Large deviation for extremes of branching random walk with regularly varying displacements

Ayan Bhattacharya

Centrum Wiskunde & Informatica, Amsterdam

February 19, 2018

Abstract

We consider discrete time branching random walk on real line where the displacements of particles coming from the same parent are allowed to be dependent and jointly regularly varying. Using the one large bunch asymptotics, we derive large deviation for the extremal processes associated to the suitably scaled positions of particles in the nth generation where the genealogical tree satisfies Kesten-Stigum condition. The large deviation limiting measure in this case is identified in terms of the cluster Poisson point process obtained in the underlying weak limit of the point processes. As a consequence of this, we derive large deviation for the rightmost particle in the nth generation giving the heavy-tailed analogue of recent work by Gantert and Höfelsauer [2018].

Key words and phrases. Branching random walk, Extreme values, Regular variation, Point process, Rightmost point, Large deviation.

2010 Mathematics Subject Classification. Primary 60G70, 60G55, 60F10; Secondary 60J80.

1 Introduction

Branching Random Walk (BRW) on real line is a generalization of a branching process which allows spatial movement of the particles on the real line. The process starts with a single particle and then each particle moves and branches independently of the other particles. A more detailed description is as follows: It starts with one particle at the origin of the real line at time 0. Each particle lives for a unit time and then gives birth to an independent copy of a point process $\mathscr{L}$. The number of atoms of the point process produced by a particle denotes the number of children of the particle and the atom(s) of the point process denote(s) displacement(s) attached to children of the particle. Position of a particle is defined to be displacement of the particle translated by position of its parent. Collection of positions of the particles in the system is known as BRW. Generation of a particle denotes time of its birth. For a more detailed mathematical description of BRW, we refer the reader to Subsection 2.2.

Since last few decades, the model BRW attracted attention of many researchers from different disciplines due to its connection to Gaussian multiplicative chaos (Ding et al. [2017]), Gaussian free field (Bramson et al. [2016a]), first and last passage percolation on Galton-Watson tree, ecology, random polymer, random algorithms (Dutta et al. [2015]) etc. We refer to Shi [2016] for a more detailed overview and references therein. Groundbreaking works include Kingman [1975], Hammersley [1974], Biggins [1976] where almost sure limit for the leftmost position is obtained for displacements with exponentially decaying tails. In this set up, the extremes for BRW on real line are extensively studied e.g. Bachmann [2000], Addario-Berry and Reed [2009], Aïdékon [2013], Bramson et al. [2016b], Madaule [2015] to mention a few. Large deviation for rightmost position is obtained in branching Brownian motion (BBM) in Chauvin and Rouault [1988]. Recently, Derrida and Shi [2016] derived large deviation for the rightmost position in different variants of BBM. The large deviation for the empirical measure in discrete time BRW is derived in Louidor and Perkins [2015], Louidor and Tsairi [2017] and Chen and He [2017] though large deviation for extremes does not seem to follow from these works. Recently, large deviation for the rightmost position in case of displacements with exponentially decaying tail has been recently addressed in Gantert and Höfelsauer [2018] using a comparison with independent collection of random walks.

It is known in literature that behavior of extremes changes dramatically if tail of displacement does not decay exponentially fast (see e.g. Durrett [1979], Durrett [1983], Gantert [2000], Maillard [2016], Bérard and Maillard [2014]). In this article, we shall focus on BRW with displacements having regularly varying tails. It has been established in Durrett [1983] that there exists a sequence $(b_n : n \geq 1)$ of scalars such that $b_n^{-1} M_n$ converges in distribution as $n \to \infty$ where M_n denotes the position of the rightmost particle at the nth generation. This result has been extended in Bhattacharya et al. [2016] and Bhattacharya et al. [2018] where weak limit of extremal processes is obtained when positions are divided by b_n. Consider another sequence of scalars $(\gamma_n : n \geq 1)$ growing faster than $(b_n : n \geq 1)$ i.e. $\lim_{n\to\infty} \gamma_n^{-1} b_n = 0$. Note that $\gamma_n^{-1} M_n$ converges to 0 in probability as $n \to \infty$. It is very natural to ask if there exists a sequence of positive scalars $(r_n : n \geq 1)$ such that $r_n \mathbf{P}(M_n > \gamma_n x)$ converges to some non-zero functions for every $x > 0$ i.e. the rate of convergence for $\mathbf{P}(M_n > \gamma_n x)$ motivated from Gantert and Höfelsauer [2018]. Note that similar questions can be asked about the kth rightmost position, joint distribution of the rightmost and leftmost position, gap statistics etc. This motivates and necessitates the investigation for large deviation of point processes (see Hult and Samorodnitsky [2010], Fasen and Roy [2016]) which can work as master key for opening many locks. Let N_n be the point process which puts unit mass to the positions of the particles in the nth generation when positions are divided by γ_n. Suppose that $\mathscr{M}$ be the space of all point measures on $\mathbb{R} \setminus \{0\}$. Then it is clear that N_n does not converge to an $\mathscr{M}$-valued random variable as $\mathscr{M}$ does not contain a point measure with atom at 0. So $\mathbf{P}(N_n \in A)$ converges to 0 for every "nice" set $A \subset \mathscr{M}$. In this work, we obtain a sequence $(r_n : n \geq 1)$ such that $r_n \mathbf{P}(N_n \in A)$ converges to a non-zero limit for "nice" subsets $A \subset \mathscr{M}$ and identify the limit in terms of a cluster Poisson point process. As a consequence, we shall derive large deviation for the rightmost particle in the nth generation i.e. limit of $r_n \mathbf{P}(M_n > \gamma_n x)$ for every $x > 0$.

We shall use tools developed in Hult and Lindskog [2006], Hult and Samorodnitsky [2010] and Lindskog et al. [2014] to derive large deviation for extremal processes. Note that the framework proposed in Hult and Samorodnitsky [2010] is aimed to study the large values occurring in stationary sequences. It is not very straightforward to adapt the framework here as positions in the nth generation is non-stationary due to dependence structure among them. The key ingredient used in this article is "principle of a bunch of large displacements" for random variables with jointly regularly varying tail. The proof is divided into four steps where we locate the large displacement in the first three steps and compute the contribution of the large displacement in the last step. BRW on real line can be viewed as a collection of dependent random walks where position of a particle at the nth generation has same distribution as position of random walker (distribution of the steps is same as that of displacements) at time n. In the first step, we show that there can be at most one large displacement along a path to cause large position of a particle in the nth generation. Then we shall cut the tree at the $(n-K)$th generation and ignore the displacements associated to members of the first $(n-K)$ generations using the fact that the large displacements can occur at the last K generations with high probability. This can be justified by the fact that total number of particles upto the first $(n-K)$th generation is negligible with respect to total number of particles upto the nth generation for large enough K. In the next step, we use a truncation technique based on progeny distribution and prune each of subtrees deleting some children with their lines of descendants. It is rare to see large displacement associated to the deleted ones. The formalization of these three steps is very similar in spirit to those in Bhattacharya et al. [2018] but worth mentioning. Then we regularize the pruned subtrees. Finally we compute limit of $r_n \mathbf{P}(N_n \in A)$ for "nice" $A \subset \mathscr{M}$ explicitly in terms of a cluster Poisson process using truncation technique for number of particles in the $(n-K)$th generation and facts from regular variation on $\mathbb{R}^{\mathbb{N}}$.

In Section 2, we discuss the tools to be used (see subsection 2.1) in this article, the model BRW with specific assumptions on genealogical structure and displacements (see subsection 2.2)and state Theorem 2.4 with its consequences (see 2.3). In Section 3, we prove Theorem 2.4 except some of the steps which are similar in spirit to those in Bhattacharya et al. [2018]. Finally, proofs of Corollaries 2.6, 2.7 are given in Section 4.

2 Preliminaries and main result

In the following subsection, we give a brief review on $\mathbb{M}$-convergence on $\mathbb{R}^{\mathbb{N}}$ and space of all point measures on $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$. This review will be helpful to state assumptions on the joint distribution of the displacements from same parent and Theorem 2.4.

2.1 $\mathbb{M}$-convergence on Polish space and regular variation on $\mathbb{R}^{\mathbb{N}}$

Let $(\mathbb{S}, d)$ is a Polish space. In Lindskog et al. [2014], $\mathbb{M}$-convergence is defined to study convergence of measures on the space $\mathbb{S} \setminus \mathbb{C}$ where $\mathbb{C}$ is a closed set in $\mathbb{S}$. In this article, we need to consider the case where the closed set $\mathbb{C}$ is a singleton. So for the sake of simplicity we shall define $\mathbb{M}$-convergence on the space $\mathbb{S} \setminus \{s_0\}$ where $s_0 \in \mathbb{S}$. A sequence of measures $(\xi_n : n \geq 1)$ converges to a measure ξ on $\mathbb{S} \setminus \{s_0\}$ if for every bounded continuous function $f : \mathbb{S} \to [0, \infty)$ which vanishes in a neighbourhood of s_0,

$$\lim_{n\to\infty} \int_{\mathbb{S}\setminus\{s_0\}} f(x)\xi_n(dx) = \int_{\mathbb{S}\setminus\{s_0\}} f(x)\xi(dx).$$

Using this notion of convergence, regular variation on the space $\mathbb{R}^{\mathbb{N}}$ can be defined as follows. For every $a > 0$ and $\mathbf{x} = (x_i : i \geq 1) \in \mathbb{R}^{\mathbb{N}}$, we define $a.\mathbf{x} = (ax_i : i \geq 1) \in \mathbb{R}^{\mathbb{N}}$ to be the scalar multiplication of $\mathbf{x}$ by a. It is clear from the definition of scalar multiplication that from $a_1, a_2 > 0$ and $a > 1$, $a_1.(a_2.\mathbf{x}) = (a_1a_2).\mathbf{x}$, $1.\mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ and $dist(a.\mathbf{x}, \mathbf{0}) > 0$ for every $\mathbf{x} \in \mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\}$. A measure ξ on $\mathbb{R}^{\mathbb{N}}$ is said to be regularly varying if there exists an increasing sequence $a_n \uparrow \infty$ such that $n\xi(a_n.\cdot)$ converge to a non-null measure ϑ on $\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\}$ where $\xi(a_n.A) = \xi\Big(\{a_n x : x \in A\}\Big)$ for every measurable $A \subset \mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\}$. It follows from the definition that the limit measure ϑ satisfies the scaling relation $\vartheta(a.A) = a^{-\alpha}\vartheta(A)$ for some $\alpha > 0$ and every A such that $\mathbf{0} \notin \bar{A}$ ($\bar{A}$ denotes the closure of A) and $\vartheta(\partial A) = 0$ (∂A denotes the boundary of the set A).

We also define the notation $\pi_j : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ such that $\pi_j\Big((x_i : i \geq 1)\Big) = x_j$ for all $j \geq 1$ and $\pi_{j_1,j_2,\ldots,j_k} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^k$ such that $\pi_{j_1,j_2,\ldots,j_k}\Big((x_i : i \geq 1)\Big) = \Big(x_{j_1}, x_{j_2}, \ldots, x_{j_k}\Big)$ for all $k \in \mathbb{N}$. This notation will turn out to be helpful to derive large deviation for the rightmost position when displacements from same parent are dependent. Suppose $\mathrm{PROJ}_j : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^j$ is a map such that $\mathrm{PROJ}_j(\mathbf{x}) = (x_1, x_2, \ldots, x_j)$ i.e. the projection to the first j coordinates of $\mathbf{x}$. Let $\mathscr{M}(\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\})$ denote the space of all measures on $\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\}$ and $\mathscr{M}_0(\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\}) = \mathscr{M}(\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\}) \setminus \{\emptyset\}$ where $\emptyset$ is the null measure on the space $\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\}$. Let $(\xi_n : n \geq 1)$ be a sequence of measures in $\mathscr{M}_0(\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\})$ and $\xi \in \mathscr{M}_0(\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\})$. Following fact will be used to prove Theorem 2.4.

Fact 2.1 (Theorem 4.1 in Lindskog et al. [2014])**.** $\xi_n \to \xi$ in $\mathscr{M}_0(\mathbb{R}^{\mathbb{N}} \setminus \{\mathbf{0}\})$ if and only if for every $j \geq 1$, $\xi_n \circ \mathrm{PROJ}_j^{-1} \to \xi \circ \mathrm{PROJ}_j^{-1}$ on the space of all measures on $\mathbb{R}^j \setminus \{\mathbf{0}\}$.

In Appendix A of Hult and Samorodnitsky [2010], tools are developed to study convergence of measures on the space of all point measures on $\mathbb{R} \setminus \{0\}$. We shall use this framework to derive the limit of sequence of measures $r_n\mathbf{P}(N_n \in \cdot)$ where $(r_n : n \geq 1)$ is chosen appropriately. Let $\mathscr{M}$ denote the space of all point measures on the space $\mathbb{R} \setminus \{0\}$ and $\mathscr{M}_0 = \mathscr{M} \setminus \{\emptyset\}$ where $\emptyset$ denotes the null measure on $\mathbb{R}_0$. Suppose that $C_c^+([-\infty, \infty] \setminus \{0\})$ denotes the space of all non-negative, bounded continuous functions $f : \mathbb{R} \to [0, \infty)$ such that f vanishes in the neighbourhood of 0. Let $(h_i : i \geq 1)$ be a countable dense collection of functions in $C_c^+([-\infty, \infty] \setminus \{0\})$ and $(\xi_n : n \geq 1)$ be a collection of countable elements in $\mathscr{M}_0$. Let $\xi_n(h_i) \to \xi(h_i)$ for all $i \geq 1$ and some $\xi \in \mathscr{M}_0$, then we say $\xi_n \stackrel{v}{\to} \xi$ as $n \to \infty$ where $\stackrel{v}{\to}$ denotes vague convergence. Note that vague convergence in $\mathscr{M}_0$ is metrizable and the metric induced by vague convergence in $\mathscr{M}_0$ is denoted by d_{vague}. It can be shown that $(\mathscr{M}_0, d_{\text{vague}})$ is a locally compact, complete, separable metric space i.e. a locally compact Polish space. Let $\mathbf{M}(\mathscr{M}_0)$ be the space of all measures on $\mathscr{M}_0$, β_0 is the null measure on $\mathscr{M}_0$ and $\mathbf{M}_0 = \mathbf{M}(\mathscr{M}_0) \setminus \{\beta_0\}$. Let $\vartheta_n \in \mathbf{M}_0$ for all $n \geq 1$ and $\vartheta \in \mathbf{M}_0$. Then we say that ϑ_n converges to ϑ in Hult-Lindskog-Samorodnitsky sense and denoted by $\vartheta_n \xrightarrow{HLS} \vartheta$ as $n \to \infty$ if ϑ_n converges to ϑ in the sense of $\mathbb{M}$-convergence with $\mathbb{S} = \mathscr{M}_0$ and $s_0 = \{\emptyset\}$. This notion of convergence will be used to state Theorem 2.4. In Hult and Samorodnitsky [2010], a convergence determining class of functions is identified for $\xrightarrow{HLS}$ which will play fundamental role in this article and will be described below briefly. Suppose that $g_i \in C_c^+([-\infty, \infty] \setminus \{0\})$, $\epsilon_i > 0$ for $i = 1, 2$. Consider a function $F_{g_1,g_2,\epsilon_1,\epsilon_2} : \mathscr{M}_0 \to [0, \infty)$ by

$$F_{g_1,g_2,\epsilon_1,\epsilon_2}(\varphi) = \prod_{i=1}^{2} \left(1 - \exp\left\{ - \left(\varphi(g_i) - \epsilon_i\right)_+ \right\}\right)$$

for every $\varphi \in \mathscr{M}_0$ which vanishes in the neighbourhood of the null measure $\emptyset \in \mathscr{M}$.

Fact 2.2 (Lemma A.1 in Hult and Samorodnitsky [2010])**.** Let $\mathbf{m}_1$ and $\mathbf{m}_2$ are two elements of $\mathbf{M}_0$. Then $\mathbf{m}_1 = \mathbf{m}_2$ if and only if for every $g_1, g_2 \in C_c^+([-\infty, \infty] \setminus \{0\})$ and $\epsilon_1, \epsilon_2 > 0$,

$$\mathbf{m}_1\Big(F_{g_1,g_2,\epsilon_1,\epsilon_2}\Big) = \mathbf{m}_2\Big(F_{g_1,g_2,\epsilon_1,\epsilon_2}\Big).$$

Fact 2.3 (Lemma A.2 in Hult and Samorodnitsky [2010])**.** Let $(m_n : n \geq 1)$ be sequence of measures in $\mathbf{M}_0$. Then $m_n \xrightarrow{HLS} m$ if and only if for every $g_1, g_2 \in C_c^+([-\infty, \infty] \setminus \{0\})$ and $\epsilon_1, \epsilon_2 > 0$,

$$\lim_{n\to\infty} \mathbf{m}_n\Big(F_{g_1,g_2,\epsilon_1,\epsilon_2}\Big) = \mathbf{m}\Big(F_{g_1,g_2,\epsilon_1,\epsilon_2}\Big). \tag{2.1}$$

It is enough to check (2.1) holds for only Lipscitz continuous functions g_1, g_2.

Note that, we have to show that $\mathbf{m}(B_r) < \infty$ where $B_r = \{\varphi \in \mathscr{M}_0 : d_{\text{vague}}(\varphi, \emptyset) > r\}$ to establish that $\mathbf{m} \in \mathbf{M}_0$.

2.2 Assumptions on $\mathscr{L}$

Recall that BRW on real line is a spatial branching process where each particle is assigned a position in the real line. It starts with one particle at the origin of the real line $\mathbb{R}$. Each particle in the system lives for a unit time and before dying, each particle produces an independent copy of the point process $\mathscr{L}$ independently of the other particles. Each particle is assigned a position which is defined to be displacement of the particle translated by position of its parent. In this article, we shall assume that

$$\mathscr{L} \stackrel{d}{=} \sum_{i=1}^{Z} \boldsymbol{\delta}_{X_i}$$

where Z is an $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$-valued random variable which is independent of the collection of real-valued random variables $(X_i : i \geq 1)$. We shall follow the convention that $\mathscr{L} = \emptyset$ if $Z = 0$ where $\emptyset$ denotes the null measure on the real line $\mathbb{R}$. We shall also assume that

$$\mathbf{E}(Z \log_+ Z) < \infty \tag{2.2}$$

where $\log_+ x = \log(x \vee 1)$ and $\mu := \mathbf{E}(Z) > 1$.

Now, we shall state assumptions on displacements. We shall assume that the random variables X_i are marginally identically distributed with regularly varying tails of index α for every $i \geq 1$ and $(X_i : i \geq 1)$ is an $\mathbb{R}^{\mathbb{N}} = \prod_{j \in \mathbb{N}} \mathbb{R}$-valued random variable with regularly varying tail i.e. jointly regularly varying on $\mathbb{R}^{\mathbb{N}}$. Precise assumptions on marginal and joint distribution of displacements are stated below.

1. **Marginal distribution**: We assume that marginal distribution of $\mathbf{X} = (X_i : i \geq 1)$ are identical and $\mu^n \mathbf{P}(b_n^{-1} X_i \in \cdot) \in RV_\alpha(\mathbb{R}, \nu_\alpha)$ where $\nu_\alpha(\cdot)$ is a measure on $\mathbb{R}$ such that

$$\nu_\alpha(dx) = \alpha p x^{-\alpha-1} \mathbb{1}(x > 0) dx + \alpha(1-p)(-x)^{-\alpha-1} \mathbb{1}(x < 0) dx \tag{2.3}$$

   where

$$p = \lim_{x \to \infty} \frac{\mathbf{P}\Big(X_1 > x\Big)}{\mathbf{P}\Big(|X_1| > x\Big)} = 1 - \lim_{x \to \infty} \frac{\mathbf{P}\Big(X_1 < -x\Big)}{\mathbf{P}\Big(|X_1| > x\Big)} \tag{2.4}$$

   for every $i \geq 1$.

2. **Joint distribution**: We assume that $\mu^n \mathbf{P}\Big(b_n^{-1} . \mathbf{X} \in \cdot\Big) \in RV_\alpha(\mathbb{R}^{\mathbb{N}}, \lambda)$ where $\lambda(\cdot)$ is a measure on $\mathbb{R}^{\mathbb{N}}$. The form of $\lambda(\cdot)$ is given in Bhattacharya et al. [2018] where X_i's are independently and identically. distributed. The expression for $\lambda(\cdot)$ is derived in Resnick and Roy [2014]

where $\mathbf{X}$ is a moving average process. It follows from Fact 2.1 that $\lambda \circ \mathrm{PROJ}_1^{-1} = \nu_\alpha$.

Under above assumptions, we can see that the underlying genealogical tree of BRW is a supercritical Galton-Watson (GW) tree with progeny distribution Z. It follows from (2.2) that the supercritical GW tree satisfies Kesten-Stigum condition. The tree will be denoted by $\mathbb{T}$. We shall use Ulam-Harris labelling to label the vertices in $\mathbb{T}$. We shall use $\mathbf{u}$ and $\mathbf{v}$ as the generic labels for vertices. Note that the displacement associated to a vertex $\mathbf{u}$ is denoted by $X(\mathbf{u})$ and position of the $\mathbf{v}$th vertex is denoted by $S(\mathbf{v})$. Generation (the time of birth) of a vertex $\mathbf{v}$ is denoted by $|\mathbf{v}|$. Z_n denotes number of particles at the generation n for every $n \geq 0$ with $Z_1 \stackrel{d}{=} Z$. Let $\mathcal{S} = \bigcap_{n>0}\{Z_n > 0\}$ denote survival of the underlying GW tree. Probability conditioned on the survival of the tree will be denoted by $\mathbf{P}^*$ and expectation corresponding to $\mathbf{P}^*$ will be denoted by $\mathbf{E}^*$.

In Bhattacharya et al. [2018], it has been shown that

$$\mathsf{N}_n = \sum_{|\mathbf{v}|=n} \boldsymbol{\delta}_{b_n^{-1} S(\mathbf{v})}$$

converges weakly (under $\mathbf{P}^*$) in the space $\mathscr{M}(\mathbb{R}_0) = \{$ space of all measures on $\mathbb{R}_0\}$ under the above assumptions, where $(b_n : n \geq 1)$ is an increasing sequence of positive real numbers introduced in (2.3). Suppose that $(\gamma_n : n \geq 1)$ is an increasing sequence of positive real numbers such that $\gamma_n^{-1} b_n \downarrow 0$ as $n \to \infty$. Then it is clear that the sequence of point processes

$$N_n = \sum_{|\mathbf{v}|=n} \boldsymbol{\delta}_{\gamma_n^{-1} S(\mathbf{v})} \tag{2.5}$$

does not converge in the space $\mathscr{M}_0 = \mathscr{M}(\mathbb{R}_0) \setminus \{\emptyset\}$ conditioned on the survival of the tree. This means that the sequence of measures $\mathbf{P}^*(N_n \in \cdot)$ converges to null-measure on the space of all measures on $\mathscr{M}_0$. Following Hult and Samorodnitsky [2010], we can say that large deviation for the point processes is to find a sequence of constants $(r_n : n \geq 1)$ such that $r_n \mathbf{P}\Big(N_n \in \cdot\Big)$ converges to some non-null measure on $\mathscr{M}_0$. The aim of this article is to find the sequence $(r_n : n \geq 1)$ and to compute limit of the sequence of measures $\Big(r_n \mathbf{P}^*(N_n \in \cdot) : n \geq 1\Big)$.

2.3 Main result and its consequences

Define

$$r_n = \Big(\mu^n \mathbf{P}(|X_1| > \gamma_n)\Big)^{-1} \tag{2.6}$$

for every $n \geq 1$. As $\mu^n \mathbf{P}(|X_1| > \gamma_n) \downarrow 0$ as $n \to \infty$, it is clear that $r_n \uparrow \infty$ as $n \to \infty$. Note that $m_n(\cdot) = r_n \mathbf{P}^*(N_n \in \cdot)$ is an element in $\mathbf{M}_0$. To describe limit of the sequence of measures

$(m_n : n \geq 1)$, we shall need following notations.

- Let U is an independent copy of Z. $\widetilde{U}$ denotes the random variable U conditioned to stay positive i.e. $\mathbf{P}(\widetilde{U} \in A) = \mathbf{P}(U \in A | U > 0)$ for every $A \subset \mathbb{N}$.
- For every $l \geq 1$, $(\widetilde{Z}_l^{(s)} : s \geq 1)$ is a collection of independent copies of the random variable $\widetilde{Z}_l$ which is the random variable Z_l conditioned to stay positive i.e. $\mathbf{P}(\widetilde{Z}_l^{(s)} \in A) = \mathbf{P}(Z_l \in A | Z_l > 0)$ for every $A \subset \mathbb{N}$.
- $\mathcal{S}$ denotes the event that $\mathbb{T}$ does not die out or become extinct. $p_e = 1 - \mathbf{P}(\mathcal{S})$ denotes the probability of extinction.
- For any set G, $|G|$ denotes the cardinality of the set G and $\mathrm{Pow}(G)$ denotes the power set of G i.e. the collection of all subsets of G.

Theorem 2.4. *Under the assumptions stated in Subsection 2.2,*

$$r_n \mathbf{P}^*(N_n \in \cdot) \xrightarrow{HLS} m^*(\cdot) \tag{2.7}$$

in the space $\mathbf{M}_0$ where m^* can be described as

$$m^*(\cdot) = (1 - p_e)^{-1} \mathbf{P}(U > 0) \sum_{l=0}^{\infty} \mu^{-(l+1)} \mathbf{E}\Big[\sum_{G \in \mathrm{Pow}(\widetilde{U}) \setminus \{\emptyset\}} \lambda\Big(\mathbf{x} \in \mathbb{R}^{\mathbb{N}} : \sum_{s \in G} \widetilde{Z}_l^{(s)} \boldsymbol{\delta}_{x_s} \in \cdot\Big)$$
$$\Big(\mathbf{P}(Z_l > 0)\Big)^{|G|} \Big(\mathbf{P}(Z_l = 0)\Big)^{\widetilde{U} - |G|}\Big]. \tag{2.8}$$

Remark 2.5. Note that the measure m^* is very similar to the measure Υ which appeared in Bhattacharya et al. [2018] (see Lemma 4.4 and display (4.14)) in order to verify that the point process associated to a typical subtree (after cutting and pruning) is regularly varying. So it seems that the approach used in the aforementioned reference to derive the weak limit of the sequence of point processes facilitates the study of large deviation of extremal process.

Corollary 2.6 (Large deviation for the rightmost position). Under the assumptions stated in Theorem 2.4,

$$\lim_{n \to \infty} r_n \mathbf{P}^* \Big(\max_{|\mathbf{v}| = n} S(\mathbf{v}) > \gamma_n x \Big) = x^{-\alpha} c_1$$

for every $x > 0$ and some positive constant c_1 given below

$$c_1 = (1 - p_e)^{-1} \mathbf{P}(U > 0) \sum_{l=0}^{\infty} \mu^{-(l+1)} \mathbf{E}\Big[\sum_{G \in \mathrm{Pow}(\widetilde{U}) \setminus \{\emptyset\}}$$

$$\Big(\lambda\Big(\cup_{s\in G} V_s\Big)\Big)\Big(\mathbf{P}(Z_l > 0)\Big)^{|G|}\Big(\mathbf{P}(Z_l = 0)\Big)^{\widetilde{U}-|G|}\Big] \tag{2.9}$$

where $V_s \subset \mathbb{R}^{\mathbb{N}}$ such that $\pi_s\Big(V_s\Big) = (1,\infty)$ and $\pi_i\Big(V_s\Big) = \mathbb{R}$ for all $i \in \mathbb{N}\setminus\{s\}$ and for all $s \in \mathbb{N}$.

Suppose that displacements associated to $\mathscr{L}$ are independently distributed. Then we can see that the limit measure λ on $\mathbb{R}^{\mathbb{N}}$ admits a special form (see Example 2.1 in Bhattacharya et al. [2018]) and we shall denote it by

$$\lambda_{iid}(\cdot) = \sum_{i=1}^{\infty} \otimes_{j=1}^{i-1}\boldsymbol{\delta}_0 \otimes \nu_\alpha \otimes_{j'=i+1}^{\infty} \boldsymbol{\delta}_0. \tag{2.10}$$

Corollary 2.7 (Large deviation for point process with IID displacements). Suppose that the assumptions in Theorem 2.4 hold and the displacements of $\mathscr{L}$ are independently and identically distributed. Then $r_n\mathbf{P}^*(N_n \in \cdot) \xrightarrow{HLS} m^*_{iid}(\cdot)$ where

$$m^*_{iid}(\cdot) = (1-p_e)^{-1}\sum_{l=1}^{\infty}\mu^{-l}\mathbf{P}(Z_l > 0)\mathbf{E}\Big(\nu_\alpha\big(x \in \mathbb{R} : \widetilde{Z}_l\boldsymbol{\delta}_x \in \cdot\big)\Big). \tag{2.11}$$

Remark 2.8 (Large deviation for rightmost position in case of IID displacements). If the displacements are independently and identically distributed then

$$\lim_{n\to\infty} r_n\mathbf{P}\big(\max_{|\mathbf{v}|=n} S(\mathbf{v}) > \gamma_n x\big) = px^{-\alpha}(1-p_e)\sum_{l=1}^{\infty}\mu^{-l}\mathbf{P}(Z_l > 0). \tag{2.12}$$

We can use the fact that $\lambda = \sum_{t=1}^{\infty}\otimes_{i=1}^{t-1}\boldsymbol{\delta}_0 \otimes \nu_\alpha \otimes_{i=t+1}^{\infty}\boldsymbol{\delta}_0$ to obtain $\lambda(\cup_{s\in G}V_s) = |G|\nu_\alpha((1,\infty)) = |G|p$. The remark follows by using this observation to right hand side of (2.9).

3 Proof of Theorem 2.4

Fix $g_i \in C_c^+(\bar{\mathbb{R}}_0)$ and $\epsilon_i > 0$ for $i = 1, 2$. In view of the Fact 2.3, it is necessary and sufficient to establish that

$$\lim_{n\to\infty} m_n(F_{g_1,g_2,\epsilon_1,\epsilon_2}) = \lim_{n\to\infty} r_n\mathbf{E}\Big[\prod_{i=1}^{2}\Big(1 - \exp\Big\{-\Big(N_n(g_i) - \epsilon_i\Big)_+\Big\}\Big)\Big] = m_*(F_{g_1,g_2,\epsilon_1,\epsilon_2}) \tag{3.1}$$

in order to prove Theorem 2.4. We shall establish (3.1) using four steps which are described as follows: 1. single large displacement on every fixed path, 2. cutting the tree $\mathbb{T}$ into subtrees, 3. pruning the subtrees and 4. regularization of the pruned subtrees. Note that these four steps have been used before to derive weak limit of the sequence $(\mathsf{N}_n : n \geq 1)$ in Bhattacharya et al. [2016]

and Bhattacharya et al. [2018] though formalization of the first three steps in the context of large deviation is a technically more challenging and needs more effort. For the sake of completeness and to keep the article self-contained, we shall describe the algorithms to prune the subtrees and regularize the pruned subtrees though these are used to derive the weak limit. The justification for the first three steps are given below without detailed proofs. The detailed proofs are given in Appendix.

Let $I(\mathbf{v})$ denote the the unique path from the root to vertex $\mathbf{v}$ for every $\mathbf{v} \in \mathbb{T}$. Define

$$\widetilde{m}_n(\cdot) = r_n \mathbf{P}^* \bigg(\sum_{|\mathbf{v}|=n} \sum_{\mathbf{u} \in I(\mathbf{v})} \boldsymbol{\delta}_{\gamma_n^{-1} X(\mathbf{u})} \in \cdot \bigg). \tag{3.2}$$

Note that displacements on the path $I(\mathbf{v})$ are independently and identically distributed with regularly varying tails of index α. As the first step, we shall use principle of a single big jump to conclude that at most one of the displacements on the path can be large enough to survive scaling by γ_n for every fixed path $I(\mathbf{v})$.

Lemma 3.1. *Under the assumptions stated in Theorem 2.4*

$$\lim_{n \to \infty} \left| m_n \Big(F_{g_1, g_2, \epsilon_1, \epsilon_2} \Big) - \widetilde{m}_n \Big(F_{g_1, g_2, \epsilon_1, \epsilon_2} \Big) \right| = 0. \tag{3.3}$$

Using this lemma, we can see that it is enough to compute the limit of

$$\widetilde{m}_n \Big(F_{g_1, g_2, \epsilon_1, \epsilon_2} \Big) = r_n \mathbf{E}^* \bigg[\prod_{i=1}^{2} \bigg(1 - \exp \bigg\{ - \Big(\sum_{|\mathbf{v}|=n} \sum_{\mathbf{u} \in I(\mathbf{v})} g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i \Big)_+ \bigg\} \bigg) \bigg]. \tag{3.4}$$

For this computation, we shall need more information about location of the large displacement. To be more specific, we shall show that large jump does not occur at the first $(n-K)$ generations with high probability for large enough $K < n$. This follows from the fact that total population size upto the $(n-K)$th generation is negligible with respect to the total population size upto the nth generation as $K \to \infty$. The next lemma formalizes this fact.

Lemma 3.2. *Under the assumptions sated in Theorem 2.4*

$$\lim_{K \to \infty} \lim_{n \to \infty} r_n \mathbf{P} \bigg[\sum_{|\mathbf{u}| \le n-K} \boldsymbol{\delta}_{\gamma_n^{-1} X(\mathbf{u})}(\theta, \infty) \ge 1 \bigg] = 0 \tag{3.5}$$

for every $\theta > 0$.

Let $\delta = \delta_1 \wedge \delta_2$. Then expression in (3.4) can be written as

$$r_n \mathbf{E}^* \Big[\prod_{i=1}^{2} \Big(1 - \exp\Big\{ - \Big(\sum_{|\mathbf{v}|=n} \sum_{\mathbf{u} \in I(\mathbf{v})} g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i \Big)_+ \Big\} \Big)$$
$$\mathbb{1}\Big(\sum_{|\mathbf{u}| \le n-K} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\delta, \infty) = 0 \Big) \Big] + r_n \mathbf{E}^* \Big[\prod_{i=1}^{2} \Big(1 - \exp\Big\{ - \Big(\sum_{|\mathbf{v}|=n} \sum_{\mathbf{u} \in I(\mathbf{v})}$$
$$g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i \Big)_+ \Big\} \Big) \mathbb{1}\Big(\sum_{|\mathbf{u}| \le n-K} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\delta, \infty) \ge 1 \Big) \Big]. \tag{3.6}$$

Let $I_K(\mathbf{v})$ denotes the subset of $I(\mathbf{v})$ containing the last K ancestors of $\mathbf{v}$ i.e. $I_K(\mathbf{v}) = \{\mathbf{u} \in I(\mathbf{v}) : |\mathbf{u} \to \mathbf{v}| \le K\}$. The the first term in (3.6) can be given following upper bound

$$r_n \mathbf{E}^* \Big[\prod_{i=1}^{2} \Big(1 - \exp\Big\{ - \Big(\sum_{|\mathbf{v}|=n} \sum_{\mathbf{u} \in I_K(\mathbf{v})} g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i \Big)_+ \Big\} \Big) \Big]. \tag{3.7}$$

It is easy to see that the term in (3.4) is larger than the expression in (3.7) and difference between these two terms is bounded by the second term in (3.6). Note that the product inside the expectation is bounded by 1 and so we get following upper bound for the second term in (3.6)

$$r_n \mathbf{P}^* \Big(\sum_{|\mathbf{v}| \le n-K} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{v})|}(\delta, \infty) \ge 1 \Big)$$

which converges to 0 as $n \to \infty$ and then $K \to \infty$.

So it is enough to compute limit of the expression in (3.7) as $n \to \infty$ and $K \to \infty$. Note that the expression does not involve any knowledge about the first $(n-K)$ generations of $\mathbb{T}$. So we cut the tree $\mathbb{T}$ in the $(n-K)$th generation and obtain a forest containing subtrees with K generations. The subtrees will be denoted by $(\mathbb{T}_i : i \ge 1)$. If $\mathbf{u} \in \cup_i \mathbb{T}_i$, then $A(\mathbf{u})$ denotes number of descendants of the particle $\mathbf{u}$ in the Kth generation of the subtree containing $\mathbf{u}$. With this notation, expression in (3.7) can be written as

$$r_n \mathbf{E}^* \Big[\prod_{i=1}^{2} \Big(1 - \exp\Big\{ - \Big(\sum_{j=1}^{Z_{n-K}} \sum_{\mathbf{u} \in \mathbb{T}_j} A(\mathbf{u}) g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i \Big)_+ \Big\} \Big) \Big]. \tag{3.8}$$

In the next step, we shall prune each of the subtrees obtained due to cutting. Fix an integer $B > 1$ large enough so that $\mu_B = \mathbf{E}(Z_1(B)) > 1$ where $Z_1(B) := Z_1 \mathbb{1}(Z_1 \le B) + B \mathbb{1}(Z_1 > B)$. The algorithm is given below:

P1. Start with the subtree $\mathbb{T}_1$ and consider its root.

P2. If the root has less than or equal to B children in the first generation, then do nothing. Otherwise, we keep the first B of them according to the Ulam-Harris labelling and delete extra children with their descendants.

P3. So we can have at most B children in the first generation. Repeat the step P2 for the children of the particles in the first generation. Follow the same algorithm until we reach to the children of particles in the $(K-1)$-th generation of the subtree $\mathbb{T}_1$.

P4. Repeat steps P2 and P3 for other subtrees.

Note that after pruning the subtrees in the forest, we delete some vertices with their lines of descendants without changing their genealogical structure. After pruning the j-th subtree will be denoted by $\mathbb{T}_j(B)$. Note that each of these subtrees has same distribution as that of a Galton-Watson tree with K generations and progeny distribution $Z_1(B)$. Let $\mathbf{u} \in \mathbb{T}_1(B)$, then number of descendants of a vertex $\mathbf{u}$ at the K-th generation of the subtree $\mathbb{T}_1(B)$ has been modified and will be denoted by $A^{(B)}(\mathbf{u})$. Following lemma formalizes that pruned subtrees contain large jump with high probability.

Lemma 3.3. *Under the assumptions given in Theorem 2.4, for every $K \geq 1$,*

$$\lim_{B\to\infty} \lim_{n\to\infty} r_n \Bigg| \mathbf{E}^* \Bigg[\prod_{i=1}^{2} \Bigg(1 - \exp\Bigg\{ - \Big(\sum_{j=1}^{Z_{n-K}} \sum_{\mathbf{u}\in\mathbb{T}_j} A(\mathbf{u}) g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i \Big)_+ \Bigg\} \Bigg) \Bigg] - \mathbf{E}^* \Bigg[\prod_{i=1}^{2} \Bigg(1 - \exp\Bigg\{ - \Big(\sum_{j=1}^{Z_{n-K}} \sum_{\mathbf{u}\in\mathbb{T}_j(B)} A^{(B)}(\mathbf{u}) g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i \Big)_+ \Bigg\} \Bigg) \Bigg] \Bigg| = 0. \tag{3.9}$$

So, it is enough to compute the limit of the expression

$$r_n \mathbf{E}^* \Bigg[\prod_{i=1}^{2} \Bigg(1 - \exp\Bigg\{ - \Big(\sum_{j=1}^{Z_{n-K}} \sum_{\mathbf{u}\in\mathbb{T}_j(B)} A^{(B)}(\mathbf{u}) g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i \Big)_+ \Bigg\} \Bigg) \Bigg]. \tag{3.10}$$

From this expression onwards, we shall give detailed proof of every step. Define

$$N_{t,n}^{(B)} = \sum_{\mathbf{u}\in\mathbb{T}_t(B)} A^{(B)}(\mathbf{u}) \boldsymbol{\delta}_{\gamma_n^{-1} X(\mathbf{u})}$$

for every $t \geq 1$. So we need to compute the limit of following expression

$$r_n \mathbf{E}^* \Bigg[\prod_{i=1}^{2} \Bigg(1 - \exp\Bigg\{ - \Big(\sum_{t=1}^{Z_{n-K}} N_{t,n}^{(B)}(g_i) - \epsilon_i \Big)_+ \Bigg\} \Bigg) \Bigg] \tag{3.11}$$

as $n \to \infty$, $B \to \infty$ and $K \to \infty$. Our first aim will be to get rid of the conditional expectation using an argument based on change of measure. We shall define S_{n-K} to be the set that at least one of Z_{n-K} subtrees does not extinct if $Z_{n-K} > 0$ otherwise S_{n-K} is empty. It is clear that $\mathcal{S} = S_{n-K} \cap \{Z_{n-K} > 0\}$ and so we get

$$d\mathbf{P}^* = \big(\mathbf{P}(\mathcal{S})\big)^{-1}\mathbb{1}(\mathcal{S})d\mathbf{P} = \big(\mathbf{P}(\mathcal{S})\big)^{-1}\mathbb{1}(Z_{n-K} > 0)\mathbb{1}(S_{n-K})d\mathbf{P}. \tag{3.12}$$

Using Radon-Nikodym derivative $\frac{d\mathbf{P}^*}{d\mathbf{P}}$ in (3.12), we can write down the expression in (3.11) as

$$r_n\big(\mathbf{P}(\mathcal{S})\big)^{-1}\mathbf{E}\bigg[\mathbb{1}(Z_{n-K} > 0)\prod_{i=1}^{2}\bigg(1 - \exp\bigg\{-\Big(\sum_{t=1}^{Z_{n-K}} N_{n,t}^{(B)}(g_i) - \epsilon_i\Big)_+\bigg\}\bigg)\bigg]$$
$$- r_n\big(\mathbf{P}(\mathcal{S})\big)^{-1}\mathbf{E}\bigg[\mathbb{1}(Z_{n-K} > 0)\mathbb{1}(S_{n-K}^c)\prod_{i=1}^{2}\bigg(1 - \exp\bigg\{-\Big(\sum_{t=1}^{Z_{n-K}} N_{(n,t)}^{(B)}(g_i) - \epsilon_i\Big)_+\bigg\}\bigg)\bigg]. \tag{3.13}$$

We shall now show that the second term in right hand side of (3.13) converges to 0 as $n \to \infty$.

Note that expression inside the expectation in the second term of (3.13) is positive and bounded by 1. Recall $\delta = \delta_1 \wedge \delta_2$ where $\delta_i = \sup\{|x| : g_i(x) > 0\}$ for $i = 1, 2$. It is clear that there must be at least one displacement $X(\mathbf{u})$ larger than $\gamma_n\delta/2$ in absolute value for $\mathbf{u} \in \cup_{t=1}^{Z_{n-K}}\mathbb{T}_t(B)$ to get a non-zero contribution from the product inside expectation. So we get following upper bound for the product inside expectation

$$\sum_{t=1}^{Z_{n-K}}\sum_{\mathbf{u}\in\mathbb{T}_t(B)}\boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\delta/2, \infty).$$

Using this upper bound and conditioning on $\mathcal{F}_{n-K}$, we get following upper bound for the second term in (3.13)

$$r_n\mathbf{E}\bigg[\mathbb{1}(Z_{n-K} > 0)\mathbb{1}(S_{n-K}^c)\sum_{t=1}^{Z_{n-K}}\sum_{\mathbf{u}\in\mathbb{T}_t(B)}\boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\delta/2, \infty)\bigg]$$
$$= r_n\mathbf{E}\bigg[\mathbb{1}(Z_{n-K} > 0)\sum_{t=1}^{Z_{n-K}}\mathbf{E}\bigg(\mathbb{1}(S_{n-K}^c)\sum_{\mathbf{u}\in\mathbb{T}_t(B)}\boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\delta/2, \infty)\bigg|\mathcal{F}_{n-K}\bigg)\bigg]. \tag{3.14}$$

Note that there can be at most $B + B^2 + \ldots + B^K = \frac{B^{K+1}-B}{B-1}$ displacements associated to each of the subtrees due to pruning step. Using the fact that branching mechanism and displacements are independently distributed and the fact sated above, we obtain following upper bound

$$\mu^{-K}\frac{\mathbf{P}(|X_1| > \gamma_n\delta/2)}{\mathbf{P}(|X_1| > \gamma_n)}\frac{B^{K+1} - B}{B - 1}\mathbf{E}\bigg[\mathbb{1}(Z_{n-K} > 0)\frac{Z_{n-K}}{\mu^{n-K}}p_e^{Z_{n-K}}\bigg]. \tag{3.15}$$

It is clear that the first and the third term does not involve n and finite in (3.15). The second term converges to a finite constant $(\delta/2)^{-\alpha}$ as $n \to \infty$. Now if we look at the expectation then the term inside expectation can be bounded by $\mu^{K-n} Z_{n-K}$ which converges to W in $\mathbb{L}^1$ (as a consequence of Kesten-Stigum condition given in (2.2)). Note that the term inside expectation converges to 0 almost surely as $n \to \infty$ ($Z_{n-K} \to \infty$ as $n \to \infty$ conditioned on $\{Z_{n-K} > 0\}$). So dominated convergence theorem applies and the expectation converges to 0 as $n \to \infty$ in (3.15).

In order to compute limit of the first term in (3.13), we shall regularize pruned subtrees according to following regularization algorithm.

1. Consider root of $\mathbb{T}_1(B)$.
2. The root can have at most B children in the next generation. If it has exactly B children, then keep it as it is. If it has l $(< B)$ children in the next generation, then we add $(B - l)$ children to it.
3. If $\mathbf{u}$ is a newly added children, then define $A^{(B)}(\mathbf{u}) = 0$, otherwise keep it as it is. Now, replace displacements of the children of the root by an independent copy of the random vector $(X_1, X_2, \ldots, X_B)$. The new displacement attached to $\mathbf{u}$ will be denoted by $X'(\mathbf{u})$.
4. Consider each particle in Generation 1 of $\mathbb{T}_1(B)$ and repeat the steps 2 and 3 until displacements attached to the children of particles in the $(K-1)$the generation are modified.
5. Repeat the steps 1,2,3 and 4 for the remaining subtrees.

After regularization, we obtain Z_{n-K} many B-ary subtrees which will be denoted by $(\widetilde{\mathbb{T}}_t(B) : 1 \le t \le Z_{n-K})$. Now we shall develop notations which will help to compute limit of the first term in (3.13). Each vertex $\mathbf{u} \in \widetilde{\mathbb{T}}_t(B)$ will be encoded by the triplet (t, l, s) where $\mathbf{u}$ is the sth vertex in the lth generation of the tth subtree. The collection of all displacements associated to the tree $\widetilde{\mathbb{T}}_t(B)$ is denoted by

$$\widetilde{X}_t = \big(X'(t,1,1), \ldots, X'(t,1,B), \ldots, X'(t,l,1), \ldots, X'(t,l,B^l), \ldots, X'(t,K,1), \ldots, X'(t,K,B^K)\big)$$

where $X'(t,l,s)$ denotes displacement attached to the (t,l,s)th vertex for every $1 \le s \le B^l$, $1 \le l \le K$ and $1 \le t \le Z_{n-K}$ conditioned on $\mathcal{F}_{n-K}$. It is clear that $\widetilde{X}_t$ is an $\tilde{R}_B$-valued random element where $\tilde{R}_B = \mathbb{R}^{B+B^2+\ldots+B^K}$. We also define

$$\begin{aligned}\widetilde{A}_t^{(B)} = \big(&A^{(B)}(t,1,1), \ldots, A^{(B)}(t,1,B), \ldots, A^{(B)}(t,l,1), \ldots, A^{(B)}(t,l,B^l), \\ &\ldots, A^{(B)}(t,K,1), \ldots, A^{(B)}(t,K,B^K)\big)\end{aligned}$$

for every $t \ge 1$. Note that $\widetilde{A}_t^{(B)}$ is an $\widetilde{S}^{(B)}$-valued random element where $\widetilde{S}^{(B)} = \prod_{l=1}^K \prod_{s=1}^{B^l} \{0, 1, 2, \ldots, B^{K-l}\} = \prod_{l=1}^K \prod_{s=1}^{B^l} [B^{K-l}]_0$ where $[B^{K-l}]_0 = \{0, 1, 2, \ldots, B^{K-l}\}$. Also note that, the random elements $(\widetilde{A}_t^{(B)} : 1 \le t \le Z_{n-K})$ are independently and identically distributed and also independent of $(\widetilde{X}_t : 1 \le t \le Z_{n-K})$ as branching mechanism and displacements are independent. Using Fact 2.1 and statement (iii) of Assumption 1.3, we obtain

$$\frac{\mathbf{P}\Big(\gamma_n^{-1}(X_1, X_2, \ldots, X_B) \in \cdot\Big)}{\mathbf{P}(|X_1| > \gamma_n)} \xrightarrow{HL} \lambda^{(B)}(\cdot) \tag{3.16}$$

on the space $\mathbb{R}^B \setminus \{\mathbf{0}\}$ where $\mathbf{0} \in \mathbb{R}$ denotes the origin of Euclidean space $\mathbb{R}^B$ and $\lambda^{(B)} = \lambda \circ \mathrm{PROJ}_B^{-1}$. Recall that PROJ_B is an operator on $\mathbb{R}^{\mathbb{N}}$ such that $\mathrm{PROJ}_B\Big((u_i : i \ge 1)\Big) = (u_1, u_2, \ldots u_B)$. Now, using the fact that displacements attached to the children coming from different parent are independent and (3.16), we obtain following convergence for joint distribution of the displacements associated to all vertices in the subtree $\widetilde{\mathbb{T}}_t(B)$

$$\frac{\mathbf{P}\Big(\gamma_n^{-1}\widetilde{X}_t \in \cdot\Big)}{\mathbf{P}\Big(|X_1| > \gamma_n\Big)} \xrightarrow{HL} \tau_t^{(B)}(\cdot) = \sum_{l=1}^{K} \sum_{s \in J_l} \tau_{t,l,s}^{(B)}(\cdot) \tag{3.17}$$

on the space $\tilde{R}_B \setminus \{\mathbf{0}\}$ where $\mathbf{0} \in \tilde{R}_B$ is the origin and

$$\tau_{t,l,s}^{(B)} = \bigotimes_{j=1}^{B+B^2+\ldots+B^{l-1}+s-1} \boldsymbol{\delta}_0 \otimes \lambda^{(B)} \bigotimes_{j'=B+B^2+\ldots+B^{l-1}+s+B}^{B+B^2+\ldots+B^K} \boldsymbol{\delta}_0 \tag{3.18}$$

for every $1 \le l \le K$ and $s \in J_l = \Big\{p \in \{1, 2, \ldots, B^l\} : p \equiv 1 \mod B\Big\}$ ($p \equiv 1 \mod B$ means p leaves remainder 1 when divided by B). Using (3.17) and the fact that $(\widetilde{X}_t : t \ge 1)$ are independently and identically distributed, we get that

$$\frac{\mathbf{P}\Big(\gamma_n^{-1}(\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_L) \in \cdot\Big)}{\mathbf{P}(|X_1| > \gamma_n)} \xrightarrow{HL} \sum_{j=1}^{L} \zeta_j^{(B)}(\cdot) \tag{3.19}$$

on $\tilde{R}_B^L \setminus \{\mathbf{0}\}$ where $\mathbf{0} \in \tilde{R}_B^L$ is the origin and

$$\zeta_j^{(B)} = \bigotimes_{j'=1}^{(j-1)(B+B^2+\ldots+B^K)} \boldsymbol{\delta}_0 \otimes \tau_j^{(B)} \bigotimes_{j''=j(B+B^2+\ldots+B^K)}^{L(B+B^2+\ldots+B^K)} \boldsymbol{\delta}_0 \tag{3.20}$$

for every finite $L \ge 1$. Also note that for every finite L, using the fact that conditioned on $\mathcal{F}_{n-K}$,

$(\widetilde{X}_t : 1 \le t \le L)$ and $(\widetilde{A}_t : 1 \le t \le L)$ are independent we get that

$$\begin{aligned}
&\frac{\mathbf{P}\Big((\widetilde{A}_1^{(B)}, \widetilde{A}_2^{(B)}, \dots, \widetilde{A}_L^{(B)}) \in \cdot, \gamma_n^{-1}(\widetilde{X}_1, \widetilde{X}_2, \dots, \widetilde{X}_L) \in \cdot\Big)}{\mathbf{P}(|X_1| > \gamma_n)} \\
&= \prod_{j=1}^{L} \mathbf{P}\Big(\widetilde{A}_j^{(B)} \in \cdot\Big) \frac{\mathbf{P}\Big(\gamma_n^{-1}(\widetilde{X}_1, \widetilde{X}_2, \dots, \widetilde{X}_L) \in \cdot\Big)}{\mathbf{P}(|X_1| > \gamma_n)} \\
&\xrightarrow{HL} \prod_{j=1}^{L} \mathbf{P}\Big(\widetilde{A}_j^{(B)} \in \cdot\Big) \sum_{j'=1}^{L} \zeta_j^{(B)}(\cdot)
\end{aligned} \tag{3.21}$$

on the space $\left(\widetilde{S}^B\right)^L \times \left(\tilde{R}_B^t \setminus \{\mathbf{0}\}\right)$ where $\left(\tilde{S}^{(B)}\right)^L = \prod_{i=1}^L \tilde{S}^{(B)}$.

Fix a large enough integer $L \in \mathbb{N}$. Then the first term in (3.13) can be written as

$$\begin{aligned}
r_n \big(\mathbf{P}(\mathcal{S})\big)^{-1} \mathbf{E}\Bigg[&\mathbb{1}(0 < Z_{n-K} \le L) \prod_{i=1}^{2} \Bigg(1 - \exp\Bigg\{ -\Big(\sum_{t=1}^{Z_{n-K}} \widetilde{N}_{n,t}^{(B)}(g_i) - \epsilon_i\Big)_+ \Bigg\}\Bigg) \\
&+ \mathbb{1}(Z_{n-K} > L) \prod_{i=1}^{2} \Bigg(1 - \exp\Bigg\{ -\Big(\sum_{t=1}^{Z_{n-K}} \widetilde{N}_{n,t}^{(B)}(g_i) - \epsilon\Big)_+ \Bigg\}\Bigg)\Bigg].
\end{aligned} \tag{3.22}$$

We shall show that the second expectation in (3.22) converges to 0 as $n \to \infty$ that is

$$\lim_{L\to\infty} \lim_{n\to\infty} r_n \mathbf{E}\Bigg[\mathbb{1}(Z_{n-K} > L) \prod_{i=1}^{2} \Bigg(1 - \exp\Bigg\{ -\Big(\sum_{t=1}^{Z_{n-K}} \widetilde{N}_{n,t}^{(B)}(g_i) - \epsilon_i\Big)_+ \Bigg\}\Bigg)\Bigg] = 0. \tag{3.23}$$

We shall first prove the claim in (3.23) and then compute the limit of the first term in (3.22).

Note that contribution of the product is non-zero if there is at least one displacement which is larger than $\gamma_n \delta/2$ in absolute value. So we can see that product inside the expectation in (3.23) is bounded by

$$\sum_{t=1}^{Z_{n-K}} \sum_{\mathbf{u} \in \widetilde{\mathbb{T}}_t^{(B)}} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\delta/2, \infty).$$

So we obtain following upper bound for the term in left hand side of (3.23)

$$\begin{aligned}
&r_n \mathbf{E}\Bigg[\mathbb{1}(Z_{n-K} > L) \sum_{t=1}^{Z_{n-K}} \sum_{\mathbf{u} \in \widetilde{\mathbb{T}}_t(B)} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\delta/2, \infty)\Bigg] \\
&= \mu^{-K} \frac{B^{K+1} - B}{B - 1} \mathbf{E}\Big[\frac{Z_{n-K}}{\mu^{n-K}} \mathbb{1}(Z_{n-K} > L)\Big] \frac{\mathbf{P}(|X_1| > \gamma_n \delta/2)}{\mathbf{P}(|X_1| > \gamma_n)}.
\end{aligned} \tag{3.24}$$

We used a conditioning on $\mathcal{F}_{n-K}$ and then unconditioning argument to derive the equality. We can see that the first two terms in (3.24) does not involve n and L. The last term converges to a finite positive constant $\big(\delta/2\big)^{-\alpha}$. So it is enough to show that

$$\lim_{L\to\infty}\lim_{n\to\infty}\mathbf{E}\bigg[\frac{Z_{n-K}}{\mu^{n-K}}\mathbb{1}(Z_{n-K}>L)\bigg]=0. \tag{3.25}$$

Under Kesten-Stigum condition, we have $\mu^{-n+K}Z_{n-K}$ converges to W in $\mathbb{L}^1$ and so we can see that the limit is finite. It is clear that $\mu^{-n+K}Z_{n-K}\mathbb{1}(Z_{n-K}>L)$ converges to 0 almost surely as $n\to\infty$ and $L\to\infty$. Using dominated convergence theorem, we can see that (3.25) holds.

It is easy to see that the first term in (3.22) can be written as

$$\mu^{-n}\big(\mathbf{P}(\mathcal{S})\big)^{-1}\sum_{i'=1}^{L}\mathbf{P}(Z_{n-K}=i')\sum_{\tilde a\in\left(\tilde S^{B}\right)^{i'}}G^{i'}(\tilde a)\int_{\tilde R_B^{i'}}\prod_{i=1}^{2}\bigg(1-\exp\bigg\{ -\Big(\sum_{t=1}^{i'}\sum_{\mathbf{u}\in\widetilde{\mathbb{T}}_t^{(B)}}a(\mathbf{u})g_i\big(x(\mathbf{u})\big)-\epsilon_i\Big)_+\bigg\}\bigg)\frac{\mathbf{P}\Big(\gamma_n^{-1}(\widetilde X_1,\widetilde X_2,\ldots,\widetilde X_j)\in d\tilde x\Big)}{\mathbf{P}(|X_1|>\gamma_n)} \tag{3.26}$$

where $G^{i'}(\tilde a)=\mathbf{P}\Big((\widetilde A_1^{(B)},\widetilde A_2^{(B)},\ldots,\widetilde A_{i'}^{(B)})=\tilde a\Big)$ for every $i'\in\{1,2,\ldots,L\}$. It is easy to see that

$$\prod_{i=1}^{2}\bigg(1-\exp\bigg\{-\Big(\sum_{t=1}^{i'}\sum_{l=1}^{K}\sum_{s=1}^{B^l}a(t,l,s)g_i\big(x(t,l,s)\big)-\epsilon_i\Big)_+\bigg\}\bigg)$$

is a bounded continuous function which vanishes in the neighbourhood of $\mathbf{0}\in\tilde R_B^{i'}$ for every $i'\in\{1,2,\ldots,L\}$. So using the HL convergence stated in (3.21), we get that

$$\lim_{n\to\infty}\sum_{\tilde a\in\left(\tilde S^{(B)}\right)^{i'}}G^{i'}(\tilde a)\int_{\tilde R_B^{i'}}\prod_{i=1}^{2}\bigg(1-\exp\bigg\{-\Big(\sum_{t=1}^{i'}\sum_{l=1}^{K}\sum_{s=1}^{B^l}a(t,l,s)g_i(x(t,l,s))-\epsilon_i\Big)_+\bigg\}\bigg)\frac{\mathbf{P}\Big(\gamma_n^{-1}(\widetilde X_1,\widetilde X_2,\ldots,\widetilde X_i)\in d\tilde x\Big)}{\mathbf{P}(|X_1|>\gamma_n)}=\sum_{\tilde a\in\left(\tilde S^{(B)}\right)^{i'}}G^{i'}(\tilde a)\int_{\tilde R_B^{i'}}\prod_{i=1}^{2}\bigg(1-\exp\bigg\{-\Big(\sum_{t=1}^{i'}\sum_{l=1}^{K}\sum_{s=1}^{B^l}a(t,l,s)g_i\big(x(t,l,s)\big)-\epsilon_i\Big)_+\bigg\}\bigg)\sum_{j=1}^{i'}\zeta_j^{(B)}(d\tilde x) \tag{3.27}$$

for every $i'\in\{1,2,\ldots,L\}$. Fix $\eta>0$. Then for large enough n (depending on η) we obtain following upper bound for the expression in (3.26)

$$\mu^{-n}\Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\sum_{i'=1}^{L}\mathbf{P}(Z_{n-K}=i')\sum_{\tilde{a}\in\left(\tilde{S}^{(B)}\right)^{i'}}G^{i'}(\tilde{a})\int_{\tilde{R}_B^{i'}}\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\sum_{t=1}^{i'}\sum_{l=1}^{K}\sum_{s=1}^{B^l}$$

$$\exp\Big\{-\Big(\sum_{t=1}^{i'}\sum_{l=1}^{K}\sum_{s=1}^{B^l}a(t,l,s)g_i\big(x(t,l,s)\big)-\epsilon_i\Big)_+\Big\}\sum_{j=1}^{i'}\zeta_j^{(B)}(d\tilde{x})$$

$$+\eta\mu^{-n}\sum_{i'=1}^{L}\mathbf{P}(Z_{n-K}=i'). \tag{3.28}$$

Note that the last term in right hand side of (3.28) is bounded by $\eta\mu^{-n}$ which converges to 0 as $n\to\infty$ uniformly in L. So it is enough to compute limit of the first term in (3.28) as $L\to\infty$ and $n\to\infty$. Using the fact that g_i's vanish in a neighbourhood of 0, we get following expression for right hand side of (3.28) can be written as

$$\mu^{-n}\Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\sum_{i'=1}^{L}\mathbf{P}(Z_{n-K}=i')\sum_{\tilde{a}\in\left(\tilde{S}^{(B)}\right)^{i'}}G^{i'}(\tilde{a})\sum_{j=1}^{i'}\int_{\tilde{R}_B}\prod_{i=1}^{2}\Big(1$$

$$-\exp\Big\{-\Big(\sum_{l=1}^{K}\sum_{s=1}^{B^l}a(j,l,s)g_i\big(x(j,l,s)\big)-\epsilon_i\Big)_+\Big\}\Big)\tau_j^{(B)}(d\tilde{x}) \tag{3.29}$$

using Fubini's theorem. Note that the product term inside integral depends only on $\widetilde{A}_j$ and $\widetilde{X}_j$ and independent of $\widetilde{A}_t$ for $t\in[i']\setminus\{j\}$. Combining above observation and the fact that $(\widetilde{A}_t:t\geq 1)$ are identically distributed, we get

$$\begin{aligned}&\sum_{\tilde{a}\in(\widetilde{S}^{(B)})^{i'}}G^{i'}(\tilde{a})\int_{\widetilde{R}_B}\prod_{i=1}^{2}\Big(1-\exp\{-\Big(\sum_{l=1}^{B}\sum_{s=1}^{B^l}a(j,l,s)g_i(x(j,l,s))-\epsilon_i\Big)_+\}\Big)\tau_j^{(B)}(d\tilde{x})\\&=\sum_{\tilde{a}\in\widetilde{S}^{(B)}}G(\tilde{a})\int_{\widetilde{R}(B)}\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\sum_{l=1}^{B}\sum_{s=1}^{B^l}a(1,l,s)g_i(x(1,l,s))-\epsilon_i\Big)_+\Big\}\Big)\tau_j^{(B)}(d\tilde{X})\end{aligned} \tag{3.30}$$

for every $j\in[i']$. Now using the fact that $\tau_j^{(B)}$ are same for every $j\geq 1$ and plugging in (3.30), we get following expression for (3.29)

$$\mu^{-n}\big(\mathbf{P}(\mathcal{S})\big)^{-1}\Big[\sum_{\tilde{a}\in\widetilde{S}^{(B)}}G(\tilde{a})\int_{\tilde{R}_B}\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\sum_{l=1}^{K}\sum_{s=1}^{B^l}a(1,l,s)g_i\big(x(1,l,s)\big)-\epsilon_i\Big)_+\Big\}\Big)$$

$$\tau_1^{(B)}(d\tilde{x})\Big] \sum_{i'=1}^{L} i' \mathbf{P}(Z_{n-K} = i'). \tag{3.31}$$

Hence we get that

$$\begin{aligned}
&\lim_{n\to\infty} \limsup_{L\to\infty} \mu^{-n} \Big(\mathbf{P}(\mathcal{S})\Big)^{-1} \mathbf{E}\Big[\mathbb{1}(0 < Z_{n-K} \le L) \prod_{i=1}^{2} \Big(1 - \exp\Big\{ - \Big(\sum_{t=1}^{Z_{n-K}} \\
&\qquad \sum_{\mathbf{u}\in\widetilde{\mathbb{T}}_t(B)} A^{(B)}(\mathbf{u}) g_i(\gamma_n^{-1} X(\mathbf{u})) - \epsilon_i\Big)_+ \Big\}\Big)\Big] \\
&\le \mu^{-K} \Big(\mathbf{P}(\mathcal{S})\Big)^{-1} \sum_{\tilde{a}\in\widetilde{S}^{(B)}} \mathbf{P}\Big(\widetilde{A}_1^{(B)} = \tilde{a}\Big) \int_{\widetilde{R}_B\setminus\mathbf{0}} \prod_{i=1}^{2} \Big(1 \\
&\qquad - \exp\Big\{ - \Big(\sum_{l=1}^{K} \sum_{s=1}^{B^l} a(1,l,s) g_i(x(1,l,s)) - \epsilon_i\Big)_+ \Big\}\Big) \tau_1^{(B)}(d\tilde{x})
\end{aligned} \tag{3.32}$$

letting $L \to \infty$ in (3.31).

It is clear that for large enough n and fixed $\eta > 0$, we get following lower bound using (3.27)

$$\begin{aligned}
&\sum_{\tilde{a}\in\left(\widetilde{S}^{(B)}\right)^{i'}} G^{i'}(\tilde{a}) \int_{\widetilde{R}_B^{i'}} \prod_{i=1}^{2} \Big(1 - \exp\Big\{ - \Big(\sum_{t=1}^{i'} \sum_{l=1}^{K} \sum_{s=1}^{B^l} a(1,l,s) g_i\big(x(1,l,s)\big) - \epsilon_i\Big)_+ \Big\}\Big) \\
&\qquad \frac{\mathbf{P}\Big(\gamma_n^{-1}(\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_{i'}) \in d\tilde{x}\Big)}{\mathbf{P}(|X|_1 > \gamma_n)} \\
&\ge \sum_{\tilde{a}\in\left(\widetilde{S}^{(B)}\right)^{i'}} G^{i'}(\tilde{a}) \int_{\widetilde{R}_B^{i'}} \prod_{i=1}^{2} \Big(1 - \exp\Big\{ - \Big(\sum_{t=1}^{i'} \sum_{l=1}^{K} \sum_{s=1}^{B^l} a(t,l,s) g_i(x(t,l,s)) - \epsilon_i\Big)_+ \Big\}\Big) \sum_{j=1}^{i'} \zeta_j^{(B)}(d\tilde{x}) - \eta.
\end{aligned}$$

Using this lower bound and same arguments as above, it can be easily obtained

$$\begin{aligned}
&\lim_{n\to\infty} \liminf_{L\to\infty} r_n \Big(\mathbf{P}(\mathcal{S})\Big)^{-1} \mathbf{E}\Big[\mathbb{1}(0 < Z_{n-K} \le L) \prod_{i=1}^{2} \Big(1 - \exp\Big\{ - \Big(\widetilde{N}_{t,n}^{(B)}(g_i) - \epsilon_i\Big)_+ \Big\}\Big)\Big] \\
&\ge \mu^{-K} \Big(\mathbf{P}(\mathcal{S})\Big)^{-1} \sum_{\tilde{a}\in\widetilde{S}^{(B)}} G(\tilde{a}) \int_{\widetilde{R}_B\setminus\mathbf{0}} \prod_{i=1}^{2} \Big(1 \\
&\qquad - \exp\Big\{ - \Big(\sum_{l=1}^{K} \sum_{s=1}^{B^l} a(1,l,s) g_i(x(1,,l,s)) - \epsilon\Big)_+ \Big\}\Big) \tau_1^{(B)}(d\tilde{x}).
\end{aligned} \tag{3.33}$$

Combining (3.32) and (3.33), we obtain

$$\lim_{n\to\infty} r_n\Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\mathbf{E}\Big[\mathbb{1}(Z_{n-K}>0)\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\sum_{t=1}^{Z_{n-K}}\widetilde{N}_{n,t}^{(B)}(g_i)-\epsilon_i\Big)_+\Big\}\Big)\Big]$$

$$=\mu^{-K}\Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\sum_{\tilde{a}\in\widetilde{S}^{(B)}}G(\tilde{a})\int_{\widetilde{R}_B\setminus\mathbf{0}}\prod_{i=1}^{2}\Big(1-\exp\Big\{$$

$$-\Big(\sum_{l=1}^{K}\sum_{s=1}^{B^l}a(1,l,s)g_i(x(1,,l,s))-\epsilon\Big)_+\Big\}\Big)\tau_1^{(B)}(d\tilde{x}). \tag{3.34}$$

Note that the expression derived in (3.34) is same as the expression obtained in (4.21) in Bhattacharya et al. [2018] except the first two terms. Also note that we can use arguments given in aforementioned reference (based on the properties of GW process) to simplify the expression and write down the sum in terms of expectation of underlying GW process (see (4.23) in aforementioned reference). To write down the simplified expression, we need the following notation.

- Suppose that $U^{(B)}$ is an independent copy of $Z_1^{(B)}$. Let $\widetilde{U}^{(B)}$ denotes the random variable $U^{(B)}$ conditioned to stay positive i.e. $\mathbf{P}\Big(\widetilde{U}^{(B)}\in A\Big)=\mathbf{P}\Big(Z_1^{(B)}\in A|Z_1^{(B)}>0\Big)$ for every $A\subset\mathbb{N}$.
- Let $(Z_l^{(B)}:l\geq 1)$ denotes the GW process with progeny distribution $Z_1^{(B)}$. Let $(Z_l^{(s,B)}:s\geq 1)$ be a collection of independent copies of $Z_l^{(B)}$ for every $l\geq 1$. $(\widetilde{Z}_l^{(s,B)}:s\geq 1)$ denotes the collection of independent copies of $\widetilde{Z}_l^{(B)}$ where $\widetilde{Z}_l^{(B)}$ denotes the random variable $Z_l^{(B)}$ conditioned to stay positive i.e. $\mathbf{P}(\widetilde{Z}_l^{(B)}\in A)=\mathbf{P}(Z_l^{(B)}\in A|Z_l^{(B)}>0)$ for every $A\subset\mathbb{N}$.
- Pow $\big(G\big)$ denotes the power set of G i.e. the collection of all possible subsets of G.

It is clear from equation (4.23) in Bhattacharya et al. [2018], the expression in (3.34) can be written as

$$\mu^{-K}\Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\sum_{l=1}^{K}\mu_B^{l-1}\int_{\mathbb{R}^B\setminus\{\mathbf{0}\}}\mathbf{E}\Big[\sum_{G\in\text{Pow}([\widetilde{U}^{(B)}])\setminus\{\emptyset\}}\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\sum_{s\in G}\widetilde{Z}_{K-l}^{(s,B)}g_i(x_s)-\epsilon_i\Big)_+\Big\}\Big)$$

$$\Big(\mathbf{P}(Z_{K-l}^{(B)}>0)\Big)^{|G|}\Big(\mathbf{P}(Z_{K-l}^{(B)}=0)\Big)^{\widetilde{U}^{(B)}-|G|}\Big]\lambda^{(B)}(d\tilde{x})$$

$$=\int_{\mathscr{M}_0}\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\nu(g_i)-\epsilon_i\Big)_+\Big\}\Big)m_{K,B}^*(d\nu) \tag{3.35}$$

where $\mathbf{P}(\mathcal{S}) = 1 - p_e$ and $m^*_{K,B}$ is a measure on $\mathscr{M}_0$ defined as

$$m^*_{K,B}(\cdot) = (1-p_e)^{-1}\mathbf{P}\big(U^{(B)} > 0\big)(\mu_B\mu^{-1})^K \sum_{l=1}^{K} \mu_B^{l-K-1}\mathbf{E}\Big[\sum_{G\in \mathrm{Pow}([\widetilde{U}^{(B)}])\setminus\{\emptyset\}} \lambda^{(B)}\Big(\mathbf{x}\in\mathbb{R}^B : \sum_{s\in G}\widetilde{Z}^{(s,B)}_{K-l}\boldsymbol{\delta}_x \in \cdot\Big)\Big(\mathbf{P}(Z^{(B)}_{K-l} > 0)\Big)^{|G|}\Big(\mathbf{P}(Z^{(B)}_{K-l} = 0)\Big)^{\widetilde{U}^{(B)}-|G|}\Big]. \tag{3.36}$$

Again using the same argument given in the last paragraph in Page 18 of Bhattacharya et al. [2018], we get that $m^*_{K,B}(B_r) < \infty$ where $B_r = \{\nu \in \mathscr{M}_0 : d_{\text{vague}}(\nu, \emptyset) > r\}$ is the subset of $\mathscr{M}_0$ which is bounded away from $\emptyset$. It is easy to see that μ_B converges to μ, $U^{(B)}$ converges almost surely to U which is an independent copy of Z_1 and $Z_l^{(s,B)}$ converges to $Z_l^{(s)}$ almost surely as $B \to \infty$ (recall the notations introduced before Theorem 2.4). As the integrand in (3.35) is bounded by 1, we get that expression derived in (3.35) converges to

$$\begin{aligned}&(1-p_e)^{-1}\mathbf{P}(U>0)\sum_{l=1}^{K}\mu^{l-K-1}\int_{\mathbb{R}^{\mathbb{N}}\setminus\{\mathbf{0}\}}\mathbf{E}\Big[\sum_{G\in\mathrm{Pow}([\widetilde{U}])\setminus\{\emptyset\}}\prod_{i=1}^{2}\Big(1-\exp\Big\{\\&\qquad -\Big(\sum_{s\in G}\widetilde{Z}^{(s)}_{K-l}g_i(x_s)-\epsilon_i\Big)_+\Big\}\Big)\Big(\mathbf{P}(Z_{K-l}>0)\Big)^{|G|}\Big(\mathbf{P}(Z_{K-l}=0)\Big)^{\widetilde{U}-|G|}\Big]\\&=\int_{\mathscr{M}_0}\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\nu(g_i)-\epsilon_i\Big)_+\Big\}\Big)m^*_K(d\nu)\end{aligned} \tag{3.37}$$

where

$$\begin{aligned}m^*_K(\cdot) &= (1-p_e)^{-1}\mathbf{P}(U>0)\sum_{l=1}^{K}\mu^{l-K-1}\mathbf{E}\Big[\sum_{G\in\mathrm{Pow}([\widetilde{U}])\setminus\{0\}}\lambda\Big(\mathbf{x}\in\mathbb{R}^{\mathbb{N}}:\sum_{s\in G}\widetilde{Z}^{(s)}_{K-l}\boldsymbol{\delta}_x\in\cdot\Big)\\&\qquad\Big(\mathbf{P}(Z_{K-l}>0)\Big)^{|G|}\Big(\mathbf{P}(Z_{K-l}=0)\Big)^{\widetilde{U}-|G|}\Big]\\&=(1-p_e)^{-1}\mathbf{P}(U>0)\sum_{l=0}^{K-1}\mu^{-l-1}\mathbf{E}\Big[\sum_{G\in\mathrm{Pow}([\widetilde{U}])\setminus\{0\}}\lambda\Big(\mathbf{x}\in\mathbb{R}^{\mathbb{N}}:\sum_{s\in G}\widetilde{Z}^{(s)}_{l}\boldsymbol{\delta}_x\in\cdot\Big)\\&\qquad\Big(\mathbf{P}(Z_l>0)\Big)^{|G|}\Big(\mathbf{P}(Z_l=0)\Big)^{\widetilde{U}-|G|}\Big]\end{aligned} \tag{3.38}$$

as $B \to \infty$ using dominated convergence theorem and Theorem 4.1 in Lindskog et al. [2014]. It can be shown that $m^*_K \in \mathbb{M}_0$. Now letting $K \to \infty$, we get that the expression converges to

$$\int_{\mathscr{M}_0}\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\nu(g_i)-\epsilon_i\Big)_+\Big\}\Big)m^*(d\nu) \tag{3.39}$$

using the form of m_K^* derived in (3.38) and dominated convergence theorem where m^* is given in Theorem 2.4.

4 Proof of corollaries

In this section, we shall give detailed proof of the corollaries 2.6 and 2.7.

4.1 Proof of Corollary 2.6

Note that

$$\begin{aligned}
&\lim_{n\to\infty} r_n \mathbf{P}^*\Big(\max_{|\mathbf{v}|=n} S(\mathbf{v}) > \gamma_n x\Big)\\
&= \lim_{n\to\infty} r_n \mathbf{P}^*\Big(N_n(x,\infty) \ge 1\Big)\\
&= \lim_{n\to\infty} r_n \mathbf{P}^*\Big(N_n \in \{\xi : \xi(x,\infty) \ge 1\}\Big)\\
&= m^*\Big(\{\xi : \xi(x,\infty) \ge 1\}\Big) \qquad (4.1)\\
&= (1-p_e)^{-1}\mathbf{P}(U>0)\sum_{l=0}^{\infty}\mu^{-(l+1)}\mathbf{E}\Big[\sum_{G\in \mathrm{Pow}([\widetilde{U}])\setminus\{\emptyset\}} \lambda\Big(\mathbf{y}\in\mathbb{R}^{\mathbb{N}} : \sum_{s\in G} \widetilde{Z}_l^{(s)} \boldsymbol{\delta}_{y_s}(x,\infty) \ge 1\Big)\\
&\qquad \Big(\mathbf{P}(Z_l>0)\Big)^{|G|}\Big(\mathbf{P}(Z_l=0)\Big)^{\widetilde{U}-|G|}\Big]\\
&= x^{-\alpha}(1-p_e)^{-1}\mathbf{P}(U>0)\sum_{l=0}^{\infty}\mu^{-(l+1)}\mathbf{E}\Big[\sum_{G\in \mathrm{Pow}([\widetilde{U}])\setminus\{\emptyset\}} \lambda\Big(\mathbf{y}\in\mathbb{R}^{\mathbb{N}} : \sum_{s\in G} \boldsymbol{\delta}_{y_s}(1,\infty) \ge 1\Big)\\
&\qquad \Big(\mathbf{P}(Z_l>0)\Big)^{|G|}\Big(\mathbf{P}(Z_l=0)\Big)^{\widetilde{U}-|G|}\Big] \qquad (4.2)
\end{aligned}$$

using the scaling relation obeyed by the measure λ on $\mathbb{R}^{\mathbb{N}}$. Note that $\Big\{\sum_{s\in G}\boldsymbol{\delta}_{y_s}\in(1,\infty)\ge 1\Big\} = \Big\{y \in \cup_{s\in G} V_s\Big\}$ and using this in (4.2), we obtain (2.9). Here we applied Theorem 2.4 to obtain the expression in right hand side of (4.1). In order to apply the theorem, we need to show that $\mathscr{A} = \{\xi \in \mathscr{M}_0 : \xi(x,\infty) \ge 1\}$ is bounded away from the null measure and $m^*(\partial\mathscr{A}) = 0$ which we skipped. Following the same arguments given in Corollary 5.1 of Hult and Samorodnitsky [2010], we can see that $\mathscr{A}^c = \{\xi : \xi(x,\infty) = 0\}$ is a closed set containing the null measure. So the $\mathscr{A}$ is bounded away from null measure. Now we need to show that $m^*(\partial\mathscr{A}) = 0$. Note that in the same corollary of the aforementioned reference, it has been show that $\widetilde{\mathscr{A}} = \{\xi : \xi[x,\infty) \ge 1\}$ is a closed set containing $\mathscr{A}$. So we get that

$$m^*(\partial\mathscr{A}) \le m^*\big(\widetilde{\mathscr{A}}\big) - m^*\big(\mathscr{A}\big)$$

$$= x^{-\alpha}(1-p_e)^{-1}\mathbf{P}(U>0)\sum_{l=0}^{\infty}\mu^{-(l+1)}\mathbf{E}\Bigg[\sum_{G\in \mathrm{Pow}([\widetilde{U}])\setminus\{0\}}\Big(\lambda\big(\cup_{s\in G}\bar{V}_s\big)$$
$$-\lambda\big(\cup_{s\in G}V_s\big)\Big)\Big(\mathbf{P}(Z_l>0)\Big)^{|G|}\Big(\mathbf{P}(Z_l=0)\Big)^{\widetilde{U}-|G|}\Bigg] \tag{4.3}$$

where $\pi_s\Big(\bar{V}_s\Big)=[1,\infty)$ and $\pi_i\Big(\bar{V}_s\Big)=\mathbb{R}$ for every $i\in\mathbb{N}\setminus\{s\}$ for every $s\in\mathbb{N}$. Our aim is to show that for every G, the difference inside the sum is 0. Recall that $\lambda\circ\pi^{-1}_{s_1,s_2,\dots,s_k}$ also satisfies the scaling relation as λ satisfies where $\pi_{s_1,s_2,\dots,s_k}(\mathbf{u})=(u_{s_1},u_{s_2},\dots,u_{s_k})$ for every $(u)=(u_i:i\geq 1)\in\mathbb{R}^{\mathbb{N}}$ for every $k\in\mathbb{N}$. Note that

$$\begin{aligned}&\lambda\Big(\cup_{s\in G}\bar{V}_s\Big)-\lambda\Big(\cup_{s\in G}V_s\Big)\\&=\sum_{s\in G}\lambda\circ\pi_s^{-1}(\{1\})-\sum_{s_1,s_2}\lambda\circ\pi^{-1}_{s_1,s_2}\big(\{1\}\times\{1\}\big)+\dots\\&\qquad+(-1)^{|G|}\lambda\circ\pi^{-1}_{s_1,s_2,\dots,s_{|G|}}\big(\{1\}\times\{1\}\times\dots\times\{1\}\big).\end{aligned}\tag{4.4}$$

Now we can see that for every $(s_1,s_2,\dots,s_k)\in\mathbb{N}^k$,

$$\begin{aligned}&\lambda\circ\pi^{-1}_{s_1,s_2,\dots,s_k}\big(\{1\}\times\{1\}\times\dots\times\{1\}\big)\\&=\lim_{n\to\infty}\Big(\lambda\circ\pi^{-1}_{s_1,\dots,s_k}\big((1-n^{-1},\infty)\times(1-n^{-1},\infty)\times\dots\times(1-n^{-1},\infty)\big)\\&\qquad-\lambda\circ\pi^{-1}_{s_1,s_2,\dots,s_k}\big((1,\infty)\times(1,\infty)\times\dots\times(1,\infty)\big)\Big)\\&=\lim_{n\to\infty}\Big((1-n^{-1})^{-\alpha}-1\Big)\lambda\circ\pi^{-1}_{s_1,s_2,\dots,s_k}\big((1,\infty)\times(1,\infty)\times\dots\times(1,\infty)\big)\\&=0\end{aligned}\tag{4.5}$$

as $\lambda\circ\pi^{-1}_{s_1,s_2,\dots,s_k}\big((1,\infty)\times(1,\infty)\times\dots\times(1,\infty)\big)<\infty$ (bounded away from $\mathbf{0}=(0,0,\dots,0)\in\mathbb{R}^k$). Using (4.5), we can see that right hand side of (4.4) is 0 for every finite set G. As $\widetilde{U}_1$ is finite almost surely, so we get that right hand side of (4.3) is 0.

4.2 Proof of Corollary 2.7

Recall that

$$\lambda_{iid}(\cdot)=\sum_{t=1}^{\infty}\otimes_{i=1}^{t-1}\boldsymbol{\delta}_0\otimes\nu_\alpha(\cdot)\otimes_{i=t+1}^{\infty}\boldsymbol{\delta}_0.$$

Then we obtain following expression for the $m^*\Big(F_{g_1,g_2,\epsilon_1,\epsilon_2}\Big)$

$$(1-p_e)^{-1}\mathbf{P}(U>0)\sum_{l=1}^{\infty}\mu^{-(l+1)}\mathbf{E}\Big[\sum_{G\in\text{Pow}([\widetilde{U}])\setminus\{\emptyset\}}\Big(\mathbf{P}(Z_l>0)\Big)^{|G|}\Big(\mathbf{P}(Z_l=0)\Big)^{\widetilde{U}-|G|}$$
$$\mathbf{E}\Big(\sum_{s\in G}\int\prod_{i=1}^{2}\Big(1-\exp\Big\{-\big(\widetilde{Z}_l^{(s)}g_i(x)-\epsilon_i\big)_+\Big\}\Big)\nu_\alpha(dx)\Big)\Big]. \tag{4.6}$$

To obtain this expression, we used the fact that $\Big(\widetilde{Z}_l^{(s)}:s\geq 1\Big)$ are independent of U and so of $\widetilde{U}$. Now we shall use that fact that $(\widetilde{Z}_l^{(s)}:s\geq 1)$ are independent copies of random variables $\widetilde{Z}_l$ for every $l\geq 1$ to obtain following expression for right hand side of (4.6) as

$$(1-p_e)^{-1}\mathbf{P}(U>0)\sum_{l=1}^{\infty}\mu^{-(l+1)}\Bigg[\mathbf{E}\Big[\int\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\widetilde{Z}_lg_i(x)-\epsilon_i\Big)_+\Big\}\Big)\nu_\alpha(dx)\Big]\Bigg]$$
$$\mathbf{E}\Big[\sum_{|G|=1}^{\widetilde{U}}|G|\binom{\widetilde{U}}{|G|}\Big(\mathbf{P}(Z_l>0)\Big)^{|G|}\Big(\mathbf{P}(Z_l=0)\Big)^{\widetilde{U}-|G|}\Big]$$
$$=(1-p_e)^{-1}\mathbf{P}(U>0)\mathbf{E}[\widetilde{U}]\sum_{l=1}^{\infty}\mu^{-(l+1)}\mathbf{P}(Z_l>0)\mathbf{E}\Big[\int\prod_{i=1}^{2}$$
$$\Big(1-\exp\Big\{-\Big(\widetilde{Z}_lg_i(x)-\epsilon_i\Big)_+\Big\}\Big)\nu_\alpha(dx)\Big]. \tag{4.7}$$

Note that $\mathbf{E}(\widetilde{U})=\mu\big(\mathbf{P}(U_1>0)\big)^{-1}$. So we get following expression for right hand side of (4.7)

$$(1-p_e)^{-1}\sum_{l=1}^{\infty}\mu^{-l}\mathbf{P}(Z_l>0)\int\mathbf{E}\Big[\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\widetilde{Z}_lg_i(s)-\epsilon_i\Big)_+\Big\}\Big)\Big]\nu_\alpha(dx). \tag{4.8}$$

It is easy to see that right hand side of (4.8) is same as $m^*_{iid}(F_{g_1,g_2,\epsilon_1,\epsilon_2})$ where m^*_{iid} is given in Corollary 2.7. Hence we conclude the proof.

Acknowledgement

The research is supported by an NWO VICI grant. The author is thankful to Rajat Subhra Hazra, Zbigniew Palmowski, Parthanil Roy, Gennady Samorodnitsky and Bert Zwart for many helpful discussions and also would like to acknowledge the warm hospitality of ISI, Bangalore for the period 15 - 26th March, 2017 where the project started. A part of the research was done when the author was a visiting scientist in ISI, Kolkata for the period May 22 - August 31, 2017.

References

L. Addario-Berry and B. Reed. Minima in branching random walks. *The Annals of Probability*, pages 1044–1079, 2009.

E. Aïdékon. Convergence in law of the minimum of a branching random walk. *Ann. Probab.*, 41(3A):1362–1426, 2013. ISSN 0091-1798. doi: 10.1214/12-AOP750. URL http://dx.doi.org/10.1214/12-AOP750.

M. Bachmann. Limit theorems for the minimal position in a branching random walk with independent logconcave displacements. *Advances in Applied Probability*, 32(1):159–176, 2000.

J. Bérard and P. Maillard. The limiting process of N-particle branching random walk with polynomial tails. *Electron. J. Probab.*, 19:no. 22, 17, 2014. ISSN 1083-6489. doi: 10.1214/EJP.v19-3111. URL http://dx.doi.org/10.1214/EJP.v19-3111.

A. Bhattacharya, R. S. Hazra, and P. Roy. Point process convergence for branching random walks with regularly varying steps. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques,* to appear. arXiv:1411.5646, 2016.

A. Bhattacharya, R. S. Hazra, and P. Roy. Branching random walks, stable point processes and regular variation. *Stochastic Processes and their Applications*, 128(1):182 – 210, 2018. ISSN 0304-4149. doi: https://doi.org/10.1016/j.spa.2017.04.009. URL http://www.sciencedirect.com/science/article/pii/S0304414917301230.

J. D. Biggins. The first- and last-birth problems for a multitype age-dependent branching process. *Advances in Appl. Probability*, 8(3):446–459, 1976. ISSN 0001-8678.

M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.*, 69(1):62–123, 2016a. ISSN 0010-3640. doi: 10.1002/cpa.21621. URL http://dx.doi.org/10.1002/cpa.21621.

M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of nonlattice branching random walk. 52(4):1897–1924, 2016b.

B. Chauvin and A. Rouault. Kpp equation and supercritical branching brownian motion in the subcritical speed area. application to spatial trees. *Probability theory and related fields*, 80(2): 299–314, 1988.

X. Chen and H. He. On large deviation probabilities for empirical distribution of branching random walks: Schr {\" o} der case and b {\" o} ttcher case. *arXiv preprint arXiv:1704.03776*, 2017.

B. Derrida and Z. Shi. Large deviations for the branching brownian motion in presence of selection or coalescence. *Journal of Statistical Physics*, 163(6):1285–1311, 2016.

J. Ding, R. Roy, and O. Zeitouni. Convergence of the centered maximum of log-correlated gaussian fields. *The Annals of Probability*, 45(6A):3886–3928, 2017.

R. Durrett. Maxima of branching random walks vs. independent random walks. *Stochastic Process. Appl.*, 9(2):117–135, 1979. ISSN 0304-4149. doi: 10.1016/0304-4149(79)90024-3. URL http://dx.doi.org/10.1016/0304-4149(79)90024-3.

R. Durrett. Maxima of branching random walks. *Z. Wahrsch. Verw. Gebiete*, 62(2):165–170, 1983. ISSN 0044-3719. doi: 10.1007/BF00538794. URL http://dx.doi.org/10.1007/BF00538794.

C. Dutta, G. Pandurangan, R. Rajaraman, and S. Roche. Coalescing-branching random walks on graphs. *ACM Transactions on Parallel Computing*, 2(3):20, 2015.

V. Fasen and P. Roy. Stable random fields, point processes and large deviations. *Stochastic Process. Appl.*, 126(3):832–856, 2016. ISSN 0304-4149. doi: 10.1016/j.spa.2015.09.020. URL http://dx.doi.org/10.1016/j.spa.2015.09.020.

N. Gantert. The maximum of a branching random walk with semiexponential increments. *Annals of probability*, pages 1219–1229, 2000.

N. Gantert and T. Höfelsauer. Large deviations for the maximum of a branching random walk. *ArXiv e-prints*, Feb. 2018. URL https://arxiv.org/abs/1802.03960.

J. M. Hammersley. Postulates for subadditive processes. *Ann. Probability*, 2:652–680, 1974.

H. Hult and F. Lindskog. Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)*, 80(94):121–140, 2006. ISSN 0350-1302. doi: 10.2298/PIM0694121H. URL http://dx.doi.org/10.2298/PIM0694121H.

H. Hult and G. Samorodnitsky. Large deviations for point processes based on stationary sequences with heavy tails. *J. Appl. Probab.*, 47(1):1–40, 2010. ISSN 0021-9002. doi: 10.1239/jap/1269610814. URL http://dx.doi.org/10.1239/jap/1269610814.

J. F. C. Kingman. The first birth problem for an age-dependent branching process. *Ann. Probability*, 3(5):790–801, 1975.

F. Lindskog, S. I. Resnick, and J. Roy. Regularly varying measures on metric spaces: hidden regular variation and hidden jumps. *Probab. Surv.*, 11:270–314, 2014. ISSN 1549-5787. doi: 10.1214/14-PS231. URL http://dx.doi.org/10.1214/14-PS231.

O. Louidor and W. Perkins. Large deviations for the empirical distribution in the branching random walk. *Electronic Journal of Probability*, 20, 2015.

O. Louidor and E. Tsairi. Large deviations for the empirical distribution in the general branching random walk. *arXiv preprint arXiv:1704.02374*, 2017.

T. Madaule. Convergence in law for the branching random walk seen from its tip. *J. Theoret. Probab.*, 2015. ISSN 0894-9840. doi: 10.1007/s10959-015-0636-6. URL http://dx.doi.org/10.1007/s10959-015-0636-6. To appear.

P. Maillard. The maximum of a tree-indexed random walk in the big jump domain. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(2):545–561, 2016. ISSN 1980-0436.

S. I. Resnick. *Extreme values, regular variation, and point processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987. ISBN 0-387-96481-9. doi: 10.1007/978-0-387-75953-1. URL http://dx.doi.org/10.1007/978-0-387-75953-1.

S. I. Resnick and J. Roy. Hidden regular variation of moving average processes with heavy-tailed innovations. *Journal of Applied Probability*, 51(A):267–279, 2014.

Z. Shi. *Branching Random Walks: École d'Été de Probabilités de Saint-Flour XLII–2012*, volume 2151. Springer, 2016.

Centrum Wiskunde & Informatica
P.O. Box 94079
1090 GB Amsterdam, Netherlands
eamil - ayan.bhattacharya@cwi.nl

Appendix

Here we shall present the proofs of Lemma 3.1, 3.2 and 3.3. Following inequality will turn out to be very useful for the proofs given later. It is clear that

$$\begin{aligned}\Big|\mathbf{E}^*(R) - \mathbf{E}^*(R')\Big| &= \Big|\mathbf{E}\Big(\mathbb{1}(\mathcal{S})(\mathbf{P}(\mathcal{S}))^{-1}R\Big) - \mathbf{E}\Big(\mathbb{1}(\mathcal{S})(\mathbf{P}(\mathcal{S}))^{-1}R'\Big)\Big| \\ &= \Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\Big|\mathbf{E}\Big(\mathbb{1}(\mathcal{S})R\Big)\mathbf{E}\Big(\mathbb{1}(\mathcal{S}R')\Big)\Big| \\ &\le \Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\mathbf{E}\Big(\mathbb{1}(\mathcal{S})|R-R'|\Big) \\ &\le \Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\mathbf{E}\Big(|R-R'|\Big) \qquad (4.9)\end{aligned}$$

for any pair of random variables R and R'.

4.3 Proof of Lemma 3.1

The key idea behind the proof of the lemma is that at most one of the displacement can survive the scaling among displacements on a typical path from the root to the vertex $\mathbf{v}$ in the n-th generation. To formalize this idea, we need following notations. Define

$$AMO_n = \Big(\bigcup_{|\mathbf{v}|=n} \Big(\sum_{\mathbf{u}\in\mathbf{v}} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(n^{-3},\infty) \geq 2 \Big) \Big)^c.$$

Our aim is to show that

$$\lim_{n\to\infty} r_n \mathbf{P}(AMO_n^c) = 0. \tag{4.10}$$

It is easy to see that

$$\begin{aligned} r_n \mathbf{P}(AMO_n^c) &= r_n \mathbf{E}\Big[\mathbf{P}\Big(\bigcup_{|\mathbf{v}|=n} \Big(\sum_{\mathbf{u}\in I(\mathbf{v})} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(n^{-3},\infty) \geq 2 \Big) \Big| \mathcal{F}_n \Big) \Big] \\ &\leq r_n \mathbf{E}\Big[\sum_{|\mathbf{v}|=n} \mathbf{P}\Big(\sum_{\mathbf{u}\in I(\mathbf{v})} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(n^{-3},\infty) \geq 2 \Big) \Big]. \end{aligned} \tag{4.11}$$

Note that for every vertex $\mathbf{v}$ in the n-th generation, the number of displacements on the path from the root to the vertex $\mathbf{v}$ is n and these n displacements are independently (as coming from different parents) and identically distributed. Let $(X_1^{(i)} : i \geq 1)$ denote the collection of independent copies of X_1. Then for every n,

$$\begin{aligned} \mathbf{P}\Big(\sum_{\mathbf{u}\in I(\mathbf{v})} \boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(n^{-3},\infty) \geq 2 \Big) &= \mathbf{P}\Big(\sum_{i=1}^{n} \boldsymbol{\delta}_{\gamma_n^{-1}|X_1^{(i)}|}(n^{-3},\infty) \geq 2 \Big) \\ &\sim n^2 \Big[\mathbf{P}\Big(|X_1| > \gamma_n n^{-3} \Big) \Big]^2. \end{aligned} \tag{4.12}$$

for large enough n. Using the upper bound obtained in (4.12), we obtain following upper bound for right hand side of (4.11) as

$$\begin{aligned} r_n n^2 \Big[\mathbf{P}\Big(|X_1| > \gamma_n n^{-3} \Big) \Big]^2 \mathbf{E}(Z_n) &= r_n \mu^n n^2 \Big[\mathbf{P}\Big(|X_1| > \gamma_n n^{-3} \Big) \Big]^2 \\ &= n^2 \mathbf{P}(|X_1| > \gamma_n) \Big[\frac{\mathbf{P}(|X_1| > \gamma_n n^{-3})}{\mathbf{P}(|X_1| > \gamma_n)} \Big]^2. \end{aligned} \tag{4.13}$$

Using Potter's bound (Lemma in Resnick [1987]), we obtain following upper bound for the ratio in (4.13) as

$$\frac{\mathbf{P}\Big(|X-1|>\gamma_n n^{-3}n^3\Big)}{\mathbf{P}(|X_1|>\gamma_n n^{-3})} \geq (1-\eta)n^{-\alpha-\eta}$$

for some $0<\eta<1$ and large enough n depending on the choice of η. So we obtain following upper bound for the expression in (4.13) as

$$n^2\mathbf{P}\Big(|X_1|>\gamma_n\Big)\Big((1-\eta)^{-1}n^{3\alpha+3\eta}\Big)^2 = \frac{n^{2+6\alpha+6\eta}}{\mu^n}\Big(\mu^n\mathbf{P}(|X_1|>\gamma_n)\Big)(1-\eta)^{-2}. \tag{4.14}$$

Note that the first and second terms in the upper bound derived in (4.14) converges to 0 as $n\to\infty$ where the third term is bounded. Hence, we see that (4.10) holds.

Now we shall prove (3.3). Define

$$\widetilde{N}_n = \sum_{\mathbf{u}\in I(\mathbf{v})}\sum_{|\mathbf{v}|=n}\delta_{\gamma_n^{-1}|X(\mathbf{u})|}.$$

It is clear that

$$\begin{aligned}
&r_n|m_n(F_{g_1,g_2,\epsilon_1,\epsilon_2})-\widetilde{m}_n(F_{g_1,g_2,\epsilon_1,\epsilon_2})|\\
&= r_n\bigg|\mathbf{E}^*\bigg[\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(N_n(g_i)-\epsilon_i\Big)_+\Big\}\Big)\bigg]-\mathbf{E}^*\bigg[\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\widetilde{N}_n(g_i)-\epsilon_i\Big)_+\Big\}\Big)\bigg]\bigg|\\
&\leq r_n\mathbf{E}\bigg[\bigg|\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(N_n(g_i)-\epsilon_i\Big)_+\Big\}\Big)-\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\widetilde{N}_n(g_i)-\epsilon_i\Big)_+\Big\}\Big)\bigg|\bigg]\\
&\leq r_n\mathbf{E}\bigg[\bigg|\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(N_n(g_i)-\epsilon_i\Big)_+\Big\}\Big)-\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\widetilde{N}_n(g_i)-\epsilon_i\Big)_+\Big\}\Big)\bigg|\mathbb{1}(AMO_n)\bigg]\\
&\qquad\qquad\qquad + 2r_n\mathbf{P}(AMO_n)
\end{aligned} \tag{4.15}$$

where the first inequality is due to (4.9) and the second inequality is derived using the fact that each products inside the modulus is bounded by 1 and $|x+y|=|x|+|y|$. Note that the last term is $o(n)$ using (4.10). So it is enough to show that the first term in the display (4.15) converges to 0 as $n\to\infty$ to prove (3.3). In following display, we derive an upper bound for the term inside the expectation in (4.15). Note that

$$\bigg|\sum_{i=1}^{2}\Big(\exp\Big\{-\Big(\widetilde{N}_n(g_i)-\epsilon_i\Big)_+\Big\}-\exp\Big\{-\Big(N_n(g_i)-\epsilon_i\Big)_+\Big\}\Big)$$

$$+\left(\exp\left\{-\sum_{i=1}^{2}\Big(N_n(g_i)-\epsilon_i\Big)_+\right\}-\exp\left\{-\sum_{i=1}^{2}\Big(\widetilde{N}_n(g_i)-\epsilon_i\Big)_+\right\}\right)\Bigg|$$

$$\leq \sum_{i=1}^{2}\left|\exp\left\{-\Big(N_n(g_i)-\epsilon_i\Big)_+\right\}-\exp\left\{-\Big(\widetilde{N}_n(g_i)-\epsilon_i\Big)_+\right\}\right|$$

$$+\left|\exp\left\{-\sum_{i=1}^{2}\Big(N_n(g_i)-\epsilon_i\Big)_+\right\}-\exp\left\{-\sum_{i=1}^{2}\Big(\widetilde{g}_i-\epsilon_i\Big)_+\right\}\right|$$

$$\leq 2\sum_{i=1}^{2}\left|\Big(\widetilde{N}_n(g_i)-\epsilon_i\Big)_+-\Big(N_n(g_i)-\epsilon_i\Big)_+\right| \tag{4.16}$$

using the facts that $|x+y|\leq|x|+|y|$ and $|e^x-e^y|\leq|x-y|$ for all $x,y\in\mathbb{R}$. Note that for all $x,y,a>0$,

$$|(x-a)_+-(y-a)_+|=\begin{cases}|x-y| & \text{if } x,y>a\\ |x-a| & \text{if } x>a>y\\ |y-a| & \text{if } y>a>x\\ 0 & \text{if } x,y<a.\end{cases}$$

So we can see that $|(x-a)_+-(y-a)_+|\leq|x-y|$ for all $x,y,a>0$. Using this fact, we obtain following upper bound for right hand side of (4.16) as

$$2\sum_{i=1}^{2}\Big|\widetilde{N}_n(g_i)-N_n(g_i)\Big|. \tag{4.17}$$

Hence, combining (4.15) with the upper bound obtained in (4.17), it is enough to show that

$$\lim_{n\to\infty} r_n\mathbf{E}\left[\mathbb{1}(AMO_n)\sum_{i=1}^{2}\Big|\widetilde{N}_n(g_i)-N_n(g_i)\Big|\right]=0. \tag{4.18}$$

Using the fact that g_1 and g_2 is any pair of functions in $C_c^+(\bar{\mathbb{R}}_0)$, to prove (4.18) it is enough to show that

$$\lim_{n\to\infty}\mathbf{E}\Big(\mathbb{1}(AMO_n)\Big|N_n(g)-\widetilde{N}_n(g)\Big|\Big)=0 \tag{4.19}$$

for any $g\in C_c^+(\bar{\mathbb{R}}_0)$. Fix a function $g\in C_c^+(\bar{\mathbb{R}}_0)$ and let $\delta_g=\inf\{|x|:g(x)>0\}$. Let $T(\mathbf{v})$ denotes the displacement $X(\mathbf{u})$ on the path $I(\mathbf{v})$ such that $|X(\mathbf{u})|=\max_{\mathbf{u}'\in I(\mathbf{v})}|X(\mathbf{u}')|$ i.e. the maximum term in modulus among all the displacements on the path $I(\mathbf{v})$ for every vertex $\mathbf{v}$ in the n-th generation. Note that on the event $\max_{|\mathbf{v}|=n}|T(\mathbf{v})|<\gamma_n\delta/2$, no displacement contributes to the sum

$\widetilde{N}_n(g) = \sum_{|\mathbf{v}|=n} \sum_{\mathbf{u}\in I(\mathbf{v})} g\Big(\gamma_n^{-1} X(\mathbf{u})\Big)$. Similarly, conditioned on AMO_n and $\max_{|\mathbf{v}|=n} |T(\mathbf{v})| < \gamma_n\delta/2$, it is clear that $|S(\mathbf{v})| \le \gamma_n\delta/2 + \gamma_n n^{-2} < \gamma_n\delta$ for large enough n and every vertex $\mathbf{v}$ in the n-th generation. So conditioned on the events AMO_n and $\max_{|\mathbf{v}|=n} |T(\mathbf{v})| < \gamma_n\delta/2$, it is clear that $N_n(g) = 0$ and so we get that the difference inside the expectation is 0. To show that (4.19), it is enough to show that

$$\lim_{n\to\infty} r_n \mathbf{E}\Big[\Big|N_n(g) - \widetilde{N}_n(g)\Big| \mathbb{1}(AMO_n)\mathbb{1}\Big(\max_{|\mathbf{v}|=n} |T(\mathbf{v})| > \gamma_n\delta/2\Big)\Big] = 0. \tag{4.20}$$

Note that the term inside the expectation can be given following upper bound

$$\sum_{|\mathbf{v}|=n} \Big| g\Big(\gamma_n^{-1} S(\mathbf{v})\Big) - \sum_{\mathbf{u}\in I(\mathbf{v})} g\Big(\gamma_n^{-1} X(\mathbf{u})\Big)\Big| \mathbb{1}(AMO_n)\mathbb{1}\Big(\max_{|\mathbf{v}|=n} |T(\mathbf{v})| > \gamma_n\delta/2\Big). \tag{4.21}$$

Conditioned on the events AMO_n and $\max_{|\mathbf{v}|=n} T(\mathbf{v}) > \gamma_n\delta/2$, we get that $\sum_{\mathbf{u}\in I(\mathbf{v})} g\Big(\gamma_n^{-1} X(\mathbf{u})\Big) = g\Big(\gamma_n^{-1} T(\mathbf{v})\Big)$ if $|T(\mathbf{v})| > \gamma_n\delta/2$ for large enough n. So the upper bound derived in (4.21) equals

$$\sum_{|\mathbf{v}|=n} \Big| g\Big(\gamma_n^{-1} S(\mathbf{v})\Big) - g\Big(\gamma_n^{-1} T(\mathbf{v})\Big)\Big| \mathbb{1}\Big(|T(\mathbf{v})| > \gamma_n\delta/2\Big)\mathbb{1}(AMO_n) \tag{4.22}$$

when $n^{-3} < \delta_g$. We have assumed that g is a Lipschitz continuous function i.e. there exists a constant $\|g\|$ such that $|g(x) - g(y)| \le \|g\||x-y|$ for all $x, y \in \mathbb{R}$. Hence we obtain following upper bound for the expression in (4.22) as

$$\|g\|\gamma_n^{-1} \sum_{|\mathbf{v}|=n} |S(\mathbf{v}) - T(\mathbf{v})| \mathbb{1}\Big(|T(\mathbf{v})| > \gamma_n\delta/2\Big)\mathbb{1}(AMO_n). \tag{4.23}$$

Note that $|S(\mathbf{v}) - T(\mathbf{v})| \le (n-1)\gamma_n/n^3 < n^{-2}\gamma_n$ conditioned on the event AMO_n. Using this fact with (4.23), we obtain following upper bound for the expectation in the left hand side of (4.20) as

$$\begin{aligned}
& r_n\|g\|\gamma_n^{-1}\gamma_n n^{-2}\mathbf{E}\Big[\sum_{|\mathbf{v}|=n} \mathbb{1}\Big(|T(\mathbf{v})| > \gamma_n\delta/2\Big)\Big] \\
&= r_n\|g\|n^{-2}\mathbf{E}\Big[\sum_{|\mathbf{v}|=n} \mathbf{E}\Big(\mathbb{1}(|T(\mathbf{v})| > \gamma_n\delta/2)\Big|\mathcal{F}_n\Big)\Big] \\
&= r_n\|g\|n^{-2}\mathbf{E}\Big[\sum_{|\mathbf{v}|=n} \mathbf{P}\Big(|T(\mathbf{v})| > \gamma_n\delta/2\Big)\Big] \\
&= r_n\|g\|n^{-2}\mathbf{P}\Big(\max_{1\le i\le n} |X_1^{(i)}| > \gamma_n\delta/2\Big)\mathbf{E}(Z_n)
\end{aligned} \tag{4.24}$$

using the fact that the displacements and branching mechanism are independent and the displacements on a typical path are independently and identically distributed. It is easy to see that that the expression in (4.24) can be given following upper bound

$$
\begin{aligned}
& r_n \|g\| n^{-2} \mu^n \mathbf{P}\Big(\bigcup_{1 \le i \le n} (|X_1^{(i)}| > \gamma_n \delta/2) \Big) \\
& \le \|g\| \Big(\mathbf{P}(|X_1| > \gamma_n) \Big)^{-1} n^2 \sum_{i=1}^{n} \mathbf{P}(|X_1| > \gamma_n \delta/2) \\
& = \|g\| n^{-1} \frac{\mathbf{P}(|X_1| > \gamma_n \delta/2)}{\mathbf{P}(|X_1| > \gamma_n \delta)}. \qquad (4.25)
\end{aligned}
$$

Note that the ratio in right hand side of (4.25) converges to $\Big(\delta/2\Big)^{-\alpha}$ as $n \to \infty$. So we see that the expression in right hand side of (4.25) converges to 0 as $n \to \infty$ and hence we conclude (4.20).

4.4 Proof of Lemma 3.2

Here, we shall formalize the fact that the probability of contribution coming from the displacements in the first $(n - K)$ generations are negligible. Let

$$
\Upsilon(n, K) = \sum_{|\mathbf{u}| \le n-K} \boldsymbol{\delta}_{\gamma_n^{-1} X(\mathbf{u})}.
$$

Note that $\mathbf{P}^*(A) = \mathbf{E}^*(\mathbb{1}(A)) = \Big(\mathbf{P}(\mathcal{S})\Big)\mathbf{E}(\mathbb{1}(A)\mathbb{1}(\mathcal{S})) \le \Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\mathbf{E}(\mathbb{1}(A)) = \Big(\mathbf{P}(\mathcal{S})\Big)^{-1}\mathbf{P}(A)$. So it is enough to show that

$$
\lim_{K \to \infty} \lim_{n \to \infty} \mathbf{P}\Big(\Upsilon(n, K)(\theta, \infty) \ge 1 \Big) = 0. \qquad (4.26)
$$

to prove Lemma 3.2. Note that the displacements in the first $(n - K)$ generations are not independent and identically distributed. It is clear that the point process coming from same particle are independent and identically distributed and so if we group the displacements according to their parents then we can use this fact. Let D_i denotes the collection of all vertices in the i-th generation. Then we can see that

$$
\Upsilon(n, K) = \sum_{i=0}^{n-K-1} \sum_{\mathbf{u} \in D_i} \sum_{j=1}^{ch(\mathbf{u})} \boldsymbol{\delta}_{\gamma_n^{-1} |X(\mathbf{u}j)|}
$$

where $ch(\mathbf{u})$ denotes the number of children of the vertex $\mathbf{u}$. Note that

$$
\Big\{ \Upsilon(n, K)(\theta, \infty) \ge 1 \Big\} = \bigcup_{i=0}^{n-K-1} \bigcup_{\mathbf{u} \in D_i} \Big\{ \sum_{j=1}^{ch(\mathbf{u})} \boldsymbol{\delta}_{\gamma_n^{-1} |X(\mathbf{u}j)|}(\theta, \infty) \ge 1 \Big\}.
$$

So we get following following upper bound for $\mathbf{P}(\Upsilon(n,K)(\theta,\infty)\geq 1)$ as follows

$$\begin{aligned}
&\sum_{i=0}^{n-K-1}\mathbf{E}\Big[\sum_{\mathbf{u}\in D_i}\mathbf{P}\Big(\sum_{j=1}^{ch(\mathbf{u})}\boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\theta,\infty)\geq 1\Big)\Big]\\
&=\sum_{i=0}^{n-K-1}\mathbf{E}\Big[\mathbf{E}\Big(\sum_{\mathbf{u}\in D_i}\mathbf{P}\Big(\sum_{j=1}^{ch(\mathbf{u})}\boldsymbol{\delta}_{\gamma_n^{-1}|X(\mathbf{u})|}(\theta,\infty)\geq 1\Big)\Big|\mathcal{F}_i\Big)\Big]\\
&=\sum_{i=0}^{n-K-1}\mathbf{E}(Z_i)\mathbf{P}\Big(\sum_{j=1}^{Z_{1,i}}\boldsymbol{\delta}_{\gamma_n^{-1}|X_j|}(\theta,\infty)\geq 1\Big) \qquad (4.27)
\end{aligned}$$

where $(Z_{1,i}: i\geq 1)$ be a collection of independent copies of Z_1, using the facts that conditioned on $\mathcal{F}_i$, each particle produces a random number of offspring according to the distribution of Z_1 and independent of what happened upto i-th generation for every $1\leq i\leq K$. So we get following expression for right hand side of (4.27)

$$\begin{aligned}
\sum_{i=0}^{n-K-1}\mu^i\mathbf{P}\Big(\sum_{i=1}^{Z_1}\boldsymbol{\delta}_{\gamma_n^{-1}|X_i|}(\theta,\infty)\geq 1\Big)&\leq \mu^{n-K}\mathbf{P}\Big(\bigcup_{i=1}^{Z_1}(|X_i|>\gamma_n\theta)\Big)\\
&=\mu^{n-K}\sum_{j=1}^{\infty}\mathbf{P}\Big(\bigcup_{i=1}^{j}(|X_j|>\gamma_n\theta)\Big)\mathbf{P}(Z_1=j)\\
&\leq \mu^{n-K}\mathbf{P}\Big(|X_1|>\gamma_n\theta\Big)\sum_{j=1}^{\infty}j\mathbf{P}(Z_1=j) \qquad (4.28)
\end{aligned}$$

using the fact that the displacements are marginally identically distributed. So we obtain following expression

$$\mu^{n-K+1}\mathbf{P}(|X_1|>\gamma_n\theta). \qquad (4.29)$$

Hence we can see that

$$r_n\mathbf{P}(\Upsilon(n,K)(\theta,\infty))\leq \mu^{-K+1}\frac{\mathbf{P}(|X_1|>\gamma_n)}{\mathbf{P}(|X_1|>\gamma_n)}. \qquad (4.30)$$

Now, letting $n\to\infty$, we can see the ratio in (4.30) converges to $\theta^{-\alpha}$. Hence letting $K\to\infty$, we conclude the proof of (4.26).

4.5 Proof of Lemma 3.3

Note that, after cutting the tree at the $(n-K)$-th generation, we obtain Z_{n-K} GW trees each with K generations and progeny distribution Z_1. Recall that the subtrees obtained after cutting are denoted by $(\mathbb{T}_i : 1 \le i \le Z_{n-K})$ conditioned on $\mathcal{F}_{n-K}$. In the pruning step, each subtree $\mathbb{T}_1$ is modified in such a way that each particle can have at most B descendants in the next generation. The modified subtrees are denoted by $(\mathbb{T}_i(B) : 1 \le i \le Z_{n-K})$. Let $\mathbf{u} \in \mathbb{T}_i$ for some i, then $A(\mathbf{u})$ denotes the number of descendants of the particle $\mathbf{u}$ in the K-th generation of the subtree $\mathbb{T}_1$. If $u \in \mathbb{T}_1(B)$, then $A^{(B)}(\mathbf{u})$ denotes the number of descendants of $\mathbf{u}$ in the K-th generation of the pruned subtree $\mathbb{T}_1(B)$. Using the inequality derived in (4.9), we get that it is enough to show that

$$\lim_{B\to\infty}\lim_{n\to\infty} r_n \mathbf{E}\Big[\Big|\prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\widetilde{N}_n^{(K)}(g_i)-\epsilon_i\Big)_+\Big\}\Big) - \prod_{i=1}^{2}\Big(1-\exp\Big\{-\Big(\widetilde{N}^{K,B}(g_i)-\epsilon_i\Big)_+\Big\}\Big)\Big|\Big] = 0 \tag{4.31}$$

to establish (3.9) where

$$\widetilde{N}_n^{(K,B)} = \sum_{i=1}^{Z_{n-K}} \sum_{\mathbf{u}\in\mathbb{T}_i(B)} A^{(B)}(\mathbf{u})\boldsymbol{\delta}_{\gamma_n^{-1}X(\mathbf{u})} \qquad \text{and} \qquad \widetilde{N}_n^{(K)} = \sum_{i=1}^{Z_{n-K}} \sum_{\mathbf{u}\in\mathbb{T}_i} A(\mathbf{u})\boldsymbol{\delta}_{\gamma_n^{-1}X(\mathbf{u})}.$$

using the same steps used in Subsection 4.3 to prove Lemma 3.1, we can see that it is enough to show that

$$\lim_{B\to\infty}\lim_{n\to\infty} r_n \mathbf{E}\Big[\Big|\widetilde{N}_n^{(K)}(g) - \widetilde{N}^{(K,B)}(g)\Big|\Big] = 0. \tag{4.32}$$

for any $g \in C_c^+(\bar{\mathbb{R}}_0)$. We need more notations to prove (4.32) which are introduced below.

Let $D_j^{(i)}$ denote the collection of all vertices in the j-th generation of the i-th subtree $\mathbb{T}_i$ for $1 \le j \le K$ and $1 \le i \le Z_{n-K}$. Similarly $D_j^{(i)}(B)$ denotes the collection of all vertices in the j-th generation of the i pruned subtree $\mathbb{T}_i(B)$ for every $1 \le j \le K$ and $1 \le i \le Z_{n-K}$. Then we can see that

$$\widetilde{N}_n^{(K,B)} = \sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \sum_{\mathbf{u}\in D_j^{(i)}(B)} A^{(B)}(\mathbf{u})\boldsymbol{\delta}_{\gamma_n^{-1}X(\mathbf{u})} \qquad \text{and} \qquad \widetilde{N}_n^{(K)} = \sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \sum_{\mathbf{u}\in D_j^{(i)}} A(\mathbf{u})\boldsymbol{\delta}_{\gamma_n^{-1}X(\mathbf{u})}.$$

In order to show that (4.32), we need assistance of another intermediate point process $N_n^{(*,K,B)}$. This trick also has been used in Bhattacharya et al. [2016] and Bhattacharya et al. [2018]. For the sake of completeness we give the proof below with a construction of the point process $N_n^{(*,K,B)}$.

Consider the subtree $\mathbb{T}_1$. Leave out the root of $\mathbb{T}_1$. We reward (mark) the particles of $\mathbb{T}_1$ according to following scheme. If the number of particles in the first generation is less than or equal to B, then we reward each particle with a $*$. If there are more than B particles in the first generation then we reward a $*$ to the first B of them (according to the Harris-Ulam) labelling and keep the others without any reward. Now we consider the particles in the second generation. We pick a particle and check whether it is one of the first B children (according to Ulam-Harris labelling) of its parent in the previous generation or not. If is so then we add a more $*$ to it than its parent. Otherwise, it will have the same reward as its parent. We follow the reward scheme until we are done with all particles at the K-th generation of $\mathbb{T}_1$. We shall do the same for the other subtrees. Employing this reward scheme, we obtain marked subtrees which will be denoted by $(\mathbb{T}_j^* : 1 \leq j \leq Z_{n-K})$.

Note that the collection of all vertices in j-th generation of i-th marked subtree $\mathbb{T}_i^*$ is same as that of $\mathbb{T}_i$ for every $1 \leq j \leq K$ and $1 \leq i \leq Z_{n-K}$. Let $\mathbf{u} \in D_j^{(i)}$, then $\mathbf{u}$ can have at most j many $*$s as award. Then for every $\mathbf{u} \in D_j^{(i)}$, we define $A^{(B)}(\mathbf{u})$ as the number of descendants of $\mathbf{u}$ in the K-th generation with $(K - j + l)$ many $*$s if $\mathbf{u}$ is awarded l many $*$s for $0 \leq l \leq j$. With these notations, we define

$$N_n^{(*,K,B)} = \sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \sum_{\mathbf{u} \in D_j^{(i)}} A^{(B)}(\mathbf{u}) \boldsymbol{\delta}_{\gamma_n^{-1} X(\mathbf{u})}. \tag{4.33}$$

This construction is lacking in Bhattacharya et al. [2016]. In that work, it was not clear that how one can keep the generations same as that of $\mathbb{T}_1$ with number of descendants of a particle same as that in the pruned subtree. Here with the help of the construction of the marked tree, we can see how one can obtain the point process $N_n^{(*,K,B)}$ from the marked subtrees. Note that it is enough to show that

$$\lim_{B\to\infty} \lim_{n\to\infty} r_n \mathbf{E}\Big(\widetilde{N}_n^{(K)}(g) - N_n^{(*,K,B)}(g)\Big) = 0 \tag{4.34}$$

and

$$\lim_{B\to\infty} \lim_{n\to\infty} r_n \mathbf{E}\Big(N_n^{(*,K,B)}(g) - \widetilde{N}_n^{(K,B)}(g)\Big) = 0. \tag{4.35}$$

Recall that $Z_1^{(B)} = Z_1 \mathbb{1}(Z_1 \leq B) + B\mathbb{1}(Z_1 > B)$. Note that the Galton-Watson process with progeny distribution $Z_1^{(B)}$ is denoted by $(Z_n^{(B)} : n \geq 1)$. Using the fact that $A^{(B)}(\mathbf{u}) \leq A(\mathbf{u})$ for all $\mathbf{u} \in \mathbb{T}_i$ for $1 \leq i \leq Z_{n-K}$, we get that the left hand side in (4.34) equals

$$r_n \mathbf{E}\Big[\sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \sum_{\mathbf{u} \in D_j^{(i)}} \Big(A(\mathbf{u}) - A^{(B)}(\mathbf{u})\Big) g\Big(\gamma_n^{-1} X(\mathbf{u})\Big)\Big]$$

$$
\begin{aligned}
&\leq \|g\| r_n \mathbf{E}\Big[\sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \sum_{\mathbf{u}\in D_j^{(i)}} \Big(A(\mathbf{u}) - A^{(B)}(\mathbf{u})\Big) \mathbb{1}\Big(|X(\mathbf{u})| > \gamma_n \delta_g\Big)\Big] \\
&= \|g\| \mathbf{E}\Big[\mathbf{E}\Big(\sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \sum_{\mathbf{u}\in D_j^{(i)}} \Big(A(\mathbf{u}) - A^{(B)}(\mathbf{u})\Big) \mathbb{1}\Big(|X(\mathbf{u})| > \gamma_n \delta_g\Big)\Big| \mathcal{F}_{n-K}\Big)\Big] \\
&= \|g\| \mathbf{E}\Big[\sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \mathbf{E}(Z_j) \mathbf{E}\Big(Z_{K-j} - Z_{K-j}^{(B)}\Big) \mathbf{P}\Big(|X(\mathbf{u})| > \gamma_n \delta_g\Big)\Big] \qquad (4.36)
\end{aligned}
$$

(recall $\delta_g = \inf\{|x| : g(x) > 0\}$) using the fact that whatever happened after $(n-K)$ generations are independent of $\mathcal{F}_{n-K}$ and displacements are independent of the branching mechanism. Using the fact that the displacements are marginally identically distributed, we get following expression for right hand side of (4.36) becomes

$$
\begin{aligned}
&r_n \|g\| \mathbf{P}\Big(|X_1| > \gamma_n \delta_g\Big) \mathbf{E}\Big[\sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \mu^j (\mu^{K-j} - \mu_B^{K-j})\Big] \\
&= \|g\| \frac{\mathbf{P}(|X_1| > \gamma_n \delta_g)}{\mathbf{P}(|X_1| > \gamma_n)} \sum_{j=0}^{K-1} \Big(1 - (\mu^{-1}\mu_B)^j\Big). \qquad (4.37)
\end{aligned}
$$

We see that the first term converges to $\delta_g^{-\alpha}$ as $n \to \infty$ and the second term converges to 0 as $B \to \infty$. So we are done with (4.34).

Similarly, the term in the left hand side of (4.35) can be given following upper bound

$$
\begin{aligned}
&r_n \|g\| \mathbf{E}\Big[\sum_{i=1}^{Z_{n-K}} \sum_{j=1}^{K} \sum_{\mathbf{u}\in D_j^{(i)} \setminus D_j^{(i)}} A^{(B)}(\mathbf{u}) \mathbb{1}\Big(|X(\mathbf{u})| > \gamma_n \delta_g\Big)\Big] \\
&= r_n \|g\| \mathbf{P}\Big(|X_1| > \gamma_n \delta_g\Big) \mu^{n-K} \sum_{j=1}^{K} (\mu^j - \mu_B^j) \mu_B^{K-j}\Big] \\
&= \|g\| \frac{\mathbf{P}(|X_1| > \gamma_n \delta_g)}{\mathbf{P}(|X_1| > \gamma_n)} \mu^n \sum_{j=1}^{K} \Big(1 - (\mu^{-1}\mu_B)^j\Big) (\mu^{-1}\mu_B)^{K-j} \qquad (4.38)
\end{aligned}
$$

using the same facts. Now, it is easy to see that the first term converges to $\delta_g^{-\alpha}$ and and the second term converges to 0 as $B \to \infty$. So we are done with (4.35).