Multitype Branching Brownian Motion and Traveling Waves

Yan-Xia Ren* and Ting Yang†

Abstract

This article studies the parabolic system of equations which is closely related to multitype branching Brownian motion. Particular attention is paid to the monotone traveling wave solutions of this system. Provided some moment conditions, we show the existence, uniqueness and asymptotic behaviors of such waves with speed greater than or equal to a critical value $\underline{c}$ and non-existence of such waves with speed smaller than $\underline{c}$.

Keywords: Multitype branching Brownian motion, Spine approach, Additive martingale, Traveling wave solution

1 Introduction and Main Results

We consider a branching particle system in which there are d $(2 \leq d < +\infty)$ different types of particles. Let $S = \{1, 2, \cdots, d\}$ be the set of types. A type i particle splits into offspring particles of possible all types according to distribution $\{p_k(i) : k \in \mathbb{Z}_+^d\}$ after a lifetime which is exponentially distributed with parameter $a_i > 0$. All particles engender independent lines of descent. In addition, each particle diffuses in space $\mathbb{R}$ independently according to a Brownian motion starting from its point of creation through its lifetime. This system is called a multitype branching Brownian motion (MBBM). For more precise configuration of this MBBM see Section 2.

In this article, we assume that each particle reproduces at least one child, which guarantees the process survives forever with probability one. Suppose $m_{ij} := \sum_{k\in\mathbb{Z}_+^d} p_k(i)k_j < +\infty$, and that the mean matrix $M = (m_{ij})_{i,j\in S}$ is irreducible, *i. e.* that there exists no permutation matrix S such that $S^{-1}MS$ is block triangular.

*The research of this author is supported in part by NNSF of China (Grant No. 10971003 and No. 11128101) and Specialized Research Fund for the Doctoral Program of Higher Education.

†Corresponding author.

We study the following parabolic system of equations which is strongly related to MBBM:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \Lambda(\psi(u) - u) \tag{1.1}$$

Here $u(t,x) = (u_1(t,x), u_2(t,x), \cdots, u_d(t,x))^T$, Λ is a diagonal matrix with diagonal entries $\{a_i : i = 1, \cdots, d\}$, and $\psi(u) = (\psi_1(u), \psi_2(u), \cdots, \psi_d(u))^T$ with $\psi_i(z_1, \cdots, z_d) = \sum_{k\in\mathbb{Z}_+^d} p_k(i) \prod_{j=1}^d z_j^{k_j}$ being the generating function of a type i particle.

Our primary concern in this paper is the solutions satisfying $u(t,x) = w(x-ct)$ where w is a monotone function connecting 0 at $-\infty$ to 1 at $+\infty$. Such solutions are called traveling waves. The analogous object to (1.1) for a single-type branching Brownian motion is called the Fisher-Kolmogorov-Petrovski-Piscounov (FKPP) equation. FKPP equation has been extensively studied both by analytic and probabilistic methods (see, for example, [3, 16, 4, 8, 12]). Among these works, [8] and [12] gave proofs for the existence, uniqueness and asymptotics of traveling wave solutions to the FKPP equation through *purely* probabilistic arguments. Recently, Kyprianou *et al.* [13] extended the probabilistic arguments to the traveling wave equations associated to super-Brownian motions with a general branching mechanism.

In this paper we outline a probabilistic study of traveling waves of system (1.1). Our work is strongly guided by the probabilistic arguments in [12] with respect to single-type branching Brownian motion. An important tool of our probabilistic arguments is a representation of the family tree in terms of a suitable size-biased tree with spine. This representation is the continuous time analogue of the size-biased tree representation introduced by [11]. This continuous time version is also used in [6] to investigate the evolution of the ancestral types of typical particles for multitype Markov branching processes.

We call u a traveling wave solution with speed c if $u(t,x)$ satisfies (1.1) and u can be written as $u(t,x) = w(x-ct) = (w_1(x-ct), \cdots, w_d(x-ct))^T$ where $w_i(\cdot)$ is a twice continuously differentiable, strictly monotone function increasing from 0 at $-\infty$ to 1 at $+\infty$. We also call w a traveling wave with speed c. Obviously, w provides a traveling wave solution to (1.1) if and only if

$$\frac{1}{2}\frac{\partial^2 w}{\partial x^2} + c\frac{\partial w}{\partial x} + \Lambda(\psi(w) - w) = 0 \tag{1.2}$$

Sometimes, we write $u_i(t,x)$ and $w_i(x)$ as $u(t,x,i)$ and $w(x,i)$, respectively.

Let $N(t) := (N_1(t), N_2(t), \cdots, N_d(t))$ be the vector denoting the population sizes of different types at time t. Let $m_{ij}(t) := E_i(N_j(t)) < +\infty$. It is known that the mean matrix $M(t) = (m_{ij}(t))_{i,j\in S}$ can be written as

$$M(t) = \exp At = \sum_{n=0}^{+\infty} \frac{A^n}{n!} t^n \quad \text{where } A = (a_{ij})_{ij\in S}, \quad a_{ij} = a_i(m_{ij} - \delta_{ij}).$$

It follows from the irreducibility of M that $M(t)$ has positive entries for some $t > 0$ (this property is also called 'positive regularity' by [2]). According to Perron-Frobenius theorem (see Theorem 2.5 in [18]), A admits a real eigenvalue $\lambda^* > 0$ larger than the real part of any other eigenvalue. The so-called Perron's root λ^* is simple, with a one-dimensional eigenspace, and there correspond left and right eigenvectors with positive coordinates. In the following we denote by π (resp. h) the associated left (resp. right) eigenvector with normalization $\langle \pi, h\rangle = \langle \pi, 1\rangle = 1$, here $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product.

For $\lambda \neq 0$, define

$$c_\lambda := \frac{\lambda}{2} + \frac{\lambda^*}{\lambda}, \tag{1.3}$$

which will serve as the speeds of the traveling waves. In the following, we deal only with the case $c_\lambda \geq 0$. Traveling waves with negative speeds can be analyzed by simple considerations of symmetry. Let $\underline{\lambda} := \sqrt{2\lambda^*}$. It is easy to see that c_λ attains a local minimum $\underline{c} = c_{\underline{\lambda}} = \sqrt{2\lambda^*}$ at $\underline{\lambda}$. We call (1.2) *subcritical*, *critical* or *supercritical* according as $c <, =,$ or $> \underline{c}$.

Let the configuration of this MBBM at time t be given by the $\mathbb{R} \times S$-valued point process $\{(X_v(t), Y_v) : v \in Z(t)\}$ where $Z(t)$ is the set of particles alive at time t, $X_v(t)$ is v's spatial location and Y_v is its type. Let the probabilities for this process be $\{P_{xi} : x \in \mathbb{R}, i \in S\}$, where P_{xi} is the law starting from a single particle of type y at spatial position x. Let E_{xi} be the expectation corresponding to P_{xi}. To state our main results, we introduce two types of additive martingales which play an important role in this paper. Define, for any $\lambda \neq 0$,

$$W_\lambda(t) := \sum_{v \in Z(t)} h_{Y_v} \mathrm{e}^{-\lambda(X_v(t) + c_\lambda t)}. \tag{1.4}$$

From the many-to-one formula (see Proposition 1 below), it is easy to see that $\{W_\lambda(t), t \geq 0\}$ is a positive martingale under P_{xi}, and consequently the almost sure limit of $W_\lambda(t)$ exists. Set $W(\lambda) := \lim_{t \to +\infty} W_\lambda(t)$. Now we define another type of additive martingale:

$$M_\lambda(t) := \sum_{v \in Z(t)} h_{Y_v} \cdot (X_v(t) + \lambda t) \mathrm{e}^{-\lambda(X_v(t) + c_\lambda t)}. \tag{1.5}$$

$\{M_\lambda(t), t \geq 0\}$ is a martingale which may take both positive and negative values. We will prove that $M(\lambda) := \lim_{t \to +\infty} M_\lambda(t)$ exists for every $\lambda \geq \underline{\lambda}$ (see Lemma 10 below).

For every $i \in S$, suppose $(\xi_{i1}, \cdots, \xi_{id})$ is a random vector with the law $\{p_k(i) : k \in \mathbb{Z}_+^d\}$. Now we are ready to state the main results of this paper:

Theorem 1. *Suppose that* $E(\xi_{ij} \log^+ \xi_{ij}) < +\infty$ *for all* $i, j \in S$.

(1) When $c > \underline{c}$, *there is a unique traveling wave at speed* c *given by*

$$w(x, i) = E_{xi}\left[\exp\{-W(\lambda)\}\right] = E_{0i}\left[\exp\{-\mathrm{e}^{-\lambda x} W(\lambda)\}\right],$$

where $0 < \lambda < \underline{\lambda}$ is the root of the equation $c_\lambda = c$. Further, for every $i \in S$, $1 - w(x,i) \sim h_i \mathrm{e}^{-\lambda x}$ as $x \to +\infty$.

(2) When $c < \underline{c}$, there is no non-trivial traveling wave solution to (1.1) with speed c.

Theorem 2. *When $c = \underline{c}$ and $E\xi_{ij}(\log^+ \xi_{ij})^2 < +\infty$ for all $i, j \in S$, there is an unique traveling wave at speed $\underline{c}$ given by*

$$\underline{w}(x,i) = E_{xi}\left[\exp\{-M(\underline{\lambda})\}\right] = E_{0i}\left[\exp\{-\mathrm{e}^{-\underline{\lambda} x} M(\underline{\lambda})\}\right].$$

Further, for every $i \in S$, $1 - \underline{w}(x,i) \sim x\, h_i \mathrm{e}^{-\underline{\lambda} x}$ as $x \to +\infty$.

Comparing the above Theorems with corresponding results for the FKPP equation (see, for example, [8] and [12]), we see that λ^* plays the role of $\beta(m-1)$ in the case of single-type branching Brownian motions, where β is the branching rate and m is the mean number of particles split by one particle.

The remainder of this paper is structured as follows. In Section 2, we recall the basic setting of family trees and the size-biased trees with spine. We also introduce some known results for MBBM, including the so-called many-to-one formula, and McKean representation of traveling wave solutions, which are necessary in the arguments afterwards. In the remaining two sections we concentrate on proofs of Theorem 1 and Theorem 2. To prove that under some moment conditions, the traveling wave solution can be given in terms of martingale limit $W(\lambda)$ or $M(\underline{\lambda})$, we first answer when $W(\lambda)$ (in supercritical case) and $M(\underline{\lambda})$ (in critical case) are non-degenerate (see Theorem 3 and Theorem 5, respectively).

2 Multitype branching Brownian motion and basic facts

Let $\mathbb{N} = \{1, 2, \cdots\}$. We use $\Gamma := \bigcup_{n=0}^{+\infty} \mathbb{N}^n \cup \{\emptyset\}$ to describe the genealogical structure of our multitype branching processes. For $u, v \in \Gamma$, we use uv to stand for the concatenation of u and v ($u\emptyset = \emptyset u = u$). And therefore Γ contains elements like (i_1, i_2, i_3) or $(\emptyset, i_1, i_2, i_3)$ which represents the i_3th child of the i_2th child of the i_1th child of the initial ancestor $\emptyset$. For each $i \in \mathbb{N}$, we write $ui = (i_1, \cdots, i_n, i)$ for the ith child of u. We use the notation $v \prec u$ to mean that v is an ancestor of u and $u \in Z(t)$ when u is alive at time t.

A subset $\tau \subset \Gamma$ is called a Galton-Watson tree if:

1. $\emptyset \in \tau$;

2. if $u, v \in \Gamma$, then $uv \in \tau$ implies $u \in \tau$;

3. for all $u \in \tau$, there exists an $r_u \in \mathbb{N}$, such that when $j \in \mathbb{N}$, $uj \in \tau$ if and only if $1 \leq j \leq r_u$.

We denote the collection of Galton-Watson trees by $\mathbb{T}$. Each $u \in \tau$ is called a node of τ or an individual in τ or just a particle.

To fully describe the multitype branching process, we need to introduce the concept of marked Galton-Watson trees. We suppose that each particle $u \in \tau$ has a mark $(Y_u, X_u, \sigma_u, A_u)$ where

1. σ_u is the life time of u, which determines the fission time or the death time of particle u as $\zeta_u = \sum_{v \prec u} \sigma_v + \sigma_u$ $(\zeta_\emptyset = \sigma_\emptyset)$ and the birth time of u as $b_u = \sum_{v \prec u} \sigma_v$ $(b_\emptyset = 0)$;

2. Y_u gives the type of u, while $X_u : [b_u, \zeta_u) \to \mathbb{R}$ gives the spatial location of u at time $t \in [b_u, \zeta_u)$. We also interpret the notation $X_u(t)$ as the spatial location of the unique ancestor of u that was alive at time $t \leq \zeta_u$;

3. $A_u = (A_u(1), A_u(2), \cdots, A_u(d))$ gives the vector of offspring born by u when it dies.

We use (τ, Y, σ, A) or simply (τ, M) to denote a marked Galton-Watson tree. Let $\mathcal{T} := \{(\tau, M) \colon \tau \in \mathbb{T}\}$. Define

$$\begin{aligned}\mathcal{F}_t := \ & \sigma\left\{[u, Y_u, \sigma_u, A_u, (X_u(s), s \in [b_u, \zeta_u)) : u \in \tau \in \mathbb{T} \text{ with } \zeta_u \leq t] \text{ and} \right. \\ & \left. [u, Y_u, (X_u(s), s \in [b_u, t)) : u \in \tau \in \mathbb{T} \text{ with } t \in [b_u, \zeta_u)]\right\}\end{aligned}$$

Set $\mathcal{F} = \bigcup_{t \geq 0} \mathcal{F}_t$. There is a unique probability measure P on $(\mathcal{T}, \mathcal{F})$ such that the system is initiated by a single ancestor, a type $i \in S$ particle splits into offspring particles of all types according to distribution $\{p_k(i) : k \in \mathbb{Z}_+^d\}$ after a lifetime which is exponentially distributed with parameter $a_i > 0$, and each particle moves according to an independent copy of standard Brownian motion from its location of creation in its lifetime. We use P_{xi} (with associated expectation operator E_{xi}) to specify that the ancestor is of type $i \in S$ and spatially located at $x \in \mathbb{R}$.

Now we extend the probability space $(\mathcal{T}, \mathcal{F}, P)$ to $(\widetilde{\mathcal{T}}, \widetilde{F}, \widetilde{P})$ defined below. For any $\tau \in \mathbb{T}$, we can select a infinite line of descent $\varepsilon = \{\varepsilon_0 = \emptyset, \varepsilon_1, \varepsilon_2, \cdots\}$, where $\varepsilon_{n+1} \in \tau$ is a child of $\varepsilon_n \in \tau$, $n = 0, 1, 2, \cdots$. Such a genealogical line is called a *spine*. We write $u \in \varepsilon$ to mean that $u = \varepsilon_k$ for some $k \in \mathbb{Z}_+$. We use

$$\widetilde{\mathcal{T}} = \{(\tau, M, \varepsilon) \colon \varepsilon \subset \tau \in \mathbb{T}\}$$

to denote the set of marked trees with distinguished spines.

We use $\widetilde{Y} = (\widetilde{Y}_t, t \geq 0)$ to denote the type process of the spine, $\widetilde{X} = (\widetilde{X}_t, t \geq 0)$ the spatial movement of the spine, and $n = (n_t, t \geq 0)$ the counting process of fission times along the spine. Let $\text{node}_t(\varepsilon) := u$ if $u \in \varepsilon$ is the node in the spine that is alive at time t. Note that if $u \in \varepsilon$, $Y_u = \widetilde{Y}_{b_u} = \widetilde{Y}_{\zeta_u -}$.

If $u \in \varepsilon$, then at the fission time ζ_u, it gives birth to $\langle A_u, 1\rangle$ offspring, one of which continuing the spine (we write this node simply as $u+1$) while the others going on to create independent subtrees. Let O_u be the set of u's children except the one belonging to the spine, then for any $j \in \{1, 2, \cdots, \langle A_u, 1\rangle\}$ such that $uj \in O_u$, we use $(\tau, M)^u_j$ to denote the marked tree rooted at uj.

Now we introduce some filtrations on $\widetilde{\mathcal{T}}$ that we shall use later. First note that $\{\mathcal{F}_t, t \geq 0\}$ is also a filtration on $\widetilde{\mathcal{T}}$. Define

$$\begin{aligned}
\widetilde{\mathcal{F}}_t :=\ & \sigma\{\mathcal{F}_t, (\text{node}_s(\varepsilon), s \leq t)\}; \\
\mathcal{G}_t :=\ & \sigma\left\{\widetilde{Y}_s, \widetilde{X}_s : 0 \leq s \leq t\right\}; \quad \mathcal{G}^{\widetilde{Y}}_t := \sigma\left\{\widetilde{Y}_s, : 0 \leq s \leq t\right\}; \quad \mathcal{G}^{\widetilde{X}}_t := \sigma\left\{\widetilde{X}_s : 0 \leq s \leq t\right\}; \\
\widehat{\mathcal{G}}_t :=\ & \sigma\left\{\mathcal{G}_t, (\text{node}_s(\varepsilon), s \leq t), (\zeta_u, u \prec \text{node}_t(\varepsilon))\right\}; \\
\widetilde{\mathcal{G}}_t :=\ & \sigma\left\{\widehat{\mathcal{G}}_t, (A_u, u \prec \text{node}_t(\varepsilon))\right\}.
\end{aligned}$$

Set $\widetilde{\mathcal{F}} = \bigcup_{t\geq 0} \widetilde{\mathcal{F}}_t$, $\mathcal{G} = \bigcup_{t\geq 0} \mathcal{G}_t$, $\widehat{\mathcal{G}} = \bigcup_{t\geq 0} \widehat{\mathcal{G}}_t$ and $\widetilde{\mathcal{G}} = \bigcup_{t\geq 0} \widetilde{\mathcal{G}}_t$.

We need to extend the probability measure P on $(\mathcal{T}, \mathcal{F})$ to a probability measure $\widetilde{P}$ on $(\widetilde{\mathcal{T}}, \widetilde{\mathcal{F}})$ such that the spine is a single genealogical line of descent chosen from the underlying tree. Enlightened by [14], when a node u of type i on the spine has offspring vector $A_u = (A_u(1), A_u(2), \cdots, A_u(d))$, we pick one of these children at random to be the successor on the spine. Specifically, children are picked with probabilities proportional to h_j when their type is j. This means, when $u \in \tau$, we have

$$Prob(u \in \varepsilon | \mathcal{F}_t) = \prod_{v \prec u} \frac{h_{Y_{v+1}}}{\langle A_v, h\rangle}.$$

It is easy to see that

$$\sum_{u \in Z(t)} \prod_{v \prec u} \frac{h_{Y_{v+1}}}{\langle A_v, h\rangle} = 1.$$

To define $\widetilde{P}$ we recall the following representation from [14].

Lemma 1. *Every $\widetilde{\mathcal{F}}_t$-measurable function f can be written as*

$$f = \sum_{u \in Z(t)} f_u \mathbf{1}_{\{u \in \varepsilon\}} \tag{2.1}$$

where f_u is $\mathcal{F}_t$-measurable.

Definition 1. *We define the probability measure $\widetilde{P}$ on $(\widetilde{\mathcal{T}}, \widetilde{\mathcal{F}})$ by*

$$\int_{\widetilde{\mathcal{T}}} f \, \mathrm{d}\widetilde{P} = \int_{\widetilde{\mathcal{T}}} \sum_{u \in Z(t)} f_u \prod_{v \prec u} \frac{h_{Y_{v+1}}}{\langle A_v, h\rangle} \, \mathrm{d}P$$

for each $f \in \widetilde{\mathcal{F}}_t$ with representation (2.1).

It follows that for any bounded $\widetilde{\mathcal{F}}_t$-measurable function f with representation (2.1),

$$\begin{aligned}\widetilde{P}(f \mid \mathcal{F}_t) &= \widetilde{P}(\sum_{u\in Z(t)} f_u \mathbf{1}_{\{u\in\varepsilon\}} \mid \mathcal{F}_t) \\ &= \sum_{u\in Z(t)} f_u \widetilde{P}(\mathbf{1}_{\{u\in\varepsilon\}} \mid \mathcal{F}_t) \\ &= \sum_{u\in Z(t)} f_u \prod_{v\prec u} \frac{h_{Y_{v+1}}}{\langle A_v, h\rangle}.\end{aligned}$$

Then we have $\widetilde{P}(f) = P\left(\sum_{u\in Z(t)} f_u \prod_{v\prec u} \frac{h_{Y_{v+1}}}{\langle A_v,h\rangle}\right)$ and $\widetilde{P}(\widetilde{\mathcal{T}}) = 1$, which means that $\widetilde{P}$ is an extension of P onto $(\widetilde{\mathcal{T}}, \widetilde{\mathcal{F}})$.

Intuitively, following the above method of choosing spine nodes, the type process of the spine $\widetilde{Y}$ is a continuous time Markov process valued in S, which stays at any state $i \in S$ for an exponential time with parameter a_i, and then transits to state j with probability $P(i,j) := \sum_{k\in\mathbb{Z}_+^d} p_k(i) \frac{k_j h_j}{\langle k,h\rangle}$. Given $\widehat{\mathcal{G}}_t$, the trajectory of $\widetilde{Y}$, the node of the spine and the birth time of each spine node before time t are determined. Then we have

$$\widetilde{P}(A_v = k_v,\ \forall v \prec \varepsilon_{n_t} \,|\, \widehat{\mathcal{G}}_t) = \prod_{v\prec\varepsilon_{n_t}} \frac{p_{k_v}(Y_v)}{P(Y_v, Y_{v+1})} \frac{k_v(Y_{v+1}) h_{Y_{v+1}}}{\langle k_v, h\rangle}$$

where $k_v = (k_v(1), k_v(2), \cdots, k_v(d)) \in \mathbb{Z}_+^d$.

Now, we can construct a probability measure $\widetilde{P}$ on $\widetilde{\mathcal{F}}_t$ by

$$\begin{aligned}\mathrm{d}\widetilde{P}(\tau, M, \varepsilon)|_{\widetilde{\mathcal{F}}_t} &= \mathrm{d}\mathbb{P}(\widetilde{Y})\mathrm{d}\mathbb{B}(\widetilde{X}) \prod_{v\prec\varepsilon_{n_t}} \frac{p_{A_v}(Y_v)}{P(Y_v, \widetilde{Y}_{v+1})} \frac{A_v(Y_{v+1}) h_{Y_{v+1}}}{\langle A_v, h\rangle} \\ &\quad \prod_{v\prec\varepsilon_{n_t}} \left[\frac{1}{A_v(Y_{v+1})} \prod_{j:\, vj\in O_v} \mathrm{d}P^{t-\zeta_v}_{\widetilde{Y}_{\zeta_v}}((\tau, M)^v_j) \right]. \qquad (2.2)\end{aligned}$$

Here $\mathbb{B}(X)$ is the law of a standard Brownian motion and $\mathbb{P}(\widetilde{Y})$ is the law of the type process $\widetilde{Y}$ which is a continuous time Markov process valued on S, and $Y_v = \widetilde{Y}_{b_v}$ for $v \in \varepsilon$.

The decomposition of $\widetilde{P}$ suggests the following intuitive description of the system under the measure $\widetilde{P}$:

- The spine's type process $\widetilde{Y}$ moves as a continuous time Markov process taking values in S according to the measure $\mathbb{P}$. The generator $G = (g_{ij})_{i,j\in S}$ of $\widetilde{Y}$ is given by $g_{ij} = a_i\,(P(i,j) - \delta_{ij})$. The spine's spatial movement is a standard Brownian motion.

- The fission time ζ_v of node v in the spine is exactly the jumping time of the spine's type process $\widetilde{Y}$, *i.e.* that the life time σ_v of v is exponentially distributed with parameter $a_{\widetilde{Y}_v}$. ($\widetilde{Y}$ may jump from i to itself at jumping time according to generator G.)

- At the fission time of node v in the spine, the single spine particle is replaced by a random vector A_v of offspring with A_v distributed according to the law $(p_k(\widetilde{Y}_{\zeta_v}))_{k\in\mathbb{Z}_+^d}$.
- At the fission time of node v in the spine, a type j child is picked to be the next spine node with probability $\frac{h_j}{\langle A_v,h\rangle}$.
- Each of the remaining $\langle A_v,1\rangle-1$ non-spine children of v gives rise to independent subtrees $(\tau,M)_j^v$ for $vj\in O_v$, which evolve as independent subtrees determined by the probability $P_{\widetilde{Y}_{\zeta_v}\widetilde{X}_{\zeta_v}}$ shifted to the time of creation.

Let $N(t):=(N_1(t),N_2(t),\cdots,N_d(t))$ be the vector denoting the population sizes of different types at time t. Note that $\{N(t),t\geq 0\}$ is a multitype branching process. Then we have the following result.

Lemma 2 ([1] Chapter V, Theorem 1)**.** $\left\{w(t)=\frac{\langle N(t),h\rangle}{e^{\lambda^* t}\langle N(0),h\rangle}: t\geq 0\right\}$ *is a non-negative martingale with respect to* $\{\mathcal{F}_t\colon t\geq 0\}$*.*

Noting that $w(t)$ is a non-negative mean one martingale, we can define a probability measure Q on $(\mathcal{T},\mathcal{F})$ by

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathcal{F}_t}=w(t). \tag{2.3}$$

In order to make the principles of measure change method clear, we introduce the following technical lemma.

Lemma 3. *Suppose* $\widetilde{\mu}_1$ *and* $\widetilde{\mu}_2$ *are two probability measures defined on the same space* $(\Omega,\widetilde{\mathcal{F}})$ *with Radon-Nikodym derivative*

$$\frac{\mathrm{d}\widetilde{\mu}_2}{\mathrm{d}\widetilde{\mu}_1}=g.$$

If $\mathcal{F}$ *is a sub-*σ*-field of* $\widetilde{\mathcal{F}}$*, then the two measures* $\mu_1:=\widetilde{\mu}_1|_{\mathcal{F}}$ $\mu_2:=\widetilde{\mu}_2|_{\mathcal{F}}$ *on* $(\Omega,\mathcal{F})$ *are related by the conditional expectation operation:*

$$\frac{\mathrm{d}\mu_2}{\mathrm{d}\mu_1}=\widetilde{\mu}_1(g\mid\mathcal{F}).$$

Proof. For any set $A\in\mathcal{F}$, we have

$$\mu_2(A)=\widetilde{\mu}_2(A)=\int_A g\,\mathrm{d}\widetilde{\mu}_1=\int_A\widetilde{\mu}_1(g\mid\mathcal{F})\,\mathrm{d}\mu_1.$$

The last equality follows from the property of conditional expectation. By the definition of Radon-Nikodym derivative we reach the conclusion. □

Lemma 3 implies that if we want to extend Q to $(\widetilde{\mathcal{T}}, \widetilde{\mathcal{F}})$, we need to construct a non-negative martingale $\widetilde{w}(t)$ with respect to $\{\widetilde{\mathcal{F}}_t \colon t \geq 0\}$ satisfying

$$\left.\frac{\mathrm{d}\widetilde{Q}}{\mathrm{d}\widetilde{P}}\right|_{\widetilde{\mathcal{F}}_t} = \widetilde{w}(t), \tag{2.4}$$

and

$$\widetilde{P}(\widetilde{w}(t) \mid \mathcal{F}_t) = w(t). \tag{2.5}$$

According to Lemma 1, we can write $\widetilde{w}(t) = \sum_{v \in Z(t)} w_v \mathbf{1}_{\{v \in \varepsilon\}}$ where w_v is $\mathcal{F}_t$-measurable. Thus

$$\begin{aligned} \widetilde{P}(\widetilde{w}(t) \mid \mathcal{F}_t) &= \widetilde{P}\left(\sum_{v \in Z(t)} w_v \mathbf{1}_{\{v \in \varepsilon\}} \mid \mathcal{F}_t\right) \\ &= \sum_{v \in Z(t)} w_v \widetilde{P}(\mathbf{1}_{\{v \in \varepsilon\}} \mid \mathcal{F}_t) \\ &= \sum_{v \in Z(t)} w_v \prod_{u \prec v} \frac{h_{Y_{u+1}}}{\langle A_u, h \rangle}. \end{aligned}$$

Since $w(t) = \sum_{v \in Z(t)} \frac{h_{Y_v}}{\mathrm{e}^{\lambda^* t} \langle N(0), h \rangle}$, in order to have (2.5), we need

$$w_v = \frac{h_{Y_v}}{\mathrm{e}^{\lambda^* t} \langle N(0), h \rangle} \left(\prod_{u \prec v} \frac{h_{Y_{u+1}}}{\langle A_u, h \rangle}\right)^{-1} = \mathrm{e}^{-\lambda^* t} \prod_{u \prec v} \frac{\langle A_u, h \rangle}{h_{Y_u}},$$

thus we get

$$\widetilde{w}(t) = \mathrm{e}^{-\lambda^* t} \prod_{v \prec \varepsilon_{n_t}} \frac{\langle A_v, h \rangle}{h_{\widetilde{Y}_{\zeta_v}}}. \tag{2.6}$$

Here we remind the reader that till now we only deduce the expression of $\widetilde{w}(t)$, but we have not proved it is a martingale yet. Next we will prove that $\{\widetilde{w}(t) \colon t \geq 0\}$ is indeed a martingale with respect to $\{\widetilde{\mathcal{F}}_t \colon t \geq 0\}$.

First of all, for each type $i \in S$, we introduce the size-biased distribution

$$\hat{p}_k(i) := \frac{p_k(i) \langle k, h \rangle}{\left(1 + \frac{\lambda^*}{a_i}\right) h_i}. \tag{2.7}$$

It is indeed a probability, since

$$\sum_{k \in \mathbb{Z}_+^d} p_k(i) \langle k, b \rangle = \sum_{j=1}^{d} m_{ij} h_j = \left(1 + \frac{\lambda^*}{a_i}\right) h_i, \quad i \in S.$$

The last equality follows from the fact that h is the right eigenvector of A with respect to λ^*. For any $i, j \in S$, Define

$$\begin{aligned} \widehat{P}(i,j) &:= \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(i) \frac{k_j h_j}{\langle k, h \rangle} \\ &= \frac{m_{ij} h_j}{\left(1 + \frac{\lambda^*}{a_i}\right) h_i}. \end{aligned} \tag{2.8}$$

It is easy to see that $\{\widehat{P}(i,j) : i,j \in S\}$ is a family of transition probabilities.

Lemma 4. *Suppose $(\widetilde{Y}, \mathbb{P})$ is defined as before. Define*

$$m_t := \mathrm{e}^{-\lambda^* t} \prod_{v \prec \varepsilon_{n_t}} \left(1 + \frac{\lambda^*}{a_{Y_v}}\right) \frac{\widehat{P}(Y_v, Y_{v+1})}{P(Y_v, Y_{v+1})}, \quad t \geq 0.$$

Then $\{m_t, t \geq 0\}$ is a non-negative mean one martingale with respect to $\{\mathcal{G}_t^{\widetilde{Y}}, t \geq 0\}$. We define another probability measure $\hat{\mathbb{P}}$ by

$$\left.\frac{\mathrm{d}\hat{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} = m_t. \tag{2.9}$$

Then under $\hat{\mathbb{P}}$, $\widetilde{Y}$ moves as a continuous time Markov process with generator $\hat{g}_{ij} := (a_i + \lambda^*)(\widehat{P}(i,j) - \delta_{ij})$.

Proof. Suppose $f : S \to R$ is a bounded measurable function.

$$u(t,x) := \mathbb{E}^x[f(\widetilde{Y}_t) m_t]$$

where $\mathbb{P}^x(\cdot) := \mathbb{P}(\cdot | \widetilde{Y}_0 = x)$ with associated expectation operator $\mathbb{E}^x$.

We define τ to be the first jumping time of $\widetilde{Y}$. Then by the strong Markov property, $u(t,x)$ can be written as

$$\begin{aligned} u(t,x) &= \mathbb{E}^x[f(\widetilde{Y}_t) m_t \mathbf{1}_{\{t<\tau\}}] + \mathbb{E}^x[f(\widetilde{Y}_t) m_t \mathbf{1}_{\{t \geq \tau\}}] \\ &= f(x)\mathrm{e}^{-(a_x+\lambda^*)t} + \int_0^t \mathrm{e}^{-(a_x+\lambda^*)s}(a_x + \lambda^*) \sum_{j \in S} \widehat{P}(x,j) u(t-s,j)\,\mathrm{d}s \\ &= f(x)\mathrm{e}^{-(a_x+\lambda^*)t} + \int_0^t \mathrm{e}^{-(a_x+\lambda^*)(t-s)}(a_x + \lambda^*) \sum_{j \in S} \widehat{P}(x,j) u(s,j)\,\mathrm{d}s \end{aligned}$$

Therefore, $u(t,x)$ satisfies

$$\frac{\partial u}{\partial t} = (a_x + \lambda^*) \sum_{j \in S} (\widehat{P}(x,j) - \delta_{xj}) u(t,j) \tag{2.10}$$

with $u(0,x) = f(x)$. In particular, if we pick $f \equiv 1$, from the uniqueness of solution to the array of ordinary partial differential equations, we obtain that $\mathbb{E}^x m_t \equiv 1$ which together with the Markov property of $\widetilde{Y}$ under $\mathbb{P}$ makes sure that m_t is a martingale. Thus the measure $\hat{\mathbb{P}}$ is well defined.

From (2.10) we see that under $\hat{\mathbb{P}}$, $\widetilde{Y}$ is a Markov Process with generator $\hat{g}_{ij}$. In other words, under probability measure $\hat{\mathbb{P}}$, $\widetilde{Y}$ can be interpreted as a Markov process which stays at any state $i \in S$ for an exponential time with parameter $a_i + \lambda^*$, and then transits to state j with probability $\widehat{P}(i,j)$. □

Just as we did before, we can construct a probability measure $\widetilde{Q}$ on $(\widetilde{\mathcal{T}}, \widetilde{\mathcal{F}})$ by

$$
\begin{aligned}
\mathrm{d}\widetilde{Q}(\tau, M, \varepsilon)|_{\widetilde{\mathcal{F}}_t} &= \mathrm{d}\hat{\mathbb{P}}(\widetilde{Y})\mathrm{d}\mathbb{B}(\widetilde{X}) \prod_{v \prec \varepsilon_{n_t}} \frac{\hat{p}_{A_v}(Y_v)}{\widehat{P}(Y_v, Y_{v+1})} \frac{A_v(Y_{v+1}) h_{Y_{v+1}}}{\langle A_v, h\rangle} \\
&\prod_{v \prec \varepsilon_{n_t}} \left[\frac{1}{A_v(Y_{v+1})} \prod_{j\colon vj \in O_v} \mathrm{d}P^{t-\zeta_v}_{\widetilde{X}_{\zeta_v} \widetilde{Y}_{\zeta_v}}((\tau, M)^v_j) \right], \qquad (2.11)
\end{aligned}
$$

which can be described as follows: under $\widetilde{Q}$

- The spine's type process $\widetilde{Y}$ moves as a continuous time Markov process valued on S according to the measure $\hat{\mathbb{P}}$. The generator of $\widetilde{Y}$ is given by $\hat{g}_{ij} = (a_i + \lambda^*)\left(\widehat{P}(i,j) - \delta_{ij}\right)$. The spine's spatial movement $\widetilde{X}$ is a standard Brownian motion.

- The fission time ζ_v of node v in the spine is exactly the jumping time of the spine's type process $\widetilde{Y}$, *i.e.* that σ_v has an exponential distribution with parameter $a_{\widetilde{Y}_v} + \lambda^*$.

- At the fission time of node v in the spine, the single spine particle is replaced by a random vector A_v of offspring with A_v distributed according to the law $(\hat{p}_k(\widetilde{Y}_{\zeta_v}))_{k \in \mathbb{Z}^d_+}$.

- At the fission time of node v in the spine, a type j particle from the offspring of v will be picked to be the next spine node with probability $\frac{h_j}{\langle A_v, h\rangle}$.

- Each of the remaining $\langle A_v, 1\rangle - 1$ non-spine children of v gives rise to independent subtrees $(\tau, M)^v_j$ for $vj \in O_v$, which evolves as independent subtrees determined by the probability $P_{\widetilde{Y}_{\zeta_v} \widetilde{X}_{\zeta_v}}$ shifted to the time of creation.

Applying (2.7), (2.9) and (2.2) into (2.11), we can easily get (2.4). Therefore $\{\widetilde{w}(t)\colon t \geq 0\}$ is a non-negative martingale with respect to $\{\widetilde{\mathcal{F}}_t\colon t \geq 0\}$.

The following formula is a byproduct of the above spine construction.

Proposition 1 (Many-to-one formula for MBBM)**.** *For any measurable function $f\colon \mathbb{R} \times S \to \mathbb{R}$, we have*

$$
E_{xy}\left(\sum_{u \in Z(t)} f(X_u(t), Y_u) \right) = \widehat{E}_{xy}\left(f(\widetilde{X}_t, \widetilde{Y}_t) \frac{h_{\widetilde{Y}_0}}{h_{\widetilde{Y}_t}} \mathrm{e}^{\lambda^* t} \right). \qquad (2.12)
$$

Here $\widehat{E}_{xy}$ denotes the law of one particle motion where the type process $\widetilde{Y}$ moves as a Markov process starting from y with generator $\hat{g}_{ij} := (a_i + \lambda^*)(\widehat{P}(i,j) - \delta_{ij})$, while the spatial location process $\widetilde{X}$ moves as a Brownian motion starting from x and is independent of $\widetilde{Y}$.

Proof. The proof is much the same as [6] Theorem 4.1 in the case of multitype Markov branching processes. We omit the details here. □

Lemma 5 (McKean representation)**.** *If $u(t,x,y) \in [0,1]$ is twice continuously differentiable in x and satisfies the parabolic system of equations* (1.1) *with initial condition $u(0,x,y) = f(x,y)$, then u has a McKean representation*

$$u(t,x,y) = E_{xy}\left(\prod_{u\in Z(t)} f(X_u(t), Y_u)\right).$$

Proof. This can be proved by a similar argument as that of [3] Theorem 1.36. We omit the details here. □

Lemma 6. *Suppose $c \in \mathbb{R}$. $w(x,y)$ is a bounded function with $0 \le w(x,y) \le 1$ for any $(x,y) \in \mathbb{R}\times S$. Let $u(t,x,y) := w(x-ct,y)$. Then u satisfies* (1.1) *if and only if*

$$w(x,y) = E_{xy}\left[\prod_{u\in Z(t)} w(X_u(t)+ct, Y_u)\right].$$

Proof. By Lemma 5, we only need to show the sufficiency. Let P_t denote the semi-group of one-dimensional Brownian motion. Let τ denote the split time of the root. We have

$$\begin{aligned} u(t,x,y) &= E_{(x-ct)\,y}\left(\prod_{u\in Z(t)} w(X_u(t)+ct, Y_u)\right) \\ &= E_{xy}\left(\prod_{u\in Z(t)} w(X_u(t), Y_u)\right) \\ &= E_{xy}\left(\prod_{u\in Z(t)} w(X_u(t), Y_u)1_{\{\tau\le t\}}\right) + E_{x\,y}\left(\prod_{u\in Z(t)} w(X_u(t), Y_u)1_{\{\tau> t\}}\right) \\ &= \int_0^t a_y \mathrm{e}^{-a_y s} P_s\psi_y(u(t-s))(x)\,\mathrm{d}s + \mathrm{e}^{-a_y t}P_t w_y(x), \end{aligned}$$

where for each $s \ge 0$, $u(s)$ is a function from $\mathbb{R}$ to $\mathbb{R}^d$ defined by $u(s)(x) := u(s,x) = (u(s,x,1),\cdots,u(s,x,d))$. Therefore, $u(t,x,y)$ solves (1.1). □

3 Proof of Theorem 1

Recall that, for any $\lambda \ne 0$,

$$W_\lambda(t) := \sum_{u\in Z(t)} h_{Y_u}\mathrm{e}^{-\lambda(X_u(t)+c_\lambda t)}.$$

It follows from Proposition 1 that $\{W_\lambda(t), t \geq 0\}$ is a positive martingale and thus it has an almost sure limit denoted by $W(\lambda)$. The following theorem answers when $W(\lambda)$ is non-degenerate, which will be used to give explicit expressions of traveling wave solutions in supercritical case.

Theorem 3. *(1) If $|\lambda| \geq \underline{\lambda}$, then $W(\lambda) = 0\, P_{xy}$-a.s..*

(2) Suppose $0 < |\lambda| < \underline{\lambda}$. If $E(\xi_{ij} \log^+ \xi_{ij}) < +\infty$ for all $i, j \in S$, then $W_\lambda(t)$ converges to $W(\lambda)$ in $L^1(P_{xy})$, and $P_{xy}(W(\lambda) = 0) = 0$. Else if $E(\xi_{ij} \log^+ \xi_{ij}) = +\infty$ for some $i, j \in S$, then $W(\lambda) = 0$ P_{xy}-a.s..

Remark 1. *It suffices to prove the claims under P_{0y}. In the following paper, we only deal with the case $\lambda > 0$. The case $\lambda < 0$ can be analyzed by simple considerations of symmetry.*

For any $\lambda \geq 0$, through the same techniques used in Section 2 , we can construct a probability measure $\widetilde{Q}^\lambda_{0y}$ on $(\widetilde{\mathcal{T}}, \widetilde{\mathcal{F}})$ such that

$$\left.\frac{\mathrm{d}Q^\lambda_{0y}}{\mathrm{d}P_{0y}}\right|_{\mathcal{F}_t} = \frac{W_\lambda(t)}{W_\lambda(0)},$$

where $Q^\lambda_{0y} := \widetilde{Q}^\lambda_{0y}|_{\mathcal{F}}$. In fact, $\widetilde{Q}^\lambda_{0y}$ has the following decomposition:

$$\begin{aligned} \mathrm{d}\widetilde{Q}^\lambda_{0y}(\tau, M, \varepsilon)|_{\widetilde{\mathcal{F}}_t} &= \mathrm{d}\hat{\mathbb{P}}_y(\widetilde{Y})\mathrm{d}\mathbb{P}^{-\lambda}(\widetilde{X}) \prod_{v \prec \varepsilon_{n_t}} \frac{\hat{p}_{A_v}(Y_v)}{\widehat{P}(Y_v, Y_{v+1})} \frac{A_v(Y_{v+1}) h_{Y_{v+1}}}{\langle A_v, h\rangle} \\ & \prod_{v \prec \varepsilon_{n_t}} \left[\frac{1}{A_v(Y_{v+1})} \prod_{j:\, vj \in O_v} \mathrm{d}P^{t-\zeta_v}_{\widetilde{X}_{\zeta_v} \widetilde{Y}_{\zeta_v}}((\tau, M)^v_j) \right]. \end{aligned} \tag{3.1}$$

Here $(\widetilde{X}, \mathbb{P}^{-\lambda})$ is a standard Brownian motion with drift $-\lambda$, and $(\widetilde{Y}, \hat{\mathbb{P}}_y)$ is a continuous time Markov chain starting from y with generator $\hat{g}_{ij} = (a_i + \lambda^*)(\widehat{P}(i,j) - \delta_{ij})$. For $vj \in O_v$, $(\tau, M)^v_j$ evolves as independent subtrees determined by the probability $P_{\widetilde{X}_{\zeta_v} \widetilde{Y}_{\zeta_v}}$ shifted to the time of creation.

Lemma 7. *We have the following spine decomposition for the martingale $W_\lambda(t)$:*

$$\widetilde{Q}^\lambda_{0y}(W_\lambda(t) \,|\, \widetilde{\mathcal{G}}) = h_{\widetilde{Y}_t} \mathrm{e}^{-\lambda(\widetilde{X}(t) + c_\lambda t)} + \sum_{j \in S} \sum_{v \prec \varepsilon_{n_t}} (A_v(j) - \delta_{Y_{v+1} j}) h_j \mathrm{e}^{-\lambda(\widetilde{X}(\zeta_v) + c_\lambda \zeta_v)}. \tag{3.2}$$

Proof. $W_\lambda(t)$ can be written as

$$\begin{aligned} W_\lambda(t) &= h_{\widetilde{Y}_t} \mathrm{e}^{-\lambda(\widetilde{X}(t) + c_\lambda t)} + \sum_{\substack{u \in Z(t) \\ u \notin \varepsilon}} h_{Y_u} \mathrm{e}^{-\lambda(X_u(t) + c_\lambda t)} \\ &= h_{\widetilde{Y}_t} \mathrm{e}^{-\lambda(\widetilde{X}(t) + c_\lambda t)} + \sum_{v \prec \varepsilon_{n_t}} \sum_{j:\, vj \in O_v} \sum_{\substack{u \in Z(t) \\ u \in (\tau, M)^v_j}} h_{Y_u} \mathrm{e}^{-\lambda(X_u(t) + c_\lambda t)}. \end{aligned}$$

The first equality is clearly true since one of the particles $u \in Z(t)$ must stay in the spine. The second one is followed by partitioning the particles into distinct subtrees that were born from the spine nodes before time t.

Recall that $\widetilde{\mathcal{G}}$ contains all information about the spine nodes, by taking the $\widetilde{Q}^{\lambda}_{0y}$ conditional expectation of $W_{\lambda}(t)$, we have

$$\begin{aligned}\widetilde{Q}^{\lambda}_{0y}(W_{\lambda}(t)\,|\,\widetilde{\mathcal{G}}) &= h_{\widetilde{Y}_t}\mathrm{e}^{-\lambda(\widetilde{X}(t)+c_{\lambda}t)} + \widetilde{Q}^{\lambda}_{0y}\left(\sum_{v\prec\varepsilon_{n_t}}\sum_{j:\,vj\in O_v}\sum_{\substack{u\in Z(t)\\ u\in(\tau,M)^v_j}} h_{Y_u}\mathrm{e}^{-\lambda(X_u(t)+c_{\lambda}t)}\,\middle|\,\widetilde{\mathcal{G}}\right)\\ &= h_{\widetilde{Y}_t}\mathrm{e}^{-\lambda(\widetilde{X}(t)+c_{\lambda}t)} + \sum_{v\prec\varepsilon_{n_t}}\sum_{j:\,vj\in O_v} h_{Y_{vj}}\mathrm{e}^{-\lambda(\widetilde{X}(\zeta_v)+c_{\lambda}\zeta_v)}\\ &\qquad \widetilde{Q}^{\lambda}_{0y}\left(\sum_{\substack{u\in Z(t)\\ u\in(\tau,M)^v_j}}\frac{h_{Y_u}}{h_{Y_{vj}}}\mathrm{e}^{-\lambda(X_u(t)-\widetilde{X}(\zeta_v)+c_{\lambda}(t-\zeta_v))}\,\middle|\,\widetilde{\mathcal{G}}\right).\end{aligned}$$

From the decomposition of $\mathrm{d}\widetilde{Q}^{\lambda}_{0y}$, we observe that under $\widetilde{Q}^{\lambda}_{0y}$, the subtrees coming off the spine evolves as if under the measure P_{0y}. Therefore

$$\widetilde{Q}^{\lambda}_{0y}\left(\sum_{\substack{u\in Z(t)\\ u\in(\tau,M)^v_j}}\frac{h_{Y_u}}{h_{Y_{vj}}}\mathrm{e}^{-\lambda(X_u(t)-\widetilde{X}(\zeta_v)+c_{\lambda}(t-\zeta_v))}\,\middle|\,\widetilde{\mathcal{G}}\right) = 1.$$

This equality is true because the additive expression being evaluated on the subtrees is just a shifted form of the martingale $W_{\lambda}(t)$. We complete the proof. □

Lemma 8 (Durret, [5], P241)**.** *Suppose μ and ν are two probability measures on a measurable space $(\Omega, \mathcal{F})$ with filtration $(\mathcal{F}_t)_{t\geq 0}$, such that*

$$\left.\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right|_{\mathcal{F}_t} = M(t)$$

for all $t \geq 0$. Let $M_{\infty} := \limsup_{t\to+\infty} M(t)$. Then for any $A \in \mathcal{F}$

$$\mu(A) = \int_A M_{\infty}\,\mathrm{d}\nu + \mu(A\cap\{M_{\infty} = +\infty\}),$$

and consequently

$$M_{\infty} = 0\ \nu\text{-a.s.} \iff M_{\infty} = +\infty\ \ \mu\text{-a.s.};$$

$$\int_{\Omega} M_{\infty}\,\mathrm{d}\nu = 1 \iff M_{\infty} < +\infty\ \ \mu\text{-a.s.}.$$

Proof of Theorem 3: (1) If $\lambda \geq \underline{\lambda} > 0$, then $\lambda \geq c_\lambda$. Obviously we have

$$W_\lambda(t) \geq h_{\widetilde{Y}_t} \mathrm{e}^{-\lambda \widetilde{X}(t) - \left(\frac{\lambda^2}{2} + \lambda^*\right)t} \geq C_0 \mathrm{e}^{-\lambda t\left(\frac{\widetilde{X}(t)}{t} + c_\lambda\right)}$$

for some constant $C_0 > 0$. Recalling that $\widetilde{X}$ moves as a Brownian motion with drift $-\lambda$ under $\widetilde{Q}^\lambda_{0y}$, we have $\lim_{t\to+\infty} \frac{\widetilde{X}(t)}{t} = -\lambda$ and $\liminf_{t\to+\infty} \widetilde{X}(t) + \lambda t = -\infty$, and

$$\limsup_{t\to+\infty} W_\lambda(t) = +\infty \quad \widetilde{Q}^\lambda_{0y}\text{-a.s.}.$$

In view of Lemma 8, we have $P_{0y}(W(\lambda) = 0) = 1$.

(2) When $0 < \lambda < \underline{\lambda}$, $\lambda < c_\lambda$. Suppose $E(\xi_{ij} \log^+ \xi_{ij}) = +\infty$ for some $i, j \in S$. First note that at each fission time of the spine, we have the lower bound

$$W_\lambda(\zeta_{\varepsilon_n}) \geq \langle A_{\varepsilon_n}, h \rangle \mathrm{e}^{-\lambda(\widetilde{X}(\zeta_{\varepsilon_n}) + c_\lambda \zeta_{\varepsilon_n})},$$

thus by Lemma 8, it suffices to show

$$\widetilde{Q}^\lambda_{0y}\left(\limsup_{n\to+\infty} \langle A_{\varepsilon_n}, h \rangle \mathrm{e}^{-\lambda(\widetilde{X}(\zeta_{\varepsilon_n}) + c_\lambda \zeta_{\varepsilon_n})} = +\infty\right) = 1. \tag{3.3}$$

Obviously we have

$$\langle A_{\varepsilon_n}, h \rangle \mathrm{e}^{-\lambda(\widetilde{X}(\zeta_{\varepsilon_n}) + c_\lambda \zeta_{\varepsilon_n})} = \exp\left\{ n \left[\frac{\log \langle A_{\varepsilon_n}, h \rangle}{n} - \lambda \frac{\zeta_{\varepsilon_n}}{n} \left(\frac{\widetilde{X}(\zeta_{\varepsilon_n})}{\zeta_{\varepsilon_n}} + c_\lambda \right) \right] \right\}.$$

Since $\widetilde{X}$ moves as a Brownian motion with drift $-\lambda$ under $\widetilde{Q}^\lambda_{oy}$, we have

$$\widetilde{Q}^\lambda_{0y}\left(\lim_{t\to+\infty} \frac{\widetilde{X}(t)}{t} + c_\lambda = c_\lambda - \lambda > 0 \right) = 1.$$

Besides, by strong law of large numbers we have

$$\widetilde{Q}^\lambda_{0y}\left(\limsup_{n\to+\infty} \frac{\zeta_{\varepsilon_n}}{n} \leq \sum_{k\in S} (a_k + \lambda^*) < +\infty \right) = 1.$$

Therefore to prove (3.3), we only need to prove

$$\widetilde{Q}^\lambda_{0y}\left(\limsup_{n\to+\infty} \frac{\log \langle A_{\varepsilon_n}, h \rangle}{n} = +\infty \right) = 1. \tag{3.4}$$

Let $N_i(n)$ denote the total number of jumps of $\widetilde{Y}$ before it hits state i for the nth time. Since $\widetilde{Y}$ moves as an irreducible Markov chain under $\widetilde{Q}^\lambda_{0y}$, $n/N_i(n)$ converges to a positive constant with probability one. Notice that $\{A_{\varepsilon_{N_i(n)}} : n \geq 0\}$ is a sequence of independent random vectors with the same distribution law $\{\hat{p}_k(i) : k \in \mathbb{Z}^d_+\}$. Immediately by the moment condition on ξ_{ij}, we have

$$\widetilde{Q}^\lambda_{0y} \log \langle A_{\varepsilon_{N_i(n)}}, h \rangle = +\infty.$$

By the Borel-Cantelli lemma,

$$\limsup_{n\to+\infty}\frac{\log\langle A_{\varepsilon_{N_i(n)}},h\rangle}{n}=+\infty\quad \widetilde{Q}^{\lambda}_{0y}\text{-a.s.},$$

and consequently

$$\limsup_{n\to+\infty}\frac{\log\langle A_{\varepsilon_{N_i(n)}},h\rangle}{N_i(n)}=+\infty\quad \widetilde{Q}^{\lambda}_{0y}\text{-a.s.}$$

which implies (3.4).

Now we suppose $E(\xi_{ij}\log^+\xi_{ij})<+\infty$ for all $i,j\in S$. Then we have for any $i\in S$,

$$\limsup_{n\to+\infty}\frac{\log\langle A_{\varepsilon_{N_i(n)}},h\rangle}{N_i(n)}=0\quad \widetilde{Q}^{\lambda}_{0y}\text{-a.s.},$$

and consequently

$$\begin{aligned}
&\widetilde{Q}^{\lambda}_{0y}\left(\sum_{n=1}^{+\infty}\langle A_{\varepsilon_n},h\rangle \mathrm{e}^{-\lambda(\widetilde{X}(\zeta_{\varepsilon_n})+c_\lambda\zeta_{\varepsilon_n})}<+\infty\right)\\
=\ &\widetilde{Q}^{\lambda}_{0y}\left(\sum_{i\in S}\sum_{n=1}^{+\infty}\langle A_{\varepsilon_{N_i(n)}},h\rangle \mathrm{e}^{-\lambda(\widetilde{X}(\zeta_{\varepsilon_{N_i(n)}})+c_\lambda\zeta_{\varepsilon_{N_i(n)}})}<+\infty\right)\\
=\ &\widetilde{Q}^{\lambda}_{0y}\left(\sum_{i\in S}\sum_{n=1}^{+\infty}\exp\left\{N_i(n)\left(\frac{\log\langle A_{\varepsilon_{N_i(n)}},h\rangle}{N_i(n)}-\lambda\frac{\zeta_{\varepsilon_{N_i(n)}}}{N_i(n)}\left(\frac{\widetilde{X}(\zeta_{\varepsilon_{N_i(n)}})}{\zeta_{\varepsilon_{N_i(n)}}}+c_\lambda\right)\right)\right\}<+\infty\right)\\
=\ &1.
\end{aligned}$$

The last equality is because $\frac{\widetilde{X}(\zeta_{\varepsilon_n})}{\zeta_{\varepsilon_n}}\to-\lambda>-c_\lambda$ as $n\to+\infty$. Therefore the second term in (3.2) is bounded from above for all $t>0$. In addition, under $\widetilde{Q}^{\lambda}_{0y}$, $-\lambda(\widetilde{X}(t)+c_\lambda t)=-\lambda t\left(\frac{\widetilde{X}(t)}{t}+c_\lambda\right)\to-\infty$ as $t\to+\infty$. Thus the first term in (3.2) is also bounded from above. So we have

$$\limsup_{t\to+\infty}\widetilde{Q}^{\lambda}_{0y}\left(W_\lambda(t)\,|\,\widetilde{\mathcal{G}}\right)<+\infty\quad \widetilde{Q}^{\lambda}_{0y}\text{-a.s.}$$

By Fatou's lemma

$$\limsup_{t\to+\infty}W_\lambda(t)<+\infty\quad \widetilde{Q}^{\lambda}_{0y}\text{-a.s.}$$

Therefore, by Lemma 8, $W_\lambda(t)$ converges to $W(\lambda)$ in $L^1(P_{0y})$ which implies that $W(\lambda)$ is non-degenerate.

Let $q_y:=P_{0y}(W(\lambda)=0)<1$. We have for any $t>s\geq 0$

$$W_\lambda(t)=\sum_{v\in Z(s)}\mathrm{e}^{-\lambda(X_v(s)+c_\lambda s)}W_\lambda(t-s,v)$$

where $W_\lambda(t-s,v)$ are independent copies of $W_\lambda(t-s)$ initiated by $v\in Z(s)$. It follows that

$$q_y=E_{0y}\left(\prod_{v\in Z(s)}q_{Y_v}\right)\leq E_{0y}\left((\max_{j\in S}q_j)^{\sharp Z(s)}\right).$$

The Kesten-Stigum theorem for MMBP (see, for example, [1]) confirms that the total population size $\sharp Z(s)$ converges to infinity almost surely on non-extinction set, thus we have $q_y = 0$ by dominated convergence theorem. $\square$

Define $L(t) := \inf\{X_u(t)\colon u \in Z(t)\}$, i.e., $L(t)$ denotes the position of the leftmost particle at time t. Then we have

Theorem 4. *For any $(x, y) \in \mathbb{R} \times S$, $P_{xy}(\lim_{t\to+\infty} L(t) + \underline{c}t = +\infty) = 1$. In particular, if $E(\xi_{ij} \log^+ \xi_{ij}) < +\infty$ for all $i, j \in S$, then $P_{xy}(\lim_{t\to+\infty} L(t)/t = -\underline{c}) = 1$.*

Proof. It is sufficient to prove the conclusion under measure P_{0y}. Note that

$$W_\lambda(t) \geq C_1 \mathrm{e}^{-\lambda(L(t)+c_\lambda t)} = C_1 \mathrm{e}^{-\lambda t\left(\frac{L(t)}{t}+c_\lambda\right)} \tag{3.5}$$

for some constant $C_1 > 0$. Since $\lim_{t\to+\infty} W_{\underline{\lambda}}(t) = 0$, in view of (3.5) we have $P_{0y}(\lim_{t\to+\infty} L(t) + \underline{c}t = +\infty) = 1$ and $P_{0y}(\liminf_{t\to+\infty} L(t)/t \geq -\underline{c}) = 1$. Recall that the spine moves as a Brownian motion with drift $-\lambda$ under the measure $\widetilde{Q}^\lambda_{0y}$, so we have

$$\widetilde{Q}^\lambda_{0y}\left(\lim_{t\to+\infty} \frac{\widetilde{X}(t)}{t} = -\lambda\right) = 1.$$

The proof of Theorem 3 shows that if $E(\xi_{ij} \log^+ \xi_{ij}) < +\infty$ for all $i, j \in S$, then for any $\lambda \in (0, \underline{\lambda})$

$$\frac{\mathrm{d}Q^\lambda_{0y}}{\mathrm{d}P_{0y}} = \frac{W(\lambda)}{h_y}$$

and $P_{0y}(W(\lambda) > 0) = 1$. This implies that $Q^\lambda_{0y}(W(\lambda) > 0) = 1$ and P_{0y} is absolutely continuous with respect to Q^λ_{0y}. Hence for any $0 < \lambda < \underline{\lambda}$

$$P_{0y}\left(\limsup_{t\to+\infty} \frac{L(t)}{t} \leq -\lambda\right) \geq P_{0y}\left(\lim_{t\to+\infty} \frac{\widetilde{X}(t)}{t} = -\lambda\right) = 1$$

Thus $P_{0y}\left(\limsup_{t\to+\infty} L(t)/t \leq -\underline{\lambda} = -\underline{c}\right) = 1$. We complete the proof. $\square$

Proof of Theorem 1(1) By Theorem 3, $w(x, y)$ is non-trivial and increasing in x. It is clear that $\lim_{x\to+\infty} w(x, y) = 1$. $P_{0y}(W(\lambda) = 0) = 0$ implies that $\lim_{x\to-\infty} w(x, y) = 0$. Besides,

$$\begin{aligned}
w(x,y) &= E_{xy}\left[\exp\{-\sum_{v\in Z(s)} \lim_{t\to+\infty} \sum_{\substack{u\in Z(t)\\ v\prec u}} h_{Y_u} \mathrm{e}^{-\lambda(X_u(t)+ct)}\}\right] \\
&= E_{xy}\left[\prod_{v\in Z(s)} E_{X_v(s)Y_v}\left(\exp\{-\mathrm{e}^{-\lambda cs} \lim_{t\to+\infty} \sum_{u\in Z(t-s)} h_{Y_u} \mathrm{e}^{-\lambda(X_u(t-s)+c(t-s))}\}\right)\right] \\
&= E_{xy}\left[\prod_{v\in Z(s)} w(X_v(s)+cs, Y_v)\right].
\end{aligned}$$

Thus it follows from Lemma 6 that $u(t, x, y) := w(x - ct, y)$ is a traveling wave solution to equation (1.1) with wave speed c.

Since $\lim_{x\to+\infty} w(x, y) = 1$ and $E_{0y}W(\lambda) = E_{0y}W_\lambda(0) = h_y$, then

$$\frac{1 - w(x, y)}{b_y \mathrm{e}^{-\lambda x}} = \frac{1 - E_{0y}\left[\exp\{-\mathrm{e}^{-\lambda x}W(\lambda)\}\right]}{E_{0y}\left[\mathrm{e}^{-\lambda x}W(\lambda)\right]} \to 1 \text{ as } x \to +\infty,$$

i.e. $1 - w(x, y) \sim h_y \mathrm{e}^{-\lambda x}$ as $x \to +\infty$.

The rest of proof is dedicated to the uniqueness. We consider the space-time barrier $\Gamma^{(x,c_\lambda)} := \{(y, t) \in \mathbb{R} \times \mathbb{R}^+ : y + c_\lambda t = x\}$ for $x \geq 0$. By arresting lines of descendants the first time they hit this barrier, we produce a random collection of particles $c(x, c_\lambda) = \bigcup_{i\in S} c_i(x, c_\lambda)$ where $c_i(x, c_\lambda)$ denotes the subset of type i particles. $\{c(x, c_\lambda) : x \geq 0\}$ is a family of stopping lines. We say $\{c(x, c_\lambda) : x \geq 0\}$ is *dissecting* in the sense that all lines of descendants will hit $\Gamma^{(x,c_\lambda)}$ with probability one for all $x > 0$. Obviously this is because $\lim_{t\to+\infty} L(t) + ct = +\infty$ for $c \geq \underline{c}$. We also observe that $\{c(x, c_\lambda) : x \geq 0\}$ is *tending to infinity* in the sense that for each $n \in \mathbb{N}$, one can choose x sufficiently large such that particles in $c(x, c_\lambda)$ are descendants of the nth generation. (For more information on general stopping lines and properties of them, we refer to [4] and [9].) Let $\mathcal{F}_{c(x,c_\lambda)}$ be the natural filtration generated by ancestral, type and spatial paths receding from particles at the moment they hit $\Gamma^{(x,c_\lambda)}$.

We use $\sharp A$ to denote the cardinal of a finite set A. Let Φ_{c_λ} be an abitrary traveling wave at speed c_λ.

$$\begin{aligned} M_x(z, c_\lambda) &:= \prod_{u\in c(x,c_\lambda)} \Phi_{c_\lambda}(z + X_u(t) + c_\lambda t, Y_u) \\ &= \exp\left\{\sum_{i\in S} \sharp c_i(x, c_\lambda) \log \Phi_{c_\lambda}(z + x, i)\right\} \end{aligned}$$

is a P_{0y}-martingale with respect to $\{\mathcal{F}_{c(x,c_\lambda)} : x \geq 0\}$. It converges to $\Phi_{c_\lambda}(z, \lambda)$ almost surely and in $L^1(P_{0y})$ (by boundedness). It follows that

$$\lim_{x\to+\infty} -\sum_{i\in S} \sharp c_i(x, c_\lambda) \log \Phi_{c_\lambda}(z + x, i) \tag{3.6}$$

exists and is positive with positive probability.

Obviously, for any $x_2 > x_1 \geq 0$ and any $v \in c(x_2, c_\lambda)$, there exists an unique $u \in c(x_1, c_\lambda)$ such that $u \prec v$. In fact, $\{(\sharp c_1(x, c_\lambda), \cdots, \sharp c_d(x, c_\lambda)) : x \geq 0\}$ forms a continuous time multitype Markov branching process (x plays the role of time). This follows from the strong Markov branching property (see, for example, [9]). The strong Markov branching property says that if $\{\sigma_u : u \in c(x, c_\lambda)\}$ are the times when particles in $c(x, c_\lambda)$ hit the barrier $\Gamma^{(x,c_\lambda)}$, then given $\mathcal{F}_{c(x,c_\lambda)}$ each of the trees relative to and rooted at the space time points

$\{(X_u(\sigma_u), \sigma_u) : u \in c(x, c_\lambda)\}$ are independent copies of multitype branching Brownian motions started by a type Y_u particle at position $X_u(\sigma_u)$. Moreover, it follows from the fact $P_{0y}(\lim_{t\to+\infty} \widetilde{X}(t) + c_\lambda t = +\infty)$ and the irreducibility of $\widetilde{Y}$ that $\{(\sharp c_1(x, c_\lambda), \cdots, \sharp c_d(x, c_\lambda)) : x \geq 0\}$ is non-extinct and positive regular. Let $M_{c_\lambda}(x) = (m_{ij}^{c_\lambda}(x))_{i,j\in S}$ where $m_{ij}^c(x) = E_{0i}\sharp c_j(x, c_\lambda)$. Let A_{c_λ} be the matrix such that $M_{c_\lambda}(x) = e^{A_{c_\lambda}x}$. By Perron-Frobenius theorem, we can find a simple positive eigenvalue $\lambda^*_{c_\lambda}$ of A_{c_λ}, and corresponding positive left and right eigenvectors $\pi_{c_\lambda} = (\pi^1_{c_\lambda}, \cdots, \pi^d_{c_\lambda})$ and $h_{c_\lambda} = (h^1_{c_\lambda}, \cdots, h^d_{c_\lambda})$ such that $\langle \pi_{c_\lambda}, h_{c_\lambda}\rangle = \langle \pi_{c_\lambda}, 1\rangle = 1$. Immediately

$$\sum_{j\in S} m_{ij}^{c_\lambda}(x) h^j_{c_\lambda}\, e^{-\lambda^*_{c_\lambda}x} = h^i_{c_\lambda} \quad \forall i \in S, \tag{3.7}$$

Define for $x \geq 0$

$$\begin{aligned} W_{c(x,c_\lambda)}(\lambda) &= \sum_{u\in c(x,c_\lambda)} h_{Y_u} e^{-\lambda(X_u(t)+c_\lambda t)} \\ &= \sum_{i\in S} \sharp c_i(x, c_\lambda) h_i e^{-\lambda x}. \end{aligned}$$

Then $\{W_{c(x,c_\lambda)}(\lambda) : x \geq 0\}$ is a P_{0y}-martingale with respect to $\{\mathcal{F}_{c(x,c_\lambda)} : x \geq 0\}$, and consequently

$$\sum_{j\in S} m_{ij}^{c_\lambda}(x) h_j\, e^{-\lambda x} = h_i \quad \forall i \in S, \tag{3.8}$$

in other words, $e^{\lambda x}$ is an eigenvalue of $M_{c_\lambda}(x)$ with corresponding right eigenvector h. Using similar arguments as in [12] Theorem 8, we can show that

$$\lim_{x\to+\infty} \sum_{i\in S} \sharp c_i(x, c_\lambda) h_i e^{-\lambda x} = W(\lambda), \quad P_{0y}\text{-a.s. and } L^1(P_{0y}). \tag{3.9}$$

On the other hand, by Kensten-Stigum theorem (see, for example, [6] Theorem 2.1) we have for any $i \in S$,

$$\lim_{x\to+\infty} \sharp c_i(x, c_\lambda)\, e^{-\lambda^*_{c_\lambda}x} = \pi^i_{c_\lambda} W_{c_\lambda} \quad P_{0y}\text{-a.s.}, \tag{3.10}$$

where $W_{c_\lambda} = \lim_{x\to+\infty} \sum_{i\in S} \sharp c_i(x, c_\lambda) \pi^i_{c_\lambda}\, e^{-\lambda^*_{c_\lambda}x} < +\infty$. Combining (3.9) and (3.10), we conclude that $\lambda^*_{c_\lambda} = \lambda$ and $P_{0y}(W_{c_\lambda} = \alpha W(\lambda)) = 1$ for some constant $\alpha > 0$. Using (3.7) and (3.8), we get $h_{c_\lambda} = \alpha h$. Thus by (3.10) we have for any $i \in S$,

$$\lim_{x\to+\infty} \sharp c_i(x, c_\lambda) e^{-\lambda x} = \alpha \pi^i_{c_\lambda} W(\lambda) \quad P_{0y}\text{-a.s.}. \tag{3.11}$$

It follows from (3.6) and (3.11) that $\lim_{x\to+\infty}(-\alpha)\sum_{i\in S} \pi^i_{c_\lambda} e^{\lambda x} \log \Phi_{c_\lambda}(x, i)$ exists and is positive. We denote this limit by β. Uniqueness (up to a multiplicative constant) is now immediate since

$$
\begin{aligned}
\Phi_{c_\lambda}(z,y) &= E_{0y}\left(\lim_{x\to+\infty} M_x(z,c_\lambda)\right)\\
&= E_{0y}\left(\exp\left\{\lim_{x\to+\infty}\sum_{i\in S}\sharp c_i(x,c_\lambda)\log\Phi_{c_\lambda}(z+x,i)\right\}\right)\\
&= E_{0y}\left(\exp\left\{\lim_{x\to+\infty}\alpha\sum_{i\in S}\pi^i_{c_\lambda}e^{\lambda x}W(\lambda)\log\Phi_{c_\lambda}(z+x,i)\right\}\right)\\
&= E_{0y}\left(\exp\left\{-W(\lambda)e^{-\lambda z}\lim_{x\to+\infty}(-\alpha)\sum_{i\in S}\pi^i_{c_\lambda}e^{\lambda(x+z)}\log\Phi_{c_\lambda}(z+x,i)\right\}\right)\\
&= E_{0y}\left(\exp\left\{-\beta\, W(\lambda)e^{-\lambda z}\right\}\right).
\end{aligned}
$$

□

Proof of Theorem 1(2): We assume $w(x,y)$ provides a monotone traveling wave solution to (1.1) with speed $c<\underline{c}$. Then by Lemma 6, $\prod_{u\in Z(t)} w(X_u(t)+x+ct,Y_u)$ is a bounded martingale under P_{0y}. It converges almost surely and in mean to some random variable. On the other hand, since $0\le w(x,y)\le 1$ and $L(t)+ct\to-\infty$ as $t\to+\infty$,

$$\prod_{u\in Z(t)} w(X_u(t)+x+ct,Y_u)\le w(L(t)+ct,Y_L(t))\to 0$$

where $Y_L(t)$ denotes the type of leftmost particle at time t. Thus $w(x,y)\equiv 0$ which contradicts the assumption. □

4 Proof of Theorem 2

Note that $M_t(\lambda)$ defined as in (1.5) is a signed martingale and therefore it does not necessarily converge almost surely. A technique used by Kyprianou [12] to get round this problem in the case of a single-type branching Brownian motion is to consider a truncated form of the devivative martingale which is a positive martingale. In order to describe the aforementioned martingale for multitype branching Brownian motion we need more notations and lemmas.

Lemma 9 ([12] Section 5). *Suppose $B=\{B_t\colon t\ge 0\}$ is a standard Brownian motion according to the law $\mathbb{P}$, $\{\mathcal{L}_t\}$ is its natural filtration. $\forall z>0$, define $\tau_\lambda:=\inf\{t>0\colon z+B_t+\lambda t\le 0\}$, then*

$$m_\lambda(t):=\frac{z+B_t+\lambda t}{z}\mathrm{e}^{-\lambda(B_t+\frac{\lambda t}{2})}1_{\{t<\tau_\lambda\}}$$

is a martingale. Define another probability measure $\hat{\mathbb{P}}^\lambda_z$ by

$$\left.\frac{\mathrm{d}\hat{\mathbb{P}}^\lambda_z}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{L}_t}=m_\lambda(t).$$

Then under measure $\hat{\mathbb{P}}^\lambda_z$, $\{z+B_t+\lambda t\colon t\geq 0\}$ *is a standard Bessel-3 process starting from* z.

Define the space-time barrier $\Gamma^{(-z,\lambda)} := \{(y,t)\in\mathbb{R}\times\mathbb{R}^+\colon y+\lambda t=-z\}$ for $z\geq 0$. $\widetilde{Z}(t)$ denotes the subset of $Z(t)$ consisting of all particles alive at t having ancestry (including themselves) whose spatial paths have not met $\Gamma^{(-z,\lambda)}$ by time t.

From the many-to-one formula, we see that

$$V_\lambda(t) := \sum_{u\in\widetilde{Z}(t)} h_{Y_u}\cdot(z+X_u(t)+\lambda t)\mathrm{e}^{-\lambda(X_u(t)+c_\lambda t)}$$

is a non-negative martingale. We want to define a new probability measure $\widetilde{R}^\lambda_{0y}$ such that if $R^\lambda_{0y} := \widetilde{R}^\lambda_{0y}|_{\mathcal{F}}$ then

$$\left.\frac{\mathrm{d}R^\lambda_{0y}}{\mathrm{d}P_{0y}}\right|_{\mathcal{F}_t} = \frac{V_\lambda(t)}{V_\lambda(0)}, \quad \forall t>0.$$

To this end, $\tilde{R}^\lambda_{0y}$ should have the following decomposition:

$$\begin{aligned}\mathrm{d}\widetilde{R}^\lambda_{0y}(\tau,M,\varepsilon)|_{\widetilde{\mathcal{F}}_t} \;&=\; \mathrm{d}\hat{\mathbb{P}}_y(\widetilde{Y})\mathrm{d}\hat{\mathbb{P}}^\lambda_z(\widetilde{X})\prod_{v\prec\varepsilon_{n_t}}\frac{\hat{p}_{A_v}(Y_v)}{\widehat{P}(Y_v,Y_{v+1})}\frac{A_v(Y_{v+1})h_{Y_{v+1}}}{\langle A_v,h\rangle}\\ &\prod_{v\prec\varepsilon_{n_t}}\left[\frac{1}{A_v(Y_{v+1})}\prod_{j\colon vj\in O_v}\mathrm{d}P^{t-\zeta_v}_{X_{\zeta_v}\widetilde{Y}_{\zeta_v}}((\tau,M)^v_j)\right].\end{aligned}\qquad(4.1)$$

Remark 2. *Under* $\widetilde{R}^\lambda_{0y}$, *the spine's spatial process* $\widetilde{X}$ *satisfies that* $\{z+\widetilde{X}(t)+\lambda t\colon t\geq 0\}$ *is a Bessel-3 process which is identically distributed to the modulus process of a three dimensional Brownian motion. Therefore it never meets the barrier* $\Gamma^{(-z,\lambda)}$.

Put

$$M_\lambda(t) := \sum_{u\in Z(t)} h_{Y_u}\cdot(z+X_u(t)+\lambda t)\mathrm{e}^{-\lambda(X_u(t)+c_\lambda t)}.$$

If we can prove that $M_\lambda(t)$ converges to a non-degenerate limit, similar analysis as in the supercritical case can be carried out to obtain the traveling wave solution to (1.1). For this purpose, we need the following lemma.

Lemma 10. *Let* $V(\lambda)=\lim_{t\to+\infty}V_\lambda(t)$. *For any* $\lambda\geq\underline{\lambda}$, $\lim_{t\to+\infty}M_\lambda(t)$ *exists and is equal to* $V(\lambda)$ *almost surely under* P_{0y}. *In addition,* $M(\lambda)$ *does not depend on* z.

Proof. Recall that $V_\lambda(t)$ is a non-negative martingale, its limit exists almost surely. Let $\gamma^{(-z,\lambda)}$ denote the event that the multitype branching Brownian motion remains entirely to the right of $\Gamma^{(-z,\lambda)}$, then

$$\lim_{t\to+\infty}M_\lambda(t)=\lim_{t\to+\infty}V_\lambda(t)\ \text{ on }\gamma^{(-z,\lambda)}\quad P_{0y}\text{-a.s.}$$

Since $P_{0y}(\lim_{t\to+\infty} L(t)+\underline{c}t=+\infty)=1$, we have $P_{0y}(\inf_{t\ge0}\{L(t)+\lambda t\}>-\infty)=1$ for all $\lambda\ge\underline{\lambda}$. Thus

$$P_{0y}(\gamma^{(-z,\lambda)})=P_{0y}\left(\inf_{t\ge0}\{L(t)+\lambda t\}>-z\right)\uparrow 1 \quad \text{as} \quad z\uparrow+\infty.$$

Therefore, we have

$$\lim_{t\to+\infty} M_\lambda(t)=\lim_{t\to+\infty} V_\lambda(t) \quad P_{0y}\text{-a.s.},$$

that is to say

$$M(\lambda):=\lim_{t\to+\infty} M_\lambda(t)=V(\lambda):=\lim_{t\to+\infty} V_\lambda(t), \quad P_{0y}\text{-a.s.},$$

for any $\lambda\ge\underline{\lambda}$. Note that

$$M_\lambda(t)=\sum_{u\in Z(t)} h_{Y_u}\cdot(X_u(t)+\lambda t)\mathrm{e}^{-\lambda(X_u(t)+c_\lambda t)}+zW_\lambda(t).$$

By Theorem 3, the second term of the right hand side converges to 0 for $\lambda\ge\underline{\lambda}$, hence the limit $M(\lambda)$ does not depend on z. □

Next,we focus on the limit theorem for the martingale $V_{\underline{\lambda}}(t)$. Hereafter, we simply write $\widetilde{R}^{\underline{\lambda}}_{0y}$ as $\widetilde{R}_{0y}$.

Theorem 5. *Suppose $\lambda=\underline{\lambda}$.*

1. *If $E\xi_{ij}(\log^+\xi_{ij})^2=+\infty$ for some $i,j\in S$, then $V(\underline{\lambda})=0$ P_{xy}-a.s.*
2. *If $E\xi_{ij}(\log^+\xi_{ij})^2<+\infty$ for all $i,j\in S$, then $V_{\underline{\lambda}}(t)$ converges to $V(\underline{\lambda})$ in $L^1(P_{xy})$ and $P_{xy}(V(\underline{\lambda})=0)=0$.*

Lemma 11. *we have the following spine decomposition for $V_{\underline{\lambda}}(t)$*

$$\begin{aligned}\widetilde{R}_{0y}\left(V_{\underline{\lambda}}(t)\,|\,\widetilde{\mathcal{G}}\right) &= h_{\widetilde{Y}_t}(z+\widetilde{X}(t)+\underline{\lambda})\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(t)+\underline{c}t)} \\ &\quad +\sum_{j\in S}\sum_{v\prec\varepsilon_{n_t}}(A_v(j)-\delta_{Y_{v+1}j})h_j(z+\widetilde{X}(\zeta_v)+\underline{\lambda}t)\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(\zeta_v)+\underline{c}\zeta_v)}.\end{aligned}$$

Lemma 12. *(1). If $E\xi_{ij}(\log^+\xi_{ij})^2=+\infty$ for some $i,j\in S$, then*

$$\limsup_{n\to+\infty}\langle A_{\varepsilon_n},h\rangle(z+\widetilde{X}(\zeta_{\varepsilon_n})+\underline{\lambda}\zeta_{\varepsilon_n})\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(\zeta_{\varepsilon_n})+\underline{c}\zeta_{\varepsilon_n})}=+\infty \quad \widetilde{R}_{0y}\text{-a.s.}$$

(2). If $E\xi_{ij}(\log^+\xi_{ij})^2<+\infty$ for all $i,j\in S$, then

$$\sum_{n=0}^{+\infty}\langle A_{\varepsilon_n},h\rangle(z+\widetilde{X}(\zeta_{\varepsilon_n})+\underline{\lambda}\zeta_{\varepsilon_n})\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(\zeta_{\varepsilon_n})+\underline{c}\zeta_{\varepsilon_n})}<+\infty \quad \widetilde{R}_{0y}\text{-a.s.}$$

Proof. (1) We want to show that for any $M \in (0, +\infty)$,

$$\sum_{n=0}^{+\infty} I_{\{\langle A_{\varepsilon_n}, h\rangle (z+\widetilde{X}(\zeta_{\varepsilon_n})+\underline{\lambda}\zeta_{\varepsilon_n})\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(\zeta_{\varepsilon_n})+\underline{c}\zeta_{\varepsilon_n})}\geq M\}} = +\infty \quad \widetilde{R}_{0y}\text{-a.s.} \tag{4.2}$$

For any set $B \in \mathcal{B}[0,\infty) \times \mathcal{B}(\mathbb{Z}_+^d)$, define

$$\phi(B) = \#\{n \geq 0 : (\zeta_{\varepsilon_n}, A_{\varepsilon_n}) \in B\}.$$

Then conditional on $\mathcal{G}^{\widetilde{Y}}$, ϕ is a Poisson random measure on $[0,\infty) \times \mathbb{Z}_+^d$ with intensity

$$(a_{\widetilde{Y}_t} + \lambda^*)dt \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(\widetilde{Y}_t)\delta_k(dy)$$

(here δ denotes the delta function.) Thus for any $T \in (0,\infty)$, given $\mathcal{G}$,

$$\#\{n \geq 0 : \zeta_{\varepsilon_n} \leq T,\ \langle A_{\varepsilon_n}, h\rangle (z + \widetilde{X}(\zeta_{\varepsilon_n}) + \underline{\lambda}\zeta_{\varepsilon_n})e^{-\underline{\lambda}(\widetilde{X}(\zeta_{\varepsilon_n})+\underline{c}\zeta_{\varepsilon_n})} \geq M\}$$

is Poisson random variable with parameter

$$\int_0^T (a_{\widetilde{Y}_t} + \lambda^*) \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(\widetilde{Y}_t) I_{\{\langle k,h\rangle (z+\widetilde{X}(t)+\underline{\lambda}t)\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(t)+\underline{c}t)}\geq M\}} \mathrm{d}t.$$

Hence to prove (4.2), we only need to show that

$$\int_0^{+\infty} (a_{\widetilde{Y}_t} + \lambda^*) \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(\widetilde{Y}_t) I_{\{\langle k,h\rangle (z+\widetilde{X}(t)+\underline{\lambda}t)\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(t)+\underline{c}t)}\geq M\}} \mathrm{d}t = +\infty \quad \widetilde{R}_{0y}\text{-a.s.}$$

Since $\min\{a_l : l \in S\} > 0$, it is sufficient to prove that

$$\widetilde{R}_{0y}\left(\int_0^{+\infty} \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(\widetilde{Y}_t) I_{\{\langle k,h\rangle (z+\widetilde{X}(t)+\underline{\lambda}t)\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(t)+\underline{c}t)}\geq M\}} \mathrm{d}t < +\infty \right) = 0.$$

For any constant $c \in (0, +\infty)$, put

$$E_c := \left\{ \int_0^{+\infty} \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(\widetilde{Y}_t) I_{\{\langle k,h\rangle (z+\widetilde{X}(t)+\underline{\lambda}t)\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(t)+\underline{c}t)}\geq M\}} \mathrm{d}t < c \right\}.$$

It is sufficient to show that $\widetilde{R}_{0y}(E_c) = 0$. In fact we have

$$\begin{aligned}
c &\geq \widetilde{R}_{0y}\left(I_{E_c} \int_0^{+\infty} \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(\widetilde{Y}_t) I_{\{\langle k,h\rangle (z+\widetilde{X}(t)+\underline{\lambda}t)\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(t)+\underline{c}t)}\geq M\}} \mathrm{d}t \right) \\
&= \int_0^{+\infty} \sum_{l \in S} \hat{\mathbb{P}}_y(\widetilde{Y}_t = l) \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(l) \widetilde{R}_{0y}\left(I_{E_c} I_{\{\mathrm{Bes}(t)\mathrm{e}^{-\underline{\lambda}\mathrm{Bes}(t)} \geq M\langle k,h\rangle^{-1}\mathrm{e}^{-\underline{\lambda}z}\}} \right) \mathrm{d}t \\
&\geq \int_0^{+\infty} \hat{\mathbb{P}}_y(\widetilde{Y}_t = i) \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(i) \widetilde{R}_{0y}\left(I_{E_c} I_{\{\mathrm{Bes}(t)\mathrm{e}^{-\underline{\lambda}\mathrm{Bes}(t)} \geq M\langle k,h\rangle^{-1}\mathrm{e}^{-\underline{\lambda}z}\}} \right) \mathrm{d}t,
\end{aligned} \tag{4.3}$$

where $\mathrm{Bes}(t) := z + \widetilde{X}(t) + \underline{\lambda} t$. It is known that under $\hat{\mathbb{P}}_y$, $\widetilde{Y}$ moves as a Q-process with the invariant distribution $\tilde{\pi}_l = h_l \pi_l,\ l \in S$. Consequently there exists some $T > 0$, such that for any $t \geq T$, $\hat{\mathbb{P}}_y(\widetilde{Y}_t = i) \geq \frac{1}{2}\tilde{\pi}_i > 0$. We continue the above domination:

$$\begin{aligned}
&\geq \frac{1}{2}\tilde{\pi}_i \sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(i) \int_T^{+\infty} \widetilde{R}_{0y}\left(I_{E_c} I_{\{\mathrm{Bes}(t)\mathrm{e}^{-\underline{\lambda}\mathrm{Bes}(t)} \geq M\langle k,h\rangle^{-1}\mathrm{e}^{-\underline{\lambda} z}\}}\right) \mathrm{d}t \\
&\geq \frac{1}{2}\tilde{\pi}_i \left(\sum_{k \in \mathbb{Z}_+^d} \hat{p}_k(i) \int_0^{+\infty} \widetilde{R}_{0y}\left(I_{E_c} I_{\{\mathrm{Bes}(t)\mathrm{e}^{-\underline{\lambda}\mathrm{Bes}(t)} \geq M\langle k,h\rangle^{-1}\mathrm{e}^{-\underline{\lambda} z}\}}\right) \mathrm{d}t - T\right).
\end{aligned} \tag{4.4}$$

We consider a process $((Q_t, W_t), \mathbb{P})$ such that $\{Q_t, t \geq 0\}$ and $\{W_t, t \geq 0\}$ are independent, $(Q_t, \mathbb{P})$ is identically distributed as $(\widetilde{Y}_t, \widetilde{R}_{0y})$, and $(W_t, \mathbb{P})$ is a standard Brownian motion on $\mathbb{R}^3$. Suppose $\hat{z}$ is a point in $\mathbb{R}^3$ with norm z. It is known that $(\mathrm{Bes}(t), \widetilde{R}_{0y})$ is a Bessel-3 process starting from z, which is identically distributed to $(|W_t + \hat{z}|, \mathbb{P})$, here $|\cdot|$ denotes the Euclidean norm.. $((Q_t, W_t), \mathbb{P})$ might be defined on another probability space. In the remaining proof we still use E_c to denote the counterpart set of E_c with respect to $((Q_t, W_t), \mathbb{P})$. Immediately we have

$$\widetilde{R}_{0y}\left(I_{E_c} I_{\{\mathrm{Bes}(t)\mathrm{e}^{-\underline{\lambda}\mathrm{Bes}(t)} \geq M\langle k,h\rangle^{-1}\mathrm{e}^{-\underline{\lambda} z}\}}\right) = \mathbb{P}\left(I_{E_c} I_{\{|W_t+\hat{z}|\mathrm{e}^{-\underline{\lambda}|W_t+\hat{z}|} \geq M\langle k,h\rangle^{-1}\mathrm{e}^{-\underline{\lambda} z}\}}\right)$$

and

$$\widetilde{R}_{0y}(E_c) = \mathbb{P}(E_c). \tag{4.5}$$

We claim that there exists $K^* > 0$ such that when $|k| \geq K^*$

$$\left\{y \in \mathbb{R}^3 : 1 + z \leq |y| \leq \frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}\right\} \subset \left\{y \in \mathbb{R}^3 : |y + \hat{z}|\mathrm{e}^{-\underline{\lambda}|y+\hat{z}|} \geq M\langle k,h\rangle^{-1}\mathrm{e}^{-\underline{\lambda} z}\right\}. \tag{4.6}$$

In fact, $1 + z \leq |y| \leq \frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}$ implies that $1 \leq |y + \hat{z}| \leq \frac{\log^+\langle k,h\rangle}{2\underline{\lambda}} + z$. Consider the function $f(x) = x\mathrm{e}^{-\underline{\lambda} x}$. On the positive half line, it increases to a supremum and then decreases to 0 as x goes to infinity. Thus we can find $K^* > 0$ large enough such that when $|k| \geq K^*$,

$$\begin{aligned}
1 + z \leq |y| \leq \frac{\log^+\langle k,h\rangle}{2\underline{\lambda}} \quad &\Rightarrow \quad f(|y + \hat{z}|) \geq f\left(\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}} + z\right) \\
&\Rightarrow \quad |y + \hat{z}|\mathrm{e}^{-\underline{\lambda}|y+\hat{z}|} \geq \left(\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}} + z\right)\langle k,h\rangle^{-1/2}\mathrm{e}^{-\underline{\lambda} z}.
\end{aligned}$$

Then we get (4.6). Now we continue the estimation of (4.4):

$$
\begin{aligned}
&= \frac{1}{2}\tilde{\pi}_i \left(\sum_{k\in\mathbb{Z}_+^d} \hat{p}_k(i) \int_0^{+\infty} \mathbb{P}\left(I_{E_c} I_{\{|W_t+\hat{z}|\mathrm{e}^{-\underline{\lambda}|W_t+\hat{z}|}\geq M\langle k,h\rangle^{-1}\mathrm{e}^{-\underline{\lambda}z}\}}\right) \mathrm{d}t - T \right) \\
&\geq \frac{1}{2}\tilde{\pi}_i \left(\sum_{k:\,|k|\geq K^*} \hat{p}_k(i) \int_0^{+\infty} \mathbb{P}\left(I_{E_c} I_{\{1+z\leq|W_t|\leq \frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}\}} \right) \mathrm{d}t - T \right) \\
&= \frac{1}{2}\tilde{\pi}_i \left(\sum_{k:\,|k|\geq K^*} \hat{p}_k(i) \mathbb{P}\left(I_{E_c} \int_0^{+\infty} I_{\left\{1+z\leq|W_t|\leq \frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}\right\}} \mathrm{d}t \right) - T \right).
\end{aligned}
\tag{4.7}
$$

Note that $(|W_t|,\mathbb{P})$ is a Bessel-3 process starting from 0. Let l^a, $a\geq 0$, be the family of its local times, then the process $\{l^a_\infty,\ a\geq 0\}$ is a BESQ$^2(0)$ process which implies that $l^a_\infty \overset{d}{=} a l^1_\infty$ and $\mathbb{P}(l^1_\infty = 0) = 0$ (see [17], P425, Exercise 2.5). Then we have the following calculations:

$$
\begin{aligned}
& \mathbb{P}\left(I_{E_c} \int_0^{+\infty} I_{\{1+z\leq|W_t|\leq \frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}\}} \mathrm{d}t \right) \\
= {} & \mathbb{P}\left(I_{E_c} \int_{1+z}^{\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}} l^a_\infty \mathrm{d}a \right) \\
= {} & \mathbb{P}\left(I_{E_c} \int_{1+z}^{\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}} a\,\mathrm{d}a \int_0^{a^{-1}l^a_\infty} \mathrm{d}u \right) \\
= {} & \int_{1+z}^{\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}} a\mathrm{d}a \int_0^{+\infty} \mathbb{P}(I_{E_c} I_{\{u\leq a^{-1}l^a_\infty\}})\mathrm{d}u \\
\geq {} & \int_{1+z}^{\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}} a\mathrm{d}a \int_0^{+\infty} \left(\mathbb{P}(E_c) - \mathbb{P}(a^{-1}l^a_\infty < u)\right)^+ \mathrm{d}u \\
= {} & \int_{1+z}^{\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}}} a\mathrm{d}a \int_0^{+\infty} \left(\mathbb{P}(E_c) - \mathbb{P}(l^1_\infty < u)\right)^+ \mathrm{d}u \\
= {} & \frac{1}{2}\left(\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}} - 1 - z \right)^2 \int_0^{+\infty} \left(\mathbb{P}(E_c) - \mathbb{P}(l^1_\infty < u)\right)^+ \mathrm{d}u.
\end{aligned}
\tag{4.8}
$$

In view of (4.3), (4.4), (4.7) and (4.8), we get

$$
\sum_{k:\,|k|\geq K^*} \hat{p}_k(i) \left(\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}} - 1 - z \right)^2 \int_0^{+\infty} \left(\mathbb{P}(E_c) - \mathbb{P}(l^1_\infty < u)\right)^+ \mathrm{d}u < +\infty.
\tag{4.9}
$$

Given that $E(\xi_{ij}(\log^+\xi_{ij})^2) = +\infty$, we have $\sum_{k\in\mathbb{Z}_+^d} \hat{p}_k(i)(\log^+\langle k,h\rangle)^2$ $= +\infty$, and then $\sum_{k:\,|k|>K^*} \hat{p}_k(i) \left(\frac{\log^+\langle k,h\rangle}{2\underline{\lambda}} - 1 - z \right)^2 = +\infty$. It follows from (4.9) that

$$
\int_0^{+\infty} \left(\mathbb{P}(E_c) - \mathbb{P}(l^1_\infty < u)\right)^+ \mathrm{d}u = 0.
$$

Thus by the fact that $\mathbb{P}(l^1_\infty = 0) = 0$ we have $\mathbb{P}(E_c) = 0$ for arbitrary $c > 0$, and then, by (4.5), $\widetilde{R}_{0y}(E_c) = 0$ for arbitrary $c > 0$. We complete the proof of part (1).

(2) Choose $\lambda \in (0, \underline{\lambda})$. We have

$$\begin{aligned} & \sum_{n=0}^{+\infty}(z + \widetilde{X}(\zeta_{\varepsilon_n}) + \underline{\lambda}\zeta_{\varepsilon_n})\langle A_{\varepsilon_n}, h\rangle \mathrm{e}^{-\underline{\lambda}(\widetilde{X}(\zeta_{\varepsilon_n})+\underline{c}\zeta_{\varepsilon_n})} \\ = & \sum_{n=0}^{+\infty}(\cdots)1_{\{\langle A_{\varepsilon_n},h\rangle \le \mathrm{e}^{\lambda(\widetilde{X}(\zeta_{\varepsilon_n})+\underline{c}\zeta_{\varepsilon_n})}\}} + \sum_{n=0}^{+\infty}(\cdots)1_{\{\langle A_{\varepsilon_n},h\rangle > \mathrm{e}^{\lambda(\widetilde{X}(\zeta_{\varepsilon_n})+\underline{c}\zeta_{\varepsilon_n})}\}} \\ \triangleq & \ \Theta + \Lambda. \end{aligned}$$

We only need to prove that both Θ and Λ are finite almost surely under $\widetilde{R}_{0y}$. Hereafter, we write "$A \lesssim B$" to mean that there exists some constant $c > 0$ such that $A \le cB$.

Recall that conditioned on $\mathcal{G}^{\widetilde{Y}}$, the split times of the spine is a Poisson point process with characteristic measure $(a_{\widetilde{Y}_t} + \lambda^*)\mathrm{d}t$. Therefore

$$\begin{aligned} & \widetilde{R}_{0y}(\Theta) \\ = & \ \widetilde{R}_{0y}\left(\int_0^{+\infty}(a_{\widetilde{Y}_s} + \lambda^*)(z + \widetilde{X}(s) + \underline{\lambda}s)\langle A_{\varepsilon_{n_s}}, h\rangle \mathrm{e}^{-\underline{\lambda}(\widetilde{X}(s)+\underline{c}s)} 1_{\{\langle A_{\varepsilon_{n_s}},h\rangle \le \mathrm{e}^{\lambda(\widetilde{X}(s)+\underline{c}s)}\}}\mathrm{d}s\right) \\ \le & \int_0^{+\infty}\sum_{i\in S}(a_i + \lambda^*)\hat{\mathbb{P}}_y(\widetilde{Y}_s = i)\sum_k \hat{p}_k(i) \cdot \\ & \hat{\mathbb{P}}^{\underline{\lambda}}_z\left(\mathrm{Bes}(s)\mathrm{e}^{-(\underline{\lambda}-\lambda)(\mathrm{Bes}(s)-z)}1_{\{\mathrm{Bes}(s)\ge \lambda^{-1}\log^+\langle k,h\rangle + z\}}\right)\mathrm{d}s \\ \lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_0^{+\infty}\mathbb{P}\left(|W_s + \hat{z}|\mathrm{e}^{-(\underline{\lambda}-\lambda)|W_s+\hat{z}|}1_{\{|W_s+\hat{z}|\ge \lambda^{-1}\log^+\langle k,h\rangle + z\}}\right)\mathrm{d}s \\ \lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_{\{|y+\hat{z}|\ge \lambda^{-1}\log^+\langle k,h\rangle + z\}}|y + \hat{z}|\mathrm{e}^{-(\underline{\lambda}-\lambda)|y+\hat{z}|}\mathrm{d}y\int_0^{+\infty}s^{-\frac{3}{2}}\mathrm{e}^{-\frac{|y|^2}{2\pi}s^{-1}}\mathrm{d}s \\ \lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_{\{|y+\hat{z}|\ge \lambda^{-1}\log^+\langle k,h\rangle + z\}}\frac{|y + \hat{z}|}{|y|}\mathrm{e}^{-(\underline{\lambda}-\lambda)|y+\hat{z}|}\mathrm{d}y \\ \lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_{\{|y|\ge \lambda^{-1}\log^+\langle k,h\rangle\}}\frac{|y| + z}{|y|}\mathrm{e}^{-(\underline{\lambda}-\lambda)|y|}\mathrm{d}y \\ \lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_{\lambda^{-1}\log^+\langle k,h\rangle}^{+\infty}(r^2 + zr)\mathrm{e}^{-(\underline{\lambda}-\lambda)r}\mathrm{d}r \\ < & +\infty. \end{aligned}$$

Thus $\widetilde{R}_{0y}(\Theta < +\infty) = 1$.

On the other side,

$$
\begin{aligned}
& \widetilde{R}_{0y}\left(\sum_{n=0}^{+\infty} 1_{\{\langle A_{\varepsilon_n}, h\rangle > \mathrm{e}^{\lambda(\widetilde{X}(\zeta_n)+\underline{c}\zeta_{\varepsilon_n})}\}}\right) \\
= & \int_0^{+\infty} \sum_{i\in S}(a_i+\lambda^*)\hat{\mathbb{P}}_y(\widetilde{Y}_s=i)\sum_k \hat{p}_k(i)\,\hat{\mathbb{P}}^{\underline{\lambda}}_z\left(\mathrm{Bes}(s)<\lambda^{-1}\log^+\langle k,h\rangle+z\right)\mathrm{d}s \\
\lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_0^{+\infty}\mathbb{P}\left(|W_s+\hat{z}|<\lambda^{-1}\log^+\langle k,h\rangle+z\right)\mathrm{d}s \\
\lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_{\{|y+\hat{z}|<\lambda^{-1}\log^+\langle k,h\rangle+z\}}\mathrm{d}y\int_0^{+\infty} s^{-\frac{3}{2}}\mathrm{e}^{-\frac{|y|^2}{2\pi s}}\mathrm{d}s \\
\lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_{\{|y+\hat{z}|<\lambda^{-1}\log^+\langle k,h\rangle+z\}}|y|^{-1}\mathrm{d}y \\
\leq & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_{\{|y|<\lambda^{-1}\log^+\langle k,h\rangle+2z\}}|y|^{-1}\mathrm{d}y \\
= & \sum_{i\in S}\sum_k \hat{p}_k(i)\int_0^{\lambda^{-1}\log^+\langle k,h\rangle+2z} r\mathrm{d}r \\
\lesssim & \sum_{i\in S}\sum_k \hat{p}_k(i)(\lambda^{-1}\log^+\langle k,h\rangle+2z)^2 \qquad (4.10)
\end{aligned}
$$

The condition that $E(\xi_{ij}(\log^+\xi_{ij})^2)<+\infty$ for all $i,j\in S$ is equivalent to that $\sum_{k\in\mathbb{Z}^d_+}\hat{p}_k(i)(\log^+\langle k,b\rangle)^2<+\infty$ for all $i\in S$. Therefore, by (4.10)

$$
\widetilde{R}_{0y}\left(\sum_{n=0}^{+\infty} 1_{\{\langle A_{\varepsilon_n}, h\rangle > \mathrm{e}^{\lambda(\widetilde{X}(\zeta_n)+\underline{c}\zeta_{\varepsilon_n})}\}}<+\infty\right)=1,
$$

which means Λ is a finite sum. Hence $\Lambda<+\infty$ $\widetilde{R}_{0y}$-a.s. We complete the proof of part (2). □

Proof of Theorem 5: Suppose $E(\xi_{ij}(\log^+\xi_{ij})^2)=+\infty$ for some $i,j\in S$. Since

$$
V_{\underline{\lambda}}(\varepsilon_n)\geq\langle A_{\varepsilon_n},b\rangle(z+\widetilde{X}(\zeta_{\varepsilon_n})+\underline{\lambda}\zeta_{\varepsilon_n})\mathrm{e}^{-\underline{\lambda}(\widetilde{X}(\zeta_{\varepsilon_n})+\underline{c}\zeta_{\varepsilon_n})},
$$

using Lemma 12(1), we have

$$
\limsup_{t\to+\infty}V_{\underline{\lambda}}(t)=+\infty\quad\widetilde{R}_{0y}\text{-a.s.}
$$

Thus $P_{0y}(V(\underline{\lambda})=0)=1$ by Lemma 8.

On the other side, suppose $E(\xi_{ij}(\log^+\xi_{ij})^2)<+\infty$ for all $i,j\in S$. Recall that under $\widetilde{R}_{0y}$, $\{z+\widetilde{X}(t)+\underline{\lambda}t\colon t\geq 0\}$ is a Bessel-3 process which is transient, i.e. $\widetilde{R}_{0y}(\lim_{t\to+\infty}(z+\widetilde{X}(t)+\underline{\lambda}t)=+\infty)=1$, then from the spine decomposition for $V_{\underline{\lambda}}(t)$ and Lemma 12(2), we have

$$
\limsup_{t\to+\infty}\widetilde{R}_{0y}(V_{\underline{\lambda}}(t)\,|\,\widetilde{\mathcal{G}})<+\infty\quad\widetilde{R}_{0y}\text{-a.s.}
$$

By Fatou's lemma, we get

$$\limsup_{t\to+\infty} V_{\underline{\lambda}}(t) < +\infty \quad \widetilde{R}_{0y}\text{-a.s.},$$

which implies that $V_{\underline{\lambda}}(t)$ converges to $V(\underline{\lambda})$ in $L^1(P_{0y})$. Thus $P_{0y}(V(\underline{\lambda})=0)<1$. Similar analysis as in the proof of Theorem 3 can be applied here to show that $P_{0y}(V(\underline{\lambda})=0)=0$. Hence we complete the proof. □

Proof of Theorem 2: Using the same techniques as in the supercritical case, we can prove that $\underline{w}(x,y)$

$$\underline{w}(x,y) = E_{0y}\left[\prod_{u\in Z(s)} \underline{w}(x+X_u(s)+\underline{c}s, Y_u)\right].$$

Obviously, $\lim_{x\to+\infty}\underline{w}(x,y)=1$, $\lim_{x\to-\infty}\underline{w}(x,y)=0$. Thus $\underline{w}(x,y)$ provides a non-trivial traveling wave solution to (1.1). Note that $E_{0y}M(\underline{\lambda}) = \lim_{t\to+\infty} E_{0y}V_{\underline{\lambda}}(t) = E_{0y}V_{\underline{\lambda}}(0) = xh_y$, and that $\lim_{x\to+\infty}\underline{w}(x,y)=1$, thus

$$\frac{1-\underline{w}(x,y)}{xh_y\mathrm{e}^{-\underline{\lambda}x}} = \frac{1-E_{0y}\left[\exp\{-\mathrm{e}^{-\underline{\lambda}x}M(\underline{\lambda})\}\right]}{E_{0y}\left[\mathrm{e}^{-\underline{\lambda}x}M(\underline{\lambda})\right]} \to 1 \quad \text{as } x\uparrow+\infty,$$

i.e. $1-\underline{w}(x,y)\sim xh_y\mathrm{e}^{-\underline{\lambda}x}$ as $x\uparrow+\infty$.

Next we prove the uniqueness. Consider the space-time barrier $\Gamma^{(z,\underline{\lambda})}$ for $z\geq 0$. By arresting lines of descendants the first time they hit this barrier we again produce a sequence of stopping lines $\{C(z,\underline{\lambda}) : z\geq 0\}$ which are dissecting and tending to infinity. Suppose $\Phi_{\underline{c}}$ is any traveling wave with speed $\underline{c}$, then

$$\begin{aligned} M_z(x,\underline{\lambda}) &:= \prod_{u\in C(z,\underline{\lambda})} \Phi_{\underline{c}}(x+X_u(t)+\underline{c}t, Y_u) \\ &= \exp\left\{\sum_{i\in S} \sharp C_i(z,\underline{\lambda}) \log\Phi_{\underline{c}}(x+z,i)\right\} \end{aligned}$$

is a P_{0y}-martingale which converges to $\Phi_{\underline{c}}(x,y)$ almost surely and in $L^1(P_{0y})$.

We turn our attention to the branching Brownian motion with a killing barrier at $\Gamma^{(-x,\underline{\lambda})}$ where $x>0$. Define $\tilde{C}(z,\underline{\lambda})$ to be the set of particles in the killed process that are stopped at the barrier $\Gamma^{(z,\underline{\lambda})}$. Obviously, $\tilde{C}(z,\underline{\lambda})$ consists of particles whose lines of descendants (including themselves) have spatial paths that have met the barrier $\Gamma^{(z,\underline{\lambda})}$ before meeting $\Gamma^{(-x,\underline{\lambda})}$. Recall that $\gamma^{(-x,\underline{\lambda})}$ denotes the event that the MBBM remains entirely to the right of $\Gamma^{(-x,\underline{\lambda})}$ and $P_{0y}(\gamma^{(-x,\underline{\lambda})})\uparrow 1$ as $x\uparrow+\infty$. On the event $\gamma^{(-x,\underline{\lambda})}$ the MBBM and the MBBM with killing barrier $\Gamma^{(-x,\underline{\lambda})}$ are the same, i.e. $\sharp C_i(z,\underline{\lambda}) = \sharp\widetilde{C}_i(z,\underline{\lambda})$ on $\gamma^{(-x,\underline{\lambda})}$. Therefore,

$$\lim_{z\to+\infty} -\sum_{i\in S}\sharp\tilde{C}_i(z,\underline{\lambda})\log\Phi_{\underline{c}}(z,i) \tag{4.11}$$

exists almost surely and is non-negative on $\gamma^{(-x,\underline{\lambda})}$. Furthermore, since the function $x \mapsto \Phi_{\underline{c}}(x,y)$ is non-trivial, an elementary argument shows that for $x > 0$ sufficiently large, $\lim_{z\to+\infty} -\sum_{i\in S} \sharp\tilde{C}_i(z,\underline{\lambda}) \log \Phi_{\underline{c}}(z,i)$ is positive with positive probability on $\gamma^{(-x,\underline{\lambda})}$.

Consider the sequence

$$\begin{aligned} V^x_{\tilde{C}(z,\underline{\lambda})} &:= (h_y x)^{-1} \sum_{u\in\tilde{C}(z,\underline{\lambda})} h_{Y_u}(x + X_u(t) + \underline{\lambda} t) e^{-\underline{\lambda}(X_u(t)+\underline{c}t)} \\ &= (h_y x)^{-1}(x+z) \sum_{i\in S} h_i \,\sharp\tilde{C}_i(z,\underline{\lambda}) e^{-\underline{\lambda} z}. \end{aligned}$$

By the property of dissecting stopping lines, $\{V^x_{\tilde{C}(z,\underline{\lambda})} : z \geq 0\}$ is a mean one P_{0y}-martingale with respect to $\{\mathcal{F}_{\tilde{C}(z,\underline{\lambda})} : z \geq 0\}$, and

$$\lim_{z\to+\infty} (x+z) e^{-\underline{\lambda} z} \sum_{i\in S} \sharp\tilde{C}_i(z,\underline{\lambda}) h_i = M(\underline{\lambda}) \quad P_{0y}\text{-a.s.} \tag{4.12}$$

The arguments on $W_{c(x,\lambda)}(\lambda)$ in proof of Theorem 1 are still work when $\lambda = \underline{\lambda}$, thus we have

$$\lim_{z\to+\infty} \sum_{i\in S} \sharp\tilde{C}_i(z,\underline{\lambda}) h_i e^{-\underline{\lambda} z} = 0 \quad P_{0y}\text{-a.s.} \tag{4.13}$$

Combining (4.12) and (4.13), we obtain

$$\lim_{z\to+\infty} z e^{-\underline{\lambda} z} \sum_{i\in S} \sharp\tilde{C}_i(z,\underline{\lambda}) h_i = M(\underline{\lambda}) \quad P_{0y}\text{-a.s.} \tag{4.14}$$

Applying similar arguments as in the supercritical case, we know that $\{(\sharp C_1(z,\underline{\lambda}), \cdots, \sharp C_d(z,\underline{\lambda})) : z \geq 0\}$ forms a non-extinct, positive regular continuous time multitype Markov branching process (z plays the role of time). By Kesten-Stigum theorem (see, for example, Theorem 2.1 of [6]), there a non-negative vector $\pi_{\underline{\lambda}} = (\pi^1_{\underline{\lambda}}, \cdots, \pi^d_{\underline{\lambda}})$ such that $\langle \pi_{\underline{\lambda}}, 1\rangle = 1$ and for all $i \in S$

$$\lim_{z\to+\infty} \frac{\sharp C_i(z,\underline{\lambda})}{\sharp C(z,\underline{\lambda})} = \pi^i_{\underline{\lambda}} \quad P_{0y}\text{-a.s.},$$

and consequently,

$$\lim_{z\to+\infty} \frac{\sharp\tilde{C}_i(z,\underline{\lambda})}{\sharp\tilde{C}(z,\underline{\lambda})} = \pi^i_{\underline{\lambda}}$$

almost surely on $\gamma^{(-x,\underline{\lambda})}$. Let $x \uparrow +\infty$, we have

$$\lim_{z\to+\infty} \frac{\sharp\tilde{C}_i(z,\underline{\lambda})}{\sharp\tilde{C}(z,\underline{\lambda})} = \pi^i_{\underline{\lambda}} \quad P_{0y}\text{-a.s.} \tag{4.15}$$

Let $\tilde{\pi} = \pi_{\underline{\lambda}}/\langle h, \pi\rangle$. Using (4.14), (4.15) and the fact that $h_i > 0$ for every $i \in S$, we get that for all $i \in S$

$$\lim_{z\to+\infty} z e^{-\underline{\lambda} z} \sharp\tilde{C}_i(z,\underline{\lambda}) = \tilde{\pi}_i M(\underline{\lambda}) \quad P_{0y}\text{-a.s.} \tag{4.16}$$

From (4.16) and (4.11), we conclude that

$$\beta := \lim_{z\to+\infty} -z^{-1}e^{\underline{\lambda}z}\sum_{i\in S}\tilde{\pi}_i \log\Phi_{\underline{c}}(z,i)$$

exists and is positive. Uniqueness (up to a multiplicative constant) is now immediate since

$$\begin{aligned}
\Phi_{\underline{c}}(x,y) &= E_{0y}\left(\lim_{z\to+\infty} M_z(x,\underline{\lambda})\right)\\
&= E_{0y}\exp\left\{\lim_{z\to+\infty}\sum_{i\in S}\sharp C_i(z,\underline{\lambda})\log\Phi_{\underline{c}}(x+z,i)\right\}\\
&= \lim_{\eta\uparrow+\infty} E_{0y}\left[\exp\left\{\lim_{z\to+\infty}\sum_{i\in S}\sharp\tilde{C}_i(z,\underline{\lambda})\log\Phi_{\underline{c}}(x+z,i)\right\};\gamma^{(-\eta,\underline{\lambda})}\right]\\
&= \lim_{\eta\uparrow+\infty} E_{0y}\left[\exp\left\{\lim_{z\to+\infty}\sum_{i\in S}\tilde{\pi}_i M(\underline{\lambda})e^{\underline{\lambda}z}z^{-1}\log\Phi_{\underline{c}}(x+z,i)\right\};\gamma^{(-\eta,\underline{\lambda})}\right]\\
&= E_{0y}\exp\left\{-M(\underline{\lambda})e^{-\underline{\lambda}x}\lim_{z\to+\infty}-\frac{x+z}{z}\sum_{i\in S}\tilde{\pi}_i(x+z)^{-1}e^{\underline{\lambda}(x+z)}\log\Phi_{\underline{c}}(x+z,i)\right\}\\
&= E_{0y}\exp\left\{-\beta\, M(\underline{\lambda})e^{-\underline{\lambda}x}\right\},
\end{aligned}$$

here in the fourth equality, we used (4.16). □

References

[1] K. Athreya (1968): Some results on multitype continuous time Markov branching processes, *Ann. Math. Statist.*, 39, 347-357.

[2] K. Athreya and P. Ney (1983): Branching Processes, Birkh*ä*user, London.

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

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

[5] R. Durret (1996): Probability: Theory and Examples (second edition), Duxbury Press.

[6] H.-O. Georgii and E. Baake (2003): Supercritical multitype branching processes: The ancestral types of typical individuals, *Adv. Appl. Prob.*, 35, 1090-1110

[7] R. Hardy and S. C. Harris (2009): A spine aproach to branching diffusions with applications to L^p-convergence of martingales, in *Séminaire de Probabilités*, XLII. C. Donati-Martin, M. mery, A. Rouault and C. Stricker (Eds), 1979, 281-330.

[8] S. C. Harris (1999): Traveling-waves for the FKPP equation via probabilistic arguments, *Proc. Roy. Soc. Edinburgh Sect. A*, 129, 503-517.

[9] P. Jagers: General branching processes as Markov fields, *Stochastic Process. Appl.*, 32, 193-212, 1989.

[10] H. Kesten and B. P. Stigum (1966): A Limit Theorem for Multidimensional Galton-Watson Process, *Ann. Math. Statist.*, 37, 1211-1223.

[11] T. G. Kurtz, R. Lyons, R. Pemantle and Y. Peres (1997): A conceptual proof of the Kesten-Stigum theorem for multitype braching processes, in *Classiscal and Modern Branching Processes*, 84, 181-186, Springer-Verlag, New York.

[12] A. E. Kyprianou (2004): Traveling wave solutions to the K-P-P equation: Alternatives to Simon Harris' probability analysis, *Ann. Inst. H. Poincaré Probab. Statist.*, 40, 50-72.

[13] A. E. Kyprianou, R.-L. Liu, A. Murillo-Salas and Y.-X. Ren (2012): Supercritical super-Brownian motion with a general branching mechnism and traveling waves, *Ann. Inst. H. Poincaré Probab. Statist.*, 48, 661C687.

[14] R. Lyons (1997): A simple path to Biggin's martingale convergence for branching random walk, in *Classiscal and Modern Branching Processes*, 84, 217-221, Springer-Verlag, New York.

[15] R. Lyons, R. Pemantle and Y. Peres (1995): Conceptual proofs of LlogL criteria for mean behavior of branching processes, *Ann. Probab.*, 23, 1125-1138

[16] H. P. McKean (1975): Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov, *Commun. Pure Appl. Math.*, 28, 323-331.

[17] D. Revuz and M. Yor (1991): *Continuous Martingales and Brownian Motion*, Springer-Verlag.

[18] E. Seneta (1973): *Non-negative matrices-An introduction to theory and applications.* Halsted Press.

Yan-Xia Ren: LMAM School of Mathematical Sciences & Center for Statistical Science, Peking University, Beijing, 100871, P. R. China, E-mail: yxren@math.pku.edu.cn

Ting Yang: Institute of Applied Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, P.R.China,
E-mail: yangt@amss.ac.cn