ALEA, Lat. Am. J. Probab. Math. Stat. 20, 1411–1432 (2023)
DOI: 10.30757/ALEA.v20-53

Conditional central limit theorem for subcritical branching random walk

Wenming Hong and Dan Yao

School of Mathematical Sciences & Laboratory of Mathematics and Complex Systems, Beijing Normal University, Beijing 100875, P.R. China
E-mail address: wmhong@bnu.edu.cn

Corresponding author. School of Mathematical Sciences & Laboratory of Mathematics and Complex Systems, Beijing Normal University, Beijing 100875, P.R. China
E-mail address: dyao@mail.bnu.edu.cn

Abstract. We consider a discrete-time subcritical branching random walk (Z_n) on the real line $\mathbb{R}$. Let $Z_n(A)$ be the number of individuals in the n-th generation located in $A \in \mathcal{B}(\mathbb{R})$, and $N_n := Z_n(\mathbb{R})$ denote the size of the n-th generation. Under some conditions, we prove that when $0 < m := \mathbf{E}N_1 < 1$, for $x \in \mathbb{R}$, as $n \to \infty$,

$$\mathcal{L}\big(Z_n\big((-\infty, \sqrt{n}x]\big)\big|N_n > 0\big) \Longrightarrow \mathcal{L}\big(\xi \mathbf{1}_{\{\mathcal{N} \le x\}}\big),$$

where $\Longrightarrow$ means weak convergence, ξ is the Yaglom limit of the associated Galton-Watson process (N_n) and $\mathcal{N}$ is a standard normal random variable independent of ξ.

1. Introduction

1.1. *Definition of the model.* Consider a discrete-time branching random walk on the real line $\mathbb{R}$. Let $\mathbb{N}_+ := \{1, 2, \cdots\}$ and $\mathbb{N} = \mathbb{N}_+ \cup \{0\}$. At time 0, there is one particle positioned at 0, which forms the 0-th generation. The initial particle (after one unit of time) splits into a random number of children according to the offspring distribution $p = \{p_j\}_{j\in\mathbb{N}}$. These children are positioned (with respect to their parent) independently according to the same probability measure G, which form the first generation. We may call G the *jump distribution*. More generally, each particle in the n-th generation gives birth independently to a random number of children with law p, who are in the $(n+1)$-th generation and are positioned in relative distances according to G. The system goes on as long as there are particles alive. The resulting system is called a *branching random walk* (BRW).

Throughout this paper, we assume that the reproduction and displacement mechanisms are independent. We denote by $\mathbf{P}$ the probability measure for this system started from a single particle at 0, and the corresponding expectation is $\mathbf{E}$.

Received by the editors February 13th, 2023; accepted October 2nd, 2023.

2010 *Mathematics Subject Classification.* Primary 60J80; secondary 60F10.

Key words and phrases. Branching random walk, subcritical Galton-Watson process, reduced process, Yaglom limit.

This work was supported in part by NSFC (NO. 11971062) and the National Key Research and Development Program of China (No. 2020YFA0712900).

Let Z_n be the point process of the positions of particles in the n-th generation. Let G^{n*} be the n-fold convolution of the jump distribution G. Let $\mathcal{B}(\mathbb{R})$ denote the Borel σ-algebra on $\mathbb{R}$. For every Borel set $A \in \mathcal{B}(\mathbb{R})$, we define the number of particles in the n-th generation located in A by $Z_n(A)$. We denote the size of the n-th generation by $N_n = Z_n(\mathbb{R})$. Of course, $(N_n, n \in \mathbb{N})$ is just a Galton-Watson process (GW-process) with offspring distribution p. Let $f(\cdot)$ be the generating function of p. By the branching property, the generating function of N_n is given by the iterate

$$f_n(s) = f(f_{n-1}(s)), \quad f_0(s) = s, \quad f_1(s) = f(s), \quad s \in [0,1].$$

For a supercritical BRW (that is, $m > 1$), it was first conjectured by Harris (1963) that if G has mean zero and variance one, then as $n \to \infty$,

$$m^{-n} Z_n\big((-\infty, \sqrt{n}x]\big) \to W\Phi(x) \text{ in probability,} \tag{1.1}$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$ is the standard normal distribution function, and W is the limit of the additive martingale $(m^{-n}N_n, n \in \mathbb{N})$ in the associated GW-process. Many authors have studied and developed these results. Kaplan and Asmussen (1976) and Stam (1966) considered the case where the displacement and the offspring reproduction of each particle are independent, and obtained that the convergence holds almost surely. Furthermore, Biggins (1990) and Klebaner (1982) removed the assumption that the branching and motion mechanisms are independent, and extended these results to the BRW in a time-varying environment. Recently, Bansaye (2019) and Bansaye and Huang (2015) considered the non-homogeneous branching Markov chain.

In this paper, we are interested in the subcritical BRW, i.e., $m < 1$. In this case, the associated GW-process dies out with probability one, so it is natural to consider the central limit theorem conditioned on non-extinction. We obtain that the limit variable reflects two parts of the randomness: the Yaglom limit comes from the subcritical branching mechanism and the normal variable comes from the space displacement (see Theorem 1.3 below).

Any $\mathbb{R}$-valued random variable $\{Y; \mathbf{P}\}$, we may consider the associated probabilities

$$\mathbf{P}(Y \in A) = (\mathbf{P} \circ Y)A, \quad A \in \mathcal{B}(\mathbb{R}).$$

The set function $\mathcal{L}(Y) = \mathbf{P} \circ Y$ of Y is a probability measure on $\mathcal{B}(\mathbb{R})$, called the *distribution* or *law* of Y. For any $B \in \mathcal{B}(\mathbb{R})$, then a probability measure $\mathcal{L}(Y \mid Y \in B)$ on $\mathcal{B}(\mathbb{R})$ is defined as follows:

$$\mathcal{L}(Y \mid Y \in B)A = \mathbf{P}(Y \in A \mid Y \in B), \quad A \in \mathcal{B}(\mathbb{R}),$$

which is called the *conditional distribution* of Y given $\{Y \in B\}$.

1.2. *Previous results.* Let us recall the known conditional limit theorems on subcritical GW-processes. Throughout the remainder of this paper, the symbol "$\Longrightarrow$" denotes weak convergence.

Theorem 1.1. *For a subcritical GW-process $(N_n, n \in \mathbb{N})$ satisfying $0 < m = \mathbf{E}N_1 < 1$, then*

(1) *As $n \to \infty$,*

$$\mathcal{L}(N_n \mid N_n > 0) \Longrightarrow \mathcal{L}(\xi), \tag{1.2}$$

where ξ is a random variable called the Yaglom limit of the subcritical GW-process (N_n). Furthermore, the sequence $(1 - f_n(s))/(m^n(1-s))$ is monotone decreasing in n and then converges to a non-decreasing function $\varphi(s)$. In particular, we have

$$\frac{1 - f_n(0)}{m^n} \downarrow \varphi(0), \quad n \to \infty. \tag{1.3}$$

(2) *For fixed $l \in \mathbb{N}$ and $j \in \mathbb{N}_+$,*

$$\lim_{n\to\infty} \mathbf{P}(N_n = j \mid N_{n+l} > 0) = b_j(l) \geq 0$$

and $\sum_{j\in\mathbb{N}_+} b_j(l) = 1$*, the generating function of* $\{b_j(l)\}_{j\in\mathbb{N}_+}$ *is* $\frac{H(s)-H(sf_l(0))}{m^l}$*, where* $H(s)$ *is the generating function of the Yaglom limit* ξ*.*

(3) *The* Q*-process associated with* (N_n) *is a Markov chain whose* n*-step transition probability is given by*

$$\pi_n(i,j) = \lim_{l\to\infty} \mathbf{P}(N_n = j | N_{n+l} > 0, N_0 = i), \quad i,j \in \mathbb{N}_+,$$

then the Q*-process is positive recurrent if and only if* $\sum (j\log j)p_j < \infty$*. Furthermore, in the positive recurrent case the stationary distribution for* Q *is*

$$\pi_j = \varphi(0) j b_j, \quad j \in \mathbb{N}_+, \tag{1.4}$$

where $b_j = \mathbf{P}(\xi = j)$ *is the distribution of the Yaglom limit* ξ*.*

Remark 1.2. Yaglom (1947) proved the existence of the Yaglom limit when $m < 1$ and N_1 has a finite second moment. This was generalized to the case without the second moment assumption in Joffe (1967); Heathcote et al. (1967); Athreya and Ney (1972); Geiger (1999). In particular, Joffe (1967) gave a sufficient condition for $\varphi(0) > 0$. Heathcote et al. (1967) showed that $\varphi(0) > 0$ if and only if $\sum (j\log j)p_j < \infty$. The second point (2) is Athreya and Ney (1972, Theorem 1, Section I.14). Some properties of the Q-process come from Joffe (1967, Another theorem) and Athreya and Ney (1972, Theorem 2, Section I.14).

1.3. *Main results.* In this section we state our main results which generalize Theorem 1.1 in the BRW. We assume that

$$0 < m = \mathbf{E}N_1 < 1, \quad \sigma^2 := \operatorname{Var} N_1 < \infty. \tag{1.5}$$

$$\int_{-\infty}^{\infty} x \, \mathrm{d}G(x) = 0, \quad \int_{-\infty}^{\infty} x^2 \, \mathrm{d}G(x) = 1. \tag{1.6}$$

Now we present the conditional central limit theorem for the subcritical BRW.

Theorem 1.3. *Assume (1.5) and (1.6) hold, then for all* $x \in \mathbb{R}$*, as* $n \to \infty$*,*

$$\mathcal{L}\big(Z_n\big((-\infty, \sqrt{n}x]\big)\big| N_n > 0\big) \Longrightarrow \mathcal{L}\big(\xi \mathbf{1}_{\{\mathcal{N}\le x\}}\big),$$

where ξ *is the Yaglom limit of the associated GW-process in (1.2) and* $\mathcal{N}$ *is a standard normal random variable independent of* ξ*.*

Remark 1.4. For the critical BRW, the conditional central limit theorem was obtained recently by Hong and Liang (2023). We also mention that $\sigma^2 < \infty$ is needed. On the one hand, it implies that $\sum (j\log j)p_j < \infty$ and then $\varphi(0) > 0$. On the other hand, it guarantees that we can use the many-to-two formula to approximate $Z_n((-\infty, \sqrt{n}x])$ given $N_n > 0$ in the L^2-norm (see Proposition 4.1 below).

It is also interesting to investigate the behavior of $\mathcal{L}(Z_n((-\infty, \sqrt{n}x])|N_{n+l} > 0)$. Fix $l \in \mathbb{N}$ and let $n \to \infty$, then we obtain the following result.

Corollary 1.5. *Under the assumptions (1.5) and (1.6), for fixed* $l, j \in \mathbb{N}$ *and* $x \in \mathbb{R}$*, we have*

$$\lim_{n\to\infty} \mathbf{P}\big(Z_n\big((-\infty, \sqrt{n}x]\big) = j \big| N_{n+l} > 0\big) = b_j(l;x) \ge 0$$

and $\sum_{j\in\mathbb{N}} b_j(l;x) = 1$*, the generating function of* $\{b_j(l;x)\}_{j\in\mathbb{N}}$ *is given by*

$$\frac{H\big(s\big) - H\big(sf_l(0)\big)}{m^l} \Phi(x) + 1 - \Phi(x),$$

where $H(s)$ *is the generating function of the Yaglom limit* ξ*.*

Remark 1.6. This corollary extends (2) in Theorem 1.1, which is a conditional limit theorem for the subcritical GW-process.

If we let $l \to \infty$ first and then $n \to \infty$, the weak limit of $\mathcal{L}(Z_n((-\infty, \sqrt{n}x])|N_{n+l} > 0)$ will evolve as the stationary distribution of the Q-process associated with the subcritical GW-process $(N_n, n \in \mathbb{N})$. In particular, if we exchange the order of $l \to \infty$ and $n \to \infty$, we get the same weak limit.

Corollary 1.7. *Assume (1.5) and (1.6) hold, for all $x \in \mathbb{R}$,*

(1) *As $l \to \infty$ first and then $n \to \infty$,*

$$\mathcal{L}\big(Z_n\big((-\infty, \sqrt{n}x]\big)\big|N_{n+l} > 0\big) \Longrightarrow \mathcal{L}\big(\zeta \mathbf{1}_{\{\mathcal{N} \le x\}}\big),$$

where ζ is a random variable with law $\{\pi_j\}_{j \in \mathbb{N}_+}$ defined in (1.4) and $\mathcal{N}$ is a standard normal random variable independent of ζ.

(2) *As $n \to \infty$ first and then $l \to \infty$,*

$$\mathcal{L}\big(Z_n\big((-\infty, \sqrt{n}x]\big)\big|N_{n+l} > 0\big) \Longrightarrow \mathcal{L}\big(\zeta \mathbf{1}_{\{\mathcal{N} \le x\}}\big),$$

which means that for each $j \in \mathbb{N}$, we have

$$b_j(l; x) \to \mathbf{P}\big(\zeta \mathbf{1}_{\{\mathcal{N} \le x\}} = j\big), \quad l \to \infty,$$

where $b_j(l; x)$ is defined as in Corollary 1.5.

Remark 1.8. The first part of this corollary extends (3) in Theorem 1.1, which describes the stationary distribution of the Q-process associated with (N_n).

Remark 1.9. The assumption (1.6) can be generalized to the case that G is in the domain of attraction of a stable law. More precisely, there exist $\{a_n\} \subset \mathbb{R}$, $\{b_n\} \subset (0, \infty)$ and a non-degenerate random variable η, such that

$$\frac{S_n - a_n}{b_n} \Longrightarrow \eta, \quad n \to \infty,$$

where S_n is the sum of n independent identically distributed variables with law G. When (1.5) and the above general assumption hold, in Theorem 1.3, Corollaries 1.5 and 1.7, with slight changes we replace the set $(-\infty, \sqrt{n}x]$, the normal random variable $\mathcal{N}$ and the distribution function $\Phi(x)$ by $(-\infty, b_n x + a_n]$, η and $\mathbf{P}(\eta \le x)$, respectively.

The rest of this paper is devoted to the proofs of Theorem 1.3, Corollaries 1.5 and 1.7. The key step is to express the conditional random variable $Z_n((-\infty, \sqrt{n}x])$ given $N_n > 0$ in terms of the reduced tree, and decompose it at the most recent common ancestor which is located close to the moment n in the subcritical case (c.f. Fleischmann and Prehn (1974) and Fleischmann and Vatutin (1999)). In Section 2 we introduce the conditional reduced GW-process. In Section 3 we give some properties of conditional reduced GW-trees and the many-to-few formula, which are the key tools in proving Theorem 1.3. Based on these preparations, we prove Theorem 1.3 in Section 4, while in Section 5 we prove Corollaries 1.5 and 1.7 as byproducts of Theorem 1.3.

2. Conditional reduced GW-processes

One of the key steps in the proof is to observe that conditional on $N_n > 0$, only the particles in generation $0 \le k \le n$ having descendants in the n-th generation make contributions to the conditional distribution $\mathcal{L}(Z_n((-\infty, \sqrt{n}x]) \mid N_n > 0)$. From the definition in Fleischmann and Siegmund-Schultze (1977), for a GW-process $(N_n, n \in \mathbb{N})$ with offspring distribution p, the random number $N_{k,n}$ of particles in the k-th generation having non-empty offspring at time n is called the *reduced GW-process*.

Let $\hat{N}_{k,n}$ be a random variable with law $\mathcal{L}(N_{k,n} \mid N_n > 0)$. Then we say $(\hat{N}_{k,n}, 0 \le k \le n)$ is the *conditional reduced GW-process* accordingly. For simplicity, we write $\hat{N}_n := \hat{N}_{n,n}$.

We now introduce some generating functions and moments of conditional reduced GW-processes. It follows from Fleischmann and Siegmund-Schultze (1977, Proposition 1.1) that the conditional reduced GW-process $(\hat{\mathbb{N}}_{k,n}, 0 \le k \le n)$ is a time-inhomogeneous GW-process, and its offspring probability generating function at time k $(0 \le k \le n-1)$ is given by

$$\begin{aligned}\hat{f}_{\mathbf{e}_k(n)}(s) :=& \mathbf{E}[s^{\hat{N}_{k+1,n}} \mid \hat{N}_{k,n} = 1] \\ =& \frac{f\big(f_{n-k-1}(0) + s(1 - f_{n-k-1}(0))\big) - f_{n-k}(0)}{1 - f_{n-k}(0)}, \quad s \in [0,1], \end{aligned} \tag{2.1}$$

which determines the offspring distribution denoted by $\{p_l(\mathbf{e}_k(n))\}_{l\in\mathbb{N}}$ (where $\mathbf{e}_k(n)$ can be viewed as an "environment"), i.e.,

$$\mathbf{P}(\hat{N}_{k+1,n} = l \mid \hat{N}_{k,n} = 1) = p_l(\mathbf{e}_k(n)), \quad l \in \mathbb{N}. \tag{2.2}$$

Now, for each $0 \le k \le n-1$ and $j \ge 0$, let

$$m_j(\mathbf{e}_k(n)) := \sum_{l\in\mathbb{N}} l^j p_l(\mathbf{e}_k(n))$$

be the j-th moment of the offspring distribution $\{p_l(\mathbf{e}_k(n))\}_{l\in\mathbb{N}}$. It is easy to calculate the first and second moments:

$$m_1\big(\mathbf{e}_k(n)\big) = m\frac{1 - f_{n-k-1}(0)}{1 - f_{n-k}(0)}, \tag{2.3}$$

$$m_2\big(\mathbf{e}_k(n)\big) = \frac{(1 - f_{n-k-1}(0))^2}{1 - f_{n-k}(0)} f''(1) + m\frac{1 - f_{n-k-1}(0)}{1 - f_{n-k}(0)}. \tag{2.4}$$

Once again by Fleischmann and Siegmund-Schultze (1977), the generating function of $\hat{N}_{k,n}$ is given by

$$\mathbf{E}\big[s^{\hat{N}_{k,n}}\big] = \frac{f_k\big(f_{n-k}(0) + s(1 - f_{n-k}(0))\big) - f_n(0)}{1 - f_n(0)}, \quad s \in [0,1]. \tag{2.5}$$

Then (2.5) implies that $\mathcal{L}(\hat{N}_n) = \mathcal{L}(N_n \mid N_n > 0)$ and the first and second moments of $\hat{N}_{k,n}$ are

$$\mathbf{E}[\hat{N}_{k,n}] = m^k \frac{1 - f_{n-k}(0)}{1 - f_n(0)}, \tag{2.6}$$

$$\mathbf{E}[\hat{N}_{k,n}^2] = f_k''(1)\frac{(1 - f_{n-k}(0))^2}{1 - f_n(0)} + m^k \frac{1 - f_{n-k}(0)}{1 - f_n(0)}. \tag{2.7}$$

Based on these, the first moment of $\hat{N}_n$ is $m^n/(1 - f_n(0))$. Hence it is obvious from (1.3) that under (1.5) we have

$$\lim_{n\to\infty} \mathbf{E}[\hat{N}_n] = \frac{1}{\varphi(0)} < \infty. \tag{2.8}$$

In addition, the second moment of $\hat{N}_n$ is $\mathbf{E}[N_n^2]/(1 - f_n(0))$. Recall that the variance of the GW-process $(N_n, n \in \mathbb{N})$ given in Athreya and Ney (1972) is

$$\operatorname{Var} N_n = \frac{\sigma^2 m^{n-1}(m^n - 1)}{m - 1}, \quad n \in \mathbb{N}. \tag{2.9}$$

Hence, by a simple calculation we have

$$\mathbf{E}[\hat{N}_n^2] = \frac{1}{1 - f_n(0)}\Big[m^{2n} + \frac{\sigma^2 m^{n-1}(m^n - 1)}{m - 1}\Big]. \tag{2.10}$$

Recall that the generating function of the Yaglom limit ξ is $H(s)$. Thanks to Joffe (1967, Yaglom's theorem), we have

$$H(s) = 1 - \lim_{n\to\infty} \frac{1-f_n(s)}{1-f_n(0)}, \quad s \in [0,1]$$

and $H'(1) = 1/\varphi(0)$. Since $f_n'(1) = m^n$ and $f_n(0) \to 1$ as $n \to \infty$, it is easy to deduce that

$$\lim_{n\to\infty} \frac{1-f_n(0)}{1-f_{n+\tau}(0)} = m^{-\tau}. \tag{2.11}$$

Then by a combination of (2.5) and the preceding results we have

$$\begin{aligned}
\mathbf{E}\big[s^{\hat{N}_{n-\tau,n}}\big] &= 1 - \frac{1-f_{n-\tau}\big(f_\tau(0)+s(1-f_\tau(0))\big)}{1-f_{n-\tau}(0)} \cdot \frac{1-f_{n-\tau}(0)}{1-f_n(0)} \\
&\to 1 - \frac{1-H\big(f_\tau(0)+s(1-f_\tau(0))\big)}{m^\tau}, \quad n\to\infty \\
&= 1 - \frac{1-H\big(f_\tau(0)+s(1-f_\tau(0))\big)}{(1-s)(1-f_\tau(0))} \cdot \frac{(1-s)\big(1-f_\tau(0)\big)}{m^\tau}.
\end{aligned}$$

It is clear that $\lim_{\tau\to\infty} f_\tau(0) \to 1$ implies that

$$\lim_{\tau\to\infty} \frac{1-H\big(f_\tau(0)+s(1-f_\tau(0))\big)}{(1-s)\big(1-f_\tau(0)\big)} = H'(1) = \frac{1}{\varphi(0)}.$$

Combining this with (1.3), first as $n \to \infty$ and then as $\tau \to \infty$ we have

$$\mathbf{E}\big[s^{\hat{N}_{n-\tau,n}}\big] \to s \tag{2.12}$$

holds when $\sum (j\log j)p_j < \infty$. This conclusion means that $\hat{N}_{n-\tau,n}$ converges to 1 in law when first $n\to\infty$ and then $\tau\to\infty$, hence it can be roughly thought that the most recent common ancestor of particles in generation n is located near n (c.f. Fleischmann and Prehn, 1974). This observation is very important in our proof for Theorem 1.3.

3. Conditional reduced GW-trees and the many-to-few formula

In order to introduce the decomposition of Z_n and the many-to-few formula in detail, we now depict the above processes in family trees.

3.1. *GW-trees.* We use the classical Ulam-Harris-Neveu notation for discrete trees. Let

$$\mathcal{U} = \{\varnothing\} \cup \bigcup_{n\ge 1} (\mathbb{N}_+)^n.$$

As before, for $u = u_1 \cdots u_n \in \mathcal{U}$, we denote by $|u| = n$ the generation of u. If $u = u_1\cdots u_n \in \mathcal{U}, v = v_1 \cdots v_l \in \mathcal{U}$, we write $uv = u_1 \cdots u_n v_1 \cdots v_l$ for the concatenation of u and v. In particular, $u\varnothing = \varnothing u = u$.

A *tree* t is a subset of $\mathcal{U}$ satisfying the following properties: (1) $\varnothing \in t$; (2) if $uj \in t$ for some $j \in \mathbb{N}_+$, then $u \in t$; (3) if $u \in t$, then $uj \in t$ if and only if $1 \le j \le N_u(t)$ for some non-negative integer $N_u(t) < \infty$. In words, $N_u := N_u(t)$ is the number of children of the vertex u in t.

Let us consider a family of $\mathbb{N}$-valued random variables $\{N_u; u \in \mathcal{U}\}$ such that under $\mathbf{P}$, they are independent of each other and have the same law p. We say a random subset $\mathbb{T}$ of $\mathcal{U}$ is a *Galton-Watson tree* (GW-tree) with offspring distribution p (rooted at $\varnothing$) if it is defined by

$$\mathbb{T} := \{u = u_1 \cdots u_n \in \mathcal{U} : 1 \le u_j \le N_{u_1\cdots u_{j-1}}, \text{ for every } 1 \le j \le n\}.$$

3.2. *Marked trees and BRWs.* The main purpose of this section is to produce an extension, by associating each vertex of the tree with the position. We define $\mathcal{V} := \{(u,x) : u \in \mathcal{U}, x \in \mathbb{R}\}$ and call $\mathbf{t}$ the *marked tree* if $\mathbf{t} := \{(u, V(u)) : u \in t\}$ is a subset of $\mathcal{V}$, see Harris and Roberts (2017).

Let $\mathcal{T}$ be the set of all marked trees. We now take a probability $\mathbf{P}_x$ on $\mathcal{T}$ such that the system evolves as a branching random walk starting with one particle at x, in which each particle has children with the total number and positions (with respect to their parent) decided by distributions p and G, respectively. When $x = 0$, this is the system described in Section 1.1, and we drop the subscript in this case. Let $\mathbb{T}$ denote the genealogical tree of this system rooted at $\varnothing$. Clearly, $\mathbb{T}$ is a GW-tree with offspring distribution p. Therefore, our BRW is defined by the random measure

$$Z_n = \sum_{u\in\mathbb{T}:|u|=n} \delta_{V(u)}.$$

Since $Z_n(A)$ is the number of particles in the n-th generation located in A, so we have

$$Z_n(A) = \#\{u \in \mathbb{T} : V(u) \in A, |u| = n\},$$

where $\#B$ denotes the cardinality of the set B.

3.3. *Conditional reduced GWM-trees and conditional reduced BRWs.* Recall that under $\mathbf{P}_x$, $\mathbb{T}$ is the GW-tree with offspring distribution p, which starts with one particle at $V(\varnothing) = x$. From the definition in Fleischmann and Siegmund-Schultze (1977), we can get the *reduced GW-tree* (rooted at $\varnothing$) by removing all branches of the original tree $\mathbb{T}$, which don't extend to generation n. We write $\hat{\mathbb{T}}_n$ for the random tree rooted at $\varnothing$, whose distribution is equal to the conditional distribution of this reduced GW-tree given $N_n > 0$.

For any $0 \le k \le n$, let

$$\hat{\mathbb{T}}_{k,n} := \{u \in \hat{\mathbb{T}}_n : |u| = k\}$$

be the set of all individuals in generation k in the tree $\hat{\mathbb{T}}_n$, namely, $\hat{\mathbb{T}}_n = \cup_{k=0}^n \hat{\mathbb{T}}_{k,n}$. Then $\hat{N}_{k,n} := \#\hat{\mathbb{T}}_{k,n}$ is the total size of individuals in generation k in $\hat{\mathbb{T}}_n$, so we are back to the conditional reduced GW-process described in Section 2. We next transform the random tree $\hat{\mathbb{T}}_n$ to the *conditional reduced Galton-Watson tree* (conditional reduced GW-tree).

From the definition of a marked tree in Section 3.2, we can get a random marked tree $\{(u, V(u)) : u \in \hat{\mathbb{T}}_n, V(u) \in \mathbb{R}\}$ by endowing each u with the position $V(u)$ in the following way: for each $u \in \hat{\mathbb{T}}_n$, children are born at distances from the parent u which are given by an independent copy of the law G. This random marked tree is called a *conditional reduced Galton-Watson marked tree* (conditional reduced GWM-tree). Accordingly, we say the random measure

$$\hat{Z}_n = \sum_{u\in\hat{\mathbb{T}}_{n,n}} \delta_{V(u)}$$

defines a *conditional reduced branching random walk* (conditional reduced BRW). Note that $\hat{Z}_n$ is the point process with law $\mathcal{L}(Z_n \mid N_n > 0)$.

3.3.1. *Subtrees in the conditional reduced GW-tree $\hat{\mathbb{T}}_n$.* For any $z \in \hat{\mathbb{T}}_n$, let $\hat{\mathbb{T}}_n(z)$ be the subtree of $\hat{\mathbb{T}}_n$ rooted at z defined by

$$\hat{\mathbb{T}}_n(z) := \{v : zv \in \hat{\mathbb{T}}_n\}.$$

For $|z| \le k \le n$, we write

$$\hat{\mathbb{T}}_{k-|z|,n}(z) := \{u \in \hat{\mathbb{T}}_n(z) : |zu| = k\},$$
$$\hat{N}_{k-|z|,n}(z) := \#\hat{\mathbb{T}}_{k-|z|,n}(z),$$

then $\hat{\mathbb{T}}_n(z) = \cup_{k=|z|}^n \hat{\mathbb{T}}_{k-|z|,n}(z)$. In particular, $\hat{\mathbb{T}}_n(\varnothing) = \hat{\mathbb{T}}_n$, $\hat{\mathbb{T}}_{k,n}(\varnothing) = \hat{\mathbb{T}}_{k,n}$, $\hat{N}_{k,n}(\varnothing) = \hat{N}_{k,n}$, $\hat{N}_{n,n}(\varnothing) = \hat{N}_n$.

It follows from Fleischmann and Siegmund-Schultze (1977, Proposition 1.1) that $\mathcal{L}(\hat{N}_{1,n}(z)) = \mathcal{L}(\hat{N}_{1,n-|z|}(\varnothing))$. We get an analogous result, i.e., Lemma 3.1, which is another important step in our proof for the main result.

Lemma 3.1. *For any $z \in \hat{\mathbb{T}}_n$ and $|z| \le k \le n$, we have*

$$\mathcal{L}\big(\hat{N}_{k-|z|,n}(z)\big) = \mathcal{L}\big(\hat{N}_{k-|z|,n-|z|}(\varnothing)\big). \tag{3.1}$$

Proof: For fixed $z \in \hat{\mathbb{T}}_n$, since $\hat{\mathbb{T}}_n(z)$ and $\hat{\mathbb{T}}_{n-|z|}$ both are time-inhomogeneous GW-trees starting with one particle, we just need to prove that their offspring distributions at time k ($|z| \le k \le n-1$) are the same. On the one hand, based on the observation that the $(k-|z|)$-th generation of the subtree $\hat{\mathbb{T}}_n(z)$ actually corresponds to the k-th generation of the tree $\hat{\mathbb{T}}_n$, so the offspring probability generating function at time $(k-|z|)$ in $\hat{\mathbb{T}}_n(z)$ is $\hat{f}_{\mathbf{e}_k(n)}(s)$ defined by (2.1). On the other hand, $\hat{f}_{\mathbf{e}_{k-|z|}(n-|z|)}(s)$ defined by (2.1) is the offspring probability generating function at time $k-|z|$ in $\hat{\mathbb{T}}_{n-|z|}$. By using (2.1) again, a simple calculation shows that $\hat{f}_{\mathbf{e}_k(n)}(s) = \hat{f}_{\mathbf{e}_{k-|z|}(n-|z|)}(s)$. This implies the desired result. □

3.3.2. *The many-to-few formula.* For fixed $r \in \mathbb{N}_+$, we now attach r additional distinguished lines of descent $\omega^1, \cdots, \omega^r$ to the conditional reduced GWM-tree $\{(u, V(u)) : u \in \hat{\mathbb{T}}_n, V(u) \in \mathbb{R}\}$, which are called *spines*.

Note that $\mathbf{P}_x$ is the probability measure on $\mathcal{T}$ such that under $\mathbf{P}_x$ the system evolves as a subcritical BRW starting from an ancestor located at x defined in Section 3.2. We next introduce the following new system, which is a conditional reduced BRW with r spines $\omega^1, \cdots, \omega^r$:

(1) Initially, there is one particle at position x which carries r marks $1, \cdots, r$.
(2) For each $1 \le i \le r$, we regard the line of descent carrying mark i as the spine ω^i. Note that $\omega^i = (\omega_1^i, \omega_2^i, \cdots)$, where ω_k^i denotes the particle carrying mark i in generation k. We write $\mathfrak{X}_k^i$ for its position, i.e., $\mathfrak{X}_k^i = V(\omega_k^i)$.
(3) A particle in generation k carrying j marks gives birth to children with the j-th *sized-biased* distribution $\big\{\frac{l^j p_l(\mathbf{e}_k(n))}{m_j(\mathbf{e}_k(n))}\big\}_{l\in\mathbb{N}}$, these children are born at distances from their parent which are given by an independent copy of the law G. Given that a particles $u_1, \cdots, u_a$ are born at such a branching event, each of the j spines chooses a particle to follow independently and uniformly at random among the a variables.
(4) Particles which carry no marks in generation k have children according to the offspring distribution $\{p_l(\mathbf{e}_k(n))\}_{l\in\mathbb{N}}$ given by (2.2) and the jump distribution G, just as under $\mathbf{P}_x$.
(5) The offspring of particles in generation k and their positions (including the marked particles) form the $(k+1)$-th generation.

Let us denote by $\mathbf{Q}_x^{[r]}$ the law of the new system defined above. In particular, we drop the subscript when $x=0$. In other words, under $\mathbf{Q}_x^{[r]}$, spine particles give birth to children according to the size-biased distribution, but the jump distribution will be unchanged. The number of children depends on how many marks the spine particle carrying, but the motion does not.

For any $u \in \hat{\mathbb{T}}_n \backslash \{\varnothing\}$, we denote by $\overleftarrow{u}$ its parent. Let $D(u)$ be the total number of marks carried by a particle u. For fixed $n \in \mathbb{N}$ and $0 \le k \le n$, let $\big(\hat{\mathcal{F}}_{k,n}, 0 \le k \le n\big)$ be the natural filtration of the conditional reduced BRW. Let $\big(\mathcal{G}_{k,n}^{[r]}, 0 \le k \le n\big)$ be the filtration containing all information about the conditional reduced BRW and the r spines. It follows from Harris and Roberts (2017) that if

Y is measurable with respect to $\mathcal{G}_{k,n}^{[r]}$, then it can be expressed as

$$Y = \sum_{u_1,\cdots,u_r \in \hat{\mathbb{T}}_{k,n}} Y(u_1,\cdots,u_r)\mathbf{1}_{\{\omega_k^1=u_1,\cdots,\omega_k^r=u_r\}},$$

where for each $u_1,\cdots,u_r \in \hat{\mathbb{T}}_{k,n}$ the random variable $Y(u_1,\cdots,u_r)$ is $\hat{\mathcal{F}}_{k,n}$-measurable. Let $\mathbf{E}_x$ be the expectation corresponding to $\mathbf{P}_x$.

The many-to-few formula is convenient to compute higher-order moments. It is standard in many other related contexts, for example, Bansaye (2019); Bansaye and Huang (2015); Shi (2015) for the many-to-one formula. We state a version of the many-to-few formula for time-inhomogeneous BRWs. The proof is similar to that in Harris and Roberts (2017) and Hong and Liang (2023). In this paper, we just need to use the many-to-one formula and the many-to-two formula.

Lemma 3.2. (Many-to-few formula) *For each* $0 \le k \le n$, $r \in \mathbb{N}_+$ *and each* $\mathcal{G}_{k,n}^{[r]}$*-measurable* Y*, we have*

$$\mathbf{E}_x\Bigg[\sum_{u_1,\cdots,u_r \in \hat{\mathbb{T}}_{k,n}} Y(u_1,\cdots,u_r)\Bigg] = \mathbf{Q}_x^{[r]}\Bigg[Y(\omega_k^1,\cdots,\omega_k^r)\prod_{\omega \in \mathrm{skel}(k)} m_{D(\overleftarrow{\omega})}\big(\mathbf{e}_{|\overleftarrow{\omega}|}(n)\big)\Bigg], \tag{3.2}$$

where $\mathrm{skel}(k)$ denotes the set of particles that have carried at least one mark up to time k in the tree $\hat{\mathbb{T}}_n$, and $D(u)$ is the total number of marks carried by a particle u.

4. Proof of Theorem 1.3

4.1. *The decomposition of* $\hat{Z}_{n+\tau}$ *and proof of Theorem 1.3.* Recall that $\hat{Z}_n$ is the point process with law $\mathcal{L}(Z_n \mid N_n > 0)$. For random variables X and Y, we write $X \overset{d}{=} Y$ if X is equal in distribution to Y. For any fixed $n,\tau \in \mathbb{N}$, we can decompose $\hat{Z}_{n+\tau}$ at generation n as

$$\begin{aligned}\hat{Z}_{n+\tau}((-\infty,\sqrt{n+\tau}x]) &\overset{d}{=} \sum_{y \in \hat{\mathbb{T}}_{n+\tau,n+\tau}} \mathbf{1}_{\{V(y)\le\sqrt{n+\tau}x\}} \\ &= \sum_{z \in \hat{\mathbb{T}}_{n,n+\tau}} \Bigg[\sum_{y \in \hat{\mathbb{T}}_{\tau,n+\tau}(z)} \mathbf{1}_{\{\Delta V(z,y)\le\sqrt{n+\tau}x - V(z)\}}\Bigg] \\ &:= A_{n+\tau} + B_{n+\tau},\end{aligned} \tag{4.1}$$

where $A_{n+\tau}$ and $B_{n+\tau}$ are given respectively in (4.2) and (4.3), and $\Delta V(z,y) := V(zy) - V(z)$. By Lemma 3.1 and the definition of $\hat{\mathbb{T}}_{n+\tau}$, under $\mathbf{P}_x$ these subtrees in (4.1) are characterized by the following properties:

(1) Given the information about $(V(z), z \in \hat{\mathbb{T}}_{n,n+\tau})$, the subtrees $\big(\Delta V(z,y), y \in \hat{\mathbb{T}}_{n+\tau}(z)\big), z \in \hat{\mathbb{T}}_{n,n+\tau}$ are independent, identically distributed copies of the random tree $\big(V(y), y \in \hat{\mathbb{T}}_\tau\big)$;

(2) For each $z \in \hat{\mathbb{T}}_{n,n+\tau}$, the subtree $\big(\Delta V(z,y), y \in \hat{\mathbb{T}}_{n+\tau}(z)\big)$ is independent of $\big(V(z), z \in \hat{\mathbb{T}}_{k,n+\tau}, 0 \le k \le n\big)$.

The first term in (4.1) can be written as

$$A_{n+\tau} := \sum_{z \in \hat{\mathbb{T}}_{n,n+\tau}} \Bigg[\Big(\sum_{y \in \hat{\mathbb{T}}_{\tau,n+\tau}(z)} \mathbf{1}_{\{\Delta V(z,y)\le\sqrt{n+\tau}x - V(z)\}}\Big) - \hat{N}_{\tau,n+\tau}(z)\mathbf{1}_{\{\widetilde{\mathcal{N}}_z \le \sqrt{n+\tau}x - V(z)\}}\Bigg], \tag{4.2}$$

where $\hat{N}_{\tau,n+\tau}(z)$ is the number of individuals in the τ-th generation of the subtree $\hat{\mathbb{T}}_{n+\tau}(z)$ rooted at the vertex z. We introduce a new sequence of random variables $\{\widetilde{\mathcal{N}}_z; z \in \hat{\mathbb{T}}_{n,n+\tau}\}$ in (4.2). Then by Lemma 3.1 and the properties of subtrees, under $\mathbf{P}_x$ these random variables in $A_{n+\tau}$ satisfy:

(1) Given the information about $\hat{\mathbb{T}}_{n,n+\tau}$, the random variables $\hat{N}_{\tau,n+\tau}(z), z \in \hat{\mathbb{T}}_{n,n+\tau}$ are independent, identically distributed copies of $\hat{N}_\tau$;
(2) Given the information about $\hat{\mathbb{T}}_{n,n+\tau}$, $\{\widetilde{\mathcal{N}}_z; z \in \hat{\mathbb{T}}_{n,n+\tau}\}$ is a sequence of independent and identically distributed random variables with the same law $G^{\tau *}$;
(3) Given the information about $\hat{\mathbb{T}}_{n,n+\tau}$, the two families of random variables $\{\widetilde{\mathcal{N}}_z; z \in \hat{\mathbb{T}}_{n,n+\tau}\}$ and $\{\hat{N}_{\tau,n+\tau}(z); z \in \hat{\mathbb{T}}_{n,n+\tau}\}$ are independent.
(4) For each $z \in \hat{\mathbb{T}}_{n,n+\tau}$, the two random variables $\widetilde{\mathcal{N}}_z$ and $\hat{N}_{\tau,n+\tau}(z)$ are independent of $(V(z), z \in \hat{\mathbb{T}}_{k,n+\tau}, 0 \le k \le n)$.

We remark that our assumption on the branching and motion mechanisms guarantees that the two families of random variables $\{\widetilde{\mathcal{N}}_z; z \in \hat{\mathbb{T}}_{n,n+\tau}\}$ and $\{\hat{N}_{\tau,n+\tau}(z); z \in \hat{\mathbb{T}}_{n,n+\tau}\}$ are independent. For the second term in (4.1),

$$B_{n+\tau} := \sum_{z \in \hat{\mathbb{T}}_{n,n+\tau}} \hat{N}_{\tau,n+\tau}(z) \mathbf{1}_{\{\widetilde{\mathcal{N}}_z \le \sqrt{n+\tau}x - V(z)\}}. \tag{4.3}$$

The intuitive idea of this decomposition of (4.1) is described as follows. Firstly, (2.12) indicates that the most recent common ancestor of particles in the $(n+\tau)$-th generation is located near n when first $n \to \infty$ and then $\tau \to \infty$, so we decompose the sum at generation n in the second equality of (4.1). Secondly, for fixed z, although random variables $\{\Delta V(z,y); y \in \hat{\mathbb{T}}_{\tau,n+\tau}(z)\}$ are correlated given $V(z)$, they all tend to 0 by the space scaling $\sqrt{n+\tau}$ as $n \to \infty$, i.e., we can roughly think that each $\Delta V(z,y)$ behaves as the same variable $\widetilde{\mathcal{N}}_z$. Consequently, when first $n \to \infty$ and then $\tau \to \infty$, the equivalent relation

$$\sum_{z \in \hat{\mathbb{T}}_{n,n+\tau}} \left[\sum_{y \in \hat{\mathbb{T}}_{\tau,n+\tau}(z)} \mathbf{1}_{\{\Delta V(z,y) \le \sqrt{n+\tau}x - V(z)\}} \right] \overset{d}{\sim} B_{n+\tau} := \sum_{z \in \hat{\mathbb{T}}_{n,n+\tau}} \hat{N}_{\tau,n+\tau}(z) \mathbf{1}_{\{\widetilde{\mathcal{N}}_z \le \sqrt{n+\tau}x - V(z)\}}$$

holds in distribution. Finally, Lemma 3.1 tells us that $\hat{N}_{\tau,n+\tau}(z) \overset{d}{=} \hat{N}_\tau$. It follows from (2.12) that $\hat{N}_{n,n+\tau} \approx 1$ when first $n \to \infty$ and then $\tau \to \infty$. Thus we obtain the following equivalent relation

$$B_{n+\tau} \overset{d}{\sim} \sum_{z \in \hat{\mathbb{T}}_{n,n+\tau}} \hat{N}_\tau \mathbf{1}_{\{\widetilde{\mathcal{N}}_z \le \sqrt{n+\tau}x - V(z)\}} \overset{d}{\sim} \hat{N}_\tau \mathbf{1}_{\left\{\frac{V(z)}{\sqrt{n+\tau}} \le x - \frac{\widetilde{\mathcal{N}}_z}{\sqrt{n+\tau}}\right\}} \overset{d}{\sim} \xi \mathbf{1}_{\{\mathcal{N} \le x\}}$$

holds in distribution, where $\mathcal{N}$ is a standard normal random variable independent of the Yaglom limit ξ.

Based on (4.1), we can prove Theorem 1.3 with the following two Propositions 4.1 and 4.2, whose proofs are postponed to Sections 4.2 and 4.3, respectively.

Proposition 4.1. *Under the assumptions (1.5) and (1.6), we have*

$$\limsup_{n \to \infty} \mathbf{E}[A_{n+\tau}^2] \le h_1(\tau),$$

where $h_1(\tau) \to 0$ as $\tau \to \infty$.

Let the generating function of $\xi \mathbf{1}_{\{\mathcal{N} \le x\}}$ be $L(s)$. Recall that ξ is independent of $\mathcal{N}$, then a simple calculation shows that

$$L(s) = \Phi(x) H(s) + 1 - \Phi(x). \tag{4.4}$$

Proposition 4.2. *Assume (1.5) and (1.6) hold, then for all $s \in [0,1]$, we have*

$$\limsup_{n \to \infty} \left| \mathbf{E}\left[s^{B_{n+\tau}} \right] - L(s) \right| \le h_2(\tau, s),$$

where $h_2(\tau, s) \to 0$ as $\tau \to \infty$.

Now we can prove Theorem 1.3 at once with Propositions 4.1 and 4.2 in hand.

Proof of Theorem 1.3. The generating function of $\hat{Z}_n((-\infty,\sqrt{n}x])$ is defined by

$$L_n(s) := \mathbf{E}\Big[s^{\hat{Z}_n((-\infty,\sqrt{n}x])}\Big], \quad s\in[0,1].$$

For $s\in(0,1)$, we again use the decomposition (4.1) at generation n to write

$$\begin{aligned}&\big|L_{n+\tau}(s)-L(s)\big|\\ \le&\Big|\mathbf{E}\big[s^{A_{n+\tau}+B_{n+\tau}}\big]-\mathbf{E}\big[s^{B_{n+\tau}}\big]\Big|+\Big|\mathbf{E}\big[s^{B_{n+\tau}}\big]-\mathbf{E}\big[s^{\xi\mathbf{1}_{\{\mathcal{N}\le x\}}}\big]\Big|\\ \le&|\ln s|\cdot\mathbf{E}\Big[\big|A_{n+\tau}\big|\Big]+\Big|\mathbf{E}\big[s^{B_{n+\tau}}\big]-\mathbf{E}\big[s^{\xi\mathbf{1}_{\{\mathcal{N}\le x\}}}\big]\Big|,\end{aligned}$$

where the last inequality comes from the fact that for integers k_1 and k_2 satisfying $k_1+k_2\ge 0, k_2\ge 0$, we have $\big|s^{k_1+k_2}-s^{k_2}\big|\le\big|\ln s\big|\cdot\big|k_1\big|$ by Lagrange's mean value theorem. By Hölder's inequality and Proposition 4.1, we have

$$\mathbf{E}\big[|A_{n+\tau}|\big]\le\big(\mathbf{E}[A_{n+\tau}^2]\big)^{1/2}\le\big(h_1(\tau)\big)^{1/2}.$$

Proposition 4.2 implies

$$\limsup_{n\to\infty}\Big|\mathbf{E}\big[s^{B_{n+\tau}}\big]-L(s)\Big|\le h_2(\tau,s).$$

Combining the last three inequalities yields

$$\limsup_{n\to\infty}\big|L_{n+\tau}(s)-L(s)\big|\le\big|\ln s\big|\cdot\big(h_1(\tau)\big)^{1/2}+h_2(\tau,s).$$

By replacing $n+\tau$ by n in the above inequality, we have

$$\limsup_{n\to\infty}\big|L_n(s)-L(s)\big|\le\big|\ln s\big|\cdot\big(h_1(\tau)\big)^{1/2}+h_2(\tau,s), \quad s\in(0,1).$$

Let $\tau\to\infty$, we obtain $L_n(s)\to L(s)$ by the properties of $h_1(\tau)$ and $h_2(\tau,s)$. Note that the continuity of $H(s)$ at $s=1$ implies the continuity of $L(s)$ at $s=1$, the proof is completed by the continuity theorem of generating functions (c.f. Feller, 1968, Continuity theorem, Section XI.6). ☐

In what follows, we focus on proving Propositions 4.1 and 4.2.

4.2. *Proof of Proposition 4.1.* To prove Proposition 4.1, we should calculate the first and second moments for the conditional reduced BRW $(V(z), z\in\hat{\mathbb{T}}_{n+\tau})$ by the many-to-few formula (c.f. Lemma 3.2). We use the convention that for any integer $k\le 0, \sum_{j=1}^{k}=0$.

Lemma 4.3. *For every $n,\tau\in\mathbb{N}$,*

(1) *For each integer $0\le k\le n+\tau$, $x\in\mathbb{R}$ and each Borel measurable function g on $\mathbb{R}$, we have*

$$\mathbf{E}_x\Bigg[\sum_{u\in\hat{\mathbb{T}}_{k,n+\tau}}g\big(V(u)\big)\Bigg]=m^k\frac{1-f_{n+\tau-k}(0)}{1-f_{n+\tau}(0)}\mathbf{E}_x\big[g(\mathfrak{X}_k^1)\big], \tag{4.5}$$

where $(\mathfrak{X}_i^1, i\in\mathbb{N})$ is the spine random walk carrying one mark with transition probability

$$\mathbf{P}_x(\mathfrak{X}_{i+1}^1\in\cdot\mid\mathfrak{X}_i^1=x)=G(\cdot-x), \quad i\in\mathbb{N}. \tag{4.6}$$

(2) *For any* $a \in \mathbb{R}$,

$$\begin{aligned}
&\mathbf{E}\Big[\hat{Z}_\tau((-\infty, \sqrt{n+\tau}x - a])^2\Big] \\
=&\frac{m^\tau}{1-f_\tau(0)}\mathbf{P}\big(\mathfrak{X}^1_\tau \le \sqrt{n+\tau}x - a\big) \\
&+ \frac{m^{2\tau-1}}{1-f_\tau(0)}f''(1)\sum_{j=1}^{\tau} m^{-j}\mathbf{P}\big(\mathfrak{X}^1_{j-1,\tau} \le \sqrt{n+\tau}x - a, \mathfrak{X}^2_{j-1,\tau} \le \sqrt{n+\tau}x - a\big),
\end{aligned} \tag{4.7}$$

where $(\mathfrak{X}^1_i, i \in \mathbb{N})$ *is defined as in* (1), the integer j is the first split time of the two spines. For each j, $(\mathfrak{X}^1_{j,i}, i \in \mathbb{N})$ and $(\mathfrak{X}^2_{j,i}, i \in \mathbb{N})$ are spine random walks, which have the same transition probability as (4.6) and satisfy

- $\mathfrak{X}^1_{j,i} = \mathfrak{X}^2_{j,i}$ for all $i \le j$;
- $(\mathfrak{X}^1_{j,j+i} - \mathfrak{X}^1_{j,j}, i \in \mathbb{N})$ and $(\mathfrak{X}^2_{j,j+i} - \mathfrak{X}^2_{j,j}, i \in \mathbb{N})$ are independent.

Proof: It can be verified by the many-to-few formula in Lemma 3.2. More specifically, let $r = 1$ and replace n by $n+\tau$ in (3.2). Then we take $Y(\mathfrak{X}^1_k) = g(\mathfrak{X}^1_k)$, which implies $Y(u) = g(V(u))$ for $u \in \hat{\mathbb{T}}_{k,n+\tau}$. Note that

$$\prod_{w\in \mathrm{skel}(k)} m_{D(\varpi)}\big(\mathbf{e}_{|\varpi|}(n+\tau)\big) = \prod_{i=0}^{k-1} m_1\big(\mathbf{e}_i(n+\tau)\big) = m^k\frac{1-f_{n+\tau-k}(0)}{1-f_{n+\tau}(0)},$$

where the last equality comes from (2.3). Note that the motion of the spine under $\mathbf{Q}^{[1]}_x$ is the same as under $\mathbf{P}_x$, it is easy to conclude (1) from (3.2).

We next give the proof of (2). In (3.2), let $r = 2$, $x = 0$, $n = k = \tau$ and

$$Y(\mathfrak{X}^1_\tau, \mathfrak{X}^2_\tau) = \mathbf{1}_{\{\mathfrak{X}^1_\tau \le \sqrt{n+\tau}x - a, \mathfrak{X}^2_\tau \le \sqrt{n+\tau}x-a\}}.$$

Clearly,

$$Y(u_1,u_2) = \mathbf{1}_{\{V(u_1)\le\sqrt{n+\tau}x-a, V(u_2)\le\sqrt{n+\tau}x-a\}}, \quad u_1,u_2 \in \hat{\mathbb{T}}_{\tau,\tau}.$$

Suppose that T denotes the first split time of the two spines, i.e., the first time at which marks 1 and 2 are carried by different particles. Consequently,

$$\begin{aligned}
&\mathbf{E}\Big[\hat{Z}_\tau((-\infty, \sqrt{n+\tau}x - a])^2\Big] \\
=&\mathbf{E}\Big[\sum_{u_1,u_2\in\hat{\mathbb{T}}_{\tau,\tau}} \mathbf{1}_{\{V(u_1)\le\sqrt{n+\tau}x-a, V(u_2)\le\sqrt{n+\tau}x-a\}}\Big] \\
=&\mathbf{Q}^{[2]}\Big[\mathbf{1}_{\{\mathfrak{X}^1_\tau\le\sqrt{n+\tau}x-a, \mathfrak{X}^2_\tau\le\sqrt{n+\tau}x-a\}} \prod_{\omega\in\mathrm{skel}(\tau)} m_{D(\varpi)}\big(\mathbf{e}_{|\varpi|}(\tau)\big)\Big] \\
=&\mathbf{Q}^{[2]}\Big[\mathbf{1}_{\{\mathfrak{X}^1_\tau\le\sqrt{n+\tau}x-a\}}\mathbf{1}_{\{T>\tau\}}\prod_{i=0}^{\tau-1} m_2\big(\mathbf{e}_i(\tau)\big)\Big] \\
&+\sum_{j=1}^{\tau}\mathbf{Q}^{[2]}\Big[\mathbf{1}_{\{\mathfrak{X}^1_{j-1,\tau}\le\sqrt{n+\tau}x-a, \mathfrak{X}^2_{j-1,\tau}\le\sqrt{n+\tau}x-a\}}\mathbf{1}_{\{T=j\}}\prod_{i=0}^{j-1} m_2\big(\mathbf{e}_i(\tau)\big)\prod_{i=j}^{\tau-1} m_1\big(\mathbf{e}_i(\tau)\big)^2\Big],
\end{aligned} \tag{4.8}$$

where $(\mathfrak{X}^1_i, i\in\mathbb{N})$, $(\mathfrak{X}^1_{j,i}, i\in\mathbb{N})$ and $(\mathfrak{X}^2_{j,i}, i\in\mathbb{N})$ are spine random walks defined in Lemma 4.3.

Under $\mathbf{Q}^{[2]}$ the particle in generation 0 which carries marks 1 and 2 splits into l particles with probability $l^2p_l(\mathbf{e}_0(\tau))/m_2(\mathbf{e}_0(\tau))$. At such a branching event, the two marks follow the same particle with probability $1/l$. Thus

$$\mathbf{Q}^{[2]}\big(T>1\big)=\frac{m_1\big(\mathbf{e}_0(\tau)\big)}{m_2\big(\mathbf{e}_0(\tau)\big)}.$$

We can get the distribution of T in a similar way. Indeed, the event $\{T>j\}$ means that the two marks don't split until time j. In addition, under $\mathbf{Q}^{[2]}$ the particle carrying two marks branches into l children at time $i<j$ with law $l^2 p_l(\mathbf{e}_i(\tau))/m_2(\mathbf{e}_i(\tau))$ then the two marks follow the same particle with probability $1/l$. Hence the law of T is given by

$$\mathbf{Q}^{[2]}\big(T>j\big)=\prod_{i=0}^{j-1}\frac{m_1\big(\mathbf{e}_i(\tau)\big)}{m_2\big(\mathbf{e}_i(\tau)\big)},\quad j=1,2,\cdots. \tag{4.9}$$

Hence, for any integer $j\geq 2$,

$$\mathbf{Q}^{[2]}\big(T=j\big)=\mathbf{Q}^{[2]}\big(T>j-1\big)-\mathbf{Q}^{[2]}\big(T>j\big)=\prod_{i=0}^{j-2}\frac{m_1\big(\mathbf{e}_i(\tau)\big)}{m_2\big(\mathbf{e}_i(\tau)\big)}\bigg(1-\frac{m_1\big(\mathbf{e}_{j-1}(\tau)\big)}{m_2\big(\mathbf{e}_{j-1}(\tau)\big)}\bigg). \tag{4.10}$$

Notice that $\mathbf{Q}^{[2]}(T=1)=1-\mathbf{Q}^{[2]}(T>1)$, by using the convention that $\prod_{i=0}^{k}=1$ for any integer $k<0$, then (4.10) holds for $j=1$. By (2.4) and (4.9), a simple calculation shows that

$$\prod_{i=0}^{\tau-1}m_2(\mathbf{e}_i(\tau))\mathbf{Q}^{[2]}(T>\tau)=\frac{m^\tau}{1-f_\tau(0)}.$$

Again using the fact that the motions of spines under $\mathbf{Q}^{[2]}$ are the same as under $\mathbf{P}$. Note that $\mathfrak{X}_\tau^1$ is independent of T under $\mathbf{Q}^{[2]}$, so we obtain

$$\mathbf{Q}^{[2]}\bigg(\mathbf{1}_{\{\mathfrak{X}_\tau^1\leq\sqrt{n+\tau}x-a\}}\mathbf{1}_{\{T>\tau\}}\prod_{i=0}^{\tau-1}m_2\big(\mathbf{e}_i(\tau)\big)\bigg)=\frac{m^\tau}{1-f_\tau(0)}\mathbf{P}(\mathfrak{X}_\tau^1\leq\sqrt{n+\tau}x-a),$$

which derives the first part on the right-hand side of (4.7). We next consider the remainder of the sum in (4.7). By (2.3) and (2.4) we get

$$m_2\big(\mathbf{e}_{j-1}(\tau)\big)-m_1\big(\mathbf{e}_{j-1}(\tau)\big)=\frac{\big(1-f_{\tau-j}(0)\big)^2}{1-f_{\tau-j+1}(0)}f''(1).$$

Moreover, it follows from (2.3), (2.4) and (4.10) that

$$\prod_{i=0}^{j-1}m_2\big(\mathbf{e}_i(\tau)\big)\prod_{i=j}^{\tau-1}m_1\big(\mathbf{e}_i(\tau)\big)^2\mathbf{Q}^{[2]}\big(T=j\big)=\frac{m^{2\tau-1}}{1-f_\tau(0)}f''(1)m^{-j}.$$

Therefore, the second part on the right-hand side of (4.8) is

$$\frac{m^{2\tau-1}}{1-f_\tau(0)}f''(1)\sum_{j=1}^{\tau}m^{-j}\mathbf{P}(\mathfrak{X}_{j-1,\tau}^1\leq\sqrt{n+\tau}x-a,\mathfrak{X}_{j-1,\tau}^2\leq\sqrt{n+\tau}x-a).$$

Thus we prove Lemma 4.3. □

Proof of Proposition 4.1. Recall that $\hat{\mathcal{F}}_{n,n+\tau}=\sigma(V(z):z\in\hat{\mathbb{T}}_{k,n+\tau},0\leq k\leq n)$ is defined in Section 3.3.2. We write $\hat{\mathbf{E}}_{n,n+\tau}[\cdot]:=\mathbf{E}[\cdot\mid\hat{\mathcal{F}}_{n,n+\tau}]$ for simplicity. From (4.2) we have

$$\begin{aligned}\mathbf{E}[A_{n+\tau}^2]=&\mathbf{E}\bigg[\hat{\mathbf{E}}_{n,n+\tau}\bigg[\bigg(\sum_{z\in\hat{\mathbb{T}}_{n,n+\tau}}\bigg(\sum_{y\in\hat{\mathbb{T}}_{\tau,n+\tau}(z)}\mathbf{1}_{\{\Delta V(z,y)\leq\sqrt{n+\tau}x-V(z)\}}-\hat{N}_{\tau,n+\tau}(z)\mathbf{1}_{\{\widetilde{\mathcal{N}}_z\leq\sqrt{n+\tau}x-V(z)\}}\bigg)\bigg)^2\bigg]\bigg]\\ :=&D_{n+\tau}^{(1)}+D_{n+\tau}^{(2)},\end{aligned}$$

where

$$D^{(1)}_{n+\tau}=\mathbf{E}\Big[\sum_{z\in\hat{\mathbb{T}}_{n,n+\tau}}\hat{\mathbf{E}}_{n,n+\tau}\Big[\Big(\sum_{y\in\hat{\mathbb{T}}_{\tau,n+\tau}(z)}\mathbf{1}_{\{\Delta V(z,y)\le\sqrt{n+\tau}x-V(z)\}}-\hat{N}_{\tau,n+\tau}(z)\mathbf{1}_{\{\widetilde{\mathcal{N}}_z\le\sqrt{n+\tau}x-V(z)\}}\Big)^2\Big]\Big],$$

$$D^{(2)}_{n+\tau}=\mathbf{E}\Big[\sum_{\substack{z_1,z_2\in\hat{\mathbb{T}}_{n,n+\tau}\\ z_1\neq z_2}}\hat{\mathbf{E}}_{n,n+\tau}\Big[\prod_{i=1,2}\Big(\sum_{y\in\hat{\mathbb{T}}_{\tau,n+\tau}(z_i)}\mathbf{1}_{\{\Delta V(z_i,y)\le\sqrt{n+\tau}x-V(z_i)\}}-\hat{N}_{\tau,n+\tau}(z_i)\mathbf{1}_{\{\widetilde{\mathcal{N}}_{z_i}\le\sqrt{n+\tau}x-V(z_i)\}}\Big)\Big]\Big].$$

Under conditions (1.5) and (1.6), we prove the following two statements in Sections 4.2.1 and 4.2.2, respectively.

$$\lim_{n\to\infty}D^{(1)}_{n+\tau}=0,\ \text{for every } \tau\in\mathbb{N}, \tag{4.11}$$

$$\limsup_{n\to\infty}\big|D^{(2)}_{n+\tau}\big|\le h^{(2)}_1(\tau),\quad \lim_{\tau\to\infty}h^{(2)}_1(\tau)=0. \tag{4.12}$$

Then we complete the proof of Proposition 4.1 by taking $h_1(\tau)=h^{(2)}_1(\tau)$. □

4.2.1. *Proof of (4.12).* Actually, (4.12) holds because of the property of conditional reduced GW-processes, which is that $\hat{N}_{n,n+\tau}$ converges to 1 in law when first $n\to\infty$ and then $\tau\to\infty$. More precisely,

$$\begin{aligned}&\hat{\mathbf{E}}_{n,n+\tau}\Bigg[\prod_{i=1,2}\Bigg|\sum_{y\in\hat{\mathbb{T}}_{\tau,n+\tau}(z_i)}\mathbf{1}_{\{\Delta V(z_i,y)\le\sqrt{n+\tau}x-V(z_i)\}}-\hat{N}_{\tau,n+\tau}(z_i)\mathbf{1}_{\{\widetilde{\mathcal{N}}_{z_i}\le\sqrt{n+\tau}x-V(z_i)\}}\Bigg|\Bigg]\\ \le&4\cdot\prod_{i=1,2}\mathbf{E}\big[\hat{N}_{\tau,n+\tau}(z_i)\big]\\ =&4\big(\mathbf{E}\hat{N}_\tau\big)^2:=C_1(\tau),\end{aligned} \tag{4.13}$$

where the last equality comes from the properties of $\{\hat{N}_{\tau,n+\tau}(z);z\in\hat{\mathbb{T}}_{n,n+\tau}\}$ in Section 4.1. Consequently, by (2.6), (2.7) and (2.9) we have

$$\begin{aligned}\big|D^{(2)}_{n+\tau}\big|&\le C_1(\tau)\cdot\Big(\mathbf{E}[\hat{N}^2_{n,n+\tau}]-\mathbf{E}[\hat{N}_{n,n+\tau}]\Big)\\ &=C_1(\tau)\cdot f''_n(1)\frac{\big(1-f_\tau(0)\big)^2}{1-f_{n+\tau}(0)}\\ &=C_1(\tau)\cdot\frac{\big(1-f_\tau(0)\big)^2}{1-f_{n+\tau}(0)}\Big(\frac{\sigma^2m^{n-1}(m^n-1)}{m-1}+(m^n)^2-m^n\Big).\end{aligned}$$

It follows from (1.3) that

$$\lim_{n\to\infty}\frac{1}{1-f_{n+\tau}(0)}\Big(\frac{\sigma^2m^{n-1}(m^n-1)}{m-1}+(m^n)^2-m^n\Big)=\frac{1}{m^\tau\varphi(0)}\Big(\frac{\sigma^2}{m(1-m)}-1\Big),$$

which yields

$$\limsup_{n\to\infty}\big|D^{(2)}_{n+\tau}\big|\le C_1(\tau)\cdot\frac{\big(1-f_\tau(0)\big)^2}{m^\tau\varphi(0)}\Big(\frac{\sigma^2}{m(1-m)}-1\Big)=:h^{(2)}_1(\tau).$$

Note that (2.8) implies $\lim_{\tau\to\infty}C_1(\tau)<\infty$ under (1.5). Moreover, recalling from (1.3) we have

$$\lim_{\tau\to\infty}\frac{\big(1-f_\tau(0)\big)^2}{m^\tau\varphi(0)}=\lim_{\tau\to\infty}1-f_\tau(0)=0,$$

then

$$\lim_{\tau\to\infty} h_1^{(2)}(\tau) = 0.$$

This implies (4.12). □

4.2.2. *Proof of (4.11).* Let

$$F(n, n+\tau, a) = \mathbf{E}\Big[\Big(\sum_{y\in\hat{\mathbb{T}}_{\tau,\tau}} \mathbf{1}_{\{V(y)\le\sqrt{n+\tau}x-a\}} - \hat{N}_\tau \mathbf{1}_{\{\widetilde{\mathcal{N}}\le\sqrt{n+\tau}x-a\}}\Big)^2\Big],$$

where under $\mathbf{P}$ the random variable $\widetilde{\mathcal{N}}$ has the same distribution as $\widetilde{\mathcal{N}}_z$. Since $\hat{N}_{\tau,n+\tau}(z)$ is independent of $\widetilde{\mathcal{N}}_z$, we assume that $\hat{N}_\tau$ and $\widetilde{\mathcal{N}}$ are independent. Then again using the properties of subtrees and the two families of random variables $\{\hat{\mathcal{N}}_{\tau,n+\tau}(z); z\in\hat{\mathbb{T}}_{n,n+\tau}\}$ and $\{\widetilde{\mathcal{N}}_z; z\in\hat{\mathbb{T}}_{n,n+\tau}\}$ in Section 4.1, we apply the many-to-one formula with $x=0$, $k=n$ and $g(\cdot)=F(n,n+\tau,\cdot)$ (see (4.5) in Lemma 4.3) to $D^{(1)}_{n+\tau}$ and get

$$D^{(1)}_{n+\tau} = m^n \frac{1-f_\tau(0)}{1-f_{n+\tau}(0)} \cdot \mathbf{E}[F(n,n+\tau,\mathfrak{X}_n)], \tag{4.14}$$

where $(\mathfrak{X}_n, n\in\mathbb{N})$ is a random walk whose transition probability is defined by (4.6). Note that the process $(V(z), z\in\hat{\mathbb{T}}_{k,n+\tau}, 0\le k\le n)$ is independent of $\hat{N}_{\tau,n+\tau}(z)$ and $\widetilde{\mathcal{N}}_z$, so we also assume that the random walk $(\mathfrak{X}_i, 0\le i\le n)$ is independent of $\hat{N}_\tau$ and $\widetilde{\mathcal{N}}$.

We first get from (1.3) that

$$\lim_{n\to\infty} m^n \frac{1-f_\tau(0)}{1-f_{n+\tau}(0)} = \frac{1-f_\tau(0)}{\varphi(0)m^\tau} =: C_2(\tau), \tag{4.15}$$

which is finite under (1.5). We next consider the integral of (4.14). Expanding the square in F yields

$$\begin{aligned} F(n,n+\tau,a) =& \mathbf{E}\Big[\hat{Z}_\tau((-\infty,\sqrt{n+\tau}x-a])^2\Big] - 2\mathbf{E}\Big[\hat{N}_\tau \sum_{y\in\hat{\mathbb{T}}_{\tau,\tau}} \mathbf{1}_{\{V(y)\le\sqrt{n+\tau}x-a\}}\mathbf{1}_{\{\widetilde{\mathcal{N}}\le\sqrt{n+\tau}x-a\}}\Big] \\ &+ \mathbf{E}\Big[\big(\hat{N}_\tau \mathbf{1}_{\{\widetilde{\mathcal{N}}\le\sqrt{n+\tau}x-a\}}\big)^2\Big] \\ :=& F_1(n,n+\tau,a) - 2F_2(n,n+\tau,a) + F_3(n,n+\tau,a). \end{aligned}$$

Therefore, we have

$$\mathbf{E}\big[F(n,n+\tau,\mathfrak{X}_n)\big] = \mathbf{E}\big[F_1(n,n+\tau,\mathfrak{X}_n)\big] - 2\mathbf{E}\big[F_2(n,n+\tau,\mathfrak{X}_n)\big] + \mathbf{E}\big[F_3(n,n+\tau,\mathfrak{X}_n)\big]. \tag{4.16}$$

Under (1.5) and (1.6), we get the following three statements:

$$\lim_{n\to\infty} \mathbf{E}\big[F_1(n,n+\tau,\mathfrak{X}_n)\big] = \mathbf{E}\big[\hat{N}_\tau^2\big]\Phi(x), \quad \forall\, \tau\in\mathbb{N}. \tag{4.17}$$

$$\lim_{n\to\infty} \mathbf{E}\big[F_2(n,n+\tau,\mathfrak{X}_n)\big] = \mathbf{E}\big[\hat{N}_\tau^2\big]\Phi(x), \quad \forall\, \tau\in\mathbb{N}. \tag{4.18}$$

$$\lim_{n\to\infty} \mathbf{E}\big[F_3(n,n+\tau,\mathfrak{X}_n)\big] = \mathbf{E}\big[\hat{N}_\tau^2\big]\Phi(x), \quad \forall\, \tau\in\mathbb{N}. \tag{4.19}$$

Then by using (4.14)-(4.19), we obtain (4.11). □

In the remainder of this section, we focus on proving (4.17)-(4.19).

(1) Proof of (4.17). To consider the first term on the right-hand side of (4.16), i.e., (4.17), we need to compute the second moment of the conditional reduced BRW. By (4.7) in Lemma 4.3, we have that $\mathbf{E}[F_1(n, n+\tau, \mathfrak{X}_n)]$ is equal to the sum of

$$\frac{m^\tau}{1-f_\tau(0)}\mathbf{P}(\mathfrak{X}_\tau^1 \leq \sqrt{n+\tau}x - \mathfrak{X}_n) \tag{4.20}$$

and

$$\frac{m^{2\tau-1}}{1-f_\tau(0)}f''(1)\sum_{j=1}^{\tau} m^{-j}\mathbf{P}(\mathfrak{X}_{j-1,\tau}^1 \leq \sqrt{n+\tau}x - \mathfrak{X}_n, \mathfrak{X}_{j-1,\tau}^2 \leq \sqrt{n+\tau}x - \mathfrak{X}_n), \tag{4.21}$$

where the random variable $\mathfrak{X}_\tau^1$ and the two random walks $(\mathfrak{X}_{j-1,\tau}^1 : 1 \leq j \leq \tau)$ and $(\mathfrak{X}_{j-1,\tau}^2 : 1 \leq j \leq \tau)$ are independent of $\mathfrak{X}_n$. Therefore, it suffices to compute the limits of (4.20) and (4.21). On the one hand, under (1.6), by the central limit theorem we have

$$\frac{\mathfrak{X}_n}{\sqrt{n+\tau}} \Longrightarrow \mathcal{N}, \quad n \to \infty, \tag{4.22}$$

where $\mathcal{N}$ is a standard normal random variable. Hence, letting $n \to \infty$ yields

$$\mathbf{P}\big(\mathfrak{X}_\tau^1 \leq \sqrt{n+\tau}x - \mathfrak{X}_n\big) = \mathbf{P}\Big(\frac{\mathfrak{X}_n}{\sqrt{n+\tau}} \leq x - \frac{\mathfrak{X}_\tau^1}{\sqrt{n+\tau}}\Big) \to \mathbf{P}(\mathcal{N} \leq x) = \Phi(x).$$

Substituting the last limit relation into (4.20) yields that under (1.6),

$$\lim_{n\to\infty} \frac{m^\tau}{1-f_\tau(0)}\mathbf{P}\big(\mathfrak{X}_\tau^1 \leq \sqrt{n+\tau}x - \mathfrak{X}_n\big) = \frac{m^\tau}{1-f_\tau(0)}\Phi(x). \tag{4.23}$$

This deals with the first term of (4.17), i.e., (4.20).

On the other hand, observe that for fixed $\tau \in \mathbb{N}$ and $1 \leq j \leq \tau$,

$$\frac{\mathfrak{X}_{j-1,\tau}^1}{\sqrt{n+\tau}} \xrightarrow{\text{a.s.}} 0, \quad \frac{\mathfrak{X}_{j-1,\tau}^2}{\sqrt{n+\tau}} \xrightarrow{\text{a.s.}} 0, \quad n \to \infty.$$

Recall that the random variable $\mathfrak{X}_n$ is independent of the two random walks $(\mathfrak{X}_{j-1,\tau}^1 : 1 \leq j \leq \tau)$ and $(\mathfrak{X}_{j-1,\tau}^2 : 1 \leq j \leq \tau)$. Then by (4.22) we see that for any $1 \leq j \leq \tau$,

$$\begin{aligned}
&\mathbf{P}\big(\mathfrak{X}_{j-1,\tau}^1 \leq \sqrt{n+\tau}x - \mathfrak{X}_n, \mathfrak{X}_{j-1,\tau}^2 \leq \sqrt{n+\tau}x - \mathfrak{X}_n\big)\\
=&\mathbf{P}\Big(\frac{\mathfrak{X}_n}{\sqrt{n+\tau}} \leq x - \frac{\mathfrak{X}_{j-1,\tau}^1}{\sqrt{n+\tau}}, \frac{\mathfrak{X}_n}{\sqrt{n+\tau}} \leq x - \frac{\mathfrak{X}_{j-1,\tau}^2}{\sqrt{n+\tau}}\Big)\\
\to&\mathbf{P}(\mathcal{N} \leq x) = \Phi(x), \quad n \to \infty.
\end{aligned} \tag{4.24}$$

It immediately proves that under (1.6), as $n \to \infty$, (4.21) converges to

$$\frac{m^\tau - m^{2\tau}}{m(1-m)\big(1-f_\tau(0)\big)}f''(1)\Phi(x), \tag{4.25}$$

if we note that $\sum_{j=1}^{\tau} m^{-j} = (1-m^{-\tau})/(m-1)$.

Therefore, recall that the second moment of $\hat{N}_\tau$ is decided by (2.10), by (4.23), (4.24) and (4.25), a simple calculation shows that

$$\begin{aligned}
\lim_{n\to\infty}\mathbf{E}\big[F_1(n, n+\tau, \mathfrak{X}_n)\big] =&\frac{\Phi(x)}{1-f_\tau(0)}\Big[m^\tau + \frac{m^\tau - m^{2\tau}}{m(1-m)}f''(1)\Big]\\
=&\frac{\Phi(x)}{1-f_\tau(0)}\Big[m^{2\tau} + \frac{\sigma^2 m^{\tau-1}\big(m^\tau - 1\big)}{m-1}\Big]\\
=&\mathbf{E}\big[\hat{N}_\tau^2\big]\Phi(x),
\end{aligned}$$

where the second equality comes from $f''(1) = \sigma^2 + m^2 - m$. This ends up the proof of (4.17) under (1.6). □

(2) Proof of (4.18). We compute the second term on the right-hand side of (4.16), i.e., (4.18). Recall that $V(y)$ is independent of $\hat{N}_\tau$, and the three random variables $\mathfrak{X}_n$, $\widetilde{\mathcal{N}}$ and $\hat{N}_\tau$ are independent of each other. So we have

$$\begin{aligned}\mathbf{E}\big[F_2(n, n+\tau, \mathfrak{X}_n)\big] &= \mathbf{E}\Big[\hat{N}_\tau \sum_{y\in\hat{\mathbb{T}}_{\tau,\tau}} \mathbf{1}_{\{V(y)\le\sqrt{n+\tau}x-\mathfrak{X}_n\}}\mathbf{1}_{\{\widetilde{\mathcal{N}}\le\sqrt{n+\tau}x-\mathfrak{X}_n\}}\Big] \\ &= \mathbf{E}\Big[\hat{N}_\tau \sum_{y\in\hat{\mathbb{T}}_{\tau,\tau}} \mathbf{E}\Big[\mathbf{1}_{\{V(y)\le\sqrt{n+\tau}x-\mathfrak{X}_n,\widetilde{\mathcal{N}}\le\sqrt{n+\tau}x-\mathfrak{X}_n\}}\Big]\Big] \\ &= \mathbf{E}\Big[\hat{N}_\tau \sum_{y\in\hat{\mathbb{T}}_{\tau,\tau}} \mathbf{E}\Big[\mathbf{1}_{\big\{\frac{\mathfrak{X}_n}{\sqrt{n+\tau}}+\frac{V(y)}{\sqrt{n+\tau}}\le x,\frac{\mathfrak{X}_n}{\sqrt{n+\tau}}+\frac{\widetilde{\mathcal{N}}}{\sqrt{n+\tau}}\le x\big\}}\Big]\Big].\end{aligned} \tag{4.26}$$

We first fix τ and then let $n\to\infty$ to get $\frac{V(y)}{\sqrt{n+\tau}} \xrightarrow{\text{a.s.}} 0$ and $\frac{\widetilde{\mathcal{N}}}{\sqrt{n+\tau}} \xrightarrow{\text{a.s.}} 0$. Consequently, by (4.22) we obtain that under (1.6),

$$\lim_{n\to\infty}\mathbf{P}\Big(\frac{\mathfrak{X}_n}{\sqrt{n+\tau}}+\frac{V(y)}{\sqrt{n+\tau}}\le x, \frac{\mathfrak{X}_n}{\sqrt{n+\tau}}+\frac{\widetilde{\mathcal{N}}}{\sqrt{n+\tau}}\le x\Big) = \mathbf{P}(\mathcal{N}\le x) = \Phi(x).$$

Recall that $\mathbf{E}[\hat{N}_\tau^2] < \infty$ under (1.5). By dominated convergence, letting $n\to\infty$ in (4.26) yields

$$\lim_{n\to\infty}\mathbf{E}\big[F_2(n,n+\tau,\mathfrak{X}_n)\big] = \mathbf{E}\big[\hat{N}_\tau^2\big]\Phi(x).$$

This is (4.18). □

(3) Proof of (4.19). Finally we discuss the third term on the right-hand side of (4.16), i.e., (4.19). Indeed, recall again that $\mathfrak{X}_n$, $\hat{N}_\tau$ and $\widetilde{\mathcal{N}}$ are independent of each other, so we have

$$\mathbf{E}\big[F_3(n,n+\tau,\mathfrak{X}_n)\big] = \mathbf{E}\Big[\hat{N}_\tau^2\mathbf{1}_{\{\widetilde{\mathcal{N}}\le\sqrt{n+\tau}x-\mathfrak{X}_n\}}\Big] = \mathbf{E}\big[\hat{N}_\tau^2\big]\mathbf{P}(\widetilde{\mathcal{N}}\le\sqrt{n+\tau}x-\mathfrak{X}_n). \tag{4.27}$$

Under (1.6), letting $n\to\infty$ in the above equality yields

$$\mathbf{P}(\widetilde{\mathcal{N}}\le\sqrt{n+\tau}x-\mathfrak{X}_n) = \mathbf{P}\Big(\frac{\mathfrak{X}_n}{\sqrt{n+\tau}}\le x-\frac{\widetilde{\mathcal{N}}}{\sqrt{n+\tau}}\Big)\to\mathbf{P}(\mathcal{N}\le x) = \Phi(x),$$

which means that (4.19) holds under (1.6). □

4.3. *Proof of Proposition 4.2.* Recall that for fixed $z\in\hat{\mathbb{T}}_{n,n+\tau}$, the two random variables $\hat{N}_{\tau,n+\tau}(z)$ and $\widetilde{\mathcal{N}}_z$ are independent of the process $(V(u), u\in\hat{\mathbb{T}}_{k,n+\tau}, 0\le k\le n)$ in Section 4.1, then the generating function of $B_{n+\tau}$ is

$$\mathbf{E}\Bigg[s^{\sum_{z\in\hat{\mathbb{T}}_{n,n+\tau}}\hat{N}_{\tau,n+\tau}(z)\mathbf{1}_{\{\widetilde{\mathcal{N}}_z\le\sqrt{n+\tau}x-V(z)\}}}\Bigg] = \mathbf{E}\Big[\prod_{z\in\hat{\mathbb{T}}_{n,n+\tau}}\hat{\mathbf{E}}_{n,n+\tau}\Big[s^{\hat{N}_{\tau,n+\tau}(z)\mathbf{1}_{\{\widetilde{\mathcal{N}}_z\le\sqrt{n+\tau}x-V(z)\}}}\Big]\Big],$$

where $\hat{\mathbf{E}}_{n,n+\tau}$ is defined as in the proof of Proposition 4.1. Since $\hat{N}_{\tau,n+\tau}(z) \stackrel{d}{=} \hat{N}_\tau$ and the variables $\hat{N}_{\tau,n+\tau}(z)$ and $\widetilde{\mathcal{N}}_z$ are independent, we have

$$\begin{aligned}
&\hat{\mathbf{E}}_{n,n+\tau}\Big[s^{\hat{N}_{\tau,n+\tau}(z)\mathbf{1}_{\{\widetilde{\mathcal{N}}_z\le\sqrt{n+\tau}x-V(z)\}}}\Big]\\
=&\hat{\mathbf{E}}_{n,n+\tau}\Big[\hat{\mathbf{E}}_{n,n+\tau}\Big[s^{\hat{N}_{\tau,n+\tau}(z)\mathbf{1}_{\{\widetilde{\mathcal{N}}_z\le\sqrt{n+\tau}x-V(z)\}}}\Big|\hat{N}_{\tau,n+\tau}(z)\Big]\Big]\\
=&\hat{\mathbf{P}}_{n,n+\tau}\big(\widetilde{\mathcal{N}}_z\le\sqrt{n+\tau}x-V(z)\big)\mathbf{E}\big[s^{\hat{N}_\tau}\big]+1-\hat{\mathbf{P}}_{n,n+\tau}\big(\widetilde{\mathcal{N}}_z\le\sqrt{n+\tau}x-V(z)\big)\\
:=&R\big(n,n+\tau,V(z)\big),
\end{aligned}$$

where $\hat{\mathbf{P}}_{n,n+\tau}(\cdot)=\mathbf{P}(\cdot\mid\hat{\mathcal{F}}_{n,n+\tau})$. Hence these above equalities imply

$$\mathbf{E}[s^{B_{n+\tau}}]=\mathbf{E}\Big[\prod_{z\in\hat{\mathbb{T}}_{n,n+\tau}}R\big(n,n+\tau,V(z)\big)\Big]. \tag{4.28}$$

Loosely speaking, according to the property of conditional reduced GW-trees, there are only one particle in generation n making contributions to the $(n+\tau)$-th generation conditioned on the event $\{N_{n+\tau}>0\}$ when first $n\to\infty$ and then $\tau\to\infty$. Thus the product on the right-hand side of (4.28) can be replaced by one factor when first $n\to\infty$ and then $\tau\to\infty$. For any fixed $z_0\in\hat{\mathbb{T}}_{n,n+\tau}$, since $0\le R(n,n+\tau,V(z_0))\le 1$, we get that for $s\in[0,1]$,

$$\begin{aligned}
&\Big|\mathbf{E}\Big[\prod_{z\in\hat{\mathbb{T}}_{n,n+\tau}}R(n,n+\tau,V(z))\Big]-L(s)\Big|\\
\le&\Big|\mathbf{E}\Big[\prod_{z\in\hat{\mathbb{T}}_{n,n+\tau}}R(n,n+\tau,V(z))\Big]-\mathbf{E}\Big[R\big(n,n+\tau,V(z_0)\big)\Big]\Big|+\Big|\mathbf{E}\Big[R\big(n,n+\tau,V(z_0)\big)\Big]-L(s)\Big|\\
\le&\mathbf{E}\Big[\Big|\prod_{z\in\hat{\mathbb{T}}_{n,n+\tau}}R\big(n,n+\tau,V(z)\big)-R\big(n,n+\tau,V(z_0)\big)\cdot\prod_{\substack{z\in\hat{\mathbb{T}}_{n,n+\tau}\\ z\ne z_0}}1\Big|\Big]+\Big|\mathbf{E}\Big[R\big(n,n+\tau,V(z_0)\big)\Big]-L(s)\Big|\\
\le&\mathbf{E}\Big[\sum_{\substack{z\in\hat{\mathbb{T}}_{n,n+\tau}\\ z\ne z_0}}\Big(1-R\big(n,n+\tau,V(z)\big)\Big)\Big]+\Big|\mathbf{E}\Big[R\big(n,n+\tau,V(z_0)\big)\Big]-L(s)\Big|,
\end{aligned} \tag{4.29}$$

where the last inequality follows from Durrett (2010, Lemma 3.4.3). We next estimate the first term on the right-hand side of (4.29). Note that $0\le 1-R(n,n+\tau,V(z))\le 1$ for each $z\in\hat{\mathbb{T}}_{n,n+\tau}$, so that

$$\mathbf{E}\Big[\sum_{\substack{z\in\hat{\mathbb{T}}_{n,n+\tau}\\ z\ne z_0}}\Big(1-R\big(n,n+\tau,V(z)\big)\Big)\Big]\le\mathbf{E}\big[\hat{N}_{n,n+\tau}-1\big].$$

By (1.3) and (2.6), a simple calculation shows that

$$\lim_{n\to\infty}\mathbf{E}\big[\hat{N}_{n,n+\tau}-1\big]=\frac{1-f_\tau(0)}{\varphi(0)m^\tau}-1.$$

We use (1.3) again to see that under (1.5),

$$\lim_{\tau\to\infty}\frac{1-f_\tau(0)}{\varphi(0)m^\tau}-1=0.$$

We finally estimate the second part of (4.29). Recall that the random variable $V(z_0)$ has the law G^{n*} and is independent of $\widetilde{\mathcal{N}}_{z_0}$ for $z_0 \in \hat{\mathbb{T}}_{n,n+\tau}$, then we obtain that under (1.6),

$$\lim_{n\to\infty} \mathbf{E}[R(n, n+\tau, V(z_0))] = \Phi(x)\mathbf{E}\big[s^{\hat{N}_\tau}\big] + 1 - \Phi(x).$$

The convergence of $\hat{N}_\tau$ and the definition of $L(s)$ in (4.4) imply that $\Phi(x)\mathbf{E}[s^{\hat{N}_\tau}] + 1 - \Phi(x) \to L(s)$ as $\tau \to \infty$. Hence, we complete the proof of Proposition 4.2 by taking

$$h_2(\tau, s) = \frac{1 - f_\tau(0)}{\varphi(0)m^\tau} - 1 + \Big|\Phi(x)\mathbf{E}\big[s^{\hat{N}_\tau}\big] + 1 - \Phi(x) - L(s)\Big|,$$

which satisfies $\lim_{\tau\to\infty} h_2(\tau, s) = 0$. □

5. Proofs of Corollaries 1.5 and 1.7

Corollaries 1.5 and 1.7 give the asymptotic behavior of $\mathcal{L}(Z_n((-\infty, \sqrt{n}x])|N_{n+l} > 0)$. To this end, we should consider the joint distribution of $(\hat{N}_n, \hat{Z}_n((-\infty, \sqrt{n}x]))$. The next lemma can be proved along the same line as the proof of Theorem 1.3.

Lemma 5.1. *If assumptions (1.5) and (1.6) are fulfilled, then for each $\lambda_1 \geq 0, \lambda_2 \geq 0$ and $x \in \mathbb{R}$, we have*

$$\lim_{n\to\infty} \mathbf{E}\Big[s^{\lambda_1 \hat{N}_n + \lambda_2 \hat{Z}_n((-\infty,\sqrt{n}x])}\Big] = \Phi(x)H\big(s^{\lambda_1+\lambda_2}\big) + \big(1 - \Phi(x)\big)H\big(s^{\lambda_1}\big), \quad s \in (0,1). \tag{5.1}$$

In other words,

$$\mathcal{L}\big(N_n, Z_n\big((-\infty, \sqrt{n}x]\big)\big|N_n > 0\big) = \mathcal{L}\big(\hat{N}_n, \hat{Z}_n((-\infty, \sqrt{n}x])\big) \Longrightarrow \mathcal{L}(\xi, \xi\mathbf{1}_{\{\mathcal{N}\leq x\}}), \quad n \to \infty,$$

where ξ and $\mathcal{N}$ are the same as in Theorem 1.3.

Proof: For each $\lambda_1 \geq 0, \lambda_2 \geq 0$ and $x \in \mathbb{R}$, we only need to decompose the random variable $\lambda_1\hat{N}_{n+\tau} + \lambda_2\hat{Z}_{n+\tau}\big((-\infty, \sqrt{n+\tau}x]\big)$ at generation n like (4.1):

$$\begin{aligned}\lambda_1\hat{N}_{n+\tau} + \lambda_2\hat{Z}_{n+\tau}\big((-\infty, \sqrt{n+\tau}x]\big) &\overset{d}{=} \sum_{z\in\hat{\mathbb{T}}_{n,n+\tau}} \Bigg[\sum_{y\in\hat{\mathbb{T}}_{\tau,n+\tau}(z)} \Big(\lambda_1 + \lambda_2\mathbf{1}_{\{\Delta V(z,y)\leq\sqrt{n+\tau}x - V(z)\}}\Big)\Bigg] \\ &= A_{n+\tau}(\lambda_1,\lambda_2) + B_{n+\tau}(\lambda_1,\lambda_2),\end{aligned}$$

where

$$A_{n+\tau}(\lambda_1,\lambda_2) = \sum_{z\in\hat{\mathbb{T}}_{n,n+\tau}} \Bigg[\sum_{y\in\hat{\mathbb{T}}_{\tau,n+\tau}(z)} \Big(\lambda_1 + \lambda_2\mathbf{1}_{\{\Delta V(z,y)\leq\sqrt{n+\tau}x - V(z)\}}\Big) - \hat{N}_{\tau,n+\tau}(z)\Big(\lambda_1 + \lambda_2\mathbf{1}_{\{\widetilde{\mathcal{N}}_z\leq\sqrt{n+\tau}x - V(z)\}}\Big)\Bigg]$$

$$B_{n+\tau}(\lambda_1,\lambda_2) = \sum_{z\in\hat{\mathbb{T}}_{n,n+\tau}} \hat{N}_{\tau,n+\tau}(z)\Big(\lambda_1 + \lambda_2\mathbf{1}_{\{\widetilde{\mathcal{N}}_z\leq\sqrt{n+\tau}x - V(z)\}}\Big)$$

Based on this decomposition, the rest proof of (5.1) is the same as that of Theorem 1.3 except for some slight changes.

Note that

$$\mathbf{E}\Big[s^{\xi(\lambda_1+\lambda_2\mathbf{1}_{\{\mathcal{N}\leq x\}})}\Big] = \Phi(x)H\big(s^{\lambda_1+\lambda_2}\big) + \big(1 - \Phi(x)\big)H\big(s^{\lambda_1}\big)$$

and $H(s)$ is continuous at $s = 1$, so the last part of the lemma follows from the continuity theorem of generating functions (c.f. Feller, 1968, Continuity theorem, Section XI.6). □

Proof of Corollary 1.5. We first recall that for $l \in \mathbb{N}$ and $s \in (0,1)$,

$$\begin{aligned}
&\lim_{n\to\infty} \mathbf{E}\Big[s^{Z_n\big((-\infty,\sqrt{n}x]\big)}\Big|N_{n+l}>0\Big]\\
&=\lim_{n\to\infty}\frac{\mathbf{P}(N_n>0)}{\mathbf{P}(N_{n+l}>0)}\cdot\mathbf{E}\Big[s^{Z_n\big((-\infty,\sqrt{n}x]\big)}\big(1-f_l(0)^{N_n}\big)\Big|N_n>0\Big]\\
&=\lim_{n\to\infty}\frac{\mathbf{P}(N_n>0)}{\mathbf{P}(N_{n+l}>0)}\cdot\Big(\mathbf{E}\Big[s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\Big]-\mathbf{E}\Big[f_l(0)^{\hat{N}_n}s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\Big]\Big).
\end{aligned}$$

We next take $\lambda_2=1$ and λ_1 satisfying $s^{\lambda_1}=f_l(0)$ in (5.1) to get that under (1.5) and (1.6),

$$\lim_{n\to\infty}\mathbf{E}\Big[f_l(0)^{\hat{N}_n}s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\Big]=\Phi(x)H\big(sf_l(0)\big)+\big(1-\Phi(x)\big)H\big(f_l(0)\big).$$

Hence, by (2.11) and Theorem 1.3 we obtain that under (1.5) and (1.6), for $s\in(0,1)$,

$$\begin{aligned}
\lim_{n\to\infty}\mathbf{E}\Big[s^{Z_n\big((-\infty,\sqrt{n}x]\big)}\Big|N_{n+l}>0\Big]&=\frac{\Phi(x)\Big[H\big(s\big)-H\big(sf_l(0)\big)\Big]+\big(1-\Phi(x)\big)\Big[1-H\big(f_l(0)\big)\Big]}{m^l}\\
&=\frac{H\big(s\big)-H\big(sf_l(0)\big)}{m^l}\Phi(x)+1-\Phi(x),
\end{aligned}\tag{5.2}$$

where the last equality comes from $1-H\big(f_l(0)\big)=m^l$ (c.f. Joffe, 1967, Yaglom's theorem). Finally, Note that $H(s)$ is continuous at $s=1$ and $1-H\big(f_l(0)\big)=m^l$, then the right-hand side of (5.2) is continuous at $s=1$. By the continuity theorem of generating functions (c.f. Feller, 1968, Continuity theorem, Section XI.6), we obtain

$$\lim_{n\to\infty}\mathbf{P}\big(Z_n\big((-\infty,\sqrt{n}x]\big)=j\big|N_{n+l}>0\big)=b_j(l;x)\geq 0,\quad j\in\mathbb{N},$$

where $\{b_j(l;x)\}_{j\in\mathbb{N}}$ is a probability law and its generating function is the right-hand side of (5.2).

□

Proof of Corollary 1.7. (1) For fixed $x\in\mathbb{R}, s\in(0,1)$ and $n\in\mathbb{N}$, we have

$$\begin{aligned}
&\lim_{l\to\infty}\mathbf{E}\Big[s^{Z_n\big((-\infty,\sqrt{n}x]\big)}\Big|N_{n+l}>0\Big]\\
&=\lim_{l\to\infty}\mathbf{E}\Big[s^{Z_n\big((-\infty,\sqrt{n}x]\big)}\big(1-f_l(0)^{N_n}\big)\Big|N_n>0\Big]\frac{\mathbf{P}(N_n>0)}{\mathbf{P}(N_{n+l}>0)}\\
&=\mathbf{P}(N_n>0)\cdot\lim_{l\to\infty}\frac{\mathbf{E}\Big[s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\big(1-f_l(0)^{\hat{N}_n}\big)\Big]}{1-f_n(f_l(0))}\\
&=\mathbf{E}\Big[\hat{N}_n s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\Big]\cdot\frac{\mathbf{P}(N_n>0)}{m^n}.
\end{aligned}$$

Recall that $\mathbf{E}[\hat{N}_n]=m^n/\mathbf{P}(N_n>0)$. Then by dominated convergence we can get that as $s\to 1$,

$$\mathbf{E}\Big[\hat{N}_n s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\Big]\cdot\frac{\mathbf{P}(N_n>0)}{m^n}\to 1.$$

Summarizing above, it thus follows from the continuity theorem of generating functions (c.f. Feller, 1968, Continuity theorem, Section XI.6) that

$$\lim_{l\to\infty}\mathbf{P}\big(Z_n\big((-\infty,\sqrt{n}x]\big)=j\big|N_{n+l}>0\big)=a_j(n;x)\geq 0$$

exists for every $j\in\mathbb{N}$ and $\sum_{j\in\mathbb{N}}a_j(n;x)=1$. In this case automatically

$$\mathbf{E}\Big[\hat{N}_n s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\Big]\cdot\frac{\mathbf{P}(N_n>0)}{m^n}=\sum_{j\in\mathbb{N}}a_j(n;x)s^j,\quad s\in[0,1].$$

Under (1.5) and (1.6), by first differentiating the both sides of (5.1) with respect to λ_1 and then taking $\lambda_1 = 0$ and $\lambda_2 = 1$, it is obvious that

$$\lim_{n\to\infty} \mathbf{E}\Big[\hat{N}_n s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\Big] = \Phi(x)sH'(s) + \big(1-\Phi(x)\big)H'(1).$$

We use (1.3) and the fact that $H'(1) = 1/\varphi(0)$ (c.f. Joffe, 1967, Theorem 1) to obtain that under (1.5),

$$\begin{aligned}\lim_{n\to\infty} \sum_{j\in\mathbb{N}} a_j(n;x)s^j = &\lim_{n\to\infty} \mathbf{E}\Big[\hat{N}_n s^{\hat{Z}_n\big((-\infty,\sqrt{n}x]\big)}\Big]\cdot \frac{\mathbf{P}(N_n>0)}{m^n}\\ =&\varphi(0)sH'(s)\Phi(x)+1-\Phi(x). \end{aligned} \tag{5.3}$$

Using again $H'(1) = 1/\varphi(0)$ yields that the limit of (5.3) is continuous at $s = 1$. Thus we can use again the continuity theorem of generating functions (c.f. Feller, 1968, Continuity theorem, Section XI.6) to get

$$\lim_{n\to\infty} a_j(n;x) = a_j(x) \geq 0 \tag{5.4}$$

exists for each $j \in \mathbb{N}$ and $\sum_{j\in\mathbb{N}} a_j(x) = 1$. Next, since the random variable ζ has the law $\{\pi_j\}_{j\in\mathbb{N}_+}$ decided by (1.4), it is easy to check that the generating function of $\zeta\mathbf{1}_{\{\mathcal{N}\leq x\}}$ is the right-hand side of (5.3). Combining this result with (5.4) and the definition of $a_j(n;x)$, we have proved that

$$\lim_{n\to\infty}\lim_{l\to\infty} \mathbf{P}\big(Z_n\big((-\infty,\sqrt{n}x]\big) = j\big|N_{n+l}>0\big) = \mathbf{P}\big(\zeta\mathbf{1}_{\{\mathcal{N}\leq x\}} = j\big), \quad j\in\mathbb{N}.$$

Then we complete the proof of the first part.

(2) This is similar to the first part of the proof. By (1.3) it is easily seen that

$$\lim_{l\to\infty} \frac{H(s)-H\big(sf_l(0)\big)}{m^l} = \lim_{l\to\infty} \frac{H(s)-H\big(sf_l(0)\big)}{1-f_l(0)}\cdot\frac{1-f_l(0)}{m^l} = \varphi(0)sH'(s).$$

By Corollary 1.5 the generating function of $\{b_j(l;x)\}_{j\in\mathbb{N}}$ is given by

$$\sum_{j\in\mathbb{N}} b_j(l;x)s^j = \frac{H(s)-H\big(sf_l(0)\big)}{m^l}\Phi(x)+1-\Phi(x).$$

Then we let $l\to\infty$ in both sides to get

$$\lim_{l\to\infty}\sum_{j\in\mathbb{N}} b_j(l;x)s^j = \varphi(0)sH'(s)\Phi(x)+1-\Phi(x),$$

which is the generating function of $\zeta\mathbf{1}_{\{\mathcal{N}\leq x\}}$. From the continuity theorem of generating functions (c.f. Feller, 1968, Continuity theorem, Section XI.6) it is natural to have

$$\lim_{l\to\infty} b_j(l;x) = \mathbf{P}\big(\zeta\mathbf{1}_{\{\mathcal{N}\leq x\}} = j\big), \quad j\in\mathbb{N}.$$

By the definition of $b_j(l;x)$ we obtain the desired result. □

Acknowledgements

The authors are deeply grateful to the anonymous referees for careful reading of the original manuscript and for helpful suggestions which improved the presentation of the paper. In particular, one of the referees suggests to exchange the order of $l\to\infty$ and $n\to\infty$ which led to (2) in Corollary 1.7.

References

Athreya, K. B. and Ney, P. E. *Branching processes*. Die Grundlehren der mathematischen Wissenschaften, Band 196. Springer-Verlag, New York-Heidelberg (1972). MR373040.

Bansaye, V. Ancestral lineages and limit theorems for branching Markov chains in varying environment. *J. Theoret. Probab.*, **32** (1), 249–281 (2019). MR3908914.

Bansaye, V. and Huang, C. Weak law of large numbers for some Markov chains along non homogeneous genealogies. *ESAIM Probab. Stat.*, **19**, 307–326 (2015). MR3412648.

Biggins, J. D. The central limit theorem for the supercritical branching random walk, and related results. *Stochastic Process. Appl.*, **34** (2), 255–274 (1990). MR1047646.

Durrett, R. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition (2010). ISBN 978-0-521-76539-8. MR2722836.

Feller, W. *An introduction to probability theory and its applications. Vol. I.* John Wiley & Sons, Inc., New York-London-Sydney, third edition (1968). MR228020.

Fleischmann, K. and Prehn, U. Ein Grenzwertsatz für subkritische Verzweigungsprozesse mit endlich vielen Typen von Teilchen. *Math. Nachr.*, **64**, 357–362 (1974). MR373042.

Fleischmann, K. and Siegmund-Schultze, R. The structure of reduced critical Galton-Watson processes. *Math. Nachr.*, **79**, 233–241 (1977). MR461689.

Fleischmann, K. and Vatutin, V. A. Reduced subcritical Galton-Watson processes in a random environment. *Adv. in Appl. Probab.*, **31** (1), 88–111 (1999). MR1699663.

Geiger, J. Elementary new proofs of classical limit theorems for Galton-Watson processes. *J. Appl. Probab.*, **36** (2), 301–309 (1999). MR1724856.

Harris, S. C. and Roberts, M. I. The many-to-few lemma and multiple spines. *Ann. Inst. Henri Poincaré Probab. Stat.*, **53** (1), 226–242 (2017). MR3606740.

Harris, T. E. *The theory of branching processes*. Die Grundlehren der mathematischen Wissenschaften, Band 119. Springer-Verlag, Berlin; Prentice Hall, Inc., Englewood Cliffs, NJ (1963). MR163361.

Heathcote, C. R., Seneta, E., and Vere-Jones, D. A refinement of two theorems in the theory of branching processes. *Teor. Verojatnost. i Primenen.*, **12**, 341–346 (1967). MR217889.

Hong, W. M. and Liang, S. L. Conditional central limit theorem for critical branching random walk. *ArXiv Mathematics e-prints* (2023). arXiv: 2311.11044.

Joffe, A. On the Galton-Watson branching process with mean less than one. *Ann. Math. Statist.*, **38**, 264–266 (1967). MR205337.

Kaplan, N. and Asmussen, S. Branching random walks. II. *Stochastic Process. Appl.*, **4** (1), 15–31 (1976). MR400430.

Klebaner, C. F. Branching random walk in varying environments. *Adv. in Appl. Probab.*, **14** (2), 359–367 (1982). MR650128.

Shi, Z. *Branching random walks*, volume 2151 of *Lecture Notes in Mathematics*. Springer, Cham (2015). ISBN 978-3-319-25371-8; 978-3-319-25372-5. Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012. MR3444654.

Stam, A. J. On a conjecture by Harris. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **5**, 202–206 (1966). MR202201.

Yaglom, A. M. Certain limit theorems of the theory of branching random processes. *Doklady Akad. Nauk SSSR (N.S.)*, **56**, 795–798 (1947). MR22045.