Long-term behavior of a critical multitype spatially homogeneous branching particle process and a related reaction-diffusion system

Klaus Fleischmann Vladimir A. Vatutin*

(WIAS preprint No. 350 of August 05, 1997)

Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstr. 39
D–10117 Berlin, Germany
e-mail: fleischmann@wias-berlin.de
fax: 49-30-204 49 75

Department of Discrete Mathematics
Steklov Mathematical Institute
8 Gubkin Street
117 966 Moscow, GSP-1, Russia
e-mail: vatutin@genesis.mi.ras.ru

1991 Mathematics Subject Classification 60J80, 60J15.

Key words and phrases extinction, survival, persistence, systems of reaction-diffusion equations, steady state, genealogical tree, reduced tree, marked particle

Running head Multitype branching systems

marked.tex typeset by LaTeX

*Supported in part by the Grant RFBR N 96–01-00338, INTAS-RFBR 95-0099, and the DFG

Abstract

A dichotomy concerning extinction and survival (even persistence) at late times is established for critical multitype spatially homogeneous branching particle systems in continuous time under natural conditions on the branching mechanism known from the "classical" processes without motion component. This generalizes results of López-Mimbela and Wakolbinger [LMW96] and others. Our simplified approach is based on some genealogical tree analysis combined with the study of the long-term behavior of L^1–norms of solutions of related systems of reaction-"diffusion" equations, which is perhaps also of some independent interest.

Contents

1 Introduction and results

1.1 Motivation

Reaction-diffusion equations usually describe the concentration of some substances or similar quantities. In the case of systems of such equations the *non-negativity* of solutions is therefore an important property which has to be established ([GH96, GH97]).

There is a class of systems of reaction-diffusion equations which is related to *multitype branching particle systems.* The relation is realized via so-called log-Laplace functionals. Here the diffusion term reflects the motion of marked particles, whereas the reaction term corresponds to the branching mechanism.

There are already a lot of papers where such connection between reaction-diffusion equations and branching systems was exploited in one or another direction. See, for instance, [Daw93, Dyn94], and references therein.

In the case of systems however, we know only a few papers: [GLM90, GRW90, GR91, GRW92, GW92, LM92, GW93, GW94, LMW96]. A common feature there is that, compared with "classical" multitype branching processes (finite particle systems without motion component), more or less strong conditions on the branching mechanism are imposed. One of our aims is to overcome this flaw in studying the long-term behavior of the particle system. This will be achieved by some genealogical tree analysis combined with an investigation of asymptotic properties of L^1–norms of solutions of the related equation system.

Our *main purpose* is to provide a simplified approach to a dichotomy between extinction and survival (even persistence) of the particle systems in low respectively high dimensions of space under natural conditions on the branching mechanism (see Corollary 10 at p.6). Such dichotomy is known from more specific models. In terms of the reaction-diffusion equation systems this concerns the extinction or survival of L^1–norms of non-negative solutions.

1.2 Systems of reaction-diffusion equations

On the equation side, typically we are concerned with systems of equations of the following type.

Definition 1 (system of reaction-diffusion equations) Fix $K \geq 1$. Consider

$$\frac{\partial}{\partial t} U_i \;=\; \Delta_{\alpha_i} U_i - \varrho_i \Big[f_i \left(\mathbf{1} - \mathbf{U} \right) - \left(1 - U_i \right) \Big], \qquad 1 \leq i \leq K. \tag{1}$$

The ingredients of this system are as follows.

(a) ("diffusion" component) For each $i \in \{1, ..., K\} =: \mathsf{K}$, the *fractional power* $-(-\Delta)^{\alpha_i/2}$ of *Laplacian* Δ in R^d is denoted by Δ_{α_i}, where the α_i are constants in $(0, 2]$. Note that Δ_{α_i} is a *differential operator only* in the boundary case $\alpha_i = 2$ (and only in this case the term *diffusion* is actually adequate).

(b) (rates ϱ_i) The constants ϱ_i, $i \in \mathsf{K}$, are (strictly) positive.

(c) (reaction term) Each component $f_i : \mathsf{R}_+^K \to [0, +\infty]$ of the vector $\mathbf{f} = (f_1, \dots, f_K)$ is a *power series* with *non-negative* coefficients. Moreover, it is assumed that the vector $\mathbf{1} = (1, \dots, 1)$ is a *fixed point* for $\mathbf{f}$, that is $\mathbf{f}(\mathbf{1}) = \mathbf{1}$. In other words, for i fixed, f_i is a *probability generating function* of a Z_+^K–valued random vector $(\zeta_{i,1}, \dots, \zeta_{i,K})$:

$$f_i(\mathbf{z}) = E z_1^{\zeta_{i,1}} \cdots z_K^{\zeta_{i,K}}, \qquad \mathbf{z} = (z_1, \dots, z_K) \in [0,1]^K. \tag{2}$$

(d) (solution) A *solution* of (1) is a vector

$$\mathbf{U} = (U_1, \dots, U_K) \text{ of functions } U_i : \mathsf{R}_+ \times \mathsf{R}^d \to [0,1]$$

satisfying (1).

(e) (initial condition) The components $U_i(0) : \mathsf{R}^d \to [0,1]$ of the vector *initial function* $\mathbf{U}(0)$ belong to $\mathcal{C}_+^{\mathrm{comp}} = \mathcal{C}_+^{\mathrm{comp}}(\mathsf{R}^d)$, the set of all non-negative continuous functions with compact support. [1)] $\diamond$

The following example will be used later to compare our results with those of other authors.

Example 2 Let $K = 3$ and, for $\mathbf{z} = (z_1, z_2, z_3) \in [0,1]^3$,

$$\begin{aligned} f_1(\mathbf{z}) &:= \tfrac{3}{4} + \tfrac{1}{4} z_1 \Big[z_2 + c_1 (1 - z_2)^{1+\beta} \Big], \\ f_2(\mathbf{z}) &:= \tfrac{1}{2} + \tfrac{1}{4} z_1^3 + \tfrac{1}{8} z_2^3 + \tfrac{1}{8} z_3^5, \\ f_3(\mathbf{z}) &:= \tfrac{5}{8} + \tfrac{3}{8} z_2 \Big[z_3 + (1 - z_3)^{1+\beta} L(1 - z_3) \Big], \end{aligned}$$

with constants $0 < \beta \le 1$, $0 \le c_1 \le (1+\beta)^{-1}$, and where $L : [0,1] \to \mathsf{R}_+$ is an appropriate (slowly varying) function satisfying

$$L(z) \sim \frac{c_3}{\log \frac{1}{z}} \quad \text{as} \quad z \downarrow 0, \quad \text{with} \quad c_3 > 0. \qquad \diamond$$

Under certain conditions on the f_i, formulated in Assumption 3 in the next subsection, the system (1) of equations has a *unique (non-negative) solution,* for each (admissible) initial vector $\mathbf{U}(0)$. In fact, existence of a non-negative

[1)] If no confusion is possible, we simplify notation and write $U(t)$ instead of $U(t, \cdot)$, for instance.

solution follows from a *probabilistic representation* in terms of a branching particle system (see (18)), and uniqueness by a standard contraction argument, we will skip. The quantities of interest are then the L^1*–norms*

$$\sum_{i=1}^{K} \lambda_i \int \mathrm{d}x \; U_i(t,x), \qquad t \geq 0, \tag{3}$$

as $t \uparrow \infty$, where the λ_i, $i \in \mathsf{K}$, are some positive constants. Under our conditions (see Assumption 3 and Hypothesis 5), the norm in (3) *converges* as $t \uparrow \infty$, and the limit vanishes or is positive in dependence on whether the dimension of space is low or high, respectively (*dichotomy*); see Remark 11. In other words, in low dimensions the reaction wins the competition with the diffusion, whereas in high dimensions the diffusion effect is dominating in this competition.

1.3 Multitype branching particle systems

Our next aim is to describe the related stochastic particle process on an intuitive level. Details on the connection between both models will follow in § 1.5.

Consider particles with types in the finite set $\mathsf{K} = \{1,\dots,K\}$. We assume first of all that they move independently in R^d. In fact, all particles of type $i \in \mathsf{K}$ move according to a *symmetric stable process* of index $\alpha_i \in (0,2]$ (that is a Markov process with generator Δ_{α_i}). Additionally, after an *exponentially distributed* time with expectation $1/\varrho_i$ (recall Definition 1 (b)) a particle of type i "dies". Upon death, it produces children of types $1,\dots,K$ according to the *offspring generating function* f_i as in (2) (that is, $\zeta_{i,j}$ children of type j are born). Offspring instantaneously start from their parent's site and continue to evolve independently and according to the same rules.

Introduce the *matrix of offspring means*

$$\mathrm{M} \;=\; (m_{i,j})_{i,j=1}^{K} \;:=\; \left(\frac{\partial f_i}{\partial z_j}(\mathbf{1})\right)_{i,j=1}^{K}. \tag{4}$$

That is, $m_{i,j}$ is the mean number of particles of type j produced from a particle of type i upon its death.

Assumption 3 (critical mean matrix) Let M be *irreducible* and have the maximal eigenvalue 1 (*criticality*). Let $\mathbf{u} = (u_1,\dots,u_K)$ and $\mathbf{v} = (v_1,\dots,v_K)$ be (positive) right and left *eigenvectors* corresponding to this eigenvalue, i.e.

$$\mathbf{M}\mathbf{u} \;=\; \mathbf{u}, \qquad \mathbf{v}\mathbf{M} \;=\; \mathbf{v}, \tag{5}$$

normalized so that

$$(\mathbf{v},\mathbf{u}) \;:=\; \sum_{i=1}^{K} v_i u_i \;=\; 1, \qquad (\mathbf{1},\mathbf{u}) \;=\; 1. \tag{6}$$

◇

Example 4 Note that $\mathbf{f}$ from Example 2 has the irreducible mean matrix

$$\mathbf{M} = \begin{pmatrix} 1/4 & 1/4 & 0 \\ 3/4 & 3/8 & 5/8 \\ 0 & 3/8 & 3/8 \end{pmatrix} \tag{7}$$

which possesses the eigenvalues 1, $\frac{1}{8}\sqrt{6}$, and $-\frac{1}{8}\sqrt{6}$, hence it meets all the requirements in Assumption 3. Moreover, the corresponding eigenvectors are $\mathbf{u} = (5, 15, 9)/29$ and $\mathbf{v} = (1, 1, 1)$. ◇

Let $N\big(t, B \times \{i\}\big)$ denote the number of particles of type $i \in \mathsf{K}$ which at time $t \geq 0$ are in the Borel set $B \subseteq \mathsf{R}^d$. Set

$$\lambda_i := v_i/\varrho_i\,, \qquad i \in \mathsf{K}, \tag{8}$$

(with v_i from the left eigenvector, and ϱ_i the death rate of a particle of type i). Assume the process starts as a *Poisson particle system* $N(0)$ in $\mathsf{R}^d \times \mathsf{K}$ with intensity measure

$$\Lambda := \lambda_1 \ell \times \cdots \times \lambda_K \ell, \tag{9}$$

with the λ_i from (8) and where ℓ is the (normalized) Lebesgue measure on R^d. That is, at time 0 the particle configurations in disjoint subsets of $\mathsf{R}^d \times \mathsf{K}$ are independent, and the number of particles of type i in the Borel set $B \subseteq \mathsf{R}^d$ is Poissonian with mean $\lambda_i\, \ell(B)$. In particular, the particles' initial intensity is finite. Write $\mathbf{P}_\Lambda$ for the law of $N = \{N(t) : t \geq 0\}$ starting with this Poisson system, and $\mathbf{E}_\Lambda$ for the related expectation. (We use the letter E always according to such a rule.)

For $\mathbf{z} \in [0,1]^K$, put

$$\Phi_i(\mathbf{z}) \;:=\; \sum_{j=1}^{K} m_{i,j}\, z_j \;-\; \big[1 - f_i(\mathbf{1} - \mathbf{z})\big], \tag{10}$$

$$\Psi(\mathbf{z}) \;:=\; \big(\mathbf{v}, \boldsymbol{\Phi}(\mathbf{z})\big) \;=\; \Big(\mathbf{v}\,,\, \mathbf{f}(\mathbf{1} - \mathbf{z}) - (\mathbf{1} - \mathbf{z})\Big). \tag{11}$$

Before we can state our results, we introduce the following *basic hypotheses* on the characteristics of the process N. The first one is exploited in low dimensions d, whereas the second one is of use for large d.

Hypothesis 5 (reproductivity conditions)

(a) (lower bound) There exist a constant $\overline{\beta} \in (0,1]$, a constant $\overline{c} > 0$, and a type $k \in \mathsf{K}$ such that

$$\Psi(\mathbf{z}) \;\geq\; \overline{c}\, z_k^{1+\overline{\beta}}, \qquad \mathbf{z} \in [0,1]^K. \tag{12}$$

(b) (upper bound) There exists a constant $\underline{\beta} \in (0,1]$ and a constant $\underline{c} > 0$ such that

$$\Psi(\mathbf{z}) \;\leq\; \underline{c}\, (\mathbf{1}, \mathbf{z})^{1+\underline{\beta}}, \qquad \mathbf{z} \in [0,1]^K. \tag{13}$$

◇

Remark 6 (interpretation of the hypotheses) Note that in **(a)** the type k is something like the *"worse"* type, which we interpret as follows: The total production of particles of type k by particles of all types is *"at most"* in the normal domain of attraction of a stable law of index $1+\overline{\beta}$ (that is, the law might have still fatter tails). On the other hand, **(b)** is an integral characteristic saying roughly that the total particle production is *"at least"* in the normal domain of attraction of a stable law of index $1+\underline{\beta}$. In particular, here the total production of particles (at death of a parent) has *moments of all orders* $1+\varepsilon < 1+\underline{\beta}$. $\diamond$

Example 7 Our Example 2 satisfies requirements **(a)** *and* **(b)** of Hypothesis 5, if we take $k=2$, assume additionally $c_1 > 0$ in the definition of f_1, and choose $\overline{\beta} \in [\beta, 1]$ and $\underline{\beta} \in (0, \beta)$, respectively. Note that necessarily $\overline{\beta} > \underline{\beta}$. $\diamond$

Set

$$\alpha := \min\{\alpha_1, ..., \alpha_K\}. \tag{14}$$

1.4 Results and discussion

From now on we *always impose the criticality Assumption* 3. Our first result concerns low dimensions:

Theorem 8 (local extinction in low dimensions) *Under Hypothesis* 5(a), *if* $d \le \alpha/\overline{\beta}$, *then for any bounded Borel set* $B \subset \mathsf{R}^d$,

$$\mathbf{P}_\Lambda\Big(N(t, B \times \mathsf{K}) \neq 0\Big) \underset{t\uparrow\infty}{\longrightarrow} 0. \tag{15}$$

In other words, in low dimensions the counting measure-valued process N suffers local extinction in the long-term limit.

We complement the previous theorem by a result concerning high dimensions.

Theorem 9 (persistent convergence in high dimensions) *Under Hypothesis* 5(b), *if* $d > \alpha/\underline{\beta}$, *then* $N(t)$ *converges in law as* $t \uparrow \infty$ *to a limiting counting measure* $N(\infty)$, *say. It is a steady state which has intensity measure* Λ *again* (*persistence*).

Both theorems combined lead immediately to the following conclusion.

Corollary 10 (necessary and sufficient criterion) *If Hypotheses* 5(a) *and* 5(b) *are satisfied for* $\overline{\beta} = \underline{\beta} =: \beta$, *then persistent convergence holds if and only if* $d > \alpha/\beta$, *whereas* N *suffers local extinction in the remaining case.*

In simple terms: In low dimensions the extinction features of critical branching dominate the diffusion of particles, whereas in high dimensions the motion is mobile enough to transport mass "from infinity" to finite regions.

The discovery of a *dichotomy* effect for spatial branching processes goes back to [Lie69], who dealt with a discrete time particle model. A multitype generalization can be found in [PR77]. A characterization in terms of parameters in the model (of the type $d \leq \alpha/\beta$ respectively $d > \alpha/\beta$) was provided in [DF85] (even in a random environment context), using a backward tree technique developed in [Kal77] and [Lie81]. See [DFFP86] for such a parametrized condition in a superprocess setting, and [GW91] in a continuous time particle version. Systems with infinite intensity are dealt with in [BCG93], [BCG97], and [Kle97]. Concerning multitype processes in continuous time under more restrictive assumptions compared with the present paper, references had already been given at the beginning of § 1.1 (p.2).

Note that for the model under consideration the results of the present note, in particular Corollary 10, cover persistence criteria recently due to [GW93] and [LMW96]. Our approach involves some genealogical tree analysis and simple equation tools rather than Palm tree constructions. In particular, we got rid of the following restrictions imposed in the mentioned papers:

(i) M is a *stochastic* matrix (that is $\mathbf{1}$ is a right eigenvector), and M has only (strictly) *positive* entries,

(ii) for *each* $i \in \mathsf{K}$, the total offspring number produced by a particle of type i is in the normal domain of attraction of a stable law of index $1+\beta_i \in (1,2]$.

Recall that the generating functions of Example 2 satisfy our Hypotheses 5 (a) and 5 (b), but note they meet neither condition **(i)** nor **(ii)**.

Theorem 8 will be proved in § 3.4, whereas the proof of Theorem 9 is postponed to § 3.5. The remaining parts of Section 2 and 3 serve as a preparation for these proofs. In particular, with Proposition 17 in § 2.5 a large deviation probability estimate is provided for ancestry lines in the reduced genealogical tree (without spatial structure).

That we start the process from a homogeneous Poisson system is a simplification. In fact, the old results on spatially homogeneous branching processes in discrete time (see [PR77]) make clear, that by approximation one should be able to pass to shift invariant initial laws of appropriate intensities.

For background on classical multitype branching theory, we refer to [AN72], on point processes and discrete-time single type spatially homogeneous branching particle systems to [MKM78, in particular, Chapter 12], and on reaction-diffusion equation systems to [Smo83].

1.5 Connection between the branching model and the equation system

In this subsection we introduce the basic relations between the multitype branching particle model under consideration and the system of reaction-diffusion equations in Definition 1.

For a measure μ on $\mathsf{R}^d \times \mathsf{K}$ and $\mathbf{h} \in \mathbf{C}_+^{\mathrm{comp}} := \left(\mathcal{C}_+^{\mathrm{comp}}\right)^K$, we write

$$\langle \mu, \mathbf{h} \rangle := \int \mu\left(\mathrm{d}[x,i]\right) h_i(x). \tag{16}$$

Similarly, we proceed in the single type case $K = 1$.

Since N starts with a Poisson system of intensity Λ (recall (9)), the log-Laplace functional of $N(t)$ has the form

$$-\log \mathbf{E}_\Lambda \exp\Big[-\langle N(t), \mathbf{h}\rangle\Big] = \Big\langle \Lambda, \mathbf{U}\left(\mathbf{h};t\right)\Big\rangle, \qquad \mathbf{h} \in \mathbf{C}_+^{\mathrm{comp}}. \tag{17}$$

In this case, the components $U_i\left(\mathbf{h};t\right) = U_i\left(\mathbf{h};t,\cdot\,\right)$ of the vector function

$$\mathbf{U}(t) = \mathbf{U}\left(\mathbf{h};t\right) = \Big(U_1\left(\mathbf{h};t\right), \ldots, U_K\left(\mathbf{h};t\right)\Big)$$

are specified by

$$0 \le U_i\left(\mathbf{h};t,x\right) := \mathbf{E}_{x,i}\left[1 - \mathrm{e}^{-\langle N(t),\mathbf{h}\rangle}\right] \le \mathbf{E}_{x,i}\left\langle N(t), \mathbf{h}\right\rangle < \infty, \tag{18}$$

$[x,i] \in \mathsf{R}^d \times \mathsf{K}$. Here $\mathbf{P}_{x,i}$ stands for the law of our branching process N if it started at time 0 from a single particle of type $i \in \mathsf{K}$ situated at site $x \in \mathsf{R}^d$. Denote by G_i the distribution function of the *exponential law* with parameter ϱ_i, and by $t \mapsto S_t^{\alpha_i}$ the *stable semigroup* with generator Δ_{α_i}. Applying standard renewal arguments one gets

$$\begin{aligned} U_i\left(\mathbf{h};t\right) \;&=\; G_i(t)\, S_t^{\alpha_i}\left[1 - \mathrm{e}^{-h_i}\right] \\ &\quad + \int_0^t G_i(\mathrm{d}s)\, S_s^{\alpha_i}\Big[1 - f_i\Big(\mathbf{1} - \mathbf{U}\left(\mathbf{h};t-s\right)\Big)\Big]. \end{aligned} \tag{19}$$

In fact, the first term at the r.h.s. of (19) concerns the case that the initial particle did not die by time t, whereas the second one results from the remaining case if that particle died at some time $s \le t$.

If $\mathbf{h}$ is additionally twice continuously differentiable (to be in the domain of the Laplacian), by a differentiation of (19) with respect to t one can verify that $\mathbf{U}(\mathbf{h})$ satisfies the equation system (1) with initial condition $\mathbf{1} - \mathrm{e}^{-\mathbf{h}}$. That is, via (18) the branching model N and the differential equation system (1) (or an integrated version as (19)) *correspond to each other.* In particular, the log-Laplace expression $\left\langle \Lambda, \mathbf{U}\left(\mathbf{h};t\right)\right\rangle$ appearing in (17) coincides with the L^1–norm in (3).

Remark 11 (convergence of L^1–norms) The L^1–norms $\left\langle \Lambda, \mathbf{U}\left(\mathbf{h};t\right)\right\rangle$ are non-increasing in t (see (86)). Under the conditions of our theorems, the limit as $t \uparrow \infty$ is zero in the low dimensional case of Theorem 8 (see (92)), whereas it is typically positive in the high dimensional situation of Theorem 9 (cf. with the persistence statement (98)). $\diamond$

Differentiating $U_i(\varepsilon\mathbf{h};t,x)$ with respect to ε at $\varepsilon=0+$, we arrive at the *expectation* $\mathbf{E}_{x,i}\langle N(t),\mathbf{h}\rangle =: V_i(\mathbf{h};t,x)$. On the other hand, if we do the same with (19) with $\mathbf{h}$ replaced by $\varepsilon\mathbf{h}$, we obtain the following *linear system of integral equations for the expectation vector* $\mathbf{V}(\mathbf{h})=\mathbf{V}$:

$$V_i(t) \;=\; G_i(t)\,S_t^{\alpha_i}h_i + \int_0^t G_i(\mathrm{d}s)\,S_s^{\alpha_i}\sum_{j\in\mathsf{K}} m_{i,j}\,V_j(t-s). \tag{20}$$

But the semigroup $t\mapsto G_i(t)\,S_t^{\alpha_i}$ has generator $\Delta_{\alpha_i}-\varrho_i\mathrm{I}$ (with I the identity operator). Hence from (20) we get the following 'variation of constants' equation system for the expectation vector $\mathbf{V}(\mathbf{h})=\mathbf{V}$:

$$V_i(t) \;=\; S_t^{\alpha_i}h_i + \varrho_i\int_0^t \mathrm{d}s\;S_s^{\alpha_i}\sum_{j\in\mathsf{K}}\left(m_{i,j}-\delta_{i,j}\right)V_j(t-s). \tag{21}$$

Integrating the space variable with Lebesgue measure ℓ, multiplying by λ_i, and summing over i give the following formula for the *intensity measure* of $N(t)$:

$$\mathbf{E}_\Lambda\langle N(t),\mathbf{h}\rangle \;=\; \int\Lambda\left(\mathrm{d}[x,i]\right)\,\mathbf{E}_{x,i}\langle N(t),\mathbf{h}\rangle \;\equiv\; \langle\Lambda,\mathbf{h}\rangle\,. \tag{22}$$

That is, after integration with Λ, time dependence of the intensity measure disappears, since the term resulting from the second summand at the r.h.s. of (21) vanishes by our choice (8) of the λ_i and the left eigenvalue property. Actually a bit care has to be taken to guarantee that after integration with ℓ the terms remain *finite*. We will skip this at this point and only refer to Lemma 19 (a) below.

Differentiating (21) with respect to t (if $\mathbf{h}$ is additionally twice continuously differentiable), for the expectation vector $\mathbf{V}(\mathbf{h})=\mathbf{V}$ we get the following linear differential equation system

$$\frac{\partial}{\partial t}V_i \;=\; \Delta_{\alpha_i}V_i + \varrho_i\sum_{j\in\mathsf{K}}\left(m_{i,j}-\delta_{i,j}\right)V_j\,,\qquad i\in\mathsf{K}, \tag{23}$$

with initial condition $\mathbf{h}$. If we assume for the moment that $\mathbf{M}$ is in addition a *stochastic* matrix, then the r.h.s. defines the generator of a Markov process in $\mathsf{R}^d\times\mathsf{K}$, which gives a good tool to study properties of the expectation vector $\mathbf{V}(\mathbf{h})=\mathbf{V}$ (see the 'basic process' in [GRW92]). But in the present case of a *general* critical $\mathbf{M}$ we are interested in, this tool is not anymore available. In the next section we will analyze instead ancestry lines in the genealogical spatial trees of our process and, in particular, at a point we will reduce the problem to a stochastic matrix $\overline{\mathbf{M}}$, by a comparison argument and some transformations.

2 Some family tree analysis

In this section we prepare for the proof of the persistence Theorem 9. For this purpose, we start our branching process only with a single particle situated at $x \in \mathsf{R}^d$ and having type i. Actually, we are looking at the whole arising (finite) *genealogical spatial tree.* Of course, this requires to work from the beginning with a finer model. However, we will not give a formal construction since it is quite intuitive and can be found in the literature. (For a systematic study of such constructions, we recommend for instance [DP91].)

The genealogical spatial tree is a detailed description of the types and paths of all the particles which arise, their birth and death times, recording this way in particular all genealogical relations. For simplicity, we again use the symbol $\mathbf{P}_{x,i}$ to denote the law of the whole tree we just discussed.

If we neglect in such description the spatial data, we denote the remaining (non-empty finite) *genealogical tree* by $\mathcal{T}$. Write $\mathbb{P}_i$ for the law of $\mathcal{T}$ if it started from a particle of type i. Moreover, if we further drop recording after time $t \geq 0$, we write $\mathcal{T}_t$. Consequently, $\mathcal{T}_t$ tells us all what happens by time t with the arising progeny of the original particle, except about their spatial data.

2.1 Moment finiteness of the embedded multitype continuous-time Galton-Watson process

Write $N^+(t,j)$ for the total number $N\left(t, \mathsf{R}^d \times \{j\}\right)$ of particles of type $j \in \mathsf{K}$ living at time t. Note that N^+ is a *critical multitype continuous-time Galton-Watson process.* We will show that under condition **(b)** of Hypothesis 5, this 'classical' process N^+ has finite moments of all orders less than $1+\underline{\beta}$. To prepare for this we introduce the following notation.

Definition 12 (function class $\mathcal{B}$) We say that a function $g : \mathsf{R}_+ \to \mathsf{R}_+$ belongs to the class $\mathcal{B}$ if

(i) g is convex,

(ii) there exists a constant C such that

$$g\,(xy) \leq C\,g\,(x)\,g(y), \quad x,y \geq 0. \tag{24}$$

$\diamond$

First we state the following lemma; compare with [Sew74, Theorems 2.3.1 and 2.3.7]. Recall that $\zeta_{i,j}$ denotes the number of offspring of type j born from a particle of type i.

Lemma 13 (expectation finiteness of convex functionals) *Let* $g \in \mathcal{B}$. *If*

$$\mathbb{E}_i\, g\,(\zeta_{i,j}) < \infty, \qquad i,j \in \mathsf{K}, \tag{25}$$

then, for each $T \geq 0$,

$$\sup_{0 \leq t \leq T} \mathbb{E}_i\, g\big(N^+(t,j)\big) \;<\; \infty, \qquad i,j \in \mathsf{K}. \tag{26}$$

Proof Introduce the probabilities

$$\Gamma_{i,j}\,(t,n) \;:=\; \mathbb{P}_i\,\big(N^+(t,j) = n\big)\,, \qquad n \geq 0,$$

the truncated expectations

$$\Delta_{i,j}\,(t,H) \;:=\; \sum_{n=0}^{H} g(n)\Gamma_{i,j}\,(t,n)\,, \qquad H \geq 0, \tag{27}$$

and

$$\Delta\,(t,H) \;:=\; \sum_{i=1}^{K}\sum_{j=1}^{K} \Delta_{i,j}\,(t,H)\,. \tag{28}$$

We have to show that

$$\sup_{t \leq T}\, \sup_{H \geq 0}\, \Delta\,(t,H) \;<\; \infty. \tag{29}$$

Recall that ϱ_i is the death rate of a particle of type i, and f_i its offspring generating function. Represent

$$\varrho_i\,\left[f_i\,(\mathbf{z}) - z_i\right] \;=:\; \sum_{\mathbf{d}\,\geq\,\mathbf{0}} p_i(\mathbf{d})\, z_1^{d_1} \cdots z_K^{d_K}, \qquad \mathbf{z} = (z_1\,, \ldots, z_K) \in [0,1]^K,$$

with $p_i(\mathbf{d}) \;\geq\; 0$ for $\mathbf{d} \neq \mathbf{e}_i \;:=\; (\delta_{i,j})_{j=1}^{K}$ (where $\delta_{i,j}$ denotes the Kronecker symbol). Put $\|\mathbf{d}\| := d_1 + \cdots + d_K$. From the Kolmogorov backward equation and the branching property we have, for $n \geq 0$,

$$\begin{aligned} \frac{\partial}{\partial t}\Gamma_{i,j}\,(t,n) \;&=\; \delta_{n,0}\; p_i(\mathbf{0}) \;+\; \sum_{k=1}^{K} p_i\,(\mathbf{e}_k)\,\Gamma_{k,j}\,(t,n) \\ &\quad +\; \sum_{\|\mathbf{d}\|\,\geq\,2} p_i(\mathbf{d}) \sum\nolimits^{n} \prod_{k=1}^{K}\prod_{q=1}^{d_k} \Gamma_{k,j}\,(t, n_{k,q}) \end{aligned} \tag{30}$$

where the symbol $\sum^n$ denotes summation over all tuples

$$(n_{1,1}\,, \ldots, n_{1,d_1}, \ldots, n_{K,1}, \ldots, n_{K,d_K})$$

of non-negative integers summing up to n :

$$(n_{1,1} + \cdots + n_{1,d_1}) + \cdots + (n_{K,1} + \cdots + n_{K,d_K}) \;=\; n. \tag{31}$$

Multiplying (30) by $g(n)$ and summing over n from 0 to H, we obtain

$$\begin{aligned}\frac{\partial}{\partial t}\Delta_{i,j}(t,H) &= p_i(\mathbf{0})\,g(0) + \sum_{k=1}^{K} p_i(\mathbf{e}_k)\,\Delta_{k,j}(t,H) \\ &\quad + \sum_{\|\mathbf{d}\|\geq 2} p_i(\mathbf{d}) \sum_{n=0}^{H} {\sum}^{n} g(n) \prod_{k=1}^{K}\prod_{q=1}^{d_k} \Gamma_{k,j}(t,n_{k,q}).\end{aligned} \tag{32}$$

Since g is convex, and $\mathbf{d}\neq\mathbf{0}$,

$$g(n) = g\left(\frac{n_{1,1}+\cdots+n_{K,d_K}}{\|\mathbf{d}\|}\,\|\mathbf{d}\|\right) \leq \frac{1}{\|\mathbf{d}\|}\sum_{k'=1}^{K}\sum_{q'=1}^{d_{k'}} g\left(n_{k',q'}\|\mathbf{d}\|\right),$$

and by the multiplicativity property (24) we obtain

$$g(n) \leq C\,\frac{g(\|\mathbf{d}\|)}{\|\mathbf{d}\|}\sum_{k'=1}^{K}\sum_{q'=1}^{d_{k'}} g(n_{k',q'}).$$

Thus, for the last term in (32) we get the estimate

$$\begin{aligned}&\sum_{\|\mathbf{d}\|\geq 2} p_i(\mathbf{d}) \sum_{n=0}^{H} {\sum}^{n} g(n) \prod_{k=1}^{K}\prod_{q=1}^{d_k} \Gamma_{k,j}(t,n_{k,q}) \\ &\leq C \sum_{\|\mathbf{d}\|\geq 2} p_i(\mathbf{d})\,\frac{g(\|\mathbf{d}\|)}{\|\mathbf{d}\|}\sum_{n=0}^{H} {\sum}^{n} \sum_{k'=1}^{K}\sum_{q'=1}^{d_{k'}} g(n_{k',q'}) \prod_{k=1}^{K}\prod_{q=1}^{d_k} \Gamma_{k,j}(t,n_{k,q}).\end{aligned} \tag{33}$$

In the latter subexpression $\sum_{n=0}^{H}\ldots$ at the r.h.s., we may interchange the order of summation, so that we first fix a pair (k',q'). Then we have to deal with the remaining internal expression

$$\sum_{n=0}^{H} {\sum}^{n} g(n_{k',q'}) \prod_{k=1}^{K}\prod_{q=1}^{d_k} \Gamma_{k,j}(t,n_{k,q}). \tag{34}$$

Again we interchange the order of summation, so that in (34) we first let run $n_{k',q'}$ from 0 to H, and withdraw $g(n_{k',q'})\,\Gamma_{k',j}(t,n_{k',q'})$. Then the remaining internal expression from (34) is a probability saying that $H-n_{k',q'}$ particles of type j are produced etc. We may estimate this probability by 1, so that for (34) we get the bound

$$\sum_{n_{k',q'}=0}^{H} g(n_{k',q'})\,\Gamma_{k',j}(t,n_{k',q'}) = \Delta_{k',j}(t,H).$$

Consequently, the r.h.s. of (33) is majorized by

$$C \sum_{\|\mathbf{d}\| \geq 2} p_i(\mathbf{d}) \frac{g(\|\mathbf{d}\|)}{\|\mathbf{d}\|} \sum_{k'=1}^{K} \sum_{q'=1}^{d_{k'}} \Delta_{k',j}(t,H).$$

But $\Delta_{k',j}(t,H)$ does not depend on q', and $d_{k'} \leq \|\mathbf{d}\|$. Hence, denoting

$$D_i := \sum_{\|\mathbf{d}\| \geq 2} p_i(\mathbf{d})\, g(\|\mathbf{d}\|)$$

(observe that $D_i < \infty$ by assumption (25)), instead of (33) we get

$$\sum_{\|\mathbf{d}\| \geq 2} p_i(\mathbf{d}) \sum_{n=0}^{H} \sum\nolimits^{n} g(n) \prod_{k=1}^{K} \prod_{q=1}^{d_k} \Gamma_{k,j}(t, n_{k,q}) \leq C\, D_i \sum_{k'=1}^{K} \Delta_{k',j}(t,H).$$

Applied to (32), together with the trivial inequality $p_i(\mathbf{e}_k) \leq 1$, we arrive at

$$\frac{\partial}{\partial t} \Delta_{i,j}(t,H) \leq p_i(\mathbf{0})\, g(0) + (1 + CD_i) \sum_{k=1}^{K} \Delta_{k,j}(t,H).$$

Summing over i and j, recalling the notation (28), and setting $D := \sum_{i=1}^{K} D_i$ gives

$$\frac{\partial}{\partial t} \Delta(t,H) \leq K\, g(0) \sum_{i=1}^{K} p_i(\mathbf{0}) + K(K + CD)\, \Delta(t,H).$$

Since $\Delta(0,H) = K\, g(1) \geq 0$, and all the constants in this linear ordinary differential inequality are non-negative and independent of H, by standard comparison it follows that the function $t \mapsto \Delta(t,H)$ is majorized by a (finite) continuous function being independent of H. This gives (29), finishing the proof. ■

The next statement follows easily from the preceding lemma and the fact that

$$\mathbb{E}_i\, \zeta_{i,j}^{1+\varepsilon} < \infty, \qquad i,j \in \mathsf{K}, \quad 0 \leq \varepsilon < \underline{\beta},$$

(recall Remark 6).

Corollary 14 (moment finiteness) *If condition* **(b)** *in Hypothesis* 5 *holds, then for each* $T \geq 0$ *and* $0 \leq \varepsilon < \underline{\beta}$,

$$\sup_{0 \leq t \leq T} \mathbb{E}_i \left(N^+(t,j)\right)^{1+\varepsilon} < \infty, \qquad i,j \in \mathsf{K}.$$

2.2 Linear bound for $(1+\varepsilon)$–moments

Along with the critical mean matrix $\mathbf{M}$ of (4), we introduce the matrices

$$\mathsf{M}(\varrho) := \left(\varrho_i\,(m_{i,j}-\delta_{i,j})\right)_{i,j=1}^{K} \tag{35}$$

and

$$\mathcal{M}(t) = \left(\mathcal{M}_{i,j}(t)\right)_{i,j=1}^{K} := \mathrm{e}^{\mathsf{M}(\varrho)t}, \qquad t \geq 0. \tag{36}$$

Actually, (see, for example, [AN72, § 5.7.2])

$$\mathcal{M}_{i,j}(t) = \mathbb{E}_i\, N^+(t,j). \tag{37}$$

Recall that $\mathbf{u}$ is the right eigenvector of $\mathbf{M}$ normalized by $(\mathbf{1},\mathbf{u}) = 1$. It is easy to check that $\mathsf{M}(\varrho)\mathbf{u} = \mathbf{0}$, hence $\mathbf{u}$ is a right [2)] eigenvector of $\mathcal{M}(t)$:

$$\mathcal{M}(t)\,\mathbf{u} = \mathbf{u}, \qquad t \geq 0. \tag{38}$$

Combined with (37) we get the well-known fact

$$\sup_{s\geq 0} \mathbb{E}_i \sum_{k=1}^{K} N^+(s,k) < \infty, \qquad i \in \mathsf{K}. \tag{39}$$

Now we are ready to majorize the moments $\mathbb{E}_i \left(N^+(t,j)\right)^{1+\varepsilon}$ as $t \to \infty$ under Hypothesis 5 (b), for fixed $0 \leq \varepsilon < \underline{\beta}$.

Lemma 15 (linear bound for $(1+\varepsilon)$–moments) *If condition* **(b)** *in Hypothesis* 5 *holds, then for any* $\varepsilon \in [0,\underline{\beta})$ *there exist a constant* c *(depending on* ε*) such that*

$$\mathbb{E}_i \left(N^+(t,j)\right)^{1+\varepsilon} \leq c\,(1+t), \qquad i,j \in \mathsf{K}, \quad t \geq 0.$$

Proof Fix ε, i. It suffices to show that

$$\mathbb{E}_i \left(N^+(t),\mathbf{u}\right)^{1+\varepsilon} \leq c\,(1+t) \qquad t > 1, \tag{40}$$

for some constant c (recall Corollary 14). To this aim, fix t (for the moment) and set $\tau := t/[t]$ and

$$X_m := \left(N^+(m\tau),\mathbf{u}\right) - \Big(N^+\left((m-1)\tau\right),\mathbf{u}\Big), \qquad m \geq 1. \tag{41}$$

Clearly, under $\mathbb{P}_i$,

$$\left(N^+(t),\mathbf{u}\right) = u_i + \sum_{m=1}^{[t]} X_m\,. \tag{42}$$

[2)] Similarly, writing λ for the vector with components $\lambda_i = v_i/\varrho_i$ (recall (8)), we get $\lambda\mathsf{M}(\varrho) = \mathbf{0}$, hence λ is a left eigenvector of the $\mathcal{M}(t)$.

Let $\mathcal{F}_t$ denote the standard σ–field generated by the (càdlàg) process N^+ up to time t. Given $\mathcal{F}_{(m-1)\tau}$, by the Markov property, time-homogeneity, and the branching property, we can represent

$$N^+(m\tau, j) = \sum_{k=1}^{K} \sum_{q=1}^{N^+\left((m-1)\tau,k\right)} N_k^{+(q)}(\tau, j) \tag{43}$$

where, for k fixed, $\left\{N_k^{+(q)} : q \geq 1\right\}$ are independent copies of N^+ under $\mathbb{P}_k$. But by (37) and (38),

$$\mathbb{E}_k \left(N_k^{+(q)}(\tau), \mathbf{u}\right) = \sum_{j=1}^{K} \mathcal{M}_{k,j}(\tau)\, u_j = u_k \tag{44}$$

is independent of q. Hence, from (43) and (44) we conclude that

$$\mathbb{E}_i \left\{\left(N^+(m\tau), \mathbf{u}\right) \,\middle|\, \mathcal{F}_{(m-1)\tau}\right\} = \left(N^+\left((m-1)\tau\right), \mathbf{u}\right).$$

Therefore, the random sequence $\left\{\left(N^+(m\tau), \mathbf{u}\right);\ m \geq 0\right\}$ is a martingale with respect to $\{\mathcal{F}_{m\tau}\,;\ m \geq 0\}$ and $\mathbb{P}_i$. In particular,

$$\mathbb{E}_i \left\{X_m \mid \mathcal{F}_{(m-1)\tau}\right\} = 0, \qquad m \geq 1. \tag{45}$$

On the other hand, from (41), (43), and the right identity in (44), given $\mathcal{F}_{(m-1)\tau}$,

$$X_m = \sum_{k=1}^{K} \sum_{q=1}^{N^+\left((m-1)\tau,k\right)} \sum_{j=1}^{K} \left[N_k^{+(q)}(\tau, j) - \mathcal{M}_{k,j}(\tau)\right] u_j\,.$$

Hence,

$$\begin{aligned} &\mathbb{E}_i \left\{|X_m|^{1+\varepsilon} \,\middle|\, \mathcal{F}_{(m-1)\tau}\right\} \\ &\leq \text{ const} \sum_{k=1}^{K}\sum_{j=1}^{K} \mathbb{E}_k^{(q)} \left| \sum_{q=1}^{N^+\left((m-1)\tau,k\right)} \left[N_k^{+(q)}(\tau, j) - \mathcal{M}_{k,j}(\tau)\right] u_j \right|^{1+\varepsilon}, \end{aligned}$$

where the latter expectation $\mathbb{E}_k^{(q)} \ldots$ means that it has to apply only to the random variables $N_k^{+(q)}(\tau, j)$. Here we used the convention to denote by *const* a positive constant which may change from term to term. Applying the von Bahr-Esseen inequality [vBE65, Theorem 2] to the zero mean i.i.d. sequence

$$\left\{\left[N_k^{+(q)}(\tau, j) - \mathcal{M}_{k,j}(\tau)\right] u_j \,:\, q \geq 1\right\},$$

the latter inequality can be continued with

$$\leq \text{ const } \sum_{k=1}^{K} \sum_{j=1}^{K} 2 \sum_{q=1}^{N^+((m-1)\tau,k)} \mathbb{E}_k^{(q)} \left| \left[N_k^{+(q)}(\tau, j) - \mathcal{M}_{k,j}(\tau) \right] u_j \right|^{1+\varepsilon} .$$

Observing that

$$\max_{j,k} \sup_{0<\tau\leq 2} \mathbb{E}_k \left| \left[N_k^+(\tau, j) - \mathcal{M}_{k,j}(\tau) \right] u_j \right|^{1+\varepsilon}$$

is finite by Corollary 14, we may summarize the previous estimates to

$$\mathbb{E}_i \left| X_m \right|^{1+\varepsilon} \leq \text{ const } \mathbb{E}_i \sum_{k=1}^{K} N^+ \big((m-1)\tau, k \big) \leq \text{ const } \sup_{s\geq 0} \mathbb{E}_i \sum_{k=1}^{K} N^+(s,k) .$$

But the latter expression is finite since N^+ is critical (recall (39)), and we arrive at

$$C := \sup_{m\geq 1} \mathbb{E}_i \left| X_m \right|^{1+\varepsilon} < \infty. \tag{46}$$

On the other hand, from (42),

$$\mathbb{E}_i \big(N^+(t), \mathbf{u} \big)^{1+\varepsilon} \leq \text{ const } \left(u_i^{1+\varepsilon} + \mathbb{E}_i \left| \sum_{m=1}^{[t]} X_m \right|^{1+\varepsilon} \right).$$

By (45) and (46), we may again exploit the von Bahr-Esseen inequality to continue the latter formula line with

$$\leq \text{ const } \left(u_i^{1+\varepsilon} + 2 \sum_{m=1}^{[t]} \mathbb{E}_i \left| X_m \right|^{1+\varepsilon} \right) \leq \text{ const } \big(1 + C[t] \big) \leq \text{ const } (1+t)$$

as required for (40). ■

2.3 An auxiliary Markov jump process

Recall that in general the critical mean matrix $\mathbf{M}$ is a non-stochastic matrix (Assumption 3). Using its (normalized) right eigenvector $\mathbf{u}$, we introduce the *stochastic* irreducible matrix $\overline{\mathbf{M}}$ with entries

$$\overline{m}_{i,j} := \frac{m_{i,j} \, u_j}{u_i}, \qquad i,j \in \mathsf{K}. \tag{47}$$

Consider a *Markov jump process* in K where a jump from i to j occurs with rate $\varrho_i \, \overline{m}_{i,j}$ (including self-transitions if $i = j$). In other words, the process spends in state i a random time with distribution G_i, and then jumps to j chosen with probability $\overline{m}_{i,j}$.

Parallel to (35) and (36), we introduce the matrices

$$\overline{\mathsf{M}}(\varrho) := \left(\varrho_i \left(\overline{m}_{i,j} - \delta_{i,j}\right)\right)_{i,j=1}^{K}$$

and

$$\overline{\mathcal{M}}(t) = \left(\overline{\mathcal{M}}_{i,j}(t)\right)_{i,j=1}^{K} := \mathrm{e}^{\overline{\mathsf{M}}(\varrho)t}, \qquad t \geq 0.$$

Note that from definition (47) of $\overline{\mathbf{M}}$, Assumption 3, and definition (8) of λ_i it follows that the vector $\overline{\lambda}$ defined by

$$\overline{\lambda} = (\overline{\lambda}_1, ..., \overline{\lambda}_K) \quad \text{with} \quad \overline{\lambda}_i = \lambda_i u_i / (\lambda, \mathbf{u}), \quad i \in K, \tag{48}$$

satisfies $\overline{\lambda}\,\overline{\mathsf{M}}(\varrho) = \mathbf{0}$. Hence, $\overline{\lambda}$ is a normalized left eigenvector of the matrices $\overline{\mathcal{M}}(t)$. In other words, $\overline{\lambda}$ is the unique invariant law for the transition matrices $\overline{\mathcal{M}}(t)$, that is for the Markov jump process under consideration. Therefore, $\overline{\lambda}_i$ may be viewed as the *approximate relative time this process spends in* i.

Let $\overline{P}_i$ denote the law of the present Markov jump process starting from i, and $\overline{t}_j(t)$ the total time this process spends at state j by time t. Set

$$\overline{r}_{i,j}(t,a) := \overline{P}_i\left(\overline{t}_j(t) \geq a\right), \qquad i,j \in \mathsf{K}, \quad t,a \geq 0. \tag{49}$$

Note that

$$\overline{r}_{i,j}(t,a) = 0 \quad \text{if} \quad 0 \leq t < a, \quad i,j \in \mathsf{K}. \tag{50}$$

The following *large deviation* estimates for occupation times are well-known (use, for instance, [Var84], Theorem 11.6, the remark on the case of compact state spaces before (1)–(5), p.34, and the contraction principle (9.4)):

$$\left.\begin{array}{l} \text{To each } \delta > 0 \text{ there are constants } c_1 \text{ and } c_2 \text{ such that} \\ \overline{r}_{i,j}\Big(t, \left(\overline{\lambda}_j + \delta\right) t\Big) \leq c_1\, \mathrm{e}^{-c_2 t}, \qquad i,j \in \mathsf{K}, \quad t \geq 0. \end{array}\right\} \tag{51}$$

By a renewal argument, for $0 \leq a \leq t$,

$$\left.\begin{array}{rcl} \overline{r}_{i,i}(t,a) & = & 1 - G_i(a) + \displaystyle\int_0^a G_i(\mathrm{d}s) \sum_{k=1}^{K} \overline{m}_{i,k}\, \overline{r}_{k,i}\left(t-s, a-s\right), \\ \overline{r}_{i,j}(t,a) & = & \displaystyle\int_0^t G_i(\mathrm{d}s) \sum_{k=1}^{K} \overline{m}_{i,k}\, \overline{r}_{k,j}\left(t-s, a\right), \qquad i \neq j. \end{array}\right\} \tag{52}$$

It is easy to see that the system (52) has a unique $[0,1]$–valued solution. (Indeed, pass for instance to a description in terms of matrices, and use that in this matrix renewal equation one has uniform bounds to reduce to a usual renewal equation.) Recall that this solution satisfies (51).

Observe now that by setting

$$r_{i,j}(t,a) := \overline{r}_{i,j}(t,a)\,\frac{u_i}{u_j} \tag{53}$$

the system (52) changes to

$$\left.\begin{aligned} r_{i,i}(t,a) &= 1 - G_i(a) + \int_0^a G_i(\mathrm{d}s)\,\sum_{k=1}^K m_{i,k}\,r_{k,i}\,(t-s,a-s)\,, \\ r_{i,j}(t,a) &= \int_0^t G_i(\mathrm{d}s)\,\sum_{k=1}^K m_{i,k}\,r_{k,j}\,(t-s,a)\,, \qquad i \neq j, \end{aligned}\right\} \tag{54}$$

$0 \le a \le t$. Then (51) immediately gives the following exponential estimates for the unique solution to (54):

$$\left.\begin{aligned} &\text{To each } \delta > 0 \text{ there are constants } c_1 \text{ and } c_2 \text{ such that} \\ &r_{i,j}\Big(t,\big(\overline{\lambda}_j + \delta\big)\,t\Big) \;\le\; c_1\,\mathrm{e}^{-c_2 t}, \qquad i,j \in \mathsf{K}, \quad t \ge 0. \end{aligned}\right\} \tag{55}$$

2.4 Exponential smallness of some ancestry line probabilities

Now we turn back to the genealogical trees $\mathcal{T}_t$ introduced in the beginning of this section. To each particle which is alive at time t there is an *ancestry line* $w : [0,t) \to \mathsf{K}$ describing the type of this particle and the types of all its ancestors. Similarly, for particles who died at a time $\tau < t$ without producing children, we may adjoin an *ancestry line* $w : [0,\tau) \to \mathsf{K}$. Since we exclude events of probability zero (the life time of particles is exponentially distributed, hence nothing may happen at a given time point as t), we can imagine $\mathcal{T}_t$ to be the collection of those ancestry lines w (of particles which are alive at time t or died before t) and we write $w \in \mathcal{T}_t$ if w belongs to $\mathcal{T}_t$ in this sense.

Let $t_j(w) \ge 0$ denote the *total time* the ancestry line $w \in \mathcal{T}_t$ spends in type $j \in \mathsf{K}$. Note that $t_1(w) + \cdots + t_K(w) \le t$ (if $w \in \mathcal{T}_t$). The key of our family tree analysis will be a *large deviation result* on $t_j(w)$ as time t tends to infinity (see Proposition 17 in the next subsection) we now prepare for.

Set

$$0 \;\le\; \sigma_j(t) \;:=\; \max_{w \in \mathcal{T}_t} t_j(w) \;\le\; t, \qquad j \in \mathsf{K}, \quad t \ge 0, \tag{56}$$

for the *maximal total time* any of the ancestry lines $w \in \mathcal{T}_t$ spent in j. Write

$$p_{i,j}(t,a) \;:=\; \mathbb{P}_i\left(\sigma_j(t) < a\right), \qquad i,j \in \mathsf{K}, \quad t,a \ge 0, \tag{57}$$

for the probability that the maximal total time $\sigma_j(t)$ an ancestry line in a tree $\mathcal{T}_t$ is smaller than a. Put

$$q_{i,j}(t,a) \;:=\; 1 - p_{i,j}(t,a). \tag{58}$$

Clearly,

$$q_{i,j}(t,a) = 0 \quad \text{if} \quad 0 \le t < a, \quad i,j \in \mathsf{K}. \tag{59}$$

By comparison with the solution to (54), we will derive the following result.

Lemma 16 (exponential smallness) *Fix $\delta > 0$. Then there are constants c_1 and c_2 such that*

$$q_{i,j}\Big(t, \big(\overline{\lambda}_j + \delta\big)\, t\Big) \le c_1 \,\mathrm{e}^{-c_2 t}, \qquad i,j \in \mathsf{K}, \quad t \ge 0.$$

Proof Write $\mathbf{p}_{\cdot,j}(t,a)$ for the vector $\big(p_{1,j}(t,a), ..., p_{K,j}(t,a)\big)$. Similarly, introduce the notation $\mathbf{q}_{\cdot,j}(t,a)$. By a renewal argument, for $0 \le a \le t$,

$$\left.\begin{aligned} p_{i,i}(t,a) &= \int_0^a G_i(\mathrm{d}s)\, f_i\Big(\mathbf{p}_{\cdot,i}(t-s,a-s)\Big), && i \in \mathsf{K}, \\ p_{i,j}(t,a) &= 1 - G_i(t) + \int_0^t G_i(\mathrm{d}s)\, f_i\Big(\mathbf{p}_{\cdot,j}(t-s,a)\Big), && i \ne j. \end{aligned}\right\} \tag{60}$$

Hence, for $0 \le a \le t$,

$$\left.\begin{aligned} q_{i,i}(t,a) &= 1 - G_i(a) + \int_0^a G_i(\mathrm{d}s)\Big[1 - f_i\Big(\big(\mathbf{1} - \mathbf{q}_{\cdot,i}(t-s,a-s)\big)\Big)\Big], \\ q_{i,j}(t,a) &= \int_0^t G_i(\mathrm{d}s)\Big[1 - f_i\Big(\mathbf{1} - \mathbf{q}_{\cdot,j}(t-s,a)\Big)\Big], \qquad i \ne j. \end{aligned}\right\} \tag{61}$$

We would like to compare $q_{i,j}(t,a)$ with $r_{i,j}(t,a)$. To this aim we approximate the solutions of (61) and (54) in a standard way by the following quantities:

$$q_{i,j}^{(0)}(t,a) := \delta_{i,j}\,\big(1 - G_i(a)\big) =: r_{i,j}^{(0)}(t,a),$$

and, for $n \ge 1$,

$$\left.\begin{aligned} q_{i,i}^{(n)}(t,a) &= 1 - G_i(a) + \int_0^a G_i(\mathrm{d}s)\Big[1 - f_i\Big(\Big(\mathbf{1} - \mathbf{q}_{\cdot,i}^{(n-1)}(t-s,a-s)\Big)\Big)\Big], \\ q_{i,j}^{(n)}(t,a) &= \int_0^t G_i(\mathrm{d}s)\Big[1 - f_i\Big(\mathbf{1} - \mathbf{q}_{\cdot,j}^{(n-1)}(t-s,a)\Big)\Big], \qquad i \ne j. \end{aligned}\right\}$$

and

$$\left.\begin{aligned} r_{i,i}^{(n)}(t,a) &= 1 - G_i(a) + \int_0^a G_i(\mathrm{d}s) \sum_{k=1}^K m_{i,k}\, r_{k,i}^{(n-1)}(t-s,a-s), \\ r_{i,j}^{(n)}(t,a) &= \int_0^t G_i(\mathrm{d}s) \sum_{k=1}^K m_{i,k}\, r_{k,j}^{(n-1)}(t-s,a), \qquad i \ne j, \end{aligned}\right\}$$

$0 \le a \le t$. Exploiting the simple estimate

$$0 \;\le\; 1 - f_i\left(\mathbf{1} - \mathbf{z}\right) \;\le\; \sum_{k=1}^{K} m_{i,k}\, z_k\,, \qquad i \in \mathsf{K}, \quad \mathbf{z} \in [0,1]^K, \tag{62}$$

(implying by the way also that Φ_i from (10) is non-negative), by induction we get

$$q_{i,j}^{(n)}(t,a) \;\le\; r_{i,j}^{(n)}(t,a), \qquad 0 \le a \le t, \quad i,j \in \mathsf{K}, \quad n \ge 1,$$

Passing to the limit as $n \uparrow \infty$ in this relation gives

$$q_{i,j}(t,a) \;\le\; r_{i,j}(t,a), \qquad 0 \le a \le t, \quad i,j \in \mathsf{K}.$$

Then by (55) the proof is complete. ■

2.5 Large deviation probabilities for ancestry lines in the reduced tree

Let $\mathcal{T}_t^{\,\mathrm{r}}$ denote the *reduced tree* obtained from $\mathcal{T}_t$ by removing all the ancestry lines $w \in \mathcal{T}_t$ which die before t. Recall the vector $\overline{\lambda}$ defined in (48) is the invariant law of the auxiliary Markov jump process.

Proposition 17 (large deviations) *Take $\delta > 0$. Then there are positive constants c_1 and c_2 such that*

$$\mathbb{P}_i\Big(\exists\, w \in \mathcal{T}_t^{\,\mathrm{r}}:\; t_j(w) < \big[\,\overline{\lambda}_j - (K-1)\delta\big]\, t\Big) \;\le\; c_1\, \mathrm{e}^{-c_2 t}, \qquad i,j \in \mathsf{K}, \quad t \ge 0.$$

Proof We may assume that $K > 1$ (otherwise the probability expression in the proposition disappears since in this case $t_1(w) \equiv t$). If $w \in \mathcal{T}_t^{\,\mathrm{r}}$ then $t_1(w) + \cdots + t_K(w) = t$. Since $\big(\mathbf{1}, \overline{\lambda}\big) = 1$, for $i,j \in \mathsf{K}$ fixed the probability in question is

$$\begin{aligned} &\mathbb{P}_i\Big(\exists\, w \in \mathcal{T}_t^{\,\mathrm{r}}:\; \sum_{k:\; k \ne j} t_k(w) > \sum_{k:\; k\ne j} \big(\,\overline{\lambda}_k + \delta\big)\, t\Big) \\ &\le\; \sum_{k:\; k \ne j} \mathbb{P}_i\Big(\exists\, w \in \mathcal{T}_t^{\,\mathrm{r}}:\; t_k(w) > \big(\,\overline{\lambda}_k + \delta\big)\, t\Big). \end{aligned}$$

Enlarging further by passing to $\mathcal{T}_t$ and to the $\ge$ relation, and using that

$$\mathbb{P}_i\Big(\exists\, w \in \mathcal{T}_t:\; t_k(w) \ge \big(\,\overline{\lambda}_k + \delta\big)\, t\Big) \;=\; q_{i,k}\Big(t, \big(\,\overline{\lambda}_k + \delta\big)\, t\Big),$$

the claim follows from Lemma 16. ■

2.6 A uniform estimate

Recall definition (18) of the 'Laplace' functionals $U_i(\mathbf{h}; t, x)$. We will use the derived linear bounds of $(1+\varepsilon)$–moments and the large deviation probabilities to estimate these functionals.

Lemma 18 (a uniform estimate) *Impose Hypothesis* 5(b). *Take* $\varepsilon > 0$ *and* $\mathbf{h} \in \mathbf{C}_+^{\mathrm{comp}}$. *Then there is a constant* c *independent of* ε *such that*

$$\sup_{(x,i) \in \mathsf{R}^d \times \mathsf{K}} U_i(\varepsilon \mathbf{h}; t, x) \leq c\,\varepsilon \left(1 \wedge t^{-d/\alpha}\right), \qquad t > 0. \tag{63}$$

Proof Without loss of generality, we may assume that the motion exponents are ordered: $\alpha_1 \leq \cdots \leq \alpha_K$. Take $\delta > 0$ such that $c^* := \overline{\lambda}_1 - (K-1)\delta > 0$, and introduce the event

$$A_{t,\delta} := \left\{\exists\, w \in \mathcal{T}_t^{\,\mathrm{r}} :\ t_1(w) < c^* t\right\}.$$

In the following, $I\{A\}$ always denotes the indicator of a set (or event) A. Using the bound in (18), for $i \in \mathsf{K}$,

$$\begin{aligned} &\sup_{x \in \mathsf{R}^d} U_i(\varepsilon \mathbf{h}; t, x) \\ &\leq\ \varepsilon\, \mathbb{E}_i \left\langle N^+(t), \mathbf{h}\right\rangle I\{A_{t,\delta}\} + \varepsilon \sup_{x \in \mathsf{R}^d} \mathbf{E}_{x,i} \left\langle N(t), \mathbf{h}\right\rangle I\{A^{\mathrm{c}}_{t,\delta}\}. \end{aligned} \tag{64}$$

Here, by an abuse of notation, $A^{\mathrm{c}}_{t,\delta}$ refers to the complement of $A_{t,\delta}$ on the set $\{N(t) \neq 0\}$. Take $\eta \in (0, \underline{\beta})$. For the first term at the r.h.s. we apply Hölder's inequality to get

$$\mathbb{E}_i \left\langle N^+(t), \mathbf{h}\right\rangle I\{A_{t,\delta}\} \leq \left(\mathbb{E}_i \left\langle N^+(t), \mathbf{h}\right\rangle^{1+\eta}\right)^{1/(1+\eta)} \left(\mathbb{P}_i(A_{t,\delta})\right)^{\eta/(1+\eta)}.$$

Then Lemma 15 and Proposition 17 give

$$\mathbb{E}_i \left\langle N^+(t), \mathbf{h}\right\rangle I\{A_{t,\delta}\} \leq \text{const}\,(1+t)^{1/(1+\eta)}\,\mathrm{e}^{-\text{const}\,t}, \tag{65}$$

which is of a smaller order than required for (63).

It remains to deal with the second term at the r.h.s. of (64). Choose a ball B in R^d centered at the origin such that

$$h_i \leq c_{\mathbf{h}} I\{B\}, \quad i \in \mathsf{K}, \quad \text{for some constant} \quad c_{\mathbf{h}}\,. \tag{66}$$

Then

$$\mathbf{E}_{x,i} \left\langle N(t), \mathbf{h}\right\rangle I\{A^{\mathrm{c}}_{t,\delta}\} \leq c_{\mathbf{h}}\, \mathbf{E}_{0,i}\, N\Big(t, (B-x) \times \mathsf{K}\Big) I\{A^{\mathrm{c}}_{t,\delta}\}.$$

Denoting by $W_j^m(t)$ the position of the mth particle of type j (in any ordering) in the (non-empty) population at time t, we have

$$\begin{aligned} &\mathbf{E}_{0,i}\, N\Big(t,(B-x)\times \mathsf{K}\Big) I\{A^{\mathrm{c}}_{t,\delta}\} \\ &= \mathbf{E}_{0,i} \sum_{j=1}^{K} \sum_{m=1}^{N^+(t,j)} I\{A^{\mathrm{c}}_{t,\delta}\}\, \mathbf{P}_{0,i}\Big\{W_j^m(t)\in B-x \,\Big|\, A^{\mathrm{c}}_{t,\delta}\Big\}. \end{aligned} \tag{67}$$

Recall that $t_k(w)$ denotes the total time the ancestry line w in the reduced tree $\mathcal{T}_t^{\,r}$ spends at type k. Write $\mathbf{t}(w) := \big(t_1(w),\dots,t_k(w)\big)$. Now we decompose the latter probability as follows:

$$\begin{aligned} &\mathbf{P}_{0,i}\Big\{W_j^m(t)\in B-x \,\Big|\, A^{\mathrm{c}}_{t,\delta}\Big\} \\ &= \int \mathbf{P}_{0,i}\Big\{\mathbf{t}(w)\in \mathrm{d}\mathbf{t} \,\Big|\, A^{\mathrm{c}}_{t,\delta}\Big\}\ \mathbf{P}_{0,i}\Big\{W_j^m(t)\in B-x \,\Big|\, \mathbf{t}(w)=\mathbf{t}\Big\}. \end{aligned} \tag{68}$$

Denoting by p^k the transition density function of the stable motion process with index α_k, the internal conditional probability can be written as

$$\begin{aligned} &\mathbf{P}_{0,i}\Big\{W_j^m(t)\in B-x \,\Big|\, \mathbf{t}(w)=\mathbf{t}\Big\} \\ &= \int \mathrm{d}y\ (\mathrm{p}^2_{t_2} * \cdots * \mathrm{p}^K_{t_K})(y) \int_{z+y\in B-x} \mathrm{d}z\ \mathrm{p}^1_{t_1}(z). \end{aligned} \tag{69}$$

In view of the unimodality and symmetry of p^1, the internal integral is bounded by

$$\int_B \mathrm{d}z\ \mathrm{p}^1_{t_1}(z) \;\le\; \mathrm{const}\ \Big(1\wedge t_1^{-d/\alpha}\Big) \;\le\; \mathrm{const}\ \Big(1\wedge t^{-d/\alpha}\Big). \tag{70}$$

Here we used the fact that by definition $t_1(w)\ge c^*t$ on the event $A^{\mathrm{c}}_{t,\delta}$. Thus the r.h.s. of (70) is a bound for (69) (on $A^{\mathrm{c}}_{t,\delta}$), hence for (68). Therefore, the r.h.s. of (67) is bounded by

$$\mathrm{const}\ \Big(1\wedge t^{-d/\alpha}\Big)\ \mathbf{E}_{0,i}\sum_{j=1}^{K} N^+(t,j) \;\le\; \mathrm{const}\ \Big(1\wedge t^{-d/\alpha}\Big)$$

since N^+ is critical ((39)). This finishes the proof. ■

3 Remaining proofs

Lemma 18 will be our main tool for the proof of the persistence Theorem 9. But first of all we prepare for the proof of the extinction Theorem 8.

3.1 On the intensity measures

For $t, L \geq 0$, set

$$C(t,L) := \Big\{ x \in \mathsf{R}^d : |x| \leq L\, t^{1/\alpha} \Big\}. \tag{71}$$

Recall the definition (9) of the initial intensity measure Λ.

Lemma 19 (intensity measures) *Let $d \geq 1$. Then the following two statements hold.*

(a) (finite intensity measures) *For all $t \geq 0$ and $\mathbf{h} \in \mathbf{C}_+^{\mathrm{comp}}$,*

$$\mathbf{E}_\Lambda \langle N(t), \mathbf{h} \rangle = \int \Lambda(\mathrm{d}[x,i])\, \mathbf{E}_{x,i} \langle N(t), \mathbf{h} \rangle = \langle \Lambda, \mathbf{h} \rangle < \infty. \tag{72}$$

(b) (negligible contribution from outside) *For each bounded Borel subset B of R^d,*

$$\sup_{t \geq 1} \sum_{i=1}^{K} \lambda_i \int_{\mathsf{R}^d \setminus C(t,L)} \mathrm{d}x\ \mathbf{E}_{x,i}\, N(t, B \times \mathsf{K}) \underset{L \uparrow \infty}{\longrightarrow} 0. \tag{73}$$

Proof Given $\mathbf{h} \in \mathbf{C}_+^{\mathrm{comp}}$, choose $c_{\mathbf{h}}$ and B satisfying (66). Assume that B is a ball of radius $r \geq 1$ (centered at the origin). Consider a fixed i.

1° (*preparation*) For $t, K \geq 0$, we want to deal with the quantity

$$\int_{\mathsf{R}^d \setminus C(t,L)} \mathrm{d}x\ \mathbf{E}_{x,i}\, N(t, B \times \mathsf{K}) = \int_{\mathsf{R}^d \setminus C(t,L)} \mathrm{d}x\ \mathbf{E}_{0,i}\, N\Big(t, (B-x) \times \mathsf{K}\Big). \tag{74}$$

As in (67), conditioning on the (non-empty) reduced tree $\mathcal{T}_t^{\,\mathrm{r}}$, the expectation expression at the r.h.s. can be dominated by

$$\mathbf{E}_{0,i}\, I\left\{ N(t) \neq 0 \right\} \sum_{j=1}^{K} \sum_{m=1}^{N^+(t,j)} \mathbf{P}_{0,i} \Big\{ W_j^m(t) \in B - x \,\Big|\, \mathcal{T}_t^{\,\mathrm{r}} \Big\}. \tag{75}$$

Integrating with respect to $\mathrm{d}x$ on $\mathsf{R}^d \setminus C(t,L)$, for (74) we obtain the bound

$$\mathbf{E}_{0,i}\, I\left\{ N(t) \neq 0 \right\} \sum_{j=1}^{K} \sum_{m=1}^{N^+(t,j)} \int_{\mathsf{R}^d \setminus C(t,L)} \mathrm{d}x \int_{B-x} \mathbf{P}_{0,i} \Big\{ W_j^m(t) \in \mathrm{d}y \,\Big|\, \mathcal{T}_t^{\,\mathrm{r}} \Big\}. \tag{76}$$

Since B is a ball of the radius r centered at the origin, the restrictions to the integration variables imply that $|x+y| \leq r$ and $|y| > Lt^{1/\alpha} - r$. Hence, interchanging the order of integration, the latter double integral can be estimated from above by

$$\int_{|y| > Lt^{1/\alpha} - r} \mathbf{P}_{0,i} \Big\{ W_j^m(t) \in \mathrm{d}y \,\Big|\, \mathcal{T}_t^{\,\mathrm{r}} \Big\} \int_{|x+y| \leq r} \mathrm{d}x. \tag{77}$$

Denoting the internal integral by c_r (which is independent of y), the double integral (77) can be written as

$$c_r\,\mathbf{P}_{0,i}\Big\{\,\big|W_j^m(t)\big| \ge Lt^{1/\alpha} - r \,\Big|\, \mathcal{T}_t^{\,\mathrm{r}}\Big\}. \tag{78}$$

2° (*proof of* **(a)**) Take now $L = 0$, estimate the latter (conditional) probability expression by 1, and combine all the previous estimates to arrive at

$$\int\!\mathrm{d}x\;\mathbf{E}_{x,i}\,N(t, B\times\mathsf{K}) \;\le\; \text{const}\;\mathbf{E}_{0,i}\sum_{j=1}^{K} N^+(t,j).$$

But the latter expectation is finite (recall (39)). This gives the finiteness of the expectation expression in (72). Combining with the identity (22), statement **(a)** follows.

3° (*proof of* **(b)**) Consider now $t \ge 1$ and $L \ge 2r$. Then the (conditional) probability expression in (78) is bounded from above by

$$\mathbf{P}_{0,i}\Big\{\,\big|W_j^m(t)\big| \;\ge\; \frac{L}{2}\,t^{1/\alpha} \;\Big|\; \mathcal{T}_t^{\,\mathrm{r}}\Big\}. \tag{79}$$

Recall that, for i, j, m, t, fixed, $W_j^m(t)$ is the position of a particular particle at time t. In the time interval $[0,t]$, this particle and its ancestors spent total time t_k in type $k \in \{1, ..., K\}$, where $t_1 + \cdots + t_K = t$. Then, evidently, given $\mathcal{T}_t^{\,\mathrm{r}}$, the following coincidence in law holds:

$$W_j^m(t) \;\overset{\mathcal{L}}{=}\; W(t_1\,,\alpha_1) + \cdots + W(t_K\,,\alpha_K)$$

where $t \mapsto W(t,\alpha_k)$ is a *symmetric stable process* with index α_k, starting from 0 at time 0, its law we denote by $P_0^{\alpha_k}$. Now the probability in (79) can be estimated from above by

$$\sum_{k=1}^{K} P_0^{\alpha_k}\Big(\big|W(t_k\,,\alpha_k)\big| \;\ge\; \frac{L}{2K}\,t_k^{1/\alpha_k}\Big),$$

since $t \ge 1$ and $t^{1/\alpha} \ge t^{1/\alpha_k} \ge t_k^{1/\alpha_k}$ (recall definition (14) of α). By the self-similarity of the stable process, we can continue with

$$= \sum_{k=1}^{K} P_0^{\alpha_k}\Big(\big|W(1,\alpha_k)\big| \ge \frac{L}{2K}\Big) \;=:\; \varepsilon_L\,.$$

Consequently, the conditional probability expression in (78) is bounded by ε_L, and for (76) hence (74) we get the upper bound

$$\text{const}\;\varepsilon_L\,\mathbf{E}_{0,i}\sum_{j=1}^{K} N^+(t,j) \;\le\; \text{const}\;\varepsilon_L \;\xrightarrow[L\uparrow\infty]{}\; 0$$

(recall (39)). This finishes the proof. ■

3.2 Some uniform continuity in time

Before we will come to some uniform continuity property of certain functionals needed for the extinction proof, we observe the following simple fact. Recall the definition (11) of Ψ.

Lemma 20 (monotonicity and non-negativity) *If the matrix* $\mathbf{M}$ *of means satisfies Assumption* 3 *then* $\Psi(\mathbf{z}) \geq 0$ *for* $\mathbf{z} \in [0,1]^K$ *and, in addition,* Ψ *is monotone non-decreasing in all of its arguments.*

Proof Recalling the identity in (11), and that $\mathbf{v}$ is the left eigenvector, we have

$$\frac{\partial}{\partial z_i}\Psi(\mathbf{z}) \;=\; v_i + \sum_{j=1}^{K} v_j \frac{\partial f_j(\mathbf{1}-\mathbf{z})}{\partial z_i} \;\geq\; v_i - \sum_{j=1}^{K} v_j m_{ji} \;=\; 0, \qquad i \in \mathsf{K},$$

proving the monotonicity. The non-negativity then follows from $\Psi(\mathbf{0}) = 0$. ■

Recalling (18) and (72), for $\mathbf{h} \in \mathbf{C}_+^{\mathrm{comp}}$ put

$$0 \;\leq\; U_i^+(\mathbf{h};t) \;:=\; \int \mathrm{d}x\, U_i(\mathbf{h};t,x) \;\leq\; \text{const}\, \langle \Lambda, \mathbf{h} \rangle \;<\; \infty. \tag{80}$$

Lemma 21 (uniform continuity in time) *Without any dimension restriction, for* $\varepsilon > 0$ *one can find* $\delta = \delta(\varepsilon)$ *such that*

$$\sup_{t \geq 0,\; i \in \mathsf{K}} \left| U_i^+(\mathbf{h};t) - U_i^+(\mathbf{h};t+\tau) \right| \;\leq\; \varepsilon \quad \text{if} \quad 0 \leq \tau \leq \delta.$$

Proof Set

$$b_i(\mathbf{h};t,x) \;:=\; E_x^{\alpha_i}\Big(1 - \exp\Big[-h_i\big(W(t,\alpha_i)\big)\Big]\Big).$$

Then, since the stable process $W(\cdot\,,\alpha_i)$ (introduced in step 3° of the proof of Lemma 19) has stationary increments,

$$\big\langle \ell, b_i(\mathbf{h};t) \big\rangle \;=\; \big\langle \ell, \mathbf{1} - \mathrm{e}^{-h_i} \big\rangle \;\leq\; \langle \ell, h_i \rangle \;<\; \infty.$$

Hence,

$$\big\langle \Lambda, \mathbf{b}(\mathbf{h};t) \big\rangle \;=\; \big\langle \Lambda, \mathbf{1} - \mathrm{e}^{-\mathbf{h}} \big\rangle \;\leq\; \langle \Lambda, \mathbf{h} \rangle \;<\; \infty. \tag{81}$$

In particular,

$$\big\langle \Lambda, \mathbf{b}(\varepsilon\mathbf{h};t) \big\rangle \;\leq\; \varepsilon \, \langle \Lambda, \mathbf{h} \rangle \underset{\varepsilon \downarrow 0}{\searrow} 0. \tag{82}$$

Integrating equation (19) with respect to ℓ we obtain

$$\begin{aligned} U_i^+(\mathbf{h};t) \;&=\; G_i(t)\,\big\langle \ell, 1 - \mathrm{e}^{-h_i} \big\rangle + \int_0^t G(\mathrm{d}s)\; U_i^+(\mathbf{h};t-s) \\ &\quad - \int_0^t G(\mathrm{d}s) \Big\langle \ell\,, f_i\Big(\mathbf{1} - \mathbf{U}(\mathbf{h};t-s)\Big) - 1 + U_i(\mathbf{h};t-s) \Big\rangle. \end{aligned} \tag{83}$$

All the terms in this equation are indeed finite by (80) and the elementary estimate (62). Introduce the function

$$H_i(t) := \sum_{n=1}^{\infty} G_i^{*n}(t), \qquad t \geq 0.$$

Note that $H_i(t) = \varrho_i t$ (use Laplace transforms). Solving the renewal equation (83) with respect to $U_i^+(\mathbf{h};t)$ (see for instance [Fel71, § 11.1]), we find

$$\begin{aligned} & U_i^+(\mathbf{h};t) \\ & = \langle \ell, 1-\mathrm{e}^{-h_i} \rangle - \varrho_i \int_0^t \mathrm{d}s \, \Big\langle \ell, f_i\left(\mathbf{1}-\mathbf{U}(\mathbf{h};s)\right) - 1 + U_i(\mathbf{h};s) \Big\rangle. \end{aligned} \tag{84}$$

Multiplying by $\lambda_i = v_i/\varrho_i$ (recall (8)), summing over $i \in \mathsf{K}$, and recalling notation (11) of $\mathbf{\Psi}$ and that $\mathbf{v}$ is the left eigenvector, we obtain

$$\big\langle \Lambda, \mathbf{U}(\mathbf{h};t) \big\rangle = \big\langle \Lambda, \mathbf{1}-\mathrm{e}^{-\mathbf{h}} \big\rangle - \int_0^t \mathrm{d}s \, \Big\langle \ell, \Psi\left(\mathbf{U}(\mathbf{h};s)\right) \Big\rangle. \tag{85}$$

But Ψ is non-negative by Lemma 20, and it follows that

$$\big\langle \Lambda, \mathbf{U}(\mathbf{h};t) \big\rangle \quad \text{is non−increasing in } t \tag{86}$$

and

$$0 \leq \big\langle \Lambda, \mathbf{U}(\mathbf{h};t) \big\rangle \leq \big\langle \Lambda, \mathbf{1}-\mathrm{e}^{-\mathbf{h}} \big\rangle \leq \langle \Lambda, \mathbf{h} \rangle < \infty. \tag{87}$$

From (84) we have

$$\begin{aligned} & \Big| U_i^+(\mathbf{h};t) - U_i^+(\mathbf{h};t+\tau) \Big| \\ & \leq \varrho_i \int_t^{t+\tau} \mathrm{d}s \, \Big\langle \ell, \Big| f_i\left(\mathbf{1}-\mathbf{U}(\mathbf{h};s)\right) - 1 + U_i(\mathbf{h};s) \Big| \Big\rangle. \end{aligned} \tag{88}$$

By (62), the term in angles can be estimated from above:

$$\Big\langle \ell, \sum_{k=1}^{K} m_{i,k} U_k(\mathbf{h};s) + U_i(\mathbf{h};s) \Big\rangle \leq \text{const} \, \big\langle \Lambda, \mathbf{U}(\mathbf{h};t) \big\rangle \leq \text{const} \, \langle \Lambda, \mathbf{h} \rangle,$$

where we used (87). Managing (88) in this way, the claim follows. ■

3.3 Extinction of some functionals

In this subsection, we impose Hypothesis 5 (a) and assume $d \leq \alpha/\overline{\beta}$.

Lemma 22 (special extinction along a subsequence) *Impose the conditions of Theorem* 8, *let* k *be such as in Hypothesis* 5 (a), *and fix* $\mathbf{h} \in \mathbf{C}_{+}^{\mathrm{comp}}$. *Then there exists a sequence* $t_n \to \infty$ *as* $n \uparrow \infty$ *such that*

$$\lim_{n\uparrow\infty} U_k^{+}(\mathbf{h};t_n) \;=\; 0.$$

Proof It follows from (85) that

$$\int_0^\infty \mathrm{d}t \int \mathrm{d}y\ \Psi\Big(\mathbf{U}(\mathbf{h};t,y)\Big) \;<\; \infty. \tag{89}$$

Hence, by the assumption (12) we conclude that

$$\infty > \int_0^\infty \mathrm{d}t \int_{C(t,L)} \mathrm{d}y\ \Psi\Big(\mathbf{U}(\mathbf{h};t,y)\Big) \;\geq\; \overline{c} \int_0^\infty \mathrm{d}t \int_{C(t,L)} \mathrm{d}y\ \Big(U_k(\mathbf{h};t,y)\Big)^{1+\overline{\beta}}.$$

Using Jensen's inequality, the estimation can be continued with

$$\geq\; \overline{c} \int_0^\infty \mathrm{d}t \left(\int_{C(t,L)} \mathrm{d}y\ U_k(\mathbf{h};t,y)\right)^{1+\overline{\beta}} |C(t,L)|^{-\overline{\beta}},$$

with $|B|$ denoting the Lebesgue measure of B. Recalling the definition (71) of $C(t,L)$, we have $|C(t,L)| = \mathrm{const}\ t^{d/\alpha}$ implying that under the conditions of Theorem 8,

$$\int_1^\infty \mathrm{d}t\ \ |C(t,L)|^{-\overline{\beta}} \;=\; \infty.$$

It follows that

$$\liminf_{t\uparrow\infty} \int_{C(t,L)} \mathrm{d}y\ U_k(\mathbf{h};t,y) \;=\; 0. \tag{90}$$

On the other hand, by the definition (18) of $U_k(\mathbf{h};t,y)$, we find a ball B in R^d such that all h_i disappear outside of B and that

$$U_k(\mathbf{h};t,y) \;\leq\; \mathbf{P}_{y,k}\Big(N(t;B\times\mathsf{K}) \neq 0\Big) \;\leq\; \mathbf{E}_{y,k}\,N(t;B\times\mathsf{K}). \tag{91}$$

Hence, by Lemma 19 (b), to $\varepsilon > 0$ we can choose an L such that

$$\sup_{t\geq 1} \int_{\mathsf{R}^d\setminus C(t,L)} \mathrm{d}y\ U_k(\mathbf{h};t,y) \;\leq\; \varepsilon.$$

Thus, combined with (90),

$$\liminf_{t\uparrow\infty} \int \mathrm{d}y\ U_k(\mathbf{h};t,y) \;\leq\; \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $\liminf_{t\uparrow\infty} U_k^{+}(\mathbf{h};t) = 0$, proving the lemma. ∎

Lemma 23 (extinction: general case) *Under the conditions of Theorem* 8, *for* $\mathbf{h} \in \mathbf{C}_{+}^{\mathrm{comp}}$,

$$\lim_{t\uparrow\infty} \big\langle \Lambda, \mathbf{U}(\mathbf{h};t) \big\rangle \;=\; 0. \tag{92}$$

Proof From the previous lemma we know that there is a sequence $t_n \to \infty$ such that

$$\lim_{n\uparrow\infty} U_i^{+}(\mathbf{h};t_n) \;=\; 0 \tag{93}$$

holds if i coincides with k from Hypothesis 5 (a). If $K > 1$, let $i \neq k$ be such that $m_{ki} > 0$ (since $\mathbf{M}$ is irreducible, such an i always exists). Then there is a $\delta > 0$ such that

$$1 - f_k(\mathbf{1} - \mathbf{z}) \;\geq\; \frac{1}{2}\, m_{ki}\, z_i, \qquad \mathbf{z} \in [0,1]^K, \quad 0 \leq z_i \leq \delta. \tag{94}$$

Recalling definition (18),

$$U_i(\mathbf{h};t,y) \;\leq\; \mathbf{P}_{0,i}\Big(N(t) \neq 0\Big) \;\underset{t\uparrow\infty}{\longrightarrow}\; 0,$$

since the critical multitype continuous-time Galton-Watson process dies. Hence, there exists a $t_0 = t_0(\delta)$ such that

$$\sup_{\mathbf{h},y} U_i(\mathbf{h};t,y) \;\leq\; \delta, \qquad t \geq t_0\,.$$

Therefore, using once more equation (19), by (94), for $\varepsilon > 0$ and $t - \varepsilon \geq t_0$,

$$\begin{aligned} U_k(\mathbf{h};t,x) \quad &\geq \quad \int_0^{\varepsilon} G_k(\mathrm{d}s)\, E_x^{\alpha_k}\Big[1 - f_k\Big(\mathbf{1} - \mathbf{U}\Big(\mathbf{h};t-s, W(s,\alpha_i)\Big)\Big)\Big] \\ &\geq \quad \frac{1}{2}\, m_{ki} \int_0^{\varepsilon} G_k(\mathrm{d}s)\, E_x^{\alpha_k}\, U_i\Big(\mathbf{h};t-s, W(s,\alpha_i)\Big). \end{aligned}$$

Integrating with respect to $\mathrm{d}x$ gives

$$\begin{aligned} U_k^{+}(\mathbf{h};t) \quad &\geq \quad \mathrm{const} \int_0^{\varepsilon} G_k(\mathrm{d}s)\, U_i^{+}(\mathbf{h};t-s) \\ &\geq \quad \mathrm{const}\; G_k(\varepsilon) \inf_{t-\varepsilon \leq s \leq t} U_i^{+}(\mathbf{h};s)\,. \end{aligned}$$

Taking now $t = t_n$ from Lemma 22, we get

$$\inf_{t_n-\varepsilon \,\leq\, s \,\leq\, t_n} U_i^{+}(\mathbf{h};s) \;\underset{n\uparrow\infty}{\longrightarrow}\; 0.$$

But $\varepsilon > 0$ is arbitrary, and from the uniform continuity established in Lemma 21 we obtain (93) for the selected i.

Since $\mathbf{M}$ is irreducible we can repeat this procedure finitely often to see that (93) holds for any i (along the same sequence t_n).

Finally, $\big\langle \Lambda, \mathbf{U}(\mathbf{h};t) \big\rangle$ is monotone in t, by (86). This, in view of (93), yields (92), finishing the proof. ■

3.4 Proof of the local extinction theorem

Since N starts with a Poisson system of intensity Λ, and the law of N is infinitely divisible, the extinction claim (15) in Theorem 8 can be reformulated as follows: For each bounded Borel subset B of R^d,

$$\int \Lambda\left(\mathrm{d}[x,i]\right)\ \mathbf{P}_{x,i}\Big(N(t,B\times\mathsf{K})\neq 0\Big)\ \underset{t\uparrow\infty}{\longrightarrow}\ 0. \tag{95}$$

(In fact, combine, for instance, Lemma 2.2.5, Theorem 4.3.3, and Proposition 2.2.14 of [MKM78].) By Lemma 19, it remains to show that

$$\int_{C(t,L)} \mathrm{d}x\ \mathbf{P}_{x,i}\Big(N(t,B\times\mathsf{K})\neq 0\Big)\ \underset{t\uparrow\infty}{\longrightarrow}\ 0 \tag{96}$$

for fixed constant $L\geq 1$, type $i\in\mathsf{K}$, and bounded Borel set B. For this we may additionally assume that B is a centered ball in R^d. We can find $\mathbf{h}=(h_1,\ldots,h_K)\in\mathbf{C}_+^{\mathrm{comp}}$ such that $h_i\geq I\{B\}$, $i\in\mathsf{K}$. For such a choice of $\mathbf{h}$, by the definition (18) of $U_i\left(\mathbf{h};t,x\right)$, for any x,i,t,

$$\mathbf{P}_{x,i}\Big(N(t;B\times\mathsf{K})\geq 1\Big)\ \leq\ \left(1-e^{-1}\right)^{-1}U_i\left(\mathbf{h};t,x\right).$$

Therefore, recalling notation (80),

$$\int \mathrm{d}x\ \mathbf{P}_{x,i}\Big(N(t;B\times\mathsf{K})\neq 0\Big)\ \leq\ \left(1-e^{-1}\right)^{-1}U_i^+\left(\mathbf{h};t\right).$$

Then (96) follows from Lemma 23, finishing the proof. ∎

3.5 Proof of the persistence theorem

In order to prove Theorem 9, take $\mathbf{h}\in\mathbf{C}_+^{\mathrm{comp}}$. Recalling (17),

$$0\ \leq\ -\log\mathbf{E}_\Lambda \mathrm{e}^{-\langle N(t),\mathbf{h}\rangle}\ =\ \langle\Lambda,\mathbf{U}\left(\mathbf{h};t\right)\rangle.$$

Using identity (85) and the finiteness (87), this equation can be continued with

$$=\ \left\langle\Lambda,\mathbf{1}-\mathbf{e}^{-\mathbf{h}}\right\rangle-\int_0^t \mathrm{d}s\ \Big\langle \ell,\Psi\left(\mathbf{U}\left(\mathbf{h};s\right)\right)\Big\rangle\ <\ \infty.$$

The monotonicity (86) implies the existence of a finite limit as $t\uparrow\infty$. Since the previous limit statement holds for all $\mathbf{h}\in\mathbf{C}_+^{\mathrm{comp}}$, by (87) we arrive at a log-Laplace functional of a limiting point field $N(\infty)$, which expectation symbol we denote by $\mathbf{E}$:

$$-\log\mathbf{E}\,\mathrm{e}^{-\langle N(\infty),\mathbf{h}\rangle}\ =\ \left\langle\Lambda,\mathbf{1}-\mathbf{e}^{-\mathbf{h}}\right\rangle-\int_0^\infty \mathrm{d}s\ \Big\langle \ell,\Psi\left(\mathbf{U}\left(\mathbf{h};s\right)\right)\Big\rangle, \tag{97}$$

$\mathbf{h} \in \mathbf{C}_+^{\text{comp}}$. Applying a continuity theorem for clustering (see e.g. [MKM78, Proposition 4.7.3]), one can easily check that $N(\infty)$ is an equilibrium state. (That is, N started from this limiting particle system leads to a time-stationary process.)

It remains to show that $N(\infty)$ has full intensity, i.e.

$$\mathbf{E}\left\langle N(\infty),\mathbf{h}\right\rangle = \langle \Lambda, \mathbf{h}\rangle . \tag{98}$$

In view of

$$\mathbf{E}\left\langle N(\infty),\mathbf{h}\right\rangle = -\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\,\mathbf{E}\exp\left[-\left\langle N(\infty),\varepsilon\mathbf{h}\right\rangle\right]\Big|_{\varepsilon=0},$$

from (97) and $\mathbf{U}(\mathbf{0}) = \mathbf{0}$ as well as $\mathbf{\Psi}(\mathbf{0}) = 0$ follows that it suffices to establish

$$\lim_{\varepsilon\downarrow 0}\int_0^\infty \mathrm{d}s\,\varepsilon^{-1}\int \mathrm{d}y\;\Psi\left(\mathbf{U}\left(\varepsilon\mathbf{h};s,y\right)\right) = 0, \qquad \mathbf{h}\in\mathbf{C}_+^{\text{comp}},$$

or, by the basic condition (13) in Hypothesis 5 (b) even that

$$\lim_{\varepsilon\downarrow 0}\int_0^\infty \mathrm{d}s\,\varepsilon^{-1}\int \mathrm{d}y\;\left(\mathbf{1},\,\mathbf{U}\left(\varepsilon\mathbf{h};s,y\right)\right)^{1+\underline{\beta}} = 0. \tag{99}$$

We know from (87) the estimate

$$0 \le \left\langle \Lambda, \mathbf{U}\left(\varepsilon\mathbf{h};s\right)\right\rangle \le \varepsilon\left\langle \Lambda,\mathbf{h}\right\rangle .$$

Therefore, up to a constant, the double integral in (99) is bounded from above by

$$\langle \Lambda,\mathbf{h}\rangle \int_0^\infty \mathrm{d}s\left(\sum_{i=1}^K \sup_{y\in\mathsf{R}^d} U_i\left(\varepsilon\mathbf{h};s,y\right)\right)^{\underline{\beta}}.$$

From the uniform estimate in Lemma 18, the integral is bounded by

$$\le \text{ const }\varepsilon^{\underline{\beta}}\int_0^\infty \mathrm{d}s\;\left(1\wedge s^{-d\underline{\beta}/\alpha}\right)$$

that vanishes as $\varepsilon\downarrow 0$, by our assumption $d > \alpha/\underline{\beta}$. This completes the proof of the theorem. ■

Remark 24 (individual survival) By a slight modification of the argument given in the latter proof, one can show that under the conditions of the persistence Theorem 9 the limit $\lim_{t\uparrow\infty}\left\langle \ell, U_i(\mathbf{h};t)\right\rangle$ exists and is positive provided that $\mathbf{h}\neq\mathbf{0}$. ◇

References

[AN72] K.B. Athreya and P.E. Ney. *Branching Processes.* Springer-Verlag, Berlin, 1972.

[BCG93] M. Bramson, J.T. Cox, and A. Greven. Ergodicity of critical spatial branching processes in low dimensions. *Ann. Probab.*, 21:1946–1957, 1993.

[BCG97] M. Bramson, J.T. Cox, and A. Greven. Invariant measures of critical spatial branching processes in high dimensions. *Ann. Probab.*, 25(1):56–70, 1997.

[Daw93] D.A. Dawson. Measure-valued Markov processes. In P.L. Hennequin, editor, *École d'été de probabilités de Saint Flour XXI–1991*, volume 1541 of *Lecture Notes in Mathematics*, pages 1–260. Springer-Verlag, Berlin, 1993.

[DF85] D.A. Dawson and K. Fleischmann. Critical dimension for a model of branching in a random medium. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 70:315–334, 1985.

[DFFP86] D.A. Dawson, K. Fleischmann, R.D. Foley, and L.A. Peletier. A critical measure-valued branching process with infinite mean. *Stoch. Analysis Appl.*, 4:117–129, 1986.

[DP91] D.A. Dawson and E.A. Perkins. Historical processes. *Mem. Amer. Math. Soc.*, 454, 1991.

[Dyn94] E.B. Dynkin. *An Introduction to Branching Measure-Valued Processes.* volume 6 of *CRM Monograph Series.* Centre de Recherches Mathématiques, Université de Montréal, Amer. Math. Soc., Providence, 1994.

[Fel71] W. Feller. *An Introduction to Probability Theory and its Applications*, volume II. John Wiley and Sons, New York, 2nd edition, 1971.

[GH96] A. Glitzky and R. Hünlich. Electro-reaction-diffusion systems for heterostructures. In M. Niezgódka and P. Strzelecki, editors, *Free Boundary Problems, Theory and Applications*, volume 363 of *Pitman Research Notes in Mathematics Series*, pages 305–313. Longman, 1996.

[GH97] A. Glitzky and R. Hünlich. Global estimates and asymptotics for electro-reaction-diffusion systems in heterostructures. *WIAS Berlin, Preprint No. 338; Appl. Anal.* to appear, 1997.

[GLM90] L.G. Gorostiza and J.A. López-Mimbela. The multitype measure branching processes. *Adv. Appl. Probab.*, 22:49–67, 1990.

[GR91] L. Gorostiza and S. Roelly. Some properties of the multitype measure branching process. *Stoch. Proc. Appl.*, 37:259–274, 1991.

[GRW90] L. Gorostiza, S. Roelly, and A. Wakolbinger. Sur la persistance du processus de Dawson-Waatanabe stable. L'interversion de la limite en temps et de la renormalisation. In *Sem. Prob. XXIV*, volume 1426 of *Lecture Notes in Mathematics*, pages 275–281, Berlin, 1990. Springer-Verlag.

[GRW92] L.G. Gorostiza, S. Roelly, and A. Wakolbinger. Persistence of critical multitype particle and measure branching processes. *Probab. Theory Relat. Fields*, 29:313–335, 1992.

[GW91] L. Gorostiza and A. Wakolbinger. Convergence to equilibrium of a class of critical branching particle systems in continuous time. *Ann. Probab.*, 19:266–288, 1991.

[GW92] L.G. Gorostiza and A. Wakolbinger. Convergence to equilibrium of critical branching particle systems and superprocesses, and related nonlinear partial differential equations. *Acta Appl. Math.*, 27:269–291, 1992.

[GW93] L.G. Gorostiza and A. Wakolbinger. Asymptotic behavior of a reaction-diffusion system. A probabilistic approach. *Random and Comp. Dynamics*, 1:445–463, 1993.

[GW94] L.G. Gorostiza and A. Wakolbinger. Long time behavior of critical branching particle systems and applications. In D.A. Dawson, editor, *Measure-Valued Processes, Stochastic Partial Differential Equations, and Interacting Systems*, 1994.

[Kal77] O. Kallenberg. Stability of critical cluster fields. *Math. Nachr.*, 77:7–43, 1977.

[Kle97] A. Klenke. Clustering and invariant measures for spatial branching models with infinite variance. Preprint, Universität Erlangen-Nürnberg, 1997.

[Lie69] A. Liemant. Invariante zufällige Punktfolgen. *Wiss. Z. Friedrich-Schiller-Universität Jena*, 18:361–372, 1969.

[Lie81] A. Liemant. Kritische Verzweigungsprozesse mit allgemeinem Phasenraum. IV. *Math. Nachr.*, 102:235–254, 1981.

[LM92] J.A. López-Mimbela. Fluctuation limits of multitype branching random fields. *J. Multivariate Analysis*, 40:56–83, 1992.

[LMW96] J.A. López-Mimbela and A. Wakolbinger. Clumping in multitype branching trees. *Adv. Appl. Probab.*, 28(4):1034–1050, 1996.

[MKM78] K. Matthes, J. Kerstan, and J. Mecke. *Infinitely Divisible Point Processes.* Wiley, Chichester, 1978.

[PR77] U. Prehn and B. Röder. Limit theorems for spatially homogeneous branching systems with a finite set of types (in Russian). *Math. Nachr.*, 80:37–86, 1977.

[Sew74] B.A. Sewastjanow. *Verzweigungsprozesse.* Akademie-Verlag, Berlin, 1974.

[Smo83] J. Smoller. *Shok Waves and Reaction-Diffusion equations*, volume 258 of *Grundlehren math. Wiss.* Springer-Verlag, New York, 1983.

[Var84] S.R.S. Varadhan. Large deviations and applications. CBMS-NSF Regional Conference Series in Applied Mathematics, volume 46. Society for Industrial and Applied Mathematics, 1984.

[vBE65] B. von Bahr and C.-G. Esseen. Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.*, 36:299–303, 1965.