon \in I$ by $\delta_{y_k^\varepsilon,1}$.

Consider the $M^1(I\times(0,\infty))$-valued family-size process $\{H_I^\varepsilon(t,dy,dx)\}_{t_0\le t}$ with branching rates equal to $1/\varepsilon$ at both levels, critical at the family level and subcritical at the individual level with subcriticality parameter εc (i.e. a mean offspring number of $1-\varepsilon c$ per branching event), and with immigration of individuals of type $y_k^\varepsilon \in I$, given by $\delta_{y_k^\varepsilon,1}$, $k \in \mathbb{N}$, at rate $ca_k^\varepsilon/\varepsilon^2$ with $\sum_k a_k^\varepsilon = a$ and $\lim_{\varepsilon\to 0}\sup_k a_k^\varepsilon = 0$. (The motivation for this comes from our geographical model, we will see in Section 5 that this setting corresponds to the situation where the surrounding population, which serves as the source of immigration, is thought to have a frozen family structure.) Consider the rescaled process $\{\eta_I^\varepsilon(t)\}_{t_0\le t}$ defined by

$$\eta_I^\varepsilon(t,\{y_k^\varepsilon\}\times(x_1,x_2)) := \varepsilon H_I^\varepsilon(t,\{y_k^\varepsilon\}\times(\frac{x_1}{\varepsilon},\frac{x_2}{\varepsilon})) \tag{3.2.1}$$

started at time t_0 with the measure

$$\eta_I^\varepsilon(t_0) = \sum_{k,j}\mu_0^\varepsilon(y_k^\varepsilon, j\varepsilon)\delta_{y_k^\varepsilon,j\varepsilon}$$

and let

$$\tilde{\alpha}^\varepsilon = \sum a_k^\varepsilon \delta_{y_k^\varepsilon,\varepsilon}.$$

Proposition 3.2.1 *Assume that as $\varepsilon\to 0$, $\eta_I^\varepsilon(t_0)\Rightarrow\mu_0\in M^1(I\times(0,\infty))$ and $\tilde{\alpha}^\varepsilon(dy,dx)\to\alpha(dy)\delta_0(dx)$ in the sense of weak convergence of finite measures on $I\times[0,\infty)$, where α is a nonatomic measure whose total mass $a=\alpha(I)$ plays the role of the overall immigration rate. Then as $\varepsilon\to 0$,*

$$\{\eta_I^\varepsilon(t,\cdot)\}_{t_0\le t}\Longrightarrow\{\xi_I(t)\}_{t_0\le t}$$

in the sense of weak convergence of $M^1(I\times(0,\infty))$-valued càdlàg processes on the time interval $[t_0,\infty)$, where $\{\xi_I(t)\}$ is the measure-valued diffusion with generator

$$\begin{aligned}\mathfrak{G}F(\mu) &= f'(\langle\mu,\phi\rangle)\langle\mu,G_c\phi\rangle+\tfrac{1}{2}f''(\langle\mu,\phi\rangle)\langle\mu,\phi^2\rangle \\ &\quad+f'(\langle\mu,\phi\rangle)c\int\tfrac{\partial\phi}{\partial x}(y,x)|_{x=0}\alpha(dy),\end{aligned} \tag{3.2.2}$$

G_c is the operator given by (3.1.3) and $D(\mathfrak{G})$ denotes the class of functions of the form $F_{f,\phi}(\mu) = f(\langle \mu, \phi \rangle)$ where $f \in C_b^3(\mathbb{R})$ and ϕ is a continuous function on $I \times (0,\infty)$ with $\frac{\partial^2 \phi(y,x)}{\partial x^2}$ bounded and continuous on $I \times [0,\infty)$ and $|\sup_y \phi(y,x)| \le \text{const} \cdot (x \wedge 1)$.

Proof. η_I^ε is given by a pregenerator $\mathfrak{G}^\varepsilon$ acting on the class of bounded continuous functions F on $M^1(I \times (0,\infty))$ given by

$$\begin{aligned}
\mathfrak{G}^\varepsilon F(\mu) \quad &= \sum_k \frac{ca_k^\varepsilon}{\varepsilon^2}[F(\mu + \varepsilon\delta_{y_k^\varepsilon,\varepsilon}) - F(\mu)] \\
&+\frac{1}{2}\sum_k\sum_{j=1}^\infty [F(\mu+\varepsilon\delta_{y_k,j\varepsilon}) - F(\mu)]\frac{\mu(y_k,j\varepsilon)}{\varepsilon^2} \\
&+\frac{1}{2}\sum_k\sum_{j=1}^\infty [F(\mu-\varepsilon\delta_{y_k,j\varepsilon}) - F(\mu)]\frac{\mu(y_k,j\varepsilon)}{\varepsilon^2} \\
&+\frac{1}{2}(1-\varepsilon c)\sum_k\sum_{j=1}^\infty [F(\mu-\varepsilon\delta_{y_k,j\varepsilon}+\varepsilon\delta_{y_k,(j+1)\varepsilon}) - F(\mu)]\frac{j\mu(y_k,j\varepsilon)}{\varepsilon^2} \\
&+\frac{1}{2}(1+\varepsilon c)\sum_k\sum_{j=1}^\infty [F(\mu-\varepsilon\delta_{y_k,j\varepsilon}+\varepsilon\delta_{y_k,(j-1)\varepsilon})) - F(\mu)]\frac{j\mu(y_k,j\varepsilon)}{\varepsilon^2}.
\end{aligned} \tag{3.2.3}$$

Here the first term comes from the immigration, the second and third from the critical branching at the family level and the fourth and fifth from the subcritical branching at the individual level with subcriticality parameter $c > 0$.

For $F \in D(\mathfrak{G})$, $\mathfrak{G}^\varepsilon F$ takes the form

$$\begin{aligned}
&\mathfrak{G}^\varepsilon F(\mu) \\
= \quad & \sum_k \frac{ca_k^\varepsilon}{\varepsilon^2}[f(\langle\mu,\phi\rangle + \varepsilon\phi(y_k^\varepsilon,\varepsilon)) - f(\langle\mu,\phi\rangle)] \\
+ \quad & \frac{1}{2}\sum_k\sum_{j=1}^\infty [f(\langle\mu,\phi\rangle + \varepsilon\phi(y_k,j\varepsilon)) - f(\langle\mu,\phi\rangle)]\frac{\mu(y_k,j\varepsilon)}{\varepsilon^2} \\
+ \quad & \frac{1}{2}\sum_k\sum_{j=1}^\infty [f(\langle\mu,\phi\rangle - \varepsilon\phi(y_k,j\varepsilon)) - f(\langle\mu,\phi\rangle)]\frac{\mu(y_k,j\varepsilon)}{\varepsilon^2} \\
+ \quad & \frac{1}{2}(1-\varepsilon c)\sum_k\sum_{j=1}^\infty [f(\langle\mu,\phi\rangle - \varepsilon\phi(y_k,j\varepsilon) + \varepsilon\phi(y_k,(j+1)\varepsilon))) - f(\langle\mu,\phi\rangle)] \\
& \qquad \cdot\frac{j\mu(y_k,j\varepsilon)}{\varepsilon^2}
\end{aligned} \tag{3.2.4}$$

$$+ \quad \frac{1}{2}(1+\varepsilon c)\sum_k \sum_{j=1}^{\infty} [f(\langle \mu, \phi\rangle - \varepsilon\phi(y_k, j\varepsilon) + \varepsilon\phi(y_k,(j-1)\varepsilon))) - f(\langle\mu,\phi\rangle)]$$
$$\cdot \frac{j\mu(y_k, j\varepsilon)}{\varepsilon^2}.$$

Tightness of the family $\{\eta_I^\varepsilon\}_{0<\varepsilon\le 1}$ is proved by a standard argument as in [Wu] or [DZ]. Using a Taylor expansion for the functions f and ϕ it can be verified that as $\varepsilon \to 0$, for $F_{f,\phi} \in D(\mathfrak{G})$, $\mathfrak{G}^\varepsilon F_{f,\phi}(\mu) \to \mathfrak{G}F_{f,\phi}$. For example, if $\sum_{k,j}\mu(y_k^\varepsilon, j\varepsilon)\delta_{y_k^\varepsilon, j\varepsilon} \Rightarrow_{\varepsilon\to 0} \mu(dy,dx)$, $\int\int x\mu(dy,dx) < \infty$ and $|\phi(y,x)| \le \text{const}\cdot(x\wedge 1)$, then applying Taylor's formula with remainder to the second and third terms leads to

$$\frac{1}{2}f''\left(\sum_{k,j}\phi(y_k,j\varepsilon)\mu(y_k,j\varepsilon)\right)\sum_{k,j}\phi^2(y_k,j\varepsilon)\mu(y_k,j\varepsilon)$$
$$+\frac{\varepsilon}{6}\|f'''\|_\infty\cdot\text{const}\left(\sup_{y,x}\frac{|\phi(y,x)|}{x}\right)^3\sum_{k,j}((j\varepsilon)^3\wedge 1)\mu(y_k,j\varepsilon)$$
$$\to \frac{1}{2}f''(\langle\mu,\phi\rangle)\langle\mu,\phi^2\rangle.$$

By Lemma 3.2.4 below, we obtain bounds on the third moments uniform in ε. Using this and the tightness it follows that any limit point of the laws of the family $\{\eta_I^\varepsilon\}$ satisfies the martingale problem associated to the generator $\mathfrak{G}$. Finally a standard argument (e.g. [DP2], proof of Theorem 1.1) shows that any solution of this martingale problem has Laplace functional given by

$$\mathbb{E}_{\mu_0}\exp\left(-\int f(u,x)\xi_I(t,dy,dx)\right) \tag{3.2.5}$$
$$= \exp\Big(-\int V_t f(y,x)\mu_0(dy,dx) - c\int_0^t\int_I \frac{\partial}{\partial x}V_{t-s}f(y,x)\Big|_{x=0}\alpha(dy)\,ds\Big),$$

where $V_t f$ is given by the unique solution of the nonlinear p.d.e.

$$\begin{aligned}\frac{\partial}{\partial t}V_t f &= \frac{1}{2}x\frac{\partial^2}{\partial x^2}V_t f - cx\frac{\partial}{\partial x}V_t f - \frac{1}{2}(V_t f)^2,\\ V_0 f &= f \in C_0^+(I\times(0,\infty)).\end{aligned} \tag{3.2.6}$$

Therefore there is a unique limit point and the proof is complete. ■

For fixed $c>0$, $a>0$ and arbitrary atomless measure α on $I=[0,1]$ with total mass a, let $\xi_I(t,dy,dx)$ be as in Proposition 3.2.1. Consider the marginal process

$$\xi(t,dx) := \int_I \xi_I(t,dy,dx). \tag{3.2.7}$$

We call $\xi(t)$ a *super subcritical Feller branching diffusion with individual immigration* (SSFBI) with initial state $\mu(dx) = \int_I \mu_0(dy, dx)$ (and parameters a, c). The expression for the Laplace functional (3.2.5) (with f only a function of x) shows that this coincides with the so-called *immigration process with immigration rate* ac *corresponding to the entrance law* κ_t (given by (6.2.12) in the Appendix) and starting from zero measure at time 0. The existence of a superprocess with immigration corresponding to an entrance law was first established by [LS] (Thm. 1.1). The resulting Laplace transform of $\xi(t)$ with zero initial measure is given by

$$\mathbb{E}\exp\left(-\langle \xi(t), f\rangle\right) = \exp\left(-ac\int_0^t \kappa_{0+}(V_s f)ds\right), \quad f \in C_1(0,\infty), \tag{3.2.8}$$

see (3.1.18).

Corollary 3.2.2 *(a) The random measure*

$$\beta(t, dy) := \int_{(0,\infty)} x\xi_I(t, dy, dx) = \sum_k b_k(t)\delta_{y_k}$$

is a purely atomic finite random measure on I *in which the atoms* $b_k(t)\delta_{y_k}$ *correspond to the aggregated mass at time* t *coming from immigrants of family type* $y_k \in I$.
(b) For the corresponding family of stationary processes $\{\bar{\eta}_I^{\varepsilon}(t)\}_{t\in\mathbb{R}}$*, the random measures* $\bar{\eta}_I^{\varepsilon}(0)$ *converge to the equilibrium for the process with generator* $\mathfrak{G}$ *given by (3.2.2). The equilibrium random measure,* ξ^a *has Laplace functional*

$$\mathbb{E}\exp\left(-\langle \xi^a, f\rangle\right) = \exp\left(-ac\int_0^\infty \kappa_{0+}(V_s f)ds\right), \quad f \in C_1(0,\infty). \tag{3.2.9}$$

Proof. (a) The random measure $\beta(t, dy)$ on I has independent increments and no fixed atoms and is therefore purely atomic (see [K], Chapt. 7).
(b) Given $t_0 < 0$, $\eta_I^{\varepsilon}(0)$ can be decomposed into two parts - one coming from the initial value at t_0 and one from the immigration in the interval $(t_0, 0)$. From Proposition 3.2.1 it follows that the immigration parts converge (in the sense of weak convergence of probability measures on the space of càdlàg functions $D([t_0, 0], M^1(I\times(0,\infty)))$ to the diffusion limit with immigration, that is the process with generator $\mathfrak{G}$. Next note that for each $\varepsilon > 0$ the contribution to the aggregated measure at time 0 from the state at time t_0 is stochastically decreasing to zero due to the subcriticality and the contribution from immigration on $(t_0, 0)$ is stochastically increasing as $t_0 \downarrow -\infty$. Moreover, using the moment bounds from Lemma 3.2.4, it follows that the family of random measures $\bar{\eta}_I^{\varepsilon}(0)$ is tight. Therefore we have convergence to $\bar{\xi}_I(0)$, the equilibrium state for the process with generator $\mathfrak{G}$. The representation for the Laplace functional of ξ^a follows by letting $t \to \infty$ in (3.2.8). ■

Remark 3.2.3 *1. Corollary 3.2.2 (a) implies that for all $\delta > 0$, asymptotically as $\varepsilon \to 0$ only a finite number of immigrant families contribute all but δ of the mass. Each atom corresponds to an excursion from zero for the SSFBI process and consists of descendants of only one immigrant family. In fact, we will see that asymptotically at the particle level each immigrant family corresponds to the descendants of one immigrating particle.*

2. Note that the assumption $\eta_I^\varepsilon(t_0) \Rightarrow \mu \in M^1(I \times (0,\infty))$ in Proposition 3.2.1 puts constraints not only on the aggregated mass but also on the family structure of the population. To understand what happens if this condition is not satisfied consider $\eta_I^\varepsilon(t_0) = \sum a_k^\varepsilon \delta_{x_k^\varepsilon}$ with $\sum_k a_k^\varepsilon x_k^\varepsilon = \theta$ but $\inf_k x_k^\varepsilon \to \infty$ as $\varepsilon \to 0$. In this case at times $t > t_0$, $\eta_I^\varepsilon(t) \to 0$ due to ultimate extinction of the critical family level branching. This observation is used below to prove by contradiction that the equilibrium populations in $B_\ell^{(N)}$ are asymptotically composed of families of size $O(N^{\ell/2})$.

3. Similarly, if the immigration mechanism is such that it feeds a few large families rather than giving small new families a chance, then in the time stationary process the family branching makes everything extinct as $\varepsilon \to 0$.

3.2.2 Moments

The following lemma was used in the proof of Corollary 3.2.2 and will also be needed below. For the ease of notation we put $t_0 = 0$, otherwise we would have to replace t by $t - t_0$.

Lemma 3.2.4 *For $t > 0$, let $\eta^\varepsilon(t,dx) := \int_I \eta_I^\varepsilon(t,dy,dx)$, where η_I^ε is as in (3.2.1). Let*

$$m_{j,k}(t) = \mathbb{E}\left[\langle \eta^\varepsilon(t), x^j\rangle^k\right],\ j,k \in \mathbb{N},$$

$$M(t) = \mathbb{E}(\langle \eta^\varepsilon(t), x\rangle \langle \eta^\varepsilon(t), x^2\rangle),$$

and $o_\varepsilon(1)$ denote a term that is uniformly bounded in ε and converges to 0 as $\varepsilon \to 0$, and $\tilde{o}(t)$ denote a term that is uniformly bounded in $t \geq 0$, and $\tilde{o}(t)$ converges exponentially fast to 0 as $t \to \infty$, and $|\tilde{o}(t)| \leq \mathrm{const} \cdot t$ for small $t > 0$. Then
(a)

$$m_{1,1}(t) = a(1 - e^{-ct}) + e^{-ct} m_{1,1}(0),$$

$$\begin{aligned} m_{2,1}(t) = &\frac{a}{2c}[1 - 2e^{-ct} + e^{-2ct}] + m_{2,1}(0)e^{-2ct} + \frac{m_{1,1}(0)}{c}(e^{-ct} - e^{-2ct}) \\ &+ \frac{\varepsilon a}{2}(1 - e^{-2ct}), \end{aligned}$$

$$m_{3,1}(t) = \frac{a}{2c^2} + o_\varepsilon(1) + \tilde{o}(t),$$

$$m_{4,1}(t) = \frac{3a}{4c^3} + o_\varepsilon(1) + \tilde{o}(t),$$

(b)

$$\begin{aligned}
&m_{1,2}(t) = m_{1,2}(0)e^{-2ct} \\
&+\frac{a}{4c^2}\left\{1 - 4e^{-ct} + 2cte^{-ct} + 3e^{-2ct}\right\} \\
&+a^2\left\{1 - 2e^{-ct} + e^{-2ct}\right\} \\
&+\frac{m_{1,1}(0)}{c^2}\left\{e^{-ct} - cte^{-ct} + 2ac^2e^{-ct} - e^{-2ct} - 2ac^2e^{-2ct}\right\} \\
&+\frac{m_{2,1}(0)}{c}(te^{-2ct}) + \frac{\varepsilon^2}{2}\left\{a(1 - e^{-2ct})\right\} \\
&+\frac{\varepsilon}{4c}(3a - 2e^{-ct} + 2ate^{-2ct} + 4e^{-ct}(-a + m_{1,1}(0)) + ae^{-2ct} - 4m_{1,1}(0)e^{-2ct}),
\end{aligned}$$

$$m_{2,2}(t) = \frac{3a}{16c^4} + o_\varepsilon(1) + \tilde{o}(t),$$

$$M(t) = \frac{a}{4c^3} + \frac{a^2}{2c} + o_\varepsilon(1) + \tilde{o}(t),$$

$$m_{1,3}(t) = \frac{3a^2}{4c^2} + a^3 + \frac{a}{12c^4} + o_\varepsilon(1) + \tilde{o}(t).$$

Proof. The proof is obtained by applying the martingale problem with the generator given by (3.2.3) and (3.2.4) to functions of the form $F(\mu) = f(\langle\mu,\phi\rangle)$ or $F(\mu) = f(\langle\mu,\phi_1\rangle, \langle\mu,\phi_2\rangle)$ to derive the following moment equations:

$$\frac{dm_{1,1}(t)}{dt} = ca - cm_{1,1}(t),$$

$$\frac{dm_{2,1}(t)}{dt} = m_{1,1}(t) - 2cm_{2,1}(t) + ca\varepsilon,$$

$$\frac{dm_{3,1}(t)}{dt} = ca\varepsilon^2 + 3m_{2,1}(t) - 3cm_{3,1}(t) + o_1(\varepsilon),$$

$$\frac{dm_{4,1}(t)}{dt} = ca\varepsilon^3 + 6m_{3,1}(t) - 4cm_{4,1}(t) + o_1(\varepsilon),$$

$$\frac{dm_{1,2}(t)}{dt} = m_{2,1}(t) - 2cm_{1,2}(t) + (2ca + \varepsilon)m_{1,1}(t) + \varepsilon^2 ca,$$

$$\frac{dM(t)}{dt} = cam_{2,1}(t) + m_{3,1}(t) + m_{1,2}(t) - 3cM(t) + o_1(\varepsilon),$$

$$\frac{dm_{2,2}(t)}{dt} = m_{4,1}(t) - 4cm_{2,2}(t) + o_1(\varepsilon),$$

$$\frac{dm_{1,3}(t)}{dt} = 3cam_{1,2}(t) + 3M(t) - 3cm_{1,3}(t) + o_1(\varepsilon).$$

Note that the coefficients of the $o_1(\varepsilon)$ terms only contain moments lower in the hierarchy and hence are asymptotically negligible. The results were obtained by solving this linear system using MAPLE. ■

Remark 3.2.5 *In the case in which we replace the constant immigration rate a by a random function of time $a(\cdot)$ the expression for $m_{1,2}(t)$ becomes*

$$\begin{aligned} m_{1,2}(t) = \quad & m_{1,2}(0)e^{-2ct} \\ & + \frac{m_{1,1}(0)}{c^2}\left\{e^{-ct} - cte^{-ct} + 2c^2ae^{-ct} - e^{-2ct} - 2ac^2e^{-2ct}\right\} \\ & + \frac{1}{c^2}\int_0^t k_1(t,s)a(s)ds + \int_0^t\int_0^{s_2} k_2(t,s_2,s_1)a(s_1)a(s_2)ds_1ds_2 \\ & + o_\varepsilon(1)\cdot\int_0^t k_3(t,s)a(s)ds, \end{aligned} \tag{3.2.10}$$

where $k_i(t,\cdot)$, $i=1,2,3$ are bounded non-negative kernels satisfying

$$\sup_t \int_0^t k_i(t,s)ds < \infty,\ i=1,3,\ \ \sup_t \int_0^t\int_0^{s_2} k_2(t,s_2,s_1)ds_1ds_2 < \infty \tag{3.2.11}$$

and $o_\varepsilon(1) \to 0$ as $\varepsilon \to 0$.

3.2.3 SPDE representation

Let ξ be an SSFBI process starting from zero measure at time 0 as in subsection 3.2.1; recall that the Laplace transform of $\xi(t)$ is given by (3.2.8). By an argument similar to that of [LS] (Thm. 1.2) it follows that there is a unique orthogonal martingale measure $M(ds\,dx)$ on $[0,\infty)\times(0,\infty)$ having quadratic variation measure $\langle M\rangle(ds\,dx) = ds\,\xi(s,dx)$ such that

$$\begin{aligned} & \langle\xi(t),f\rangle - \langle\xi(0),f\rangle \\ & = \int_0^t [\langle\xi(s),G_cf\rangle + ac\,\kappa_{0+}(f)]ds + \int_0^t\int_{(0,\infty)} f(x)M(ds\,dx), \quad f\in D_\rho(G_c). \end{aligned} \tag{3.2.12}$$

Proceeding as in the proof of Theorem 1.7 of [LS] and tracing the arguments of [KS] one infers that $\xi(t)$ has absolutely continuous states, that is, $\xi(t,dx) = \xi(t,x)dx$, and that one can define a time-space white noise $\dot{W}_t(x)$ on an extension of the original probability space such that

$$M(ds\,dx) = \sqrt{\xi(s,x)}\,\dot{W}_s(x)ds\,dx. \tag{3.2.13}$$

Moreover, $\xi(t,x)$ is almost surely jointly continuous in $(t,x) \in [0,\infty) \times (0,\infty)$ and is a solution of the SPDE (3.2.18) below. Note however that in contrast to [LS], $\xi(t,x)$ does not have a finite limit as $x \downarrow 0$. Indeed, putting

$$Z_t(x) = \int_0^t \int_0^\infty p_{t-s}(y,x) M(ds,dy), \tag{3.2.14}$$

where p is the transition density of the c-FBD process (see section 6.2 in the Appendix), we obtain as in [LS] ((4.10), (4.11)):

$$\xi(t,x) = Z_t(x) + \int_0^t ac\kappa_s(x)ds, \quad x > 0. \tag{3.2.15}$$

From (6.2.16) we obtain

$$\int_0^t ac\kappa_s(x)ds = \frac{2ac}{x} \exp\left(\frac{-2cx}{1-e^{-ct}}\right). \tag{3.2.16}$$

Thus, $\mathbb{E}(\int_\varepsilon^1 \xi(t,x)dx) = C(t,\varepsilon)|\log\varepsilon|$ where $C(t,\varepsilon)$ is uniformly bounded away from 0, and

$$\begin{aligned} &\mathrm{Var}\left(\int_\varepsilon^1 \xi(t,x)dx\right) \\ &= \mathbb{E}\left[\int_0^t \int_0^\infty f(s,y)M(ds,dy)\right]^2 \\ &= \mathbb{E}\left[\int_0^t \int_0^\infty f(s,y)^2 \xi(s,y)dy\right] \end{aligned}$$

where

$$f(s,y) = \int_\varepsilon^1 p_{t-s}(y,x)dx \le \frac{y}{\varepsilon} \wedge 1, \qquad \text{and } \mathbb{E}\xi(s,y) \le \frac{ac}{y} \exp\left(\frac{-2cy}{1-e^{-cs}}\right),$$

where we have used a stochastic Fubini theorem (cf. [IW] Chapt. 3, Lemma 4.1).

Therefore

$$\mathrm{Var}\left(\int_\varepsilon^1 \xi(t,x)dx\right) \le C_1(t) + C_2(t)|\log\varepsilon|$$

for some positive constants $C_1(t), C_2(t)$. We then obtain for any $\delta > 0$

$$P\left(\left|\int_\varepsilon^1 \xi(t,x)dx - C(t)|\log\varepsilon|\right| > \delta C(t)|\log\varepsilon|\right) \le \frac{C_1(t) + C_2(t)|\log\varepsilon|}{(\delta C(t)|\log\varepsilon|)^2} \tag{3.2.17}$$

and therefore $\int_\varepsilon^1 \xi(t,x)dx$ converges in probability to ∞ as $\varepsilon \to 0$. Since $\int_\varepsilon^1 \xi(t,x)dx$ is monotone in ε this convergence must be a.s. , and thus $\xi(t,x)$ is a.s. unbounded as $x \to 0$.

Finally, recalling (3.1.13) and (3.2.13), we see that (3.2.12) is the integral form of the SPDE

$$\frac{\partial}{\partial t}\xi(t,x) = \sqrt{\xi(t,x)}\,\dot{W}_t(x) + G_c^*\xi(t,x) - ca\delta_0'(x). \tag{3.2.18}$$

Let us comment on the meaning of the three terms on the right hand side of (3.2.18) viewed as the limiting family size process as $N \to \infty$ in a ball $B_\ell^{(N)}$. The first one comes from the family branching, the second one incorporates the individual branching and individual emigration at rate $c = c_\ell$ from $B_\ell^{(N)}$ (recall that G_c is the generator of a c-subcritical Feller branching diffusion), and the third term describes immigration of small families into $B_\ell^{(N)}$ from the surrounding medium at a large rate. In fact, δ_0' can be viewed as the limit as $\varepsilon \to 0$ of a large number $1/\varepsilon$ of small families of size ε; note that $(1/\varepsilon)\delta_\varepsilon$ converges to $-\delta_0'$ in the sense of Schwartz distributions on the smooth functions vanishing at 0.

Since any solution of (3.2.18) satisfies the martingale problem with generator (3.2.2), weak uniqueness follows from Proposition 3.2.1; however, as in the case of [KS] it is an open question whether strong uniqueness holds.

The total (or aggregated) population size

$$\zeta(t) = \int_{(0,\infty)} x\,\xi(t,x)\,dx \tag{3.2.19}$$

solves the equation

$$d\zeta(t) = \sqrt{\int_{(0,\infty)} x^2\xi(t,x)dx}\,dW_t - c(\zeta(t) - a)\,dt, \tag{3.2.20}$$

which is a one-dimensional projection of equation (3.2.18). Note that the process ζ is not Markov.

3.2.4 Equilibrium canonical moments

As $t \to \infty$, the SSFBI process $\xi(t)$ with parameters c and a converges in distribution to the infinitely divisible equilibrium random measure ξ^a, cf. Corollary 3.2.2. Writing

$$\zeta^a := \int_{(0,\infty)} x\xi^a(dx) \tag{3.2.21}$$

for the "aggregation" of ξ^a, and ν_c for the canonical measure of ζ^a, we obtain from (3.2.8) and (3.1.18):

$$\nu_c([b_1, b_2]) := c\int_0^\infty m([b_1, b_2])K_t(dm)dt, \quad 0 < b_1 < b_2, \tag{3.2.22}$$

According to Lemma 3.2.4, ζ^a has first and second moments

$$\mathbb{E}[\zeta^a] = a, \qquad \mathbb{E}[(\zeta^a)^2] = a\left(a + \frac{1}{4c^2}\right). \tag{3.2.23}$$

Definition 3.2.6 *Let us write $\hat{\nu}_c$ for the size-biasing of ν_c (cf. subsection 6.1).*

Because of the well-known relations (cf. Remark 4.5.2)

$$\mathbb{E}[\zeta^a] = \int_{(0,\infty)} x\nu_c(dx), \quad \mathbb{E}[(\zeta^a)^2] = \mathbb{E}[\zeta^a]\left(\mathbb{E}[\zeta^a] + \int_{(0,\infty)} x\hat{\nu}_c(dx)\right)$$

we obtain immediately from (3.2.23):

Remark 3.2.7 *a) $\int_{(0,\infty)} x\nu_c(dx) = 1$*
b) $\int_{(0,\infty)} x^2\nu_c(dx) = \frac{1}{4c^2}$.

Remark 3.2.8 *We also note that infinitely many immigrant families contribute to ζ^a. This follows from*

$$\nu_c(0,\infty) = \infty. \tag{3.2.24}$$

To see this note that from (3.2.22) and (3.1.12),

$$\nu_c(0,\infty) = c\int_0^\infty (-\log P(\xi_t^0 = 0))dt.$$

For $\delta > 0,\ x > 0$ let $f_\delta(x) = \frac{x}{\delta} \wedge 1$. Recalling (3.1.7) and (3.1.10) note that

$$P(\xi_t^0 = 0) = \lim_{\theta\to\infty}\lim_{\varepsilon\to 0} e^{V_t(\theta f_\delta)(\varepsilon)/\varepsilon}.$$

Then by a simple modification of [DH](6.10), for any $\delta > 0$ and $\theta > 1$

$$\lim_{\varepsilon\to 0}\frac{V_t(\theta f_\delta)(\varepsilon)}{\varepsilon} \geq e^{-t}\lim_{\varepsilon\to 0}\frac{T_t f_\delta(\varepsilon)}{\varepsilon} \geq e^{-t}\int_\delta^\infty \kappa_t(x)dx.$$

Therefore by (6.2.16)

$$\nu_c(0,\infty) \geq c\lim_{\delta\downarrow 0}\int_0^\infty e^{-t}(\int_\delta^\infty \kappa_t(x)dx)dt = \infty.$$

4 The genealogy of jumps in a cascade of subordinators

In this section we will carry out the program outlined in Subsection 2.1.1 to obtain a representation for the sequence $\{\zeta_\ell^\theta\}$ in terms of a cascade of subordinators and then use this representation to obtain a genealogical description of the population.

4.1 Propagation of equilibria

In the preceding section we encountered the equilibrium distribution for an $\mathbb{R}_+$-valued process $\zeta(t)$, whose dynamics is given by (3.2.20). A simpler situation is the one corresponding to one-level branching where we have the equilibrium, ζ^a of

$$d\zeta(t) = \sqrt{\zeta(t)}dW(t) - c(\zeta(t) - a)dt, \tag{4.1.1}$$

recall subsection 2.1. In this section we will derive structural results which are common to *both* situations. We therefore denote the equilibrium states of (4.1.1) and (3.2.20) by the same symbol ζ^a.

In both situations the dynamics has two parameters a and c, and the equilibrium distribution is infinitely divisible with expectation a. Therefore as in (2.1.12) this equilibrium distribution has a representation as

$$\mathcal{L}(\zeta^a) = \mathcal{L}(S(a)), \tag{4.1.2}$$

where $S(\tau), \tau \geq 0$, is a subordinator with $\mathbb{E}S(\tau) \equiv \tau$, $\tau \geq 0$.

Let us denote the Lévy measure of S by μ, and note that μ has expectation 1. Note that $\mu = \nu_c$ (defined in subsection 3.2.4) if ζ follows the dynamics (3.2.20), and $\mu = \gamma_c$ (given by (6.2.14)) if ζ follows the dynamics (4.1.1).

$S(a)$ has a Lévy-Khinchin representation as a Poissonian superposition

$$S(a) =^d \sum_{i:\tau_i \leq a} y_i, \tag{4.1.3}$$

where $\sum_i \delta_{(\tau_i, y_i)}$ is a Poisson population on $\mathbb{R}_+ \times (0, \infty)$ with intensity measure $d\tau\, \mu(dy)$. Since by (6.2.14), resp. (3.2.24),

$$\mu(0, \infty) = \infty \tag{4.1.4}$$

in the two cases, $P(S(a) = 0) = 0$ if $a > 0$ and $S(\cdot)$ has infinitely many jumps in any open interval.

Since ζ^a is a Poisson superposition of immigrant clusters (recall (3.2.22) and (6.2.13)), the representation (4.1.3) has a genealogical interpretation: the summands y_i measure the size of those parts of ζ^a which trace back to one and the same immigrant.

Our aim in this section is to study the hierarchy (2.1.1) into which (4.1.1) and (3.2.20) are embedded. In both situations, the parameters of the hierarchy are a sequence $(c_\ell)_{\ell=1,2,\ldots}$

of positive numbers. Recall that, for different levels ℓ in the hierarchy, the dynamics of $\zeta_\ell(t)$ run at separated time scales, and the equilibrium state $\zeta_{\ell+1}$ at level $\ell+1$ acts as (random) parameter a for the dynamics at level ℓ. Instead of one Lévy measure μ, we now have a sequence of Lévy measures (μ_ℓ) (which is either (γ_{c_ℓ}) or (ν_{c_ℓ})).

In this way we obtain a Markovian dynamics which transports the equilibria down the levels:

$$\text{given } \zeta_{\ell+1} = a\,, \zeta_\ell \text{ is infinitely divisible with canonical measure } a\mu_\ell. \tag{4.1.5}$$

Since μ has expectation 1, (ζ_ℓ) constitutes a backward martingale. We now turn to the following problems:

a) Find a condition on (μ_ℓ) which guarantees the existence of an entrance law, denoted by ζ_ℓ^θ, for (ζ_ℓ) starting in $\theta > 0$ "at level ∞" and having constant expectation θ.

b) Describe the "branching genealogy" underlying such an entrance law.

We will answer these questions in the next subsections. Later on, we will give a relation with the asymptotics of the genealogy of the equilibrium branching populations on Ω_N as $N \to \infty$.

4.2 An entrance law from infinity

Let S_k, $k = 1, 2, \ldots$ be independent subordinators with Lévy measures μ_k. We denote the second moment of μ_k by m_k. For $j > \ell$ define the random variables

$$S_\ell^j(a) := S_\ell(S_{\ell+1}(\ldots(S_{j-1}(a)))), \tag{4.2.1}$$

and write $\Pi_\ell(a, .)$ for the distribution of $S_\ell(a)$, $\ell = 1, 2, ...$, $a > 0$.

Proposition 4.2.1 *If*

$$\sum_{k=1}^\infty m_k < \infty \tag{4.2.2}$$

then, for each $\theta > 0$, the sequence of processes

$$(S_{j-1}^j(\theta),\, S_{j-2}^j(\theta), ...,\, S_2^j(\theta),\, S_1^j(\theta))$$

converges as $j \to \infty$ (component-wise) in probability to a sequence

$$(...,\zeta_2^\theta,\, \zeta_1^\theta)$$

which obeys

$$S_\ell(\zeta_{\ell+1}^\theta) = \zeta_\ell^\theta \ \textit{a.s. for all } \ell \tag{4.2.3}$$

and

$$\lim_{\ell\to\infty} \zeta_\ell^\theta = \theta \ \textit{a.s.} \tag{4.2.4}$$

In particular, the distributions $\pi_\ell = \mathcal{L}(\zeta_\ell^\theta)$ are an entrance law for the backward Markov chain with probability transition function

$$P(\zeta_\ell \in A | \zeta_{\ell+1} = a) = \Pi_\ell(a, A), \tag{4.2.5}$$

and they are its unique entrance law with the property

$$\pi_\ell \Rightarrow \delta_\theta \text{ as } \ell \to \infty. \tag{4.2.6}$$

Proof. Since $S_k(a)$ is infinitely divisible with canonical measure $a\mu$, we have

$$\mathrm{Var}(S_k(a)) = am_k.$$

Hence we obtain

$$\begin{aligned}\mathrm{Var}(S_k^{k+2}(a)) &= \mathrm{Var}[\mathbb{E}[S_k(S_{k+1}(a))|S_{k+1}(a)]] + \mathbb{E}[\mathrm{Var}[S_k(S_{k+1}(a))|S_{k+1}(a)]] \\ &= a(m_{k+1} + m_k).\end{aligned}$$

In the same way we get for all $j > k > \ell$:

$$\mathrm{Var}(S_k^j(a)) = a(m_{j-1} + ... + m_k) \tag{4.2.7}$$

and

$$\mathbb{E}(S_\ell^j(a) - S_\ell^k(a))^2 = a(m_{j-1} + ... + m_k).$$

Thus, due to (4.2.2), for fixed ℓ the sequence $(S_\ell^j(\theta))_{j>\ell}$ is Cauchy in L^2. We define

$$\zeta_\ell^\theta \equiv L^2 - \lim_{j\to\infty} S_\ell^j(\theta). \tag{4.2.8}$$

Since $a \mapsto S_\ell(a)$ is continuous in L^1, we have

$$S_\ell(\zeta_{\ell+1}^\theta) = S_\ell(\lim_{j\to\infty} S_{\ell+1}^j(\theta)) = \lim_{j\to\infty} S_\ell^j(\theta) = \zeta_\ell^\theta \text{ a.s. }, \tag{4.2.9}$$

which proves (4.2.3) and, a fortiori, implies (4.2.5). From (4.2.8) and (4.2.7) it is clear that

$$\mathrm{Var}\,\zeta_\ell^\theta = \theta \sum_{k=\ell}^\infty m_k.$$

Since $\mathbb{E}\zeta_\ell^\theta \equiv \theta$, this together with (4.2.2) implies that ζ_ℓ^θ converges to θ in probability as $\ell \to \infty$. Moreover, since because of (4.2.3) ζ_ℓ^θ is a backwards martingale, this convergence is even a.s., and we have (4.2.4).

It remains to show the claimed uniqueness statement. For this let (π_k) be an entrance law for (Π_k) obeying (4.2.6), and let X_k, $k = 1, 2, ...$ be random variables, independent of the subordinators S_ℓ, with

$$\mathcal{L}(X_j) = \pi_j.$$

From the entrance law property of (π_k) and the definition of (Π_k) we have for all $j > \ell$:

$$\pi_\ell = \mathcal{L}(S_\ell^j(X_j)). \tag{4.2.10}$$

On the other hand we have by monotonicity of $\tau \mapsto S_\ell^j(\tau)$:

$$\mathbb{E}|S_\ell^j(X_j) - S_\ell^j(\theta)| \le \mathbb{E}|X_j - \theta|.$$

From this, the claimed identity $\pi_\ell = \mathcal{L}(\zeta_\ell^\theta)$ follows by Markov's inequality together with (4.2.8) and (4.2.10). ■

Proof of Proposition 2.1.1

Recall that with the notation of Proposition 2.1.1, $\Pi_\ell^{(1)}(a,.) = \mathcal{L}(S_\ell^{(1)}(a))$ and $\Pi_\ell^{(2)}(a,.) = \mathcal{L}(S_\ell^{(2)}(a))$, where $S_\ell^{(1)}$ and $S_\ell^{(2)}$ are subordinators with Lévy measures γ_ℓ and ν_ℓ, respectively. From Remark 6.2.3 we have that the second moment of γ_k equals $1/(2c_k)$, and Remark 3.2.7 shows that the second moment of ν_k equals $1/(4c_k^2)$. The proof of Proposition 2.1.1 is thus immediate from Proposition 4.2.1.

4.3 The genealogy of jumps in an iteration of subordinators

The composition of subordinators gives rise to a "genealogy" of their jumps. To illustrate this, consider the two subordinators S_1, S_2, where

$$S_1(b) = \sum_{t_i \le b} (S_1(t_i) - S_1(t_i-)),$$

$$S_2(a) = \sum_{\tau_n \le a} (S_2(\tau_n) - S_2(\tau_n-)).$$

Then

$$S_1(S_2(a)) = \sum_{\tau_n \le a} \quad \sum_{S_2(\tau_n-) < t_i \le S_2(\tau_n)} (S_1(t_i) - S_1(t_i-)).$$

In this way, the jumps of S_1 are coagulated into families of jumps stemming from one and the same jump of S_2.

Iterating this, we obtain from the flow property (4.2.3) that

$$\zeta_\ell^\theta = \sum_{\tau_i \in [0, \zeta_j^\theta]} (S_\ell^j(\tau_i) - S_\ell^j(\tau_i-)) \text{ a.s. }, \tag{4.3.1}$$

where $\{\tau_i\}$ is the set of all points in $[0, \zeta_j^\theta]$ in which $\tau \mapsto S_\ell^j(\tau)$ has a jump. The representation (4.3.1) induces a partition of $[0, \zeta_\ell^\theta]$ which we denote by $\mathfrak{P}_{j,\ell}$. Note that for fixed ℓ the $\mathfrak{P}_{j,\ell}$ are coalescing (i.e. becoming coarser) as j increases.

The sequence of coalescing partitions $\mathfrak{P}_{j,\ell}$ induces a graph $\mathfrak{G}_\ell$ as follows: The set of nodes of $\mathfrak{G}_\ell$ is the union $\bigcup_{j\geq\ell}\{j\} \times \mathfrak{P}_{j,\ell}$. For $n \in \mathfrak{G}_\ell$ we call its first component the *level* of n. For two nodes $n_1 = (j_1, I_1)$, $n_2 = (j_2, I_2)$ of $\mathfrak{G}_\ell$ we say that n_1 is an *ancestor* of n_2 if $j_1 > j_2$ and $I_2 \subseteq I_1$, and we say that n_1 is the *parent* of n_2 if n_1 is the ancestor of n_2 with $j_1 = j_2 + 1$. The (directed) edges of $\mathfrak{G}_\ell$ then are all the parent-child pairs in $\mathfrak{G}_\ell \times \mathfrak{G}_\ell$. Say that two nodes in $\mathfrak{G}_\ell$ are *related* if they have a common ancestor. Then by construction of the sequence $(\mathfrak{P}_{j,\ell})$ each equivalence class of $\mathfrak{G}_\ell$ is a *tree*, i.e. a directed connected graph without cycles. Therefore, $\mathfrak{G}_\ell$ is a *forest*, i.e. a union of pairwise disconnected trees. Finally, we label each node of $\mathfrak{G}_\ell$ with the length of the subinterval of $[0, \zeta_\ell^\theta]$ to which it corresponds, thus arriving at the random *labelled forest* $\mathfrak{F}_\ell$ which we associate with the random sequence of coalescing partitions $\mathfrak{P}_{j,\ell}$, and which encodes the genealogy of the jumps of the process $(..., \zeta_2^\theta, \zeta_1^\theta)$ constructed in Proposition 4.2.1.

Intuitively, viewing $[0, \zeta_\ell^\theta]$ as a continuum of individuals, this means that two individuals a_1, $a_2 \in [0, \zeta_\ell^\theta]$ belong to the same element of $\mathfrak{P}_{j,\ell}$ if and only if they descend from a common ancestor (or equivalently, from one subordinator jump) at some level less or equal than j. Furthermore two individuals a_1, $a_2 \in [0, \zeta_\ell^\theta]$ belong to the same element of the minimal partition $\mathfrak{P}_{\infty,\ell}$ if and only if they descend from a common ancestor at any level higher than ℓ. Using the independent increments property in θ, (4.1.4) and (4.2.9) it can be shown that there are countably many distinct elements in $\mathfrak{P}_{\infty,\ell}$ each corresponding to an infinite tree. Therefore we have a decomposition of the equilibrium population into a countable set of subpopulations each consisting of individuals having a common ancestor.

4.4 The genealogy in the hierarchichal mean field limit

With the special choice of (S_ℓ) described at the beginning of this section, we have all reasons to conjecture that the random labelled forests $\mathfrak{F}_\ell$ defined in the previous subsection describe the genealogy of the (one- or two-level) branching population in equilibrium as $N \to \infty$.

To make this more precise, consider a fixed sequence $B_\ell^{(N)}$, $\ell = 1, 2, ...$ of nested balls in Ω_N, and let $\mathcal{P}_\ell^{(N)}$ be that part of the equilibrium population which lives in $B_\ell^{(N)}$. (Here and below we suppress the notation of $\theta > 0$ which we keep fixed.) Fix N and ℓ for the moment. For two individuals I_1, I_2 in $\mathcal{P}_\ell^{(N)}$ and $j \geq \ell$ we say that

$$I_1 \sim_j I_2$$

if I_1 and I_2 have a common ancestor in $B_j^{(N)}$. This induces a partition on $\mathcal{P}_\ell^{(N)}$ which we denote by $\mathfrak{P}_{j,\ell}^{(N)}$. The sequence $\mathfrak{P}_{j,\ell}^{(N)}$, $j = \ell, \ell+1, ...$ is coalescing, and we can associate with it a labelled forest $\mathfrak{F}_\ell^{(N)}$ in the same way as we associated $\mathfrak{F}_\ell$ with $\mathfrak{P}_{j,\ell}$ in the previous subsection, the only difference being that now we label the nodes of $\mathfrak{F}_\ell^{(N)}$ by the cardinalities of their corresponding sub-populations of $\mathcal{P}_\ell^{(N)}$, divided by N^ℓ.

Our main result (Theorem 2.4.1) suggests to conjecture that, in a suitable topology,

$$\mathfrak{F}_\ell^{(N)} \to \mathfrak{F}_\ell \text{ as } N \to \infty.$$

4.5 Size-biasing iterated subordinators

With a view towards the genealogy of a sampled individual (see subsection 4.6) we will now prove a representation of the size-biasing of $\mathcal{L}(S_\ell^j(a))$, where S_ℓ^j is the composition of subordinators defined in (4.2.1). To this purpose we first consider a single subordinator S evaluated at a random argument.

Proposition 4.5.1 *Let $S(\tau), \tau \geq 0$, be a subordinator with Lévy measure denoted by μ, and let A be an $\mathbb{R}_+$-valued random variable independent of S and with finite expectation. Then the size-biasing of $\mathcal{L}(S(A))$ arises as the distribution of $S(\hat{A}) + \hat{Y}$, where $\mathcal{L}(\hat{A})$ is the size-biasing of $\mathcal{L}(A)$, $\mathcal{L}(\hat{Y})$ is the size-biasing of μ, and $\hat{A}$ and $\hat{Y}$ are independent.*

Proof. We write $\Pi_{\tau\mu}$ for the distribution of a Poisson point configuration on $\mathbb{R}_+$ with intensity measure $\tau\mu$. Then

$$\mathcal{L}(S(\tau)) = \mathcal{L}_\tau(\langle \Psi, \mathrm{id}_{\mathbb{R}_+}\rangle),$$

where

$$\mathcal{L}_\tau(\Psi) = \Pi_{\tau\mu}.$$

Writing σ for the distribution of A, we thus have

$$\mathcal{L}(S(A)) = \mathcal{L}_\sigma(\langle \Psi, \mathrm{id}_{\mathbb{R}_+}\rangle), \tag{4.5.1}$$

where

$$\mathcal{L}_\sigma(\Psi) = \int \Pi_{\tau\mu}(.)\sigma(d\tau).$$

Our task is to compute the size-biasing of $\mathcal{L}_\sigma(\langle \Psi, \mathrm{id}_{\mathbb{R}_+}\rangle)$ with $\mathfrak{s}(\psi) = \langle \psi, \mathrm{id}_{\mathbb{R}_+}\rangle$, cf. Definition 6.1.1. To this end let us first compute the size-biasing of $\mathcal{L}_\sigma(\Psi)$ with $\langle \psi, \mathrm{id}_{\mathbb{R}_+}\rangle$, and then project. It follows from Corollary 6.1.4 in the Appendix that the size-biasing of $\mathcal{L}_\sigma(\Psi)$ with $\langle \psi, \mathrm{id}_{\mathbb{R}_+}\rangle$ is $\mathcal{L}(\Phi + \delta_{\hat{Y}})$, where

$$\mathcal{L}(\Phi) = \int \Pi_{\tau\mu}(.)\hat{\sigma}(d\tau),$$

$\hat{\sigma}$ is the size-biasing of σ, $\mathcal{L}(\hat{Y})$ is the size-biasing of μ, and Φ and $\hat{Y}$ are independent. Consequently, the size-biasing of $\mathcal{L}_\sigma(\langle \Psi, \mathrm{id}_{\mathbb{R}_+}\rangle)$ with $\langle \psi, \mathrm{id}_{\mathbb{R}_+}\rangle$ is

$$\mathcal{L}(\langle \Phi + \delta_{\hat{Y}}, \mathrm{id}_{\mathbb{R}_+}\rangle) = \mathcal{L}(S(\hat{A}) + \hat{Y})),$$

where $\mathcal{L}(\hat{A}) = \hat{\sigma}$, and S, $\hat{A}$ and $\hat{Y}$ are independent. Together with (4.5.1) this proves the claim. ∎

Remark 4.5.2 *For deterministic A, Proposition 4.5.1 renders the well-known fact that the size-biasing of an infinitely divisible distribution π on $\mathbb{R}_+$ is the convolution of π with the size-biasing of the canonical measure of π.*

Corollary 4.5.3 *Let $S_\ell^j(a)$ be the iteration of subordinators defined in (4.2.1), where the S_k are independent subordinators with Lévy measures μ_k. Then the size-biasing of $\mathcal{L}(S_\ell^j(a))$ arises as the distribution of $\widehat{S}_\ell^j(a)$ defined by*

$$\widehat{S}_\ell^j(a) = S_\ell^j(a) + \widetilde{S}_\ell^{j-1}(\widehat{Y}_{j-1}) + \widetilde{S}_\ell^{j-2}(\widehat{Y}_{j-2}) + ... + \widetilde{S}_\ell^{\ell+1}(\widehat{Y}_{\ell+1}) + \widehat{Y}_\ell, \tag{4.5.2}$$

where $\mathcal{L}(\hat{Y}_k)$ is the size-biasing of μ_k, $\widetilde{S}_\ell^k$ is distributed as S_ℓ^k and all random variables occurring on the r.h.s. of (4.5.2) are independent.

Proof. For $\ell = j-1$, Proposition 4.5.1 shows that the size-biasing of $\mathcal{L}(S_{j-1}^j(a)) = \mathcal{L}(S_{j-1}(a))$ arises as the distribution of

$$S_{j-1}(a) + \hat{Y}_{j-1},$$

where both summands are independent. One more application of Proposition 4.5.1 thus gives that the size-biasing of $\mathcal{L}(S_{j-2}^j(a)) = \mathcal{L}(S_{j-2}(S_{j-1}^j(a))$ arises as the distribution of

$$S_{j-2}(S_{j-1}(a) + \hat{Y}_{j-1}) + \hat{Y}_{j-2}$$

which due to the independence of $S_{j-1}(a)$ and $\hat{Y}_{j-1}$ equals in distribution to

$$S_{j-2}(S_{j-1}(a)) + S'_{j-2}(\hat{Y}_{j-1}) + \hat{Y}_{j-2},$$

S'_{j-2} being an independent copy of S_{j-2}.

Iterating the argument we arrive at our assertion. ∎

Remark 4.5.4 *As before, let us denote the second moment of μ_k (or equivalently the first moment of $\hat{Y}_k$) by m_k. From (4.5.2) it follows that*

$$\mathbb{E}\widehat{S}_\ell^j(a) = a + \sum_{k=\ell}^{j-1} m_k. \tag{4.5.3}$$

Hence the summability of the m_k is a sufficient condition for tightness of the $\widehat{S}_\ell^j(a)$.

We now turn to the entrance law constructed in Proposition 4.2.1.

Corollary 4.5.5 *From (4.2.8) and Corollary 4.5.3 we obtain that the size-biasing of $\mathcal{L}(\zeta_\ell^\theta)$ arises as the distribution of*

$$\hat{\zeta}_\ell^\theta \equiv \widehat{Y}_\ell + \widetilde{S}_\ell^{\ell+1}(\widehat{Y}_{\ell+1}) + \widetilde{S}_\ell^{\ell+2}(\widehat{Y}_{\ell+2}) + ... + \zeta_\ell^\theta, \tag{4.5.4}$$

where the random variables $\widehat{Y}_k$ and $\widetilde{S}_\ell^k$ are as in Corollary 4.5.2, and all random variables occurring on the right hand side of (4.5.4) are independent.

We can go one step further and study the genealogical relationships underlying the representation (4.5.4).

To this purpose, let us study the branching dynamics on the "populations of jumps" induced by the composition of the subordinators S_k, resuming the reasoning of subsection 4.3. For each k we consider a branching dynamics which takes a counting measure ϕ_k on $\mathbb{R}_+$ into a random counting measure Φ_{k-1} in the following way: If $\phi_k = \sum_{i\in I_k} \delta_{y_i}$, then

$$\Phi_{k-1} = \sum_{i\in I_k} \Psi_i,$$

where Ψ_i is a Poisson counting measure on $\mathbb{R}_+$ with intensity measure $y_i\mu_{k-1}$, and the Ψ_i are independent.

Now fix two levels $j, \ell \in \mathbb{N}$ with $j > \ell$. Starting with $\phi_j = \delta_a$, and iterating the branching dynamics from level j down to level ℓ, we construct a random path of counting measures on $\mathbb{R}_+$ which we denote by

$$(\delta_a, \Phi_{j-1}^j(a), \dots, \Phi_\ell^j(a)) =: H_\ell^j(a).$$

By keeping track which atom in $\Phi_{k-1}^j(a)$ stems from which atom in $\Phi_k^j(a)$, we can enrich the history $H_\ell^j(a)$ to a tree $T_\ell^j(a)$, each of whose nodes is marked by a non-negative real number. For example, if $\Phi_k^j(a) = \sum_{i\in I_k} \delta_{y_i}$, $j \ge k \ge \ell$, then the set of nodes of $T_\ell^j(a)$ at level k corresponds to the index set I_k, and y_i is the mark (or "size") of the node with index i. We write

$$\mathfrak{s}(T_\ell^j(a)) = \langle \Phi_\ell^j(a), \mathrm{id}_{\mathbb{R}_+} \rangle$$

for the *total size of the tree $T_\ell^j(a)$ at level ℓ*, and note that

$$\mathfrak{s}(T_\ell^j(a)) =^d A_\ell^j(a),$$

where $S_\ell^j(a)$ is defined in (4.2.1). Proceeding in a similar way as in the proof of Corollary 4.5.3 we obtain a "spinal decomposition" of the size-biased tree, which we state here without proof.

Proposition 4.5.6 *The size-biasing of $\mathcal{L}(T_\ell^j(a))$ with $\mathfrak{s}(T_\ell^j(a))$ arises as the distribution of the superposition of $T_\ell^j(a)$ and*

$$(\delta_{\hat{Y}_{j-1}}, \dots, \delta_{\hat{Y}_k}, \Phi_{k-1}^k(\hat{Y}_k), \dots, \Phi_\ell^k(\hat{Y}_k)), \quad k = j-1, j-2, \dots, \ell,$$

where $\mathcal{L}(\hat{Y}_k)$ is the size-biasing of μ_k, $T_\ell^j(a)$, $\hat{Y}_{j-1}, \ldots, \hat{Y}_\ell$ are independent, and given $\hat{Y}_{j-1}, \ldots, \hat{Y}_\ell$, the $\Phi_r^k(\hat{Y}_k)$, $k > r$, are independent.

Under the assumption (4.2.2) of summability of the second moments of μ_k, $k \in \mathbb{N}$, the size-biasing of the random labelled forest $\mathfrak{F}^\ell = \mathfrak{F}_\ell^\theta$ (constructed in subsection 4.3) with respect to its " size" $\mathfrak{s}(\mathfrak{F}_\ell^\theta) = \zeta_\ell^\theta$ arises as the independent superposition of $\mathfrak{F}_\ell^\theta$ and $\hat{T}_\ell^{\infty,\mathrm{can}}$, where $\hat{T}_\ell^{\infty,\mathrm{can}}$ is constructed as follows:

First build a "spine" $(\ldots, \hat{Y}_{\ell-2}, \hat{Y}_{\ell-1}, \hat{Y}_\ell)$, and given the spine, superimpose independently the trees $T_\ell^k(\hat{Y}_k), k \geq \ell$.

4.6 The genealogy of relatives of a sampled individual

In subsection 4.3 we fixed a ball (or ℓ-block) $B_\ell^{(N)}$ from the beginning. Now we take a different viewpoint and think of an individual sampled from the equilibrium population within a union of many ℓ-blocks in Ω_N from the beginning. Denote the chosen individual by I, and the ℓ-block by $\hat{B}_\ell$. Recall that for large N the total number of individuals in an ℓ-block is approximately distributed like $N^\ell \zeta_\ell$ (see Theorem 2.4.1), we see that the number of individuals in the chosen block is approximately distributed like $N^\ell \hat{\zeta}_\ell^\theta$, where $\mathcal{L}(\hat{\zeta}_\ell^\theta)$ is the size-biasing of $\mathcal{L}(\zeta_\ell^\theta)$.

The block $\hat{B}_\ell$ sits in a nested sequence of blocks of levels $\ell+1, \ell+2, \ldots$ which we denote by $\hat{B}_{\ell+1}, \hat{B}_{\ell+2}, \ldots$ Consider the population $\hat{\mathcal{P}}_\ell$ of all those individuals in $\hat{B}_\ell$ which have an ancestral family in common with the individual I. The population $\hat{\mathcal{P}}_\ell$ can be decomposed in a natural way according to its immigration history into the $\hat{B}_j$, $j \geq \ell$.

For $j > \ell$, denote by $\hat{\mathcal{P}}_\ell^j$ the subpopulation of all those individuals in $\hat{\mathcal{P}}_\ell$ that have some common ancestor with I who lived in $\hat{B}_j$ but none who lived in $\hat{B}_{j-1}$. In other words, $\hat{\mathcal{P}}_\ell^j$ consists of all those individuals J which obey $J \sim_j I$ but not $J \sim_{j-1} I$.

In this way we obtain a decomposition of $\hat{\mathcal{P}}_\ell$ according to the hierarchical distance of the (geographically) closest ancestors common with the chosen individual I: for $j > \ell$, the subpopulation $\hat{\mathcal{P}}_\ell^j$ consists of those individuals in $\hat{B}_\ell$ whose geographically closest common ancestor with I has hierarchical distance j from I.

The size (i.e. the total number of individuals) of $\hat{\mathcal{P}}_\ell^j$ is approximately distributed as

$$N^\ell S_\ell^{j-1}(\hat{Y}_{j-1}) = N^\ell S_\ell(S_{\ell-1} \ldots (S_{j-2}(\hat{Y}_{j-1})))$$

and thus has approximate expectation $N^\ell \frac{1}{4c_{j-1}^2}$. Hence the summability condition (4.2.2) (which corresponds to the condition for transience resp. strong transience of the hierarchical random walk) amounts precisely to an expected finite number of relatives of the chosen individual in the block $\hat{B}_\ell$.

5 The hierarchical mean field limit of two-level branching systems in equilibrium

In this section we investigate the two-level branching equilibrium $\Psi^{(N,\theta)}(0)$ described in Proposition 2.3.5 and its limiting behavior as $N \to \infty$. We assume that the underlying random walk is a $(2, (c_\ell), N)$-random walk on Ω_N and (c_ℓ) satisfies the strong transience conditions (2.3.3) and (2.3.4). Recall that $(B_\ell^{(N)})$ denotes a sequence of nested blocks in Ω_N.

We will see in Lemma 5.2.1 that in equilibrium, asymptotically as $N \to \infty$, $\Psi^{(N,\theta)}(0)$ consists of the order of $N^{\ell/2}$ families in $B_\ell^{(N)}$, a typical such family having a random multiple of $N^{\ell/2}$ individuals.

5.1 A spatial ergodic theorem

In this subsection we first collect some basic facts about the two-level branching systems $\Psi^{(N)}(t)$ on Ω_N which were introduced in section 2.2.6 and their equilibria. The main result will be a spatial ergodic theorem for the aggregated equilibria; this will also be an ingredient in the proof of Theorem 2.4.1.

Let $M_c(\Omega_N) \setminus \{o\}$ denote the space of non-zero counting measures on Ω_N such that finite sets have finite measure, and let $M^{\mathrm{fi}}(M(\Omega_N) \setminus \{o\})$ be the set of measures ν on $M_c(\Omega_N) \setminus \{o\}$ such that $\int_{M(\Omega_N)\setminus\{o\}} \mu(B)\nu(d\mu) < \infty$ if B is a finite set. Then let $M_c^1(M_c(\Omega_N) \setminus \{o\})$ denote the subspace of counting measures in $M^{\mathrm{fi}}(M_c(\Omega_N) \setminus \{o\})$. The set $M_c(\Omega_N) \setminus \{o\}$ carries a complete separable metric ρ which generates the restriction of the vague topology on $M_c(\Omega_N)$ to $M_c(\Omega_N) \setminus \{o\}$ and makes a subset of $M_c(\Omega_N)$ bounded if and only if it is contained in $\{\mu | \mu(B) > 0$ for some finite $B \subset \Omega_N\}$, see [MKM] Proposition 3.3.2. The set $M^{\mathrm{fi}}(M_c(\Omega_N) \setminus \{o\})$ is equipped with the topology generated by $\nu \mapsto \int_{M_c(\Omega_N)\setminus\{o\}} \mu(B)\nu(d\mu)$, where B is a finite subset of Ω_N, and $\nu \mapsto \int F(\mu)\nu(d\mu)$, where $F : M_c(\Omega_N) \setminus \{o\} \to \mathbb{R}$ is continuous, bounded and with ρ-bounded support.

We can represent $\Psi^{(N)}(t)$ as a càdlàg Markov process with state space $M_c^1(M_c(\Omega_N) \setminus \{o\})$. On $M_c^1(M_c(\Omega_N) \setminus \{o\})$, we consider the class of functions $D(\mathfrak{G}^{\Omega_N}) = \{H(\nu) = h\left(\int F(\mu)\nu(d\mu)\right)\}$ with $F(\mu) = f(\langle \mu, \varphi \rangle)$ where φ has finite support and h, f are continuous functions on $\mathbb{R}$ with bounded second derivatives. We also define

$$\frac{\delta F(\mu)}{\delta \mu(x)} = \frac{d}{d\varepsilon} F(\mu + \varepsilon\delta_x)|_{\varepsilon=0}, \quad \frac{\delta^2 F(\mu)}{\delta\mu(x)\delta\mu(y)} = \frac{\partial^2}{\partial\varepsilon_1\partial\varepsilon_2} F(\mu + \varepsilon_1\delta_x + \varepsilon_2\delta_y)|_{\varepsilon_1=\varepsilon_2=0}.$$

Then $\Psi^{(N)}$ is the unique solution to the martingale problem given by the generator

$$\mathfrak{G}^{\Omega_N} H(\nu) \tag{5.1.1}$$

$$
\begin{aligned}
&= \sum_{x\in\Omega_N} \int \nu(d\mu) h'(\int F(\mu)\nu(d\mu)) \\
&\qquad \cdot \left[\frac{\delta F(\mu)}{\delta\mu(x)} \left(\sum_{k=1}^{\infty} \frac{c_{k+1}}{N^{k/2}} (\bar{\mu}_k(x) - \mu(x)) \right) + \frac{\delta^2 F(\mu)}{\delta\mu(x)\delta\mu(x)} \mu(x) \right] \\
&+ \int h''(\int F(\mu)\nu(d\mu)) F^2(\mu)\nu(d\mu)
\end{aligned}
$$

where

$$
\bar{\mu}_k(x) = \frac{1}{N^k} \sum_{y\in\Omega_N : d_N(y,x)\le k} \mu(y).
$$

Proposition 2.3.5 establishes the existence of a non-trivial equilibrium $\Psi^{(N,\theta)}$ that is spatially homogeneous (that is, with law invariant under translations in Ω_N) and mean θ. The *equilibrium random field* $\{\psi_x^{(N,\theta)}(t)\}_{x\in\Omega_N}$ is defined by

$$
\psi_x^{(N,\theta)}(t) = \int \mu(x)\Psi^{(N,\theta)}(t, d\mu). \tag{5.1.2}
$$

This gives the total number of individuals at site x irrespective of their family memberships.

Recall that $\{\eta_\ell^{(N,\theta)}(t, dx)\}$ defined in (2.4.1) describes the equilibrium normalized family size process in the block $B_\ell^{(N)}$ with mean number θ individuals per site.

Also recall that the normalized equilibrium population mass in $B_\ell^{(N)}$ is given by (see (2.4.3),(2.4.4))

$$
\begin{aligned}
\zeta_\ell^{(N,\theta)}(t) \quad &= \int x\eta_\ell^{(N,\theta)}(t, dx) = \frac{1}{N^\ell} \sum_{x\in B_\ell^{(N)}} \psi_x^{(N,\theta)}(N^{\ell/2}t) \\
&= \frac{1}{N^\ell} \sum_{x\in B_\ell^{(N)}} \int \mu(x)\Psi^{(N,\theta)}(N^{\ell/2}t, d\mu).
\end{aligned} \tag{5.1.3}
$$

Lemma 5.1.1 *Let G_N denote the Green operator (defined by (5.1.9)) of the $(2,(c_j),N)$-random walk with (c_j) satisfying condition (2.3.4) for strong transience, and write, for $\varphi_1, \varphi_2 :$ $\Omega_N \to \mathbb{R}_+$,*

$$
\langle \varphi_1, \varphi_2 \rangle = \sum_{x\in\Omega_N} \varphi_1(x)\varphi_2(x).
$$

Then

$$
\langle \varphi_1, G_N\varphi_2 \rangle \le \text{const} \sum_{x,y\in\Omega_N} \frac{\varphi_1(x)\varphi_2(y)}{N^{|x-y|/2}} \tag{5.1.4}
$$

and

$$\langle\varphi_1, G_N^2\varphi_2\rangle \le \text{const} \sum_{x,y\in\Omega_N} \varphi_1(x)\varphi_2(y) \sum_{j=|x-y|}^{\infty} \frac{1}{c_j^2} \tag{5.1.5}$$

where the constants do not depend on N. Hence

$$\langle\varphi_1, G_N\varphi_2\rangle \le \text{const}\langle 1,\varphi_1\rangle\langle 1,\varphi_2\rangle, \tag{5.1.6}$$
$$\langle\varphi_1, G_N^2\varphi_2\rangle \le \text{const}\langle 1,\varphi_1\rangle\langle 1,\varphi_2\rangle. \tag{5.1.7}$$

Proof. The transition probability $p_t(x,y)$ of the $(2,(c_j),N)$-random walk is given by

$$p_t(x,y) = (\delta_{0,|x-y|} - 1)\frac{e^{-h_{|x-y|}^{(N)}t}}{N^{|x-y|}} + (N-1)\sum_{j=|x-y|+1}^{\infty} \frac{e^{-h_j^{(N)}t}}{N^j} \tag{5.1.8}$$

where the $h_j^{(N)}$ are positive numbers (depending on N) such that

$$h_j^{(N)} \ge \text{const}\frac{c_{j-1}}{N^{(j-1)/2}}, \quad j \ge 1$$

(see [DGW2], subsection 3.1). Hence

$$\begin{aligned} G_N\varphi(x) &= \int_0^\infty \sum_y p_t(x,y)\varphi(y)dt \qquad (5.1.9)\\ &\le (N-1)\sum_y \varphi(y) \sum_{j=|x-y|+1}^{\infty} \frac{1}{N^j h_j^{(N)}}\\ &\le \text{const}(N-1)\sum_y \varphi(y) \sum_{j=|x-y|+1}^{\infty} \frac{N^{(j-1)/2}}{N^j c_{j-1}}\\ &\le \text{const}\sum_y \varphi(y) \sum_{j=|x-y|}^{\infty} \frac{1}{N^{j/2}c_j}\\ &\le \text{const}\sum_y \varphi(y) \sum_{j=|x-y|}^{\infty} \frac{1}{N^{j/2}}\\ &\le \text{const}\sum_y \varphi(y)\frac{1}{N^{|x-y|/2}}, \end{aligned}$$

and then (5.1.4) follows.

Similarly (see [DGW2])

$$\begin{aligned}G_N^2\varphi(x) &= \int_0^\infty t\left(\sum_y p_t(x,y)\varphi(y)\right)dt \\ &\le (N-1)\sum_y \varphi(y) \sum_{j=|x-y|+1}^{\infty} \frac{1}{N^j (h_j^{(N)})^2} \\ &\le \mathrm{const}(N-1)\sum_y \varphi(y) \sum_{j=|x-y|+1}^{\infty} \frac{N^{j-1}}{N^j c_{j-1}^2} \\ &\le \mathrm{const}\sum_y \varphi(y) \sum_{j=|x-y|}^{\infty} \frac{1}{c_j^2}\end{aligned}$$

and (5.1.5) follows. ■

Remark 5.1.2 *In [DGW2] the exponential waiting time of the hierarchical random walk has parameter* 1*, whereas here this parameter is* $\sum_\ell q_\ell^{(N)}$*, see subsection 2.2.1. In the present case,* $\sum_\ell q_\ell^{(N)} = \sum_{j=0}^\infty \frac{c_j}{N^{j/2}} =: L_N$*. This amounts to a time change* $t \to L_N t$ *in the calculations above, which produces* $L_N h_j^{(N)}$ *in place of* $h_j^{(N)}$ *and does not change the results of Lemma 5.1.1 (assuming* $c_0 > 0$*).*

The following is the analogue of [DGW1](2.3.3).

Proposition 5.1.3 *Under the conditions of Lemma 5.1.1, the first and second moments of* $\zeta_\ell^{(N,\theta)}(0)$ *are given by*

$$\mathbb{E}\zeta_\ell^{(N,\theta)}(0) = \theta$$

and

$$\mathbb{E}\left(\zeta_\ell^{(N,\theta)}(0)\right)^2 = \theta^2 + \theta(\langle\varphi_{N,\ell},\varphi_{N,\ell}\rangle + \langle\varphi_{N,\ell}, G_N\varphi_{N,\ell}\rangle + \frac{1}{4}\langle\varphi_{N,\ell}, G_N^2\varphi_{N,\ell}\rangle) \tag{5.1.10}$$

where $\varphi_{N,\ell}(x) = \frac{1}{N^\ell}1_{B_\ell^{(N)}}(x)$ *and* G_N *is the Green operator of the (strongly transient)* $(2,(c_\ell),N)$*-random walk.*

Corollary 5.1.4 *If the* $(2,(c_j),N)$*-random walk satisfies conditions (2.2.2) and (2.3.4), then for each* ℓ
(a)

$$\sup_N E((\zeta_\ell^{(N,\theta)}(0))^2) < \infty. \tag{5.1.11}$$

(b)

$$\mathrm{Var}(\zeta_\ell^{(N,\theta)}(0)) \to 0 \ \ \textit{as } \ell \to \infty \tag{5.1.12}$$

uniformly in N.

Proof. (a) follows immediately from Proposition 5.1.3 and Lemma 5.1.1.
(b) From Proposition 5.1.3

$$\mathrm{Var}(\zeta_\ell^{(N,\theta}(0)) = \theta(\langle \varphi_{N,\ell}, \varphi_{N,\ell}\rangle + \langle \varphi_{N,\ell}, G_N \varphi_{N,\ell}\rangle + \frac{1}{4}\langle \varphi_{N,\ell}, G_N^2 \varphi_{N,\ell}\rangle). \tag{5.1.13}$$

We will show that each of the three terms on the r.h.s. of (5.1.13) converges to 0 as $\ell \to \infty$ uniformly in N. First $\langle \varphi_{N,\ell}, \varphi_{N,\ell}\rangle \le \frac{1}{N^\ell} \to 0$ as $\ell \to \infty$ uniformly in N.

We have from (5.1.4) and using the ultrametric property of $|\cdot|$,

$$\begin{aligned}
\langle \varphi_{N,\ell}, G_N \varphi_{N,\ell}\rangle \quad & \le \mathrm{const}\frac{1}{N^{2\ell}} \sum_{|x|,|y|\le \ell} \frac{1}{N^{|x-y|/2}} \\
& \le \mathrm{const}\frac{1}{N^{2\ell}} \left(N^\ell + \sum_{x\neq y,\, |x|\le|y|\le\ell} \frac{1}{N^{|y|/2}} \right) \\
& \le \mathrm{const}\frac{1}{N^{2\ell}} \left(N^\ell + \sum_{k=1}^{\ell} \frac{N^{2k}}{N^{k/2}} \right) \\
& = \mathrm{const}\frac{1}{N^{2\ell}} \left(N^\ell + \frac{N^{(\ell+1)3/2} - N^{3/2}}{N^{3/2}-1} \right) \\
& \le \mathrm{const} \left(\frac{1}{N^\ell} + \frac{1}{N^{\ell/2}} \right) \longrightarrow 0 \ \text{ as } \ell \to \infty
\end{aligned}$$

uniformly in N, and from (5.1.5), again using the ultrametric property,

$$\begin{aligned}
\langle \varphi_{N,\ell}, G_N^2 \varphi_{N,\ell}\rangle \quad & \le \mathrm{const}\frac{1}{N^{2\ell}} \sum_{|x|,|y|\le\ell} \sum_{j=|x-y|}^{\infty} \frac{1}{c_j^2} \\
& \le \mathrm{const}\frac{1}{N^{2\ell}} \left(N^\ell \sum_{j=0}^{\infty} \frac{1}{c_j^2} + \sum_{x\neq y,\, |x|\le|y|\le\ell} \sum_{j=|y|}^{\infty} \frac{1}{c_j^2} \right) \\
& \le \mathrm{const}\frac{1}{N^{2\ell}} \left(N^\ell + \sum_{k=1}^{\ell} N^{2k} \sum_{j=k}^{\infty} \frac{1}{c_j^2} \right)
\end{aligned}$$

$$\leq \text{const}\frac{1}{N^{2\ell}}\left(N^\ell + \sum_{j=1}^{\infty}\frac{1}{c_j^2}\sum_{k=1}^{j\wedge\ell}N^{2k}\right)$$
$$\leq \text{const}\frac{1}{N^{2\ell}}\left(N^\ell + \sum_{j=1}^{\infty}\frac{1}{c_j^2}N^{2(j\wedge\ell)}\right)$$
$$= \text{const}\left(\frac{1}{N^\ell} + \sum_{j=1}^{\ell-1}\frac{1}{c_j^2}\frac{1}{N^{2(\ell-j)}} + \sum_{j=\ell}^{\infty}\frac{1}{c_j^2}\right). \tag{5.1.14}$$

The first term goes to zero as $\ell \to \infty$, the second term goes to zero as $\ell \to \infty$ by dominated convergence, and clearly the last term goes to zero as $\ell \to \infty$. Therefore

$$\langle \varphi_{N,\ell}, G_N^2 \varphi_{N,\ell}\rangle \to 0 \text{ as } \ell \to \infty \quad \text{uniformly in } N. \tag{5.1.15}$$

■

Remark 5.1.5 *(a)*

$$\lim_{N\to\infty} \text{Var}(\zeta_\ell^{(N,\theta)}(0)) = \theta\sum_{j=\ell}^{\infty}\frac{1}{c_j^2}, \tag{5.1.16}$$

(b)

$$\langle \varphi_{N,\ell}, G_N^2 \varphi_{N,\ell}\rangle \leq \text{const}\left(\frac{1}{N^\ell} + \frac{1}{N^2} + \sum_{j=\ell}^{\infty}\frac{1}{c_j^2}\right). \tag{5.1.17}$$

Corollary 5.1.6 *Consider the "exterior function"*

$$\varphi_{N,\ell,\text{ext}}(x) = \sum_{k=1}^{\infty}\frac{c_{\ell+k-1}}{N^{\ell+2k-1}}\mathbf{1}_{B_{\ell+k}^{(N)}}(x). \tag{5.1.18}$$

Then

$$\mathbb{E}\left[\left(\sum_{x\in\Omega_N}\varphi_{N,\ell,\text{ext}}(x)\int \mu(x)\Psi^{(N,\theta)}(0,d\mu)\right)^2\right]$$
$$= \mathbb{E}\left[\left(\sum_{k=1}^{\infty}\frac{c_{\ell+k-1}}{N^{k-1}}\zeta_{\ell+k}^{(N,\theta)}(0)\right)^2\right] < \infty \tag{5.1.19}$$

uniformly in N.

Proof. As in Proposition 5.1.3, (from [DGW1], comment 2.3.5)

$$\mathbb{E}\left[\left(\sum_{x\in\Omega_N}\varphi_{N,\ell,\mathrm{ext}}(x)\int\mu(x)\Psi^{(N,\theta)}(0,d\mu)\right)^2\right] \tag{5.1.20}$$
$$=\theta^2\langle 1,\varphi_{N,\ell,\mathrm{ext}}\rangle^2+\theta(\langle\varphi_{N,\ell,\mathrm{ext}},\varphi_{N,\ell,\mathrm{ext}}\rangle$$
$$+\langle\varphi_{N,\ell,\mathrm{ext}},G_N\varphi_{N,\ell,\mathrm{ext}}\rangle+\frac{1}{2}\langle\varphi_{N,\ell,\mathrm{ext}},G_N^2\varphi_{N,\ell,\mathrm{ext}}\rangle).$$

We will prove that each term on the r.h.s. of (5.1.20) is bounded uniformly in N.

$$\begin{aligned}\langle 1,\varphi_{N,\ell,\mathrm{ext}}\rangle &= \sum_{k=1}^{\infty}\frac{c_{\ell+k-1}}{N^{\ell+2k-1}}\\ &+\sum_{k=1}^{\infty}\frac{c_{\ell+k-1}}{N^{\ell+2k-1}}\sum_{n=1}^{\ell+k}(N-1)N^{n-1}\\ &\le\sum_{k=1}^{\infty}\frac{c_{\ell+k-1}}{N^{\ell+2k-1}}+\sum_{k=1}^{\infty}\frac{c_{\ell+k-1}}{N^{\ell+2k-1}}N^{\ell+k}\\ &\le \mathrm{const}\sum_{k=1}^{\infty}\frac{c_{\ell+k-1}}{N^k}\\ &<\infty\quad\text{uniformly in } N.\end{aligned}$$

Then, from (5.1.6) the terms $\langle\varphi_{N,\ell,\mathrm{ext}},G_N\varphi_{N,\ell,\mathrm{ext}}\rangle$ and $\langle\varphi_{N,\ell,\mathrm{ext}},G_N^2\varphi_{N,\ell,\mathrm{ext}}\rangle$ are bounded uniformly in N and since $\varphi_{N,\ell,\mathrm{ext}}$ has a uniform bound in N as well, using (2.2.2) we see that $\langle\varphi_{N,\ell,\mathrm{ext}},\varphi_{N,\ell,\mathrm{ext}}\rangle$ is also bounded uniformly in N (with $N>\sup c_{\ell+1}/c_\ell$). ■

Theorem 5.1.7 (Spatial ergodic theorem) *The pointwise ergodic theorem holds on Ω_N, that is,*

$$\lim_{j\to\infty}\zeta_j^{(N,\theta)}(0)=\theta \ a.s. \tag{5.1.21}$$

Proof. First note that the equilibrium random field $\{\psi_x^{(N,\theta)}\}_{x\in\Omega_N}$ defined by (5.1.2) is invariant under the action of the group Ω_N, and $E(\psi_x^{(N,\theta)})=\theta$. Moreover by (5.1.12) the spatial averages $\zeta_\ell^{(N,\theta)}(0)=\frac{1}{N^\ell}\sum_{x\in B_\ell^{(N)}}\psi_x^{(N,\theta)}(0)$ satisfy

$$\lim_{j\to\infty}\mathrm{Var}(\zeta_j^{(N,\theta)}(0))=0,$$

so that the convergence in probability follows. To complete the proof note that Ω_N is an amenable group and the collection of balls $\{B_\ell^{(N)}\}_{\ell\in\mathbb{N}}$ is a tempered Følner sequence. The a.s. pointwise convergence then follows by [Lin](Theorem 1.2). ■

5.2 Reduction to two successive scales

In the section we show that the analysis of the multiscale behavior can be reduced to the case of two successive scales.

5.2.1 The equilibrium family size process

Let $\{\zeta_\ell^{(N,\theta)}(t),\ t\in\mathbb{R}\}_{\ell=\dots,2,1}$ denote the collection of equilibrium block average processes in the nested sequence of blocks $B_\ell^{(N)}$ (see (5.1.3)).

Lemma 5.2.1 *The fixed time marginal distributions $\{\zeta_\ell^{(N,\theta)}(0)\}$ at time $t=0$ satisfy:*
(a) For each $\ell\in\mathbb{N}$ the family $\{\zeta_\ell^{(N,\theta)}(0)\}_{N\in\mathbb{N}}$ is tight and every limit point has expected value θ.
(b) The family-size constraint

$$\mathbb{E}\left[\int_K^\infty x\eta_\ell^{(N,\theta)}(0,dx)\right]\le\frac{\text{const}}{K}\tag{5.2.1}$$

is satisfied for a suitable constant not depending on $K>0$ and $N\ge 2$.

Proof. (a) follows immediately from Corollary 5.1.4.
(b) Let $\varphi_\ell^N(x)=\frac{1}{N^\ell}1_{B_\ell^{(N)}}(x)$. Then

$$\begin{aligned}
&\mathbb{E}\left[\int_K^\infty x\eta_\ell^{(N,\theta)}(0,dx)\right]\le\frac{1}{K}\mathbb{E}\left[\int_0^\infty x^2\eta_\ell^{(N,\theta)}(0,dx)\right]\\
&\le\frac{1}{K}\mathbb{E}\left[\int_0^\infty x^2\eta_\ell^{(N,\theta)}(0,dx)\right]\\
&=\frac{1}{K}\frac{1}{N^{3\ell/2}}\mathbb{E}\left[\int(\mu(B_\ell^{(N)}))^2\Psi^{(N,\theta)}(0,d\mu)\right]\\
&=\frac{1}{K}N^{\ell/2}\mathbb{E}\left[\int(\mu(\varphi_\ell^N))^2\Psi^{(N,\theta)}(0,d\mu)\right]\\
&=\frac{N^{\ell/2}}{K}\left[\langle\varphi_\ell^N,\varphi_\ell^N\rangle+\frac{1}{2}\langle\varphi_\ell^N,G_N\varphi_\ell^N\rangle\right]\quad\text{(by [DGW1](2.3.1))}\\
&\le\frac{N^{\ell/2}}{K}\text{const}\left(\frac{1}{N^\ell}+\frac{1}{N^{\ell/2}}\right)\quad\text{(by the proof of Corollary 5.1.4)}\\
&\le\frac{\text{const}}{K}.
\end{aligned}$$

■

Remark 5.2.2 *Part (b) of the Lemma 5.2.1 implies that asymptotically the restriction of $\Psi^{(N,\theta)}(0)$ to the ball $B_\ell^{(N)}$ consists of families of size $O(N^{\ell/2})$ or smaller.*

5.2.2 Distant immigrants

In order to establish that there is a unique limit law and to identify it we now return to the dynamical picture. In the next lemma we show that the expected contribution to the equilibrium population in a ball $B_\ell^{(N)}$ coming from immigrants who immigrate directly from outside the ball $B_{\ell+1}^{(N)}$ and the descendants of the population in the ball $B_\ell^{(N)}$ in the distant past are both negligible.

Lemma 5.2.3 *(a) Let $\{\zeta_{\ell,\text{ext}}^{(N,\theta)}(0)\}$ denote the contribution to the equilibrium block average in the ball $B_\ell^{(N)}$ at time 0 coming from individuals immigrating directly from $(B_{\ell+1}^{(N)})^c$. Then*

$$E\left(\zeta_{\ell,\text{ext}}^{(N,\theta)}(0)\right) \leq \frac{\text{const}}{N^{\frac{1}{2}}}.$$

(b) The expected mass to enter $B_{\ell+1}^{(N)}$ from $(B_{\ell+1}^{(N)})^c$ in a time interval of length $N^{\ell/2}$ is $O(\frac{1}{N^{1/2}})$.
(c) The expected contribution to the population in the ball $B_\ell^{(N)}$ at time 0 from the descendants of individuals alive at time t_0 is of order $O(e^{-c_\ell|t_0|})$ as $t_0 \to -\infty$.

Proof. (a) To verify this, we first note that the total number of individuals to immigrate from $B_{\ell+k}^{(N)}$ with $k \geq 2$ to $B_\ell^{(N)}$ in the time interval $[-N^{\ell/2}t, -N^{\ell/2}(t+dt))$ is of the order

$$(\zeta_{\ell+k}^{(N,\theta)}(-t)N^{\ell+k}) \times (\frac{c_{\ell+k-1}}{N^{(\ell+k-1)/2}}) \times \frac{1}{N^k} \times dtN^{\ell/2}$$

where the first factor is the number of particles in $B_{\ell+k}^{(N)}$, the second factor is the rate of migration of each of these particles to a point chosen randomly in $B_{\ell+k}^{(N)}$, the third factor is the probability this point falls in the tagged ball $B_\ell^{(N)}$ and the last factor is the length of the time period. Recalling that the mass process in $B_\ell^{(N)}$ is subcritical with parameter $\frac{c_\ell}{N^{\ell/2}}$, the expected total mass at time 0 coming from immigration from outside $B_{\ell+1}^{(N)}$ in the time interval $[-N^{\ell/2}T, 0)$, $T > 0$, is of the order

$$\begin{aligned}&\sum_{k\geq 2}\int_{-T}^{0} \frac{c_{\ell+k-1}}{N^{(k-1)/2}} \times N^\ell \times e^{c_\ell t}\mathbb{E}(\zeta_{\ell+k}^{(N,\theta)}(-t))dt \qquad (5.2.2)\\ &\leq \text{const} \times \frac{N^\ell}{N^{1/2}},\end{aligned}$$

uniformly in T and ℓ, where we have used the assumption (2.2.2). Therefore the expected mass (normalized by $\frac{1}{N^\ell}$) at time 0 in $B_\ell^{(N)}$ that immigrated during the time interval $[N^{\ell/2}t_0, 0)$ directly from outside $B_{\ell+1}^{(N)}$ is $O(N^{-\frac{1}{2}})$ as $N \to \infty$.

(b) and (c) follow from a first moment calculation.

■

5.3 Diffusion limit of the family size process in two spatial scales

In this section we consider the asymptotic (as $N \to \infty$) time development of the population occupying $B_{\ell+1}^{(N)}$ and in particular the subpopulation obtained by considering individuals that occupy a tagged ball $B_\ell^{(N)}$ in the natural time scale for the population in $B_\ell^{(N)}$ and assuming that the initial family size processes in $B_{\ell+1}^{(N)}$ satisfy the family size constraint of Lemma 5.2.1.

As a result of the previous lemma, asymptotically as $N \to \infty$ and $t_0 \to -\infty$ the equilibrium population in $B_\ell^{(N)}$ consists of the descendants of immigrants coming from $B_{\ell+1}^{(N)}$ during the time interval $(-N^{-\ell/2}t_0, 0]$ (recall Remark 2.3.4). Another key point is that, as we verify below, in this time interval the family size process in $B_{\ell+1}^{(N)}$ is asymptotically constant. Moreover from Lemma 5.2.1 all but an arbitrarily small proportion of the population in $B_{\ell+1}^{(N)}$ is structured into $O(N^{(\ell+1)/2})$ families containing $O(N^{(\ell+1)/2})$ individuals.

We next show that in the $N^{\ell/2}$-time scale the total population structure in the ball $B_{\ell+1}^{(N)}$ is essentially constant.

Lemma 5.3.1 *Let $\{\zeta_{\ell+1}^{(N,\theta)}(s)\}$ be the equilibrium normalized process (see (5.1.3)) in its natural time scale $N^{(\ell+1)/2}$.*
For $t_0 = t_0(N) < 0$ such that $\frac{|t_0|}{N^{\frac{1}{2}}} < c$,

$$P\left(\sup_{t_0 \le t \le 0} |\zeta_{\ell+1}^{(N,\theta)}(N^{-1/2}t) - \zeta_{\ell+1}^{(N,\theta)}(N^{-1/2}t_0)| > K\right)$$
$$\le \left(\frac{\text{const}}{K^2}\right) c + o(1), \tag{5.3.1}$$

where const does not depend on t_0 and N, and $o(1)$ converges to 0 as $N \to \infty$.

Proof. Recall that $\Psi^{(N)}(t)$ is characterized as the unique solution of the martingale problem with generator (5.1.1). Applying the generator to the function $F(\nu) = \int \langle \mu, \frac{1}{N^{\ell+1}} 1_{B_{\ell+1}^{(N)}} \rangle \nu(d\mu)$, it follows that $\{M_{\ell+1}(t)\}_{t_0 \le t \le 0}$, defined by

$$\begin{aligned} M_{\ell+1}(t) = \quad & \zeta_{\ell+1}^{(N,\theta)}(t) - \zeta_{\ell+1}^{(N,\theta)}(t_0) \\ & - \int_{t_0}^{t} \sum_{k=1}^{\infty} \frac{c_{\ell+k}}{N^{k-1}} \left(\zeta_{\ell+1+k}^{(N,\theta)}(N^{-k/2}s) - \zeta_{\ell+1}^{(N,\theta)}(s)\right) ds \end{aligned}$$

is a martingale.

Then

$$P\left(\sup_{t_0 \le t \le 0} |\zeta_{\ell+1}^{(N,\theta)}(N^{-1/2}t) - \zeta_{\ell+1}^{(N,\theta)}(N^{-1/2}t_0)| > K\right)$$

$$
\begin{aligned}
&\leq P\left(\int_{\frac{t_0}{N^{1/2}}}^{0} |\sum_{k=1}^{\infty}\frac{c_{\ell+k}}{N^{k-1}}\left(\zeta_{\ell+1+k}^{(N,\theta)}(N^{-k/2}s)-\zeta_{\ell+1}^{(N,\theta)}(s)\right)|ds > \frac{K}{2})\right)\\
&+P\left(\sup_{\frac{t_0}{N^{1/2}}\leq t\leq 0}|M_{\ell+1}(t)|>\frac{K}{2}\right). \qquad (5.3.2)
\end{aligned}
$$

Recall from Corollary 5.1.4 that $\mathbb{E}\left(\zeta_\ell^{(N,\theta)}(t)\right)^2$ is bounded uniformly in N and t. Let

$$
g_N(s) := |\sum_{k=1}^{\infty}\frac{c_{\ell+k}}{N^{k-1}}\left(\zeta_{\ell+1+k}^{(N,\theta)}(N^{-k/2}s)-\zeta_{\ell+1}^{(N,\theta)}(s)\right)|
$$

and note that by Corollary 5.1.6 $\mathbb{E}(g_N^2(s))$ is uniformly bounded in s and N. Then

$$
\begin{aligned}
&P\left(\int_{\frac{t_0}{N^{1/2}}}^{0} g_N(s)ds > \frac{K}{2}\right)\\
&\leq \frac{\text{const}}{K^2}(\frac{t_0}{N^{1/2}})^2\mathbb{E}\left[\frac{1}{\frac{t_0}{N^{1/2}}}\int_{\frac{t_0}{N^{1/2}}}^{0} g_N(s)ds\right]^2\\
&\leq \frac{\text{const}}{K^2}(\frac{t_0}{N^{1/2}})^2\left[\frac{1}{\frac{t_0}{N^{1/2}}}\int_{\frac{t_0}{N^{1/2}}}^{0} \mathbb{E}(g_N(s)^2)ds\right]\\
&\leq \frac{\text{const}}{K^2}\left(\frac{t_0}{N^{1/2}}\right)^2.
\end{aligned}
$$

Next we note that by Lemma 3.2.4(b) and Remark 3.2.5 (with $\varepsilon = \frac{1}{N}$) we get for $t_0 \leq t \leq 0$

$$
\begin{aligned}
&\mathbb{E}\Big([\zeta_{\ell+1}^{(N,\theta)}(t)-\zeta_{\ell+1}^{(N,\theta)}(t_0)]^2\Big|\zeta_{\ell+1}^{(N,\theta)}(t_0)=m,\sum_{k=1}^{\infty}\frac{c_{\ell+k}}{N^{k-1}}\zeta_{\ell+1+k}^{(N,\theta)}(\cdot)=a(\cdot)\Big)\\
&=(m^2e^{-2c_{\ell+1}(t-t_0)}-m^2)\\
&+\frac{m}{c_{\ell+1}^2}\{c_{\ell+1}(t-t_0)e^{-c_{\ell+1}(t-t_0)}+e^{-2c_{\ell+1}(t-t_0)}+2(t-t_0)c_{\ell+1}^3ae^{-c_{\ell+1}(t-t_0)}\\
&\qquad -e^{-c_{\ell+1}(t-t_0)}+(t-t_0)c_{\ell+1}^2\varepsilon e^{-c_{\ell+1}(t-t_0)}\}\\
&+\frac{1}{c_{\ell+1}^2}\int_{t_0}^{t}k_1(s,t)a(s)ds \quad +\int_{t_0}^{t}\int_{t_0}^{s_2}k_2(t,s_2,s_1)a(s_1)a(s_2)ds_1ds_2\\
&+o(1)\cdot\int_{t_0}^{t}k_3(t,s)s(s)ds,
\end{aligned}
$$

where $k_i(t,\cdot)$ $i=1,3$ and $k_2(t,\cdot,\cdot)$ are bounded non-negative kernels that satisfy (3.2.11) and $o(1)$ converges to 0 as $N\to\infty$. Using Corollary 5.1.4 we conclude that

$$\mathbb{E}\Big([\zeta_{\ell+1}^{(N,\theta)}(N^{-1/2}t)-\zeta_{\ell}^{(N,\theta)}(N^{-1/2}t_0)]^2\Big)\le \text{const}\cdot\frac{(t-t_0)}{N^{1/2}}+o_1(\frac{1}{N}). \tag{5.3.3}$$

We now apply the L^2-martingale inequality,

$$\begin{aligned}&P\left(\sup_{\frac{t_0}{N^{1/2}}\le t\le 0}|M_{\ell+1}(t)|>\frac{K}{2}\Big|M_{\ell+1}(t_0/N^{1/2})=0\right)\\ &\le\frac{1}{K^2}\mathbb{E}\Big(|M_{\ell+1}(0)|^2\Big|M_{\ell+1}(t_0/N^{1/2})=0\Big)\\ &\le\frac{\text{const}}{K^2}\frac{|t_0|}{N^{1/2}}+o_1(\frac{1}{N}).\end{aligned}$$

■

Before stating the main result we introduce the required spaces of measures and functions. Let M^{TL} denote the set of Radon measures on $[0,\infty)\times(0,\infty)$ such that $\int_{0+}^{\infty}\int_0^{\infty}y\mu(dx,dy)+\int_{0+}^{\infty}y\mu(\{0\},dy)<\infty$ and $\int_{0+}^{\infty}\int_0^{\infty}x\mu(dx,dy)<\infty$. Then let $M^{TL+}=\{\mu\in M^{TL}:\int_{0+}^{\infty}y\mu(\{0\},dy)=0\}$. Let $C^{TL,2}$ denote the space of continuous functions on $[0,\infty)\times(0,\infty)$ with bounded first and second partial derivatives and which satisfy $|\phi(x,y)|\le\text{const}|x\wedge 1|$.

Proposition 5.3.2 *Let ℓ be fixed and consider the M^{TL}-valued process*

$$\mathcal{Y}^{(N)}(t;dx,dy)=\frac{1}{N^{(\ell+1)/2}}\int \mathbf{1}_{\{\frac{\nu(B_\ell^{(N)})}{N^{\ell/2}}\in dx,\frac{\nu(B_{\ell+1}^{(N)})}{N^{(\ell+1)/2}}\in dy\}}\Psi^{(N,\theta)}(N^{\ell/2}t,d\nu).$$

Assume that

$$\mathcal{Y}^{(N)}(t_0;dx,dy)\overset{N\to\infty}{\Longrightarrow}\mu_{\ell+1}^0(dy)\delta_0(dx).$$

Then

$$\mathcal{Y}^{(N)}(t;dx,dy)\overset{N\to\infty}{\Longrightarrow}\mu_{\ell+1}^0(dy)\delta_0(dx)\quad\text{for all }\ t\ge t_0,$$

and $\{N^{1/2}\mathcal{Y}^{(N)}(t;dx,dy)\mathbf{1}_{x>0}\}_{t_0<t\le 0}$ *converges weakly as* $N\to\infty$ *to the* M^{TL+}*-valued diffusion* $\{\eta(t)\}_{t_0<t\le 0}$ *with generator* $\mathfrak{G}_\ell$ *defined as follows. Let* $F(\mu_\ell)=f(\langle\mu_\ell,\varphi\rangle)$ *with* $\varphi\in C^{TL,2}$*. Then*

$$\begin{aligned}\mathfrak{G}_\ell F(\mu_\ell)\ &=f'(\langle\mu_\ell,\varphi\rangle)\langle\mu_\ell,G_{c_\ell}^{(2)}\varphi\rangle+\tfrac{1}{2}f''(\langle\mu_\ell,\varphi\rangle)\langle\mu_\ell,\varphi^2\rangle\\ &\quad+f'(\langle\mu_\ell,\varphi\rangle)c_\ell\int\tfrac{\partial\varphi}{\partial x}(x,y)|_{x=0}y\mu_{\ell+1}^0(dy)\end{aligned}\tag{5.3.4}$$

where $G_{c_\ell}^{(2)}$ denotes the application of the operator G_{c_ℓ} to the x variable. That is, $\eta(t)$ is a two-level branching process with constant multitype immigration source with immigration from zero of type y at rate given by $y\mu_{\ell+1}^0(dy)$ and total immigration rate of $\int_0^{\infty}y\mu_{\ell+1}^0(dy)$.

Proof. The proof is a refinement of the proof of Proposition 3.2.1. We begin by noting that Lemma 5.2.3 implies that asymptotically as $N \to \infty$ the contribution of immigrants into $B_{\ell+1}^{(N)}$ from $B_{\ell+k}^{(N)}$, $k \geq 2$ is negligible. Moreover the contribution of immigrants into $B_{\ell+1}^{(N)}$ from $B_{\ell+2}^{(N)}$ in the time scale $N^{\ell/2}$ is also asymptotically negligible. Therefore in the time scale $N^{\ell/2}$ we can restrict attention to the population in $B_{\ell+1}^{(N)}$. More precisely, asymptotically as $N \to \infty$ the population $\zeta_{\ell}^{(N,\theta)}(t) = \int\int x N^{1/2} \mathcal{Y}^{(N)}(t, dx, dy) \mathbf{1}_{x>0}$ consists of descendants of immigrants entering from $B_{\ell+1}^{(N)}$ in the time interval $[N^{\ell/2} t_0(N), 0]$ provided that we take $t_0(N) \to -\infty$.

As in the proof of Proposition 3.2.1 a standard argument yields the tightness of the processes $\{\zeta_{\ell}^{(N,\theta)}(t)\}_{t_0 \leq t \leq 0}$ and $\{\zeta_{\ell+1}^{(N,\theta)}(N^{-1/2} t)\}_{t_0 \leq t \leq 0}$. The tightness in $C([t_0, 0], M^{TL,+})$ of $N^{1/2} \mathcal{Y}^{(N)}(\cdot) \mathbf{1}_{x>0}$ is also obtained as in the proof of Proposition 3.2.1. One difference is the presence of additional terms in the expressions for the moments. However these expressions tend to zero as $N \to \infty$.

A first moment calculation shows that the expected contribution of individuals who leave $B_{\ell}^{(N)}$ and then re-immigrate is $O(\frac{1}{N})$ and therefore asymptotically negligible. This means that we can treat the population in $B_{\ell}^{(N)}$ in the time interval $[t_0, 0]$ as a two-level branching system critical at the family level and subcritical at the individual level with subcriticality parameter c_ℓ and with immigration of individuals from $B_{\ell+1}^{(N)}$. Moreover by Lemma 5.3.1 the source of immigrant individuals $\zeta_{\ell+1}^{(N,\theta)}(t)$ is asymptotically constant in the time interval $(t_0(N), 0)$ provided that

$$\frac{|t_0(N)|}{N^{\frac{1}{2}}} \to 0$$

as $N \to \infty$.

The main difference from the proof of Proposition 3.2.1 is that we cannot assume that each immigrant belongs to a different family. Since the individuals immigrating into $B_{\ell}^{(N)}$ come from families in $B_{\ell+1}^{(N)}$ subject to family branching, it is necessary to keep track of the family structure in the ball $B_{\ell+1}^{(N)}$. The reason for this is that in principle an individual immigrating into $B_{\ell}^{(N)}$ could be a member of a $B_{\ell+1}^{(N)}$-family already represented in $B_{\ell}^{(N)}$ and then could not be viewed as the founder of an independent family in $B_{\ell}^{(N)}$. Part of the argument below is to verify that this effect is negligible.

By the family size constraint at time t_0 (see Lemma 5.2.1(b)), the population in $B_{\ell+1}^{(N)}$ consists of a collection of families whose sizes are $O(N^{(\ell+1)/2})$. We index these families at time t_0 by $i \in \mathbb{N}$ with masses $y_i N^{(\ell+1)/2}$.

Recall from section 2.3.2 that $\Psi^{(N,\theta)}(t) = \sum_j \psi_j(t, \cdot)$ where $\psi_j(t, \cdot)$ is a counting measure on Ω_N corresponding to the spatial distribution of the family indexed by j at time t. We now give a precise formulation to the time development of the families simultaneously in $B_{\ell+1}^{(N)}$ and

$B_\ell^{(N)}$. To each family ψ_j and $t \in \mathbb{R}$ we associate a couple $(x_j(t), y_j(t))$ where

$$x_j(t) = \frac{\psi_j(t, B_\ell^{(N)})}{N^{\ell/2}}, \quad y_j(t) = \frac{\psi_j(t, B_{\ell+1}^{(N)})}{N^{(\ell+1)/2}}.$$

Then

$$\mathcal{Y}^{(N)}(t; dx, dy) = \varepsilon_2 \sum_i \delta_{(x_i(N^{\ell/2}t), y_i(N^{\ell/2}t))}(dx, dy),$$

where $\varepsilon_2 = N^{-(\ell+1)/2}$. Also, let $\varepsilon_0 = \frac{1}{N}$, $\varepsilon_1 = \frac{1}{N^{\ell/2}}$.

By assumption, the family size distribution in $B_{\ell+1}^{(N)}$ at time t_0, asymptotically as $N \to \infty$, is given by μ^0 with $\int y\mu^0(dy) < \infty$.

It suffices to show that for $t \geq t_0$, as $N \to \infty$

$$\mathcal{Y}^{(N)}(t; dx, dy)\mathbf{1}_{\{x < N^{-\frac{1}{4}}\}} \Rightarrow \mu_{\ell+1}^0(dy)\delta_0(dx) \tag{5.3.5}$$

$$\mathcal{Y}^{(N)}(t; dx, dy)\mathbf{1}_{\{x > 0\}} \Rightarrow 0 \tag{5.3.6}$$

$$\{\widetilde{\mathcal{Y}}^{(N)}(t)\}_{t \geq t_0} \Rightarrow \{\eta(t)\}_{t \geq t_0} \tag{5.3.7}$$

where $\widetilde{\mathcal{Y}}^{(N)}(t)$ is the renormalized family measure given by

$$\widetilde{\mathcal{Y}}^{(N)}(t; dx, dy) \quad = \quad N^{1/2}\mathcal{Y}^{(N)}(t; dx, dy)\mathbf{1}_{\{x > 0\}}$$

and where $\eta(t)$ is a M^{TL+}-valued diffusion with generator $\mathfrak{G}_\ell$.

To verify (5.3.5), first using Lemma 5.2.3 and Proposition 3.2.1 we can verify that

$$\sup_{t \in (-t_0, 0]} \int\int (1 \wedge y)|\mathcal{Y}^{(N)}(t; dx, dy) - \mathcal{Y}^{(N)}(t_0; dx, dy)| \to 0 \text{ in probability.}$$

Moreover

$$\int\int_{N^{-1/4}}^{\infty} N^{1/2}\mathcal{Y}^{(N)}(t; dx, dy) < N^{1/4}\int\int_0^\infty xN^{1/2}\mathcal{Y}^{(N)}(t; dx, dy).$$

Therefore by the tightness of $\int\int_0^\infty x(N^{1/2}\mathcal{Y}^{(N)}(\cdot; dx, dy)\mathbf{1}_{\{x>0\}})$ and $\{\zeta_{\ell+1}^{(N,\theta)}\}$, as $N \to \infty$,

$$\int\int_{N^{-1/4}}^{\infty} \mathcal{Y}^{(N)}(1 \wedge y)(t; dx, dy) \to 0 \text{ in probability}$$

and (5.3.5) and (5.3.6) follow.

We now turn to the proof of (5.3.7). In order to implement the rescaling we introduce the class of functions of the form

$$F(\mu) = f(\langle \mu, \varphi \rangle) = f(\int\int \varphi(x, y)\mu(dx, dy)) = f(\varepsilon_2 \sum_i \varphi(x_i, y_i)),$$

where

$$\varphi(x,y)=\varepsilon_0^{-1/2}\mathbf{1}_{\{x>0\}}\varphi_\ell(x,y)+\varphi_{\ell+1}(y)\mathbf{1}_{\{x<N^{-\frac{1}{8}}\}}.$$

We assume that $|\varphi_\ell|,|\frac{\partial\varphi_\ell(x,y)}{\partial x}|,|\varphi_{\ell+1}|$ are bounded, C^2, and $|\varphi_\ell(x,y)|\le \text{const}\cdot x$ for $x\ge 0$ and $|\varphi_{\ell+1}(y)|\le \text{const}\cdot y$ for $y>0$.

Now define

$$\mu_\ell(dx,dy)=N^{1/2}\mu(dx,dy)\mathbf{1}_{\{x>0\}}$$

and note that

$$\int\varphi(x,y)\mu(dx,dy)=\int\varphi_\ell(x,y)\mu_\ell(dx,dy)+\int\varphi_{\ell+1}(x,y)\mathbf{1}_{\{x<N^{-\frac{1}{8}}\}}\mu(dx,dy).$$

The generator of $\mathcal{Y}^{(N)}$ acting on F is given by

$$
\begin{aligned}
& \mathfrak{G}^{(N)} F(\mu) \\
= \quad & c_\ell \varepsilon_0 \sum_{j_1=0}^{N^{\frac{2\ell-1}{4}}} \sum_{j_2>j_1} [f(\langle \mu, \varphi \rangle + \varepsilon_1(\varphi_\ell((j_1+1)\varepsilon_1, j_2\varepsilon_2) - \varphi_\ell(j_1\varepsilon_1, j_2\varepsilon_2)) - f(\langle \mu, \varphi \rangle)] \\
& \qquad\qquad \cdot \frac{(j_2-j_1)\varepsilon_2 \mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_2^2} \\
+ \quad & c_\ell \varepsilon_0 \sum_{j_1>N^{\frac{2\ell-1}{4}}} \sum_{j_2>j_1} [f(\langle \mu, \varphi \rangle + \varepsilon_1(\varphi_\ell((j_1+1)\varepsilon_1, j_2\varepsilon_2) - \varphi_\ell(j_1\varepsilon_1, j_2\varepsilon_2)) - f(\langle \mu, \varphi \rangle)] \\
& \qquad\qquad \cdot \frac{(j_2-j_1)\varepsilon_2 \mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_2^2} \\
+ \quad & \frac{1}{2} \sum_{j_1=0}^{\infty} \sum_{j_2 \geq j_1}^{\infty} [f(\langle \mu, \varphi \rangle + \varepsilon_2 \varphi_\ell(j_1\varepsilon_1, j_2\varepsilon_2))) - f(\langle \mu, \varphi \rangle)] \frac{\mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_1\varepsilon_2} \\
+ \quad & \frac{1}{2} \sum_{j_1=0}^{\infty} \sum_{j_2 \geq j_1}^{\infty} [f(\langle \mu, \varphi \rangle - \varepsilon_2 \varphi_\ell(j_1\varepsilon_1, j_2\varepsilon_2))) - f(\langle \mu, \varphi \rangle)] \frac{\mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_1\varepsilon_2} \\
+ \quad & \frac{1}{2}(1-\varepsilon_1 c_\ell) \sum_{j_1=1}^{\infty} \sum_{j_2 \geq j_1}^{\infty} \\
& \cdot [f(\langle \mu, \varphi \rangle - \varepsilon_1 \varphi_\ell(j_1\varepsilon_1, j_2\varepsilon_2) + \varepsilon_1 \varphi_\ell((j_1+1)\varepsilon_1, (j_2+1)\varepsilon_2)) - f(\langle \mu, \varphi \rangle)] \frac{j_1 \mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_1\varepsilon_2} \\
+ \quad & \frac{1}{2}(1+\varepsilon_1 c_\ell) \sum_{j_1=1}^{\infty} \sum_{j_2 \geq j_1}^{\infty} \\
& \cdot [f(\langle \mu, \varphi \rangle - \varepsilon_1 \varphi_\ell(j_1\varepsilon_1, j_2\varepsilon_2) + \varepsilon_1 \varphi_\ell((j_1-1)\varepsilon_1, (j_2-1)\varepsilon_2))) - f(\langle \mu, \varphi \rangle)] \frac{j_1 \mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_1\varepsilon_2} \\
+ \quad & \frac{1}{2} \sum_{j_1=0}^{N^{\frac{2\ell-1}{4}}} \sum_{j_2 \geq j_1}^{\infty} \\
& \cdot [f(\langle \mu, \varphi \rangle - \varepsilon_2 \varphi_{\ell+1}(j_2\varepsilon_1) + \varepsilon_2 \varphi_{\ell+1}((j_2+1)\varepsilon_2)))) - f(\langle \mu, \varphi \rangle)] \frac{j_2 \mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_1\varepsilon_2} \\
+ \quad & \frac{1}{2} \sum_{j_1=0}^{N^{\frac{2\ell-1}{4}}} \sum_{j_2 \geq j_1}^{\infty} \\
& \cdot [f(\langle \mu, \varphi \rangle - \varepsilon_2 \varphi_{\ell+1}(j_2\varepsilon_2) + \varepsilon_2 \varphi_{\ell+1}((j_2-1)\varepsilon_2))) - f(\langle \mu, \varphi \rangle)] \frac{j_2 \mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_1\varepsilon_2} \\
+ \quad & \frac{1}{2} \sum_{j_1>N^{\frac{2\ell-1}{4}}}^{\infty} \sum_{j_2 \geq j_1}^{\infty}
\end{aligned}
$$

$$\cdot [f(\langle \mu, \varphi \rangle - \varepsilon_1 \varphi_\ell(j_1\varepsilon_1, j_2\varepsilon_2) + \varepsilon_1 \varphi_\ell(j_1\varepsilon_1, (j_2+1)\varepsilon_2)) - f(\langle \mu, \varphi \rangle)] \frac{j_2 \mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_1\varepsilon_2}$$

$$+ \quad \frac{1}{2} \sum_{j_1 > N^{\frac{2\ell-1}{4}}}^{\infty} \sum_{j_2 \geq j_1}^{\infty}$$

$$\cdot [f(\langle \mu, \varphi_\ell \rangle - \varepsilon_1 \varphi_\ell(j_1\varepsilon_1, j_2\varepsilon_2) + \varepsilon_1 \varphi_\ell(j_1\varepsilon_1, (j_2-1)\varepsilon_2))) - f(\langle \mu, \varphi \rangle)] \frac{j_2 \mu(j_1\varepsilon_1, j_2\varepsilon_2)}{\varepsilon_1\varepsilon_2}.$$

We give a brief explanation of these terms.

1. This term corresponds to the migration into $B_\ell^{(N)}$ of individuals from families currently minimally occupying $B_\ell^{(N)}$. Asymptotically as $N \to \infty$, this is

$$c_\ell f'(\langle \mu, \varphi \rangle) \int_{[0,\varepsilon_0^{1/4})} \int (\frac{\partial}{\partial x} \varphi_\ell(x,y))\, y\, \mu_{\ell+1}(dx, dy),$$

whose limit (by (5.3.5)) is

$$c_\ell f'(\langle \mu, \varphi_\ell \rangle) \int (\frac{\partial}{\partial x} \varphi_\ell(x,y)|_{x=0})\, y\, \mu^0_{\ell+1}(dy).$$

2. The second term is similar to the first except that here only families having more than minimal mass in $B_\ell^{(N)}$ appear. In the limit $N \to \infty$ we have asymptotically

$$c_\ell f'(\langle \mu, \varphi \rangle) \int_{x > \varepsilon_0^{1/4}} \int \frac{\partial}{\partial x} \varphi_\ell(x,y) \frac{y}{N^{1/2}}\, \mu_{\ell,N}(dx, dy) \tag{5.3.8}$$

$$\leq c_\ell |f'(\langle \mu, \varphi \rangle)| \sup |\frac{\partial}{\partial x} \varphi_\ell(x,y)|$$

$$\Big(\int \int_{y > N^{\frac{1}{8}}} y\mu(dx, dy) + \frac{N^{1/8}}{N^{1/2}} \int_{N^{-1/4}}^{\infty} \int_0^{N^{1/8}} \mu_{\ell,N}(dx, dy) \Big)$$

$$\leq c_\ell |f'(\langle \mu, \varphi \rangle)| \sup |\frac{\partial}{\partial x} \varphi_\ell(x,y)|$$

$$\Big(\int \int_{y > N^{\frac{1}{8}}} y\mu(dx, dy) + N^{-1/8} \int_0^{\infty} \int_0^{\infty} x\mu_{\ell,N}(dx, dy) \Big)$$

$$= o(1) + O(\frac{1}{N^{1/8}}) \tag{5.3.9}$$

since $\int \int_{x>0} \int x\, \mu_{\ell,N}(dx, dy) = O(1)$.

3. and 4. These two terms arise from the family level branching. If $|\varphi_\ell(x,y)| \leq \text{const} \cdot x$, then asymptotically they yield

$$\frac{1}{2N^{1/2}} f''(\langle \mu, \varphi \rangle) \int \int \varphi^2(x,y) \mu(dx, dy)$$

$$
\begin{aligned}
&= \frac{1}{2N^{1/2}} f''(\langle \mu, \varphi \rangle) \int\int \varphi_{\ell+1}^2(y)\mu(dx,dy) \\
&+ f''(\langle \mu, \varphi \rangle) \int_0^{N^{-1/4}} \int \varphi_{\ell+1}(y)\varphi_\ell(x,y)\mu(dx,dy) \\
&+ f''(\langle \mu, \varphi \rangle) \int \varphi_\ell^2(x,y)\mu_\ell(dx,dy).
\end{aligned}
$$

5. and 6. This corresponds to the critical birth and death of individuals in $B_\ell^{(N)}$. The limiting term is

$$f'(\langle \mu, \varphi_\ell \rangle)\langle \mu, G_{c_\ell}^{(2)}\varphi_\ell \rangle,$$

where $G_{c_\ell}^{(2)}$ denotes the application of the operator G_{c_ℓ} to the x variable.

7. and 8. These terms correspond to the birth and death of individuals in $B_{\ell+1}^{(N)}$ in families that do not or sparsely occupy $B_\ell^{(N)}$. Asymptotically we obtain

$$\frac{1}{N^{1/2}} f'(\langle \mu, \varphi \rangle)\langle \mu, y\frac{\partial^2}{\partial y^2}\varphi_{\ell+1}(y)\rangle.$$

9. and 10. These terms correspond to the birth and death of individuals in $B_{\ell+1}^{(N)}$ in families which have at least a minimal number of members in $B_\ell^{(N)}$. Asymptotically we obtain from the proof of (5.3.5),

$$f'(\langle \mu, \varphi \rangle) \int_{x > \varepsilon_0^{1/4}} \frac{1}{N^{1/2}} \mu_\ell(dx,dy) y \frac{\partial^2}{\partial y^2}\varphi_\ell(x,y) = O(\frac{1}{N^{1/4}}).$$

Collecting the limiting terms as $N \to \infty$, we obtain $\mathfrak{G}^{(N)}F(\mu) \to \mathfrak{G}_\ell F(\mu)$ where

$$
\begin{aligned}
\mathfrak{G}_\ell F(\mu) \quad &= f'(\langle \mu, \varphi \rangle)\langle \mu_\ell, G_{c_\ell}^{(2)}\varphi_\ell \rangle + \tfrac{1}{2} f''(\langle \mu, \varphi_\ell \rangle)\langle \mu_\ell, \varphi_\ell^2 \rangle \\
&+ f'(\langle \mu, \varphi_\ell \rangle) c_\ell \textstyle\int \frac{\partial \varphi_\ell}{\partial x}(x,y)|_{x=0} y \mu_{\ell+1}^0(dy).
\end{aligned} \tag{5.3.10}
$$

We conclude that for any limit point of the probability laws of $\tilde{\mathcal{Y}}^{(N)}$,

$$\tilde{F}(\eta(t)) - \int_{t_0}^t \mathfrak{G}_\ell \tilde{F}(\eta(s))ds$$

is a martingale where $\tilde{F} \in C(M^{TL+})$ is defined by $\tilde{F}(\mu) = f(\langle \mu, \varphi_\ell \rangle + \langle \mu^0, \varphi_{\ell+1} \rangle)$ and f, φ_ℓ and $\varphi_{\ell+1}$ satisfy the same conditions as above. But then for $0 \le a < b \le \infty$, $\eta(t, dx, (a,b))$ satisfies with the martingale problem of the two-level branching diffusion with constant multitype immigration source with immigration from 0 of type $y \in (a,b)$ at rate given by $y\mu_{\ell+1}^0(dy)$ and total immigration rate of $\int_a^b y\mu_{\ell+1}^0(dy)$ (recall Proposition 3.2.1) which is well posed and determines a $M^1((0,\infty)$-valued diffusion process. This completes the proof of the proposition.

■

5.4 The limiting multiscale transition function

The main result of the section is the identification of the limiting multiscale structure, which serves to complete the proof of Theorem 2.4.1.

Theorem 5.4.1 *Fix $j \geq 1$ and $\zeta^*_{j+1} > 0$. Conditioned on $\zeta^{(N,\theta)}_{j+1}(0) = \zeta^{*(N)}_{j+1}$, and provided that $\zeta^{*(N)}_{j+1} \to \zeta^*_{j+1}$ as $N \to \infty$, then the normalized equilibrium masses $\{\zeta^{(N,\theta)}_\ell(0)\}_{\ell=j+1,j,\ldots,1}$ in a sequence of nested blocks $\{B^{(N)}_\ell\}_{\ell=1,\ldots,j+1}$ converge as $N \to \infty$ in distribution to the backward Markov chain $\{\zeta_\ell\}_{\ell=j+1,j,\ldots,1}$ with transition kernel*

$$P(\zeta_\ell \in A|\zeta_{\ell+1} = a) = \Pi^{(2)}_\ell(a, A),$$

where $\Pi^{(2)}_\ell$ is as in Proposition 2.1.1 b), and $\zeta_{j+1} = \zeta^*_{j+1}$.

Proof. The principal step in proving this is given by Proposition 5.3.2. Then to verify that the limiting equilibrium distribution as $N \to \infty$ of $\zeta^{(N,\theta)}_\ell$ conditioned on $\zeta^{(N,\theta)}_{\ell+1} = a$ is given by $\Pi^{(2)}_\ell(a, dx)$ we then follow the argument in Corollary 3.2.2(b). The convergence of $\{\zeta_\ell(0)\}_{\ell=j+1,j,\ldots,0}$ to the Markov chain then follows by recursion and the continuity of the mappings $a \to \Pi^{(2)}_\ell(a, dx)$.

■

Combining Theorem 5.4.1 with the spatial ergodic theorem (Theorem 5.1.7) and Proposition 2.1.1 b), we see that the interchange of the limits $N \to \infty$ and $j \to \infty$ is justified and completes the proof of Theorem 2.4.1.

5.5 The particle level picture

In the previous section the main result is obtained using the convergence of the solutions of the appropriate sequence of martingale problems. Some additional understanding of the limiting process can be obtained by examining the particle picture in both $B^{(N)}_{\ell+1}$ and $B^{(N)}_\ell$ which we now briefly sketch in an informal way. First note that the change in the population in $B^{(N)}_{\ell+1}$ due to movement of particles between $B^{(N)}_{\ell+1}$ and its exterior in time scale $N^{\ell/2}$ has expected value $o(\frac{1}{N})$ and therefore is negligible.

For the moment we fix $t_0 < 0$ and consider the contributions to $\zeta^{(N,\theta)}_\ell(0)$ coming from immigrants arriving in the interval $[t_0 N^{\ell/2}, 0)$. Next we recall that by (5.2.1) all but $O(\frac{1}{K})$ of the population in $B^{(N)}_{\ell+1}$ at time 0 is contained in families with sizes in $(0, KN^{(\ell+1)/2})$. Therefore we can subdivide the population at time 0 into $O(N^{(\ell+1)/2})$ independent subpopulations of size $O(N^{(\ell+1)/2})$ where if necessary we group together smaller families to form subpopulations of size $O(N^{(\ell+1)/2})$. (This ensures independent level two branching in $B^{(N)}_\ell$ for clusters coming from distinct families in this subdivision.) Since the families undergo critical branching these are the descendants of $O(N^{1/2})$ ancestral families at time $t_0 N^{\ell/2}$ and in the time interval $(t_0 N^{\ell/2}, 0)$ these undergo binary branching with $n_N(s)$ branches at time $sN^{\ell/2}$. Now consider a single family in $B_{\ell+1}^{(N)}$ containing $O(N^{(\ell+1)/2})$ individuals. By the analogue of Lemma 5.3.1 the normalized population size of this family is constant in the time scale $N^{\ell/2}$.

We subdivide the interval $[N^{\ell/2}t_0, 0)$ into $M(\in \mathbb{N})$ subintervals of equal length $\frac{1}{M}t_0 N^{\ell/2}$. Let $t_k := \frac{M-k}{M}t_0$ and consider an interval $(t_{k+1}N^{\ell/2}, t_k N^{\ell/2})$. Recall that the rate of migration of individuals into $B_\ell^{(N)}$ from $B_{\ell+1}^{(N)}$ is $c_\ell N^{-\ell/2}$. Therefore in the time interval the number of individuals to immigrate into the ball $B_\ell^{(N)}$ from a family of size $mN^{(\ell+1)/2}$ in $B_{\ell+1}^{(N)}$ is Poisson with mean

$$\begin{aligned} &mN^{(\ell+1)/2} \times \frac{c_\ell}{N^{\ell/2}}((t_k - t_{k+1})N^{\ell/2}) \times \frac{1}{N} \\ &= c_\ell m (t_k - t_{k+1}) N^{(\ell-1)/2}. \end{aligned} \tag{5.5.1}$$

Now consider the question of which of these have descendants alive in $B_\ell^{(N)}$ at time zero. Let $U := \min\{s : n_N(s) = 2\}$. In order to determine this recall that from the structure of the genealogy of the critical branching cluster (see e.g. [Du],[F]) the random variable U is uniform distributed on $[0,1]$. Moreover the probability that an initial individual produces a non-empty set of descendants in a time interval of length $U|t_k|N^{\ell/2}$ is asymptotically as $N \to \infty$

$$\frac{2c_\ell e^{-c_\ell U t_k}}{N^{\ell/2}(1 - e^{-c_\ell U t_k})}.$$

(see e.g. [AN](Chapt. 3)). Therefore the probability that any of the $O(N^{(\ell-1)/2})$ immigrants arriving from a family of size $N^{(\ell+1)/2}$ in $[t_{k+1}N^{\ell/2}, t_k N^{\ell/2})$ have descendants alive at time 0 is no larger than

$$\frac{2c_\ell e^{-c_\ell u_0(N) t_k}}{N^{1/2}(1 - e^{-c_\ell u_0(N) t_k})} + P(U < u_0(N))$$

for any $0 < u_0(N) < 1$. Again by ([Du],[F]) the number of (family) branches $n_N(t_k)$ at time $t_k N^{\ell/2}$ converges in distribution as $N \to \infty$ for each of the $O(N^{1/2})$ ancestral families and the family trees from the different ancestral families are independent. Choosing $u_0(N) \to 0$ such that $N^{1/2}u_0(N) \to \infty$ we conclude that there is a Poisson [with mean of order $O((t_k - t_{k+1})e^{-c_\ell t_k})$] number of individuals who immigrate in $(N^{\ell/2}t_0, N^{\ell/2}t_k]$ producing descendants at time 0, and these all come from different branching trees.

Therefore the population at time zero coming from immigrants arriving in $(N^{\ell/2}t_0, N^{\ell/2}t_k)$ is asymptotically composed of $O(1)$ two-level families each originating from one individual and these all come from different independent subpopulations in $B_{\ell+1}^{(N)}$. In particular the expected mass coming from clusters containing two or more immigrants from the same subpopulation in $B_{\ell+1}^{(N)}$ is $O(\frac{1}{N^{1/2}})$. Each two-level cluster coming from an immigrating individual develops by two-level branching, namely, family level branching inherited from the family branching in $B_{\ell+1}^{(N)}$ and subcritical individual level branching in $B_\ell^{(N)}$. Finally, in both cases the total migration rate of individuals into $B_\ell^{(N)}$ from $B_{\ell+1}^{(N)}$ is

$$(\zeta_{\ell+1}^{(N,\theta)}(N^{\ell/2}t_0)N^{\ell+1}) \times (c_\ell N^{-\ell/2}) \times N^{-1} = c_\ell N^{\ell/2}\zeta_{\ell+1}^{(N,\theta)}(N^{\ell/2}t_0)$$

where on the left hand side the first factor is the number of particles in $B_{\ell+1}^{(N)}$, the second is the individual migration rate to a point chosen randomly in $B_{\ell+1}^{(N)}$ and the last factor is the probability that the tagged ball is chosen. Therefore, asymptotically as $N \to \infty$, the population at time zero consists of clusters of descendants of individuals that immigrate into $B_\ell^{(N)}$ during the time interval $(-\infty, 0]$ at rate $c_\ell N^{\ell/2}\zeta_{\ell+1}^{(N,\theta)}(N^{\ell/2}t_0)$ and subsequently undergo two-level branching. In the limit $N \to \infty$ these clusters correspond to the jumps of the subordinator $S_\ell(\cdot)$ defined in section 4.2.

Remark 5.5.1 *One can also gain some understanding of the convergence to equilibrium from a spatially homogeneous initial population (more general than that addressed in Proposition 2.3.5) with intensity $\theta > 0$. Two ingredients are involved in the convergence to an equilibrium with intensity θ. The first is the strong transience condition on the random walk. The other feature is the structure of the local family sizes. We see from the above that the property that the contribution of families in $B_{\ell+1}^{(N)}$ containing a number of individuals larger than $O(N^{(\ell+1)/2})$ to the equilibrium population in $B_\ell^{(N)}$ is asymptotically negligible and this property is then inherited by $B_\ell^{(N)}$. However if the initial family sizes are too large this iteration can degenerate due to the family level critical branching and the limiting population is locally degenerate. For a more detailed analysis of this phenomenon, see [GH].*

6 Appendix

6.1 Size-biasing and Palm distributions

Definition 6.1.1 *If π is a measure on some measurable space M, and $\mathfrak{s}$ is a nonnegative measurable function on M with $0 < \bar{s} := \int \mathfrak{s}(z)\pi(dz) < \infty$, then we call the probability measure $\hat{\pi}$ given by*

$$\hat{\pi}(dz) := \frac{\mathfrak{s}(z)}{\bar{s}}\pi(dz)$$

the size-biasing of π with respect to $\mathfrak{s}$. (Here, we think of $\mathfrak{s}(z)$ as measuring the size of the object z.)

Remark 6.1.2 *a) An example of size-biasing which is important in our context arises as follows. Let* $\pi = \mathcal{L}(\eta)$ *be the distribution of a random measure* η *(on some Polish space* E*, say). Denote the intensity measure of* η *by* λ*, and fix a nonnegative measurable function* f *on* E *with* $0 < \langle \lambda, f \rangle < \infty$*. Define the size of a measure* m *on* E *by* $\mathfrak{s}(m) := \int f(x)m(dx) = \langle m, f \rangle$*, and denote by* π_f *the size-biasing of* $\pi(dm)$ *with* $\langle m, f \rangle$*.*

b) Assuming that λ *is locally finite, choosing* $f = 1_B$ *where* B *is a ball in* E*, and letting* B *shrink gives the family of Palm distributions* $\pi_x, x \in S$*. Formally, these arise as the disintegration of the measure* $\pi(dm)m(dx)$ *with respect to its second marginal* $\mathbb{E}\eta$*, that is*

$$\mathbb{E}G(\eta)\langle \eta, h \rangle = \int h(x)G(m)\pi_x(dm)\lambda(dx). \tag{6.1.1}$$

See [K], chapter 10, for more background; there, a random measure whose distribution is the size-biasing of $(\mathcal{L}(\eta))(dm)$ *with* $\langle m, f \rangle$ *is denoted by* η_f*.*

c) The following fact ([K], formula (10.6)) is immediate from (6.1.1):

The size-biasing π_f *of* $\pi(dm)$ *with* $\langle m, f \rangle$ *is*

$$\frac{1}{\langle \lambda, f \rangle} \int f(x)\pi_x(.)\lambda(dx) = \mathbb{E}\pi_{\hat{X}}, \tag{6.1.2}$$

where $\hat{X}$ *is a random element in* E *whose distribution is the size-biasing of* λ *with* f*.*

d) If in the just described situation E *consists of one element only, then the finite random measures on* E *are* $\mathbb{R}_+$*-valued random variables. When speaking of the size-biasing of a measure* π *on* $\mathbb{R}_+$ *without specifying the size function, we always mean the size-biasing of* $\pi(dx)$ *with* x*.*

Let us write Π_λ for the distribution of a Poisson random counting measure on E with intensity measure λ. It is well known (see e.g. [K], beginning of chapter 11) that the Palm distributions of Π_λ arise as the distributions of $\Phi + \delta_x$, $x \in E$, where $\mathcal{L}(\Phi) = \Pi_\lambda$.

Now let σ be a probability measure on $\mathbb{R}_+$ with $m_\sigma := \int \tau\sigma(d\tau) \in (0, \infty)$, and write

$$\Pi_{\sigma,\lambda} := \int \Pi_{\tau\lambda}\sigma(d\tau)$$

for the *mixed Poisson distribution* with mixing measure σ.

The following lemma, whose proof we include for convenience, is part of a characterization theorem ([K], Theorem 11.5) of mixed Poisson processes.

Lemma 6.1.3 *The Palm distributions of* $\Pi_{\sigma,\lambda}$ *arise as the distributions of* $\Phi + \delta_x$*,* $x \in E$*, where* $\mathcal{L}(\Phi) = \Pi_{\hat{\sigma},\lambda}$*, and* $\hat{\sigma}$ *is the size-biasing of* σ*.*

Proof. Using the above mentioned form of the Palm distributions of $\Pi_{\tau\lambda}$, we obtain for all nonnegative measurable h and G defined on E and $M_c(E)$, the space of locally finite counting measures on E, respectively:

$$\begin{aligned}\int G(\psi)\langle\psi,h\rangle\Pi_{\sigma,\lambda}(d\psi) &= \int\int G(\psi)\langle\psi,h\rangle\Pi_{\tau\lambda}(d\psi)\sigma(d\tau) \qquad (6.1.3)\\ &= \int\int \tau\lambda(dx)h(x)G(\psi+\delta_x)\Pi_{\tau\lambda}(d\psi)\sigma(d\tau)\\ &= \int m_\sigma\lambda(dx)h(x)\int G(\psi+\delta_x)\Pi_{\tau\lambda}(d\psi)\frac{1}{m_\sigma}\tau\sigma(d\tau).\end{aligned}$$

■

Corollary 6.1.4 *Assume $0 < \langle\lambda,h\rangle < \infty$. Then the size-biasing of $\Pi_{\sigma,\lambda}(d\psi)$ with $\langle\psi,h\rangle$ arises as the distribution of $\Phi+\delta_{\hat{X}}$, where $\mathcal{L}(\Phi) = \int \Pi_{\tau\lambda}\hat{\sigma}(d\tau)$, $\mathcal{L}(\hat{X})$ is the size-biasing of λ with h, and Φ and $\hat{X}$ are independent.*

Proof. This can be seen either by combining Remark 6.1.2 c) and Lemma 6.1.3, or by dividing (6.1.3) through $\mathbb{E}_{\sigma,\lambda}\langle\Psi,h\rangle = m_\sigma\langle\lambda,h\rangle$. ■

6.2 Subcritical Feller branching

Let us fix $c > 0$. In the following, X will denote a *c-subcritical Feller branching diffusion (c-FBD) process.* In other words, X is an $[0,\infty)$-valued diffusion process satisfying

$$dX_t = \sqrt{X_t}\,dW_t - cX_t dt \qquad (6.2.1)$$

where W_t is Brownian motion. For $\varepsilon > 0$, let X^ε be the c-FBD process starting in ε at time 0.

From well-known results on Galton-Watson processes conditioned to survival [AN, Ge] and cluster sizes in continuous-state branching processes [D], one expects that the conditional law $\mathcal{L}(X_t^\varepsilon | X_t^\varepsilon > 0)$ converges to an exponential distribution as $\varepsilon \to 0$. The following calculation verifies this and identifies the parameter.

The Laplace functional of X_t^ε is given by

$$\mathbb{E}(e^{-\lambda X_t^\varepsilon}) = e^{-\varepsilon v(t,\lambda)}, \quad \lambda \geq 0, \qquad (6.2.2)$$

where $v = v(t,\lambda)$ is the solution of

$$\frac{\partial v(t,\lambda)}{\partial t} = -cv(t,\lambda) - \frac{1}{2}v^2(t,\lambda), \quad v(0,\lambda) = \lambda. \qquad (6.2.3)$$

The solution of (6.2.3) is given by

$$v(t,\lambda) = \frac{2\lambda c e^{-ct}}{\lambda(1-e^{-ct})+2c}. \qquad (6.2.4)$$

Combining (6.2.2) and (6.2.4) one obtains by a straightforward calculation

$$\mathbb{P}(X_t^\varepsilon \neq 0) = 1 - \exp\left(\frac{-2\varepsilon c e^{-ct}}{1-e^{-ct}}\right), \tag{6.2.5}$$

$$\mathbb{E}X_t^\varepsilon = \varepsilon e^{-ct}, \quad \varepsilon > 0 \tag{6.2.6}$$

and

$$\lim_{\varepsilon\to 0} \mathbb{E}(e^{-\lambda X_t^\varepsilon} | X_t^\varepsilon \neq 0) = \frac{2c}{\lambda(1-e^{-ct})+2c}. \tag{6.2.7}$$

Writing $\mathrm{Exp}(u)$ for the exponential distribution with parameter u, we obtain immediately:

Lemma 6.2.1 *Fix $t > 0$.*
a)

$$\mathcal{L}(X_t^\varepsilon | X_t^\varepsilon > 0) \Rightarrow \mathrm{Exp}\left(\frac{2c}{1-e^{-ct}}\right) \qquad \textit{as } \varepsilon \to 0. \tag{6.2.8}$$

b)

$$\varepsilon^{-1}\mathbb{P}[X_t^\varepsilon > 0] \to \frac{2ce^{-ct}}{1-e^{-ct}} \qquad \textit{as } \varepsilon \to 0. \tag{6.2.9}$$

For $\varepsilon > 0$, let

$$\bar{X}^\varepsilon = \sum_{i=1}^{N_\varepsilon} X^{\varepsilon,i}, \tag{6.2.10}$$

where N_ε is a Poisson(ε^{-1})-random variable and $X^{\varepsilon,1}, X^{\varepsilon,2}, \ldots$ are independent copies of X^ε. The following lemma is an easy consequence of Lemma 6.2.1.

Lemma 6.2.2 *Fix $t > 0$. Then*

$$\bar{X}_t^\varepsilon \Rightarrow \bar{X}_t^0 \;\; \textit{as } \varepsilon \to 0, \tag{6.2.11}$$

where $\bar{X}_t^0$ is infinitely divisible with canonical measure

$$\kappa_t := \frac{2ce^{-ct}}{1-e^{-ct}} \mathrm{Exp}\left(\frac{2c}{1-e^{-ct}}\right). \tag{6.2.12}$$

Let us now define the measure

$$\gamma_c := c\int_0^\infty \kappa_s \, ds. \tag{6.2.13}$$

An elementary calculation based on (6.2.12) and (6.2.13) shows that

$$\gamma_c(dx) = 2c\frac{1}{x}e^{-2cx}dx, \quad x > 0. \tag{6.2.14}$$

This identifies γ_c as the canonical measure of the Gamma$(2c, 2c)$-distribution, and goes along with the well-known fact that the equilibrium distribution of (4.1.1) is the Gamma$(2ca, 2c)$-distribution. We note that γ_c is the Lévy measure of the Gamma subordinator $S(\tau)$, $\tau \geq 0$, with scale parameter $2c$ and $\mathbb{E}S(1) = 1$. The following is obvious from (6.2.14):

Remark 6.2.3 *a)* $\int_0^\infty x\gamma_c(dx) = 1$.
b) $\int_0^\infty x^2\gamma_c(dx) = 1/2c$.

Finally we consider the semigroup (T_t) be the semigroup of the c-FBD process. Recalling (6.2.8), (6.2.9) and (6.2.12), the (T_t)-entrance law (κ_t) from 0 is given by

$$\kappa_t = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \delta_\varepsilon T_t = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathcal{L}(X_t^\varepsilon; X_t^\varepsilon \neq 0), \quad t > 0, \tag{6.2.15}$$

where (X_t^ε) is the c-FBD-process starting in ε at time 0. Because of (6.2.12), $\kappa_t(dx)$ has density

$$\kappa_t(x) = \frac{(2c)^2 e^{-ct}}{(1 - e^{-ct})^2} \exp\left(-\frac{2cx}{1 - e^{-ct}}\right), \quad x \in (0, \infty). \tag{6.2.16}$$

Then for $f \in C_1((0,\infty))$ and $t > 0$,

$$\frac{d}{dx} T_t f(x)\Big|_{x=0} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} T_t f(\varepsilon) = \int f(y)\kappa_t(dy). \tag{6.2.17}$$

Acknowledgements The authors thank the hospitality of The Fields Institute (Toronto, Canada), Carleton University (Ottawa, Canada), the Center for Mathematical Research (CIMAT, Guanajuato, Mexico), Erwin Schrödinger Institute (Vienna, Austria) and the Johann Wolfgang Goethe University (Frankfurt, Germany), where mutual working visits took place. L.G.G. also thanks the Institute of Mathematics, National University of Mexico (UNAM), where he spent a sabbatical during 2002.

References

[AN] K. B. Athreya and P. E. Ney, Branching processes, Springer 1972.

[BLG] J. Bertoin and J.F. Le Gall, The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. Probab. Theory and Related Fields 117 (2000), 249–266.

[D] D. A. Dawson, Measure-valued Markov Processes, in: Ecole d'Été de Probabilités de Saint-Flour 1991, Lecture Notes in Mathematics 1541, pp. 1-260, Springer 1993

[DGW1] D. A. Dawson, L. G. Gorostiza and A.Wakolbinger, Occupation time fluctuations in branching systems. J. Theor. Probab. 14 (2001), 729-796.

[DGW2] D. A. Dawson, L. G. Gorostiza and A. Wakolbinger, Degrees of transience and recurrence and hierarchichal random walks, Potential Analysis, to appear.

[DGW3] D.A. Dawson, L.G. Gorostiza and A. Wakolbinger, Hierarchical random walks, in "Asymptotic Methods in Stochastics", Fields Institute Communications and Monograph Series, Amer. Math. Soc., to appear.

[DG1] D. A. Dawson and A. Greven, Multiple time scale analysis of interacting diffusions. Probab. Theory Related Fields 95 (1993), 467-508.

[DG2] D. A. Dawson and A. Greven, Multiple space–time scale analysis for interacting branching models. Electronic J. Probab. 1 (1996), paper no 14.

[DH] D. A. Dawson and K. J. Hochberg, A multilevel branching model. Adv. Appl. Prob. 23 (1991), 701-715.

[DLM] D.Dawson, Y. Li and C. Mueller, The support of measure-valued branching processes in a random environment, Ann. Probab. 23 (1995), 1692-1718.

[DP1] D. A. Dawson and E. Perkins, Historical Processes. Memoirs of the AMS 454 (1991).

[DP2] D. A. Dawson and E. Perkins, Measure-valued processes and renormailzation of branching particle systems. In: Carmona, R.A., and Rozovskii, B. (eds.), Stochastic Partial Differential Equations: Six Perspectives, pp. 45-106. Math. Surveys and Monographs, Vol. 64, AMS 1999.

[DZ] D.A. Dawson and X. Zheng, Law of large numbers and central limit theorem for unbounded jump mean-field models, Adv. in Applied Math. 12 (1991), 293-326.

[Du] R. Durrett, The genealogy of critical branching processes, Stochastic Process. Appl. 8 (1978), 101-116.

[EK] S. N. Ethier and T. G. Kurtz, Markov Processes: Characterization and Convergence. Wiley, New York 1986.

[F] K. Fleischmann and R. Siegmund-Schultze, The structure of reduced critical Galton-Watson processes, Math. Nachr. 79 (1977) , 233-241.

[Ge] J. Geiger, Elementary new proofs of classical limit theorems for Galton-Watson processes, J. Appl. Prob. 36 (1999), 301-309.

[Gr] A. Greven, A phase transition for the coupled branching process, I. The ergodic theory in the range of finite second moments. Probab. Theory Related Fields, 87 (1991), 417-458.

[GH] A. Greven and K.J. Hochberg, New Behavioral Patterns for Two-Level Branching Systems. In: Stochastic Models, (L. G. Gorostiza and G. Ivanov, eds.), pp. 205-215, CMS Conference Proceedings and Lecture Notes, Vol. 26, AMS 2000.

[GHW] L. Gorostiza, K. J. Hochberg and A. Wakolbinger, Persistence of a critical super-2 process. J. Appl. Probab. 32 (1995), 534-540.

[GW] L. Gorostiza and A. Wakolbinger, Persistence criteria for a class of critical branching particle systems in continuous time. Ann. Probab. 19 (1995), 266-288.

[IW] N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes, North Holland 1989.

[K] O. Kallenberg, Random Measures. 4th ed., Akademie-Verlag, Berlin and Academic Press, London, 1986.

[K2] O. Kallenberg, Foundations of Modern Probability, 2nd ed., Springer, 2002.

[KS] N. Konno and T. Shiga, Stochastic partial differential equations for some measure-valued diffusions. Probab. Th. Rel. Fields 79 (1988), 34-51.

[LS] Z. Li and T. Shiga, Measure-valued branching diffusions: immigrations, excursions and limit theorems. J. Math. Kyoto Univ. 35 (1995), 233-274.

[LMW] A. Liemant, K. Matthes and A. Wakolbinger, Equilibrium Distributions of Branching Processes, Akademie Verlag, Berlin, and Kluwer Academic Publishers, Dordrecht, 1988.

[Lin] E. Lindenstrauss, Pointwise theorems for amenable groups, Invent. Math. 146 (2001), 259-295.

[MKM] K. Matthes, J. Kerstan and J. Mecke, Infinitely divisible point processes, Wiley, 1978.

[SaW] K. I. Sato and T. Watanabe. Moments of last exit times for Lévy processes, Ann. Inst. H. Poincaré, Probab. et Stat., 40(2004), 207-225.

[SF] S. Sawyer and J. Felsenstein, Isolation by distance in a hierarchically clusteres population, J. Appl. Prob. 20 (1983), 1-10.

[StW] A. Stöckl and A. Wakolbinger, On clan-recurrence and -transience in time stationary branching Brownian particle systems. In: Dawson, D.A. (ed.), Measure-Valued Processes, Stochastic Partial Differential Equations, and Interacting Systems, pp. 213 - 219, CMS Conference Proceedings and Lecture Notes, Vol. 5, AMS 1994.

[Wu] Y. Wu, A multilevel birth-death particle system and its continuous diffusion, Adv. Appl. Prob. 25(1993), 549-569.