Conditional central limit theorem for critical branching random walk

Wenming Hong and Shengli Liang

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`

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

Abstract. Consider a critical branching random walk on $\mathbb{R}$. Let $Z^{(n)}(A)$ be the number of individuals in the n-th generation located in $A \in \mathcal{B}(\mathbb{R})$ and $Z_n := Z^{(n)}(\mathbb{R})$ denote the population of the n-th generation. We prove that, under some conditions, for all $x \in \mathbb{R}$, as $n \to \infty$,

$$\mathcal{L}\left(\frac{Z^{(n)}(-\infty, \sqrt{n}x]}{n} \;\middle|\; Z_n > 0 \right) \Longrightarrow \mathcal{L}\big(Y(x)\big),$$

where $\Rightarrow$ means weak convergence and $Y(x)$ is a random variable whose distribution is specified by its moments.

1. Introduction and main result

Consider a discrete-time *branching random walk* (BRW) on the real line, which can be viewed as a natural extension of the Galton–Watson (G–W) process for the additional spatial structure. More precisely, the process starts with an initial particle positioned at the origin. It dies at time 1 and gives birth to a random number of children who form the first generation according to the offspring distribution $p = \{p_k\}_{k \geq 0}$, meanwhile, each of the individuals are independently positioned (with respect to their parent) according to the same probability measure ν. By induction, each particle alive at generation $n - 1$ dies at time n, and gives birth independently of all others to its own children who are in the n-th generation and are located in a relative position according to p and ν respectively. The process goes on as described above if there are particles alive. We assume that the reproduction and displacement mechanisms are independent.

Let $Z^{(n)}(-\infty, x]$ be the number of particles of the n-th generation whose positions are in $(-\infty, x]$ and $Z_n := Z^{(n)}(\mathbb{R})$ be the total number of particles in the n-th generation. Z_n is a *Galton–Watson process*, which called *supercritical, critical, or subcritical* according to the mean number of offspring $m := \mathbf{E} Z_1 > 1$, $m = 1$, or $m < 1$ respectively.

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

$$\lim_{n \to \infty} \frac{Z^{(n)}(-\infty, \sqrt{n}x]}{m^n} = \Phi(x)W \text{ in probability,}$$

where

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy, \ W = \lim_{n \to \infty} \frac{Z_n}{m^n}.$$

Results on central limit theorem of this type have been established and extended by many authors. Stam (1966) and Kaplan and Asmussen (1976) proved the conjecture and obtained that the convergence holds almost surely when $E(Z_1^2) < \infty$ and $E(Z_1 \log^{1+\epsilon} Z_1) < \infty$, respectively. Klebaner (1982) and Biggins (1990) extended these results to the branching random walk in varying environment (generation dependent). Gao et al. (2014) considered this central limit theorem for the BRW with random environment in time. Recently, Bansaye (2019) extended it to the model of branching Markov chain in random environment.

In this paper, we will focus on critical case, i.e. $m = 1$. The asymptotic behaviour for the BRW with critical branching mechanism is significantly different from that for supercritical case, since the corresponding G–W process becomes extinct with probability one. For instance, the asymptotic behaviour for the maximal displacement of critical BRW has been investigated by Kesten (1995) and Lalley and Shao (2015), the former obtained the asymptotic law for the rightmost position at time n conditioned on survival up to time βn or on extinction occurring at time $k_n \geqslant n$ with $k_n/n \to \gamma$, while the latter is devoted to studying the distribution of the rightmost position ever reached by the BRW. These asymptotic law are very different from those in the supercritical case that have been extensively studied over the past decades, see for example Hammersley (1974); Kingman (1975); Biggins (1976); Hu and Shi (2009); Aïdékon (2013). Here, since we will consider the central limit theorem in the critical case, it is more natural to describe this limit theorem by conditioning on survival until time n.

Now, we assume that

$$m = \mathbf{E} Z_1 = 1, \quad \sigma^2 := \text{Var } Z_1 \in (0, \infty). \tag{1.1}$$

$$\int_{-\infty}^{\infty} x\nu(dx) = 0, \quad \int_{-\infty}^{\infty} x^2 \nu(dx) = 1. \tag{1.2}$$

We obtain the following conditional central limit theorem for the critical BRW.

Theorem 1.1. *Under the assumptions* (1.1), (1.2) *and that p has finite moments $m_r = \sum_{j=0}^{\infty} j^r p_j$ of all orders, we have, for all $x \in \mathbb{R}$, as $n \to \infty$,*

$$\mathcal{L}\left(\frac{Z^{(n)}(-\infty, \sqrt{n}x]}{n} \ \Big| \ Z_n > 0 \right) \Longrightarrow \mathcal{L}(Y(x)),$$

where $\Rightarrow$ means weak convergence and $Y(x)$ is a random variable whose distribution is determined by its moments $\mu_r(x)$:

$$r = 1, \ \mu_1(x) = \frac{\sigma^2}{2}\Phi(x),$$

$$r = 2, \ \mu_2(x) = \binom{2}{1}\int_0^1 \mathbf{E}\big[(\mu_1^t(x - B_t))^2\big]dt = \frac{\sigma^4}{2}\int_0^1 \mathbf{E}\big[\Phi_{1-t}^2(x - B_t)\big]dt,$$

$$r \geq 3, \ \mu_r(x) = \sum_{i=1}^{r-1}\binom{r}{i}\int_0^1 \mathbf{E}\big[\mu_i^t(x - B_t)\mu_{r-i}^t(x - B_t)\big]dt,$$

where B_t is the standard Brownian motion starting from zero,

$$\Phi_{1-t}(x) := \int_{-\infty}^x \frac{1}{\sqrt{2\pi(1-t)}}e^{-\frac{y^2}{2(1-t)}}\, dy,$$

$$\mu_1^t(x) := \frac{\sigma^2}{2}\Phi_{1-t}(x), \ \mu_2^t(x) := \frac{\sigma^4}{2}\int_t^1 \mathbf{E}\big[\Phi_{1-s}^2(x - B_s)\big]ds,$$

and for $i \geq 3$, μ_i^t is defined by recursion, i.e.,

$$\mu_i^t(x) := \sum_{j=1}^{i-1}\binom{i}{j}\int_t^1 \mathbf{E}\big[\mu_j^s(x - B_s)\mu_{i-j}^s(x - B_s)\big]ds.$$

Remark 1.2. 1. Taking $x \to \infty$, we have $\mu_r(\infty) = \frac{r!\sigma^{2r}}{2^r}$, Theorem 1.1 indicates that

$$\mathcal{L}\left(\frac{Z_n}{n} \ \Big| \ Z_n > 0\right) \Longrightarrow \mathcal{L}(Y),$$

where Y is an exponential random variable with mean $\sigma^2/2$. Indeed, taking $\sigma^2 = 2$ for simplicity, we have $\mu_r^t(\infty) = r!(1-t)^{r-1}$ by induction. In fact, $\mu_1^t(\infty) = 1$, $\mu_2^t(\infty) = 2(1-t)$. Assume that, for any $r \geq 3$, $\mu_{r-1}^t(\infty) = (r-1)!(1-t)^{r-2}$. Then

$$\mu_r^t(\infty) = \sum_{i=1}^{r-1}\binom{r}{i}\int_t^1 \mu_i^s(\infty)\mu_{r-i}^s(\infty)ds$$

$$= \sum_{i=1}^{r-1}\frac{r\cdots(r-i+1)}{i!}\int_t^1 i!(1-s)^{i-1}(r-i)!(1-s)^{r-i-1}ds$$

$$= \sum_{i=1}^{r-1}\int_t^1 r!(1-s)^{r-2}ds$$

$$= r!(1-t)^{r-1}.$$

It follows from the same argument as above that $\mu_r(\infty) = r!$. This is the classical theorem proved by Yaglom (1947) under a third moment assumption (removed by Kesten et al., 1966). We refer to Lyons et al. (1995), Geiger (2000) and Ren et al. (2018) for elegant probabilistic proofs of Yaglom's theorem.

2. It is worth noting that the limit distribution $\mathcal{L}(Y(x))$ is related to some binary branching Brownian motion obtained by Fleischmann and Siegmund-Schultze (1978). More specifically, the binary branching Brownian motion $(X_t, 0 \leq t < 1)$ evolves as follows. At time zero, an initial particle moves according to a Brownian motion and splits into two at a time which is uniformly distributed on $(0, 1)$. Then the two particles independently branch and move. Each particle born at time t moves relatively to its position according to a Brownian motion up to a time which is

uniformly distributed on $(t, 1)$ and this process continues. The number of particles at time t with position in $(-\infty, x]$ of the process X_t, rescaled by $1 - t$, converges in law to $Y(x)$ as $t \to 1$.

3. The assumption (1.2) can be generalized to the case that ν is in the domain of attraction of a stable law with subtle changes on the proof of Theorem 1.1. More explicitly, we assume that there exists two positive constant sequences $\{a_n\}$ and $\{b_n\}$ such that, as $n \to \infty$,

$$\frac{S_n - a_n}{b_n} \Rightarrow \eta,$$

where S_n is a sum of n independent random variables with common law ν, η has a non-degenerate distribution function $G(x)$ and $G(x)$ is continuous. In this case, we have, for all $x \in \mathbb{R}$, as $n \to \infty$,

$$\mathcal{L}\left(\frac{Z^{(n)}(-\infty, b_n x + a_n]}{n} \;\middle|\; Z_n > 0 \right) \implies \mathcal{L}(L(x)),$$

where the distribution of $L(x)$ is determined by its moments in a similar way as Theorem 1.1 (with the function Φ and Brownian motion being replaced by function G and stable process, respectively).

4. For the subcritical case, we refer to the recent work Hong and Yao (2023).

2. Reduced Galton–Watson Process

To prove our main theorem, we calculate the moments of $\left(\frac{Z^{(n)}(-\infty, \sqrt{n}x]}{n} \;\middle|\; Z_n > 0 \right)$ and use the result regarding the moment convergence problem, which ensures that the limit distribution is uniquely determined by its moments. By adapting the many-to-few formula to the critical branching random walk, we can calculate explicitly its moments. The key step to formulate the many-to-few lemma in the critical case is based on the observation that to count $Z^{(n)}(A)$, the number of the particles in the n-th generation positioned in $A \in \mathcal{B}(\mathbb{R})$, only the ancestors at generation k ($k \leqslant n$) having at least one offspring at generation n make contributions. The process by which these ancestors are composed is called "reduced branching process" in the literature.

Recall that $\{Z_n, n \geq 0\}$ is a G–W process. For each fixed n, we denote by $Z_{k,n}$ the number of particles in the original process at time $k \leq n$ having a positive number of descendants at time n. That is, $\{Z_{k,n}, 0 \leqslant k \leqslant n\}$ is the so-called *reduced G–W process*, which is obtained by removing all individuals of the original process that do not have descendants at time n. In Fleischmann and Siegmund-Schultze (1977), the authors proved a functional conditional limit theorem for the reduced critical G–W process, which says that $Z_{nt,n}(0 \leqslant t < 1)$, given that $Z_n > 0$, has a limit distribution specified by a Yule process with a suitable transformation of time. Also in Fleischmann and Siegmund-Schultze (1978), they obtained a conditional invariance principle establishing convergence of the reduced critical branching random walk (under contractions of time and space) to a certain binary branching Brownian motion on $[0, 1)$. In our proof, we would rather make use of the behaviour of the reduced process at time n than that at time nt.

Let us denote by $\hat{Z}^{(n)}$ a point process with law

$$\mathcal{L}\left(Z^{(n)} \mid Z_n > 0 \right)$$

and $\hat{Z}_{k,n}$ a random variable with law

$$\mathcal{L}\left(Z_{k,n} \mid Z_n > 0 \right).$$

Let f be the generating function of offspring distribution p, that is

$$f(s) := \sum_{j=0}^{\infty} p_j s^j, \quad s \in [0, 1].$$

Let $f_{(n)}$ be the generating function of Z_n, by the branching property, we have

$$f_{(n)}(s) = f(f_{(n-1)}(s)),$$

with the convention that $f_{(0)}(s) := s$, $f_{(1)}(s) := f(s)$.

Fleischmann and Siegmund-Schultze (1977) investigated the reduced critical G–W process and showed that, for any $0 \leq k \leq n$, the generating function of $\hat{Z}_{k,n}$ was given by

$$\hat{f}_{(k),n}(s) := \mathbf{E}\big[s^{\hat{Z}_{k,n}}\big] = \frac{f_{(k)}\big(f_{(n-k)}(0) + s\big(1 - f_{(n-k)}(0)\big)\big) - f_{(n)}(0)}{1 - f_{(n)}(0)}, \quad s \in [0,1].$$

In Proposition 1.1, they showed that $\{\hat{Z}_{k,n}; 0 \leq k \leq n\}$ is a inhomogeneous G–W process, and the offspring probability generating function of a particle at time $k-1$ $\big(1 \leq k \leq n\big)$ is given by

$$\hat{f}_{k,n}(s) := \mathbf{E}\big[s^{\hat{Z}_{k,n}} \mid \hat{Z}_{k-1,n} = 1\big] = \frac{f\big(f_{(n-k)}(0) + s\big(1 - f_{(n-k)}(0)\big)\big) - f_{(n-k+1)}(0)}{1 - f_{(n-k+1)}(0)}, s \in [0,1].$$

Then, we get the following lemma by simple calculations.

Lemma 2.1. *Assume that the offspring distribution p has finite moments of all orders, let $\hat{f}_{k,n}^{(r)}$ be the r $(r \geq 1)$-order derivatives of $\hat{f}_{k,n}$, then*

$$\hat{f}_{k,n}^{(r)}(1) = \frac{\big(1 - f_{(n-k)}(0)\big)^r}{1 - f_{(n-k+1)}(0)} f^{(r)}(1). \tag{2.1}$$

3. Conditioned reduced spatial trees and multiple spines

In order to introduce conditioned reduced family trees of spatial branching processes and the many-to-few formula, we recall the standard Ulam–Harris–Neveu formalism for rooted, ordered, locally finite trees.

3.1. *Family trees.* We introduce the space

$$\mathcal{U} := \{\varnothing\} \cup \bigcup_{n=1}^{\infty} \big(\mathbb{N}^*\big)^n,$$

where $\mathbb{N}^* := \{1, 2, \cdots\}$. An element in $\mathcal{U}$ is a finite sequence of integers and we think of the elements of $\mathcal{U}$ as the labels of nodes (particles) in a tree. For example, $u = 236$ is the sixth child of the third child of the second child of the initial ancestor $\varnothing$. For any two elements u and v of $\mathcal{U}$, let uv be the concatenated element, with the convention that $u\varnothing = \varnothing u = u$. For a given vertex $u = u_1 \cdots u_n \in \mathcal{U}$, we denote by $|u| := n$ its generation (with $u = \varnothing$ if $n = 0$) and $\overleftarrow{u} := u_1 \cdots u_{n-1}$ its parent.

A tree τ is a subset of $\mathcal{U}$ satisfying the following properties:

- $\varnothing \in \tau$;
- if $u \in \tau \backslash \{\varnothing\}$, then $\overleftarrow{u} \in \tau$;
- for each $u \in \tau$, there exists $N_u(\tau) \in \{0, 1, 2, \cdots\}$ such that for every $j \in \mathbb{N}^*$, $uj \in \tau$ if and only if $1 \leq j \leq N_u(\tau)$.

The integer $N_u := N_u(\tau)$ represents the number of offspring of the particle $u \in \tau$. These rooted, ordered, locally finite trees are often called *family trees*. We denote by $\mathcal{T}$ the set of all trees.

3.2. *G–W trees and G–W processes.* Let $\big(N_u, u \in \mathcal{U}\big)$ be a collection of independent random variables with distribution p indexed by $\mathcal{U}$. Denote by $\mathbb{T}$ the random subset of $\mathcal{U}$ defined by

$$\mathbb{T} := \big\{u = u_1 \cdots u_n \in \mathcal{U} : n \geq 0, u_j \leq N_{u_1 \cdots u_{j-1}}, \text{ for all } 1 \leq j \leq n\big\}.$$

Then, $\mathbb{T}$ is a *G–W tree* with offspring distribution p and

$$Z_n := \#\big\{u \in \mathbb{T} : |u| = n\big\}$$

is the associated G–W process.

3.3. *Random spatial trees and branching random walks.* A *spatial tree* is a tree $\tau \in \mathcal{T}$ enriched with additional spatial motion: for each $u \in \tau$, $V(u) \in \mathbb{R}$ is the position of u. Formally, a spatial tree is a set of the form

$$\mathbf{t} := \big\{\big(u, V(u), u \in \tau\big)\big\}.$$

We denote by $\mathbf{T}$ the set of all spatial trees. For $\mathbf{t} \in \mathbf{T}$, we let

$$G_n := \big\{u : \big(u, V(u)\big) \in \mathbf{t}, |u| = n\big\}$$

be the set of all individuals at generation n.

A branching random walk with reproduction law p and displacement distribution ν is a random spatial tree such that

- τ is G–W tree with reproduction law p;
- for each $u \in \tau$, the offspring of u are born at distances from $V(u)$ which are given by a family of independent random variables with common law ν.

Thus, the random measure

$$Z^{(n)}(\cdot) := \sum_{u \in G_n} \delta_{V(u)}(\cdot)$$

defines our branching random walk. For every Borel set A, the number of particles located in A in the n-th generation is defined by

$$Z^{(n)}(A) := \#\big\{u \in G_n : V(u) \in A\big\}.$$

3.4. *Reduced random spatial trees conditioned on* $\big\{Z_n > 0\big\}$. Given $n \in \mathbb{N}$, the reduced random spatial tree is defined by removing all nodes of the original random tree that have no branches at time n. By the fact stated in Section 2, the reduced random spatial tree conditioned on the event $\big\{Z_n > 0\big\}$ is a random spatial tree with inhomogeneous branching mechanism: the reproduction law of a particle u with $|u| = k$ $\big(0 \leq k \leq n-1\big)$, denote by $\big\{p_l\big(k+1, n\big)\big\}_{l \geq 0}$, is determined by the generating function $\hat{f}_{k+1,n}$, the offspring's displacements are given by the same step law ν.

Let $\hat{\mathbb{T}}_n$ denote the reduced random spatial tree conditioned on the event $\big\{Z_n > 0\big\}$, which we call *conditioned reduced random spatial tree.* For each $0 \leq k \leq n$, we denote by

$$\hat{G}_{k,n} := \big\{u : \big(u, V(u)\big) \in \hat{\mathbb{T}}_n, |u| = k\big\}$$

the set of all individuals at generation k of conditioned reduced random spatial tree $\hat{\mathbb{T}}_n$. Then, the random measure

$$\hat{Z}^{(k,n)}(\cdot) := \sum_{u \in \hat{G}_{k,n}} \delta_{V(u)}(\cdot)$$

is the conditioned reduced BRW and the random process $\hat{Z}_{k,n} := \#\hat{G}_{k,n}$ is the conditioned reduced G–W process. Let

$$\hat{G}_n := \hat{G}_{n,n}, \ \hat{Z}^{(n)} := \hat{Z}^{(n,n)}, \ \hat{Z}_n := \hat{Z}_{n,n}.$$

3.5. *Conditioned reduced random spatial trees with spines.* In this subsection, given $n \in \mathbb{N}$, we attach to $\left((u, V(u)), u \in \hat{\mathbb{T}}_n \right)$ an extra random path $(w, V(w)) = \{w_k, V(w_k)\}_{0 \leq k \leq n}$ called spine and write $\xi_k := V(w_k)$. For any $r \geq 1$, we denote by $\hat{\hat{\mathbb{T}}}_n^{[r]}$ the conditioned reduced random spatial trees with spines. To give a version of many-to-few formula for the conditioned reduced BRW, we need to define a new probability $\mathbf{Q}^{[r]}$. Under $\mathbf{Q}^{[r]}$ the particles in the conditioned reduced random spatial tree behave as follows:

- At time 0, an initial particle positioned at the origin is carrying r marks $1, 2, \cdots, r$. We think of each mark as distinguishing a spine and write $w^1, \cdots, w^r$ for the r spines and $\xi^1, \cdots, \xi^r$ for the positions of these spines.
- This particle dies at time 1 and has children whose number and displacements are determined by the r-th sized-biased distribution $\left\{ \frac{l^r p_l(1,n)}{\sum_{l=0}^{\infty} l^r p_l(1,n)} \right\}_{l \geq 0}$ (with respect to the law $\{p_l(1,n)\}_{l \geq 0}$) and the law ν, respectively. Then, the r marks each choose a particle among its children to follow independently and uniformly at random.
- For any $1 \leq k \leq n - 1$, a particle with j marks at generation k dies at time $k + 1$ and gives birth independently of all others to its children whose number and displacements are determined by the j-th sized-biased distribution (with respect to the law $\{p_l(k+1, n)\}_{l \geq 0}$) and the law ν, respectively, and the j marks each choose a particle among its children to follow independently and uniformly at random; particles carrying no marks behave just as under $\mathbf{P}$.

Note that, under $\mathbf{Q}^{[r]}$, the particles in the spines give birth to size-biased numbers of children and move to a relative position according to the law ν. The reproduction law depends on how many marks the spine particle is carrying, whereas the position does not.

Let $\mathcal{F}_n^{[r]}$ be the filtration containing all information about the conditioned reduced random spatial tree and the r spines up to time n. For any $\mathcal{F}_n^{[r]}$-measurable F, we can express it as the sum

$$F = \sum_{u^1, \cdots, u^r \in \hat{G}_n} F(u^1, \cdots, u^r) \mathbf{1}_{\{w_n^1 = u^1, \cdots, w_n^r = u^r\}},$$

where the random variable $F(u^1, \cdots, u^r)$ is $\mathcal{F}_n$-measurable and $\mathcal{F}_n$ is the natural filtration of the conditioned reduced random spatial tree. Indeed, one can show that if

$$A \in \mathcal{F}_n \cup \{\{w_k^i = u\} : u \in \mathcal{U}, i \leq r, k \leq n\},$$

then

$$A = \cup_{u^1, \cdots, u^r \in \mathcal{U}} \left(A(u^1, \cdots, u^r) \cap \{w_n^1 = u^1, \cdots, w_n^r = u^r\} \right)$$

for some collection of sets $A(u^1, \cdots, u^r) \in \mathcal{F}_n$. This property is retained on taking countable unions or complements, so it holds for any $A \in \mathcal{F}_n$. And then it is easy to check that this representation holds for any $\mathcal{F}_n^{[r]}$-measurable function by standard argument.

4. Proofs

4.1. *Many-to-few formula for conditioned reduced BRW.* We assume that the assumptions in Theorem 1.1 hold throughout this section. For $r \geq 1$, we define the r-th moment of the offspring distribution of a particle in the conditional reduced random spatial tree $\hat{\mathbb{T}}_n$ at time $k - 1$ $(1 \leq k \leq n)$ as

$$m_r(k) := \mathbf{E}\left[(\hat{Z}_{k,n})^r \big| \hat{Z}_{k-1,n} = 1 \right] = \sum_{l=0}^{\infty} l^r p_l(k, n).$$

By (2.1), we have

$$m_1(k) = \frac{1 - f_{(n-k)}(0)}{1 - f_{(n-k+1)}(0)} \tag{4.1}$$

and

$$m_2(k) = \frac{\left(1 - f_{(n-k)}(0)\right)^2}{1 - f_{(n-k+1)}(0)}\sigma^2 + \frac{1 - f_{(n-k)}(0)}{1 - f_{(n-k+1)}(0)}. \tag{4.2}$$

Now we are ready to state a version of many-to-few formula for the conditioned reduced BRW. We begin with the many-to-one and many-to-two formulas.

Lemma 4.1 (Many-to-one formula). *For any $n \in \mathbb{N}$ and any $\mathcal{F}_n^{[1]}$ measurable function F, we have*

$$\mathbf{E}\Big[\sum_{u \in \hat{G}_n} F(u)\Big] = \frac{1}{1 - f_{(n)}(0)}\mathbf{E}_{\mathbf{Q}^{[1]}}\big[F(w_n^1)\big]. \tag{4.3}$$

Proof: The many-to-one formula has been used in many situations to compute the first moment, cf. Lyons et al. (1995); Hu and Shi (2009); Aïdékon (2013); Harris and Roberts (2017); Bansaye (2019). In our setting, we can prove by induction that

$$\mathbf{E}\Big[\sum_{u \in \hat{G}_n} F(u)\Big] = \mathbf{E}_{\mathbf{Q}^{[1]}}\Big[F(w_n^1)\prod_{k=1}^{n} m_1(k)\Big].$$

Then, by (4.1), we obtain (4.3). □

If we take $F(w_n^1) = \mathbf{1}_{\{V(w_n^1) \le \sqrt{n}x\}} = \mathbf{1}_{\{\xi_n^1 \le \sqrt{n}x\}}$, then $F(u) = \mathbf{1}_{\{V(u) \le \sqrt{n}x\}}$, it follows from (4.3) that

$$\mathbf{E}\Big[\frac{\hat{Z}^{(n)}(-\infty, \sqrt{n}x]}{n}\Big] = \frac{1}{n}\mathbf{E}\Big[\sum_{u \in \hat{G}_n} \mathbf{1}_{\{V(u) \le \sqrt{n}x\}}\Big] = \frac{1}{n(1 - f_{(n)}(0))}\mathbf{Q}^{[1]}\big(\xi_n^1 \le \sqrt{n}x\big),$$

where ξ_n^1 is the random walk with the following step law

$$\mathbf{Q}^{[1]}\big(\xi_1^1 \in \cdot \mid \xi_0^1 = x\big) = \nu(\cdot - x). \tag{4.4}$$

Note that

$$n\big(1 - f_{(n)}(0)\big) \to \frac{2}{\sigma^2}$$

(cf. Athreya and Ney, 1972, Chapter I, Section 9, Theorem 1) and

$$\lim_{n \to \infty} \mathbf{Q}^{[1]}\big(\xi_n^1 \le \sqrt{n}x\big) = \Phi(x).$$

Then, we immediately obtain

Corollary 4.2 (Limit of first moment).

$$\lim_{n \to \infty} \mathbf{E}\Big[\frac{\hat{Z}^{(n)}(-\infty, \sqrt{n}x]}{n}\Big] = \frac{\sigma^2}{2}\Phi(x). \tag{4.5}$$

To compute the second moment, we need the following many-to-two formula.

Lemma 4.3 (many-to-two formula). *For any $n \in \mathbb{N}$ and any $\mathcal{F}_n^{[2]}$ measurable function F, we have*

$$\mathbf{E}\bigg[\sum_{u^1, u^2 \in \hat{G}_n} F(u^1, u^2)\bigg]$$

$$= \frac{1}{1 - f_{(n)}(0)} \mathbf{E}_{\mathbf{Q}^{[2]}}\Big[F(w_n^1, w_n^1)\Big] + \frac{\sigma^2}{1 - f_{(n)}(0)} \sum_{j=1}^{n} \mathbf{E}_{\mathbf{Q}^{[2]}}\Big[F(w_{j,n}^1, w_{j,n}^2)\Big], \tag{4.6}$$

where w_n^1 is the spine carrying two marks, $w_{j,n}^1$ and $w_{j,n}^2$ are two spines having the same trajectory before time j and splitting at time j.

Proof: For a version of the many-to-two formula for the conditioned reduced BRW, we apply the many-to-two formula with martingale $\zeta \equiv 1$ in Lemma 8 of Harris and Roberts (2017). In fact, the result in Harris and Roberts (2017) is for time-homogeneous branching random walk and the conditioned reduced BRW is time-inhomogeneous. Checking that the proof of Harris and Roberts (2017), one only needs to replace the moment of offspring distribution $\{p_l\}_{l \geq 0}$ with that of $\{p_l(k, n)\}_{l \geq 0}$ at time $k - 1$.

Let $T^{[2]}$ denote the first split time at which marks 1 and 2 are carried by different particles. Under $\mathbf{Q}^{[2]}$ the initial particle which carries marks 1 and 2 branches at time 1 into l particles with probability $l^2 p_l(1, n)/m_2(1)$. At such a branching event, the two marks follow the same particle with probability $1/l$. Thus

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

The distribution of $T^{[2]}$ is determined by first and second moments of offspring distribution. Indeed, under $\mathbf{Q}^{[2]}$ the particle carrying two marks branches according to $\{l^2 p_l(k, n)/m_2(k)\}$ at time k and the two marks follow the same particle with probability $1/l$. Hence,

$$\mathbf{Q}^{[2]}(T^{[2]} > i) = \prod_{k=1}^{i} \frac{m_1(k)}{m_2(k)}, \quad i \geq 1, \tag{4.7}$$

and

$$\mathbf{Q}^{[2]}(T^{[2]} = i) = \mathbf{Q}^{[2]}(T^{[2]} > i - 1) - \mathbf{Q}^{[2]}(T^{[2]} > i) = \prod_{k=1}^{i-1} \frac{m_1(k)}{m_2(k)}\bigg(1 - \frac{m_1(i)}{m_2(i)}\bigg), \tag{4.8}$$

with the convention that $\prod_{k=1}^{0} := 1$.

Since the reproduction and position of an individual are independent, writing out the many-to-two formula of Harris and Roberts (2017) and simplifying, we get

$$\mathbf{E}\bigg[\sum_{u^1, u^2 \in \hat{G}_n} F(u^1, u^2)\bigg] = \mathbf{E}_{\mathbf{Q}^{[2]}}\bigg[F(w_n^1, w_n^1) \prod_{k=1}^{n} m_2(k)\bigg]\mathbf{Q}^{[2]}(T^{[2]} > n)$$

$$+ \sum_{j=1}^{n} \mathbf{E}_{\mathbf{Q}^{[2]}}\bigg[F(w_{j,n}^1, w_{j,n}^2) \prod_{k=1}^{j} m_2(k) \prod_{k=j+1}^{n} (m_1(k))^2\bigg]\mathbf{Q}^{[2]}(T^{[2]} = j),$$

where w_n^1 is the spine carrying two marks, $w_{j,n}^1$ and $w_{j,n}^2$ are two spines having the same trajectory before time j and splitting at time j.

It follows from (4.1), (4.2), (4.7) and (4.8) that

$$\prod_{k=1}^{n} m_2(k)\mathbf{Q}^{[2]}(T^{[2]} > n) = \frac{1}{1 - f_{(n)}(0)}$$

and

$$\prod_{k=1}^{j} m_2(k) \prod_{k=j+1}^{n} \big(m_1(k)\big)^2 \mathbf{Q}^{[2]}\big(T^{[2]} = j\big) = \frac{\sigma^2}{1 - f_{(n)}(0)}.$$

Thus, we prove (4.6). □

Corollary 4.4 (Limit of second moment).

$$\lim_{n\to\infty} \mathbf{E}\left[\left(\frac{\hat{Z}^{(n)}(-\infty, \sqrt{n}x]}{n}\right)^2\right] = \frac{\sigma^4}{2} \int_0^1 \mathbf{E}\Big(\Phi_{1-t}^2\big(x - B_t\big)\Big)dt, \tag{4.9}$$

where B_t is a Brownian motion and

$$\Phi_{1-t}(x) := \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi(1-t)}} e^{-\frac{y^2}{2(1-t)}} \, dy.$$

Proof: Taking

$$F\big(w_n^1, w_n^2\big) = \mathbf{1}_{\big\{V\big(w_n^1\big)\le\sqrt{n}x, V\big(w_n^2\big)\le\sqrt{n}x\big\}} = \mathbf{1}_{\big\{\xi_n^1\le\sqrt{n}x, \xi_n^2\le\sqrt{n}x\big\}},$$

then

$$F\big(u^1, u^2\big) = \mathbf{1}_{\big\{V(u^1)\le\sqrt{n}x, V(u^2)\le\sqrt{n}x\big\}}.$$

It follows from (4.6) that

$$\mathbf{E}\left[\left(\frac{\hat{Z}^{(n)}(-\infty, \sqrt{n}x]}{n}\right)^2\right] = \frac{1}{n^2}\mathbf{E}\left[\sum_{u^1,u^2\in\hat{G}_n} \mathbf{1}_{\big\{V(u^1)\le\sqrt{n}x, V(u^2)\le\sqrt{n}x\big\}}\right]$$

$$= \frac{1}{n^2\big(1 - f_{(n)}(0)\big)}\mathbf{Q}^{[2]}\big(\xi_n^1 \le \sqrt{n}x\big)$$

$$+ \frac{\sigma^2}{n^2\big(1 - f_{(n)}(0)\big)}\sum_{j=1}^{n}\mathbf{Q}^{[2]}\big(\xi_{j,n}^1 \le \sqrt{n}x, \xi_{j,n}^2 \le \sqrt{n}x\big),$$

where $\xi_{j,n}^1$ and $\xi_{j,n}^2$ are two random walks having the same trajectory before time j and splitting at time j. For the first term in last equation, taking limit as $n \to \infty$, we have

$$\lim_{n\to\infty} \frac{1}{n^2\big(1 - f_{(n)}(0)\big)}\mathbf{Q}^{[2]}\big(\xi_n^1 \le \sqrt{n}x\big) = 0.$$

Taking limit for the second term, we get

$$\lim_{n\to\infty} \frac{\sigma^2}{n^2\big(1 - f_{(n)}(0)\big)}\sum_{j=1}^{n}\mathbf{Q}^{[2]}\big(\xi_{j,n}^1 \le \sqrt{n}x, \xi_{j,n}^2 \le \sqrt{n}x\big)$$

$$= \lim_{n\to\infty} \frac{\sigma^2}{n^2\big(1 - f_{(n)}(0)\big)}\sum_{j=1}^{n}\mathbf{Q}^{[2]}\left[\mathbf{Q}^{[2]}\Big(\xi_j^1 + \tilde{\xi}_{n-j}^1 \le \sqrt{n}x, \xi_j^1 + \tilde{\xi}_{n-j}^2 \le \sqrt{n}x \mid \sigma\big(\xi_j^1\big)\Big)\right]$$

$$= \lim_{n\to\infty} \frac{\sigma^2}{n^2\big(1 - f_{(n)}(0)\big)}\sum_{j=1}^{n}\mathbf{Q}^{[2]}\left[\left(\mathbf{Q}^{[2]}\Big(\xi_j^1 + \tilde{\xi}_{n-j}^1 \le \sqrt{n}x \mid \sigma\big(\xi_j^1\big)\Big)\right)^2\right]$$

$$= \lim_{n\to\infty} \frac{\sigma^2}{n\big(1 - f_{(n)}(0)\big)}\frac{1}{n}\sum_{j=1}^{n}\mathbf{Q}^{[2]}\left[\left(\mathbf{Q}^{[2]}\Big(\frac{\xi_j^1}{\sqrt{n}} + \frac{\tilde{\xi}_{n-j}^1}{\sqrt{n}} \le x \mid \sigma\big(\xi_j^1\big)\Big)\right)^2\right]$$

$$= \frac{\sigma^4}{2}\int_0^1 \mathbf{E}\big(\Phi_{1-t}^2\big(x - B_t\big)\big)dt, \tag{4.10}$$

where $\tilde{\xi}^1_{n-j}$ and $\tilde{\xi}^2_{n-j}$ are independent and identically distributed, and they are independent with ξ^1_j. The last equality follows by the central limit theorem. Indeed, for any fixed $\epsilon > 0$,

$$\frac{1}{n}\sum_{j=1}^{n}\mathbf{Q}^{[2]}\left[\left(\mathbf{Q}^{[2]}\left(\frac{\xi^1_j}{\sqrt{n}} + \frac{\tilde{\xi}^1_{n-j}}{\sqrt{n}} \le x \mid \sigma(\xi^1_j)\right)\right)^2\right] =: A_{n,1}(x) + A_{n,2}(x) + A_{n,3}(x),$$

where

$$A_{n,1}(x) = \frac{1}{n}\sum_{j=[\epsilon n]}^{[(1-\epsilon)n]}\mathbf{Q}^{[2]}\left[\left(\mathbf{Q}^{[2]}\left(\frac{\xi^1_j}{\sqrt{n}} + \frac{\tilde{\xi}^1_{n-j}}{\sqrt{n}} \le x \mid \sigma(\xi^1_j)\right)\right)^2\right],$$

$$A_{n,2}(x) = \frac{1}{n}\sum_{j=1}^{[\epsilon n]-1}\mathbf{Q}^{[2]}\left[\left(\mathbf{Q}^{[2]}\left(\frac{\xi^1_j}{\sqrt{n}} + \frac{\tilde{\xi}^1_{n-j}}{\sqrt{n}} \le x \mid \sigma(\xi^1_j)\right)\right)^2\right],$$

$$A_{n,3}(x) = \frac{1}{n}\sum_{j=[(1-\epsilon)n]+1}^{n}\mathbf{Q}^{[2]}\left[\left(\mathbf{Q}^{[2]}\left(\frac{\xi^1_j}{\sqrt{n}} + \frac{\tilde{\xi}^1_{n-j}}{\sqrt{n}} \le x \mid \sigma(\xi^1_j)\right)\right)^2\right].$$

Note that $A_{n,2}(x) \le \frac{1}{n}\sum_{j=1}^{[\epsilon n]-1} 1 \le \epsilon$ and $A_{n,3}(x) = \frac{1}{n}\sum_{j=[(1-\epsilon)n]+1}^{n} 1 \le \epsilon$. And by the central limit theorem, we get

$$A_{n,1}(x) = \frac{1}{n}\sum_{j=[\epsilon n]}^{[(1-\epsilon)n]}\mathbf{E}_{\mathbf{Q}^{[2]}}\left[\left(\Phi\left(\left(x - \frac{\xi^1_j}{\sqrt{n}}\right)\left(1 - \frac{j}{n}\right)^{-\frac{1}{2}}\right) + o(1)\right)^2\right].$$

Due to Sakhanenko (2006), for every $n \ge 1$, we can reconstruct a random walk, still denote by $\{\xi^1_n\}_{n\in\mathbb{N}}$, and a Brownian motion $\{B^n_t\}_{t\in\mathbb{R}_+}$ on a common probability space such that

$$\mathbf{Q}^{[2]}\left(\sup_{0\le t\le 1}\left|\xi^1_{(nt)} - B^n_{nt}\right| > \sqrt{n}\pi_n\right) \le \pi_n,$$

where π_n denotes the Prokhorov distance between the distribution on $C[0,1]$ of the process $\xi^1_{(\cdot)}$ (the linear interpolation of ξ^1) and the Brownian motion $B^n_\cdot$, and $\pi_n \to 0$ as $n \to \infty$. It follows from this, $\frac{\xi^1_j}{\sqrt{n}} - \frac{B^n_j}{\sqrt{n}} \ge -\sup_{0\le t\le 1}\left|\xi^1_{(nt)} - B^n_{nt}\right|/\sqrt{n}$ and $\Phi(x+y) - \Phi(x) \le y$ for $y \ge 0$, that

$$A_{n,1}(x)$$

$$= \frac{1}{n}\sum_{j=[\epsilon n]}^{[(1-\epsilon)n]}\mathbf{E}_{\mathbf{Q}^{[2]}}\left[\left(\Phi\left(\left(x - \frac{\xi^1_j}{\sqrt{n}}\right)\left(1 - \frac{j}{n}\right)^{-\frac{1}{2}}\right) + o(1)\right)^2 \mathbf{1}_{\left\{\sup_{0\le t\le 1}\left|\xi^1_{(nt)} - B^n_{nt}\right| \le \sqrt{n}\pi_n\right\}}\right]$$

$$+ \frac{1}{n}\sum_{j=[\epsilon n]}^{[(1-\epsilon)n]}\mathbf{E}_{\mathbf{Q}^{[2]}}\left[\left(\Phi\left(\left(x - \frac{\xi^1_j}{\sqrt{n}}\right)\left(1 - \frac{j}{n}\right)^{-\frac{1}{2}}\right) + o(1)\right)^2 \mathbf{1}_{\left\{\sup_{0\le t\le 1}\left|\xi^1_{(nt)} - B^n_{nt}\right| > \sqrt{n}\pi_n\right\}}\right]$$

$$\le \frac{1}{n}\sum_{j=[\epsilon n]}^{[(1-\epsilon)n]}\mathbf{E}_{\mathbf{Q}^{[2]}}\left[\left(\Phi\left(\left(x - B^n_{j/n} + \pi_n\right)\left(1 - \frac{j}{n}\right)^{-\frac{1}{2}}\right) + o(1)\right)^2\right]$$

$$+ \frac{1}{n}\sum_{j=[\epsilon n]}^{[(1-\epsilon)n]}\mathbf{Q}^{[2]}\left(\sup_{0\le t\le 1}\left|\xi^1_{(nt)} - B^n_{nt}\right| > \sqrt{n}\pi_n\right)$$

$$\le \frac{1}{n}\sum_{j=[\epsilon n]}^{[(1-\epsilon)n]}\mathbf{E}_{\mathbf{Q}^{[2]}}\left[\left(\Phi\left(\left(x - B^n_{j/n}\right)\left(1 - \frac{j}{n}\right)^{-\frac{1}{2}}\right) + o(1)\right)^2\right] + \pi_n^2 + 2\pi_n(1 + o(1)) + \pi_n,$$

the last sum converges, as $n \to \infty$, to the integral

$$\int_{\epsilon}^{1-\epsilon} \mathbf{E}\left[\Phi^2\left((x - B_t)(1 - t)^{-\frac{1}{2}}\right)\right] dt = \int_{\epsilon}^{1-\epsilon} \mathbf{E}\left[\Phi_{1-t}^2(x - B_t)\right] dt.$$

On the other hand, $\frac{\xi_j^1}{\sqrt{n}} - \frac{B_j^n}{\sqrt{n}} \leq \sup_{0 \leq t \leq 1} |\xi_{(nt)}^1 - B_{nt}^n|/\sqrt{n}$, we have

$$A_{n,1}(x)$$

$$\geq \frac{1}{n} \sum_{j=[\epsilon n]}^{[(1-\epsilon)n]} \mathbf{E}_{\mathbf{Q}^{[2]}}\left[\left(\Phi\left((x - \frac{\xi_j^1}{\sqrt{n}})(1 - \frac{j}{n})^{-\frac{1}{2}}\right) + o(1)\right)^2 \mathbf{1}_{\left\{\sup_{0 \leq t \leq 1} |\xi_{(nt)}^1 - B_{nt}^n| \leq \sqrt{n}\pi_n\right\}}\right]$$

$$\geq \frac{1}{n} \sum_{j=[\epsilon n]}^{[(1-\epsilon)n]} \mathbf{E}_{\mathbf{Q}^{[2]}}\left[\left(\Phi\left((x - B_{j/n}^n - \pi_n)(1 - \frac{j}{n})^{-\frac{1}{2}}\right) + o(1)\right)^2\right] - \pi_n$$

$$\geq \frac{1}{n} \sum_{j=[\epsilon n]}^{[(1-\epsilon)n]} \mathbf{E}_{\mathbf{Q}^{[2]}}\left[\left(\Phi\left((x - B_{j/n}^n)(1 - \frac{j}{n})^{-\frac{1}{2}}\right) + o(1)\right)^2\right] + \pi_n^2 - 2\pi_n(1 + o(1)) - \pi_n.$$

Since $\pi_n \to 0$ as $n \to \infty$, we obtain $\lim_{n \to \infty} A_{n,1}(x) = \int_{\epsilon}^{1-\epsilon} \mathbf{E}\left[\Phi_{1-t}^2(x - B_t)\right] dt$. Then taking $\epsilon \to 0$, the last equality in display (4.10) holds. Thus, (4.9) follows. $\square$

Now, we state the general many-to-few formula (Lemma 8 of Harris and Roberts, 2017) in our setting for computing higher-order moments.

Lemma 4.5 (many-to-few formula). *For any $n \in \mathbb{N}$, $r \geq 3$, and any $\mathcal{F}_n^{[r]}$ measurable function F, we have*

$$\mathbf{E}\left[\sum_{u^1,\cdots,u^r \in \hat{G}_n} F\left(u^1, \cdots, u^r\right)\right] = \mathbf{E}_{\mathbf{Q}^{[r]}}\left[F\left(w_n^1, \cdots, w_n^r\right) \prod_{w \in skel^{(r)}(n) \setminus \{\varnothing\}} m_{D\left(\overleftarrow{w}\right)}\left(|w|\right)\right], \qquad (4.11)$$

where $skel^{(r)}(n)$ is the set of all particles up to time n that have carried at least one mark and $D(w)$ is the number of marks carried by particle w.

Corollary 4.6 (Limit of r-th moment). *For any $r \geq 3$, let*

$$\mu_r(x) := \lim_{n \to \infty} \mathbf{E}\left[\left(\frac{\hat{Z}^{(n)}(-\infty, \sqrt{n}x]}{n}\right)^r\right],$$

then $\mu_r(x)$ satisfies the following recursive formula:

$$\mu_r(x) = \sum_{i=1}^{r-1} \binom{r}{i} \int_0^1 \mathbf{E}\left[\mu_i^t(x - B_t)\mu_{r-i}^t(x - B_t)\right] dt, \qquad (4.12)$$

where $\mu_1^t(x) := \frac{\sigma^2}{2}\Phi_{1-t}(x)$, $\mu_2^t(x) := \frac{\sigma^4}{2}\int_t^1 \mathbf{E}\left[\Phi_{1-s}^2(x - B_s)\right] ds$ and for $i \geq 3$, μ_i^t is defined by recursion, i.e.,

$$\mu_i^t(x) := \sum_{j=1}^{i-1} \binom{i}{j} \int_t^1 \mathbf{E}\left[\mu_j^s(x - B_s)\mu_{i-j}^s(x - B_s)\right] ds.$$

Proof: Let

$$F\left(w_n^1, \cdots, w_n^r\right) = \mathbf{1}_{\left\{V\left(w_n^1\right) \leq \sqrt{n}x, \cdots, V\left(w_n^r\right) \leq \sqrt{n}x\right\}} = \mathbf{1}_{\left\{\xi_n^1 \leq \sqrt{n}x, \cdots, \xi_n^r \leq \sqrt{n}x\right\}},$$

then

$$F\left(u^1, \cdots, u^r\right) = \mathbf{1}_{\left\{V(u^1) \leq \sqrt{n}x, \cdots, V(u^r) \leq \sqrt{n}x\right\}}.$$

It follows from (4.11) that

$$
\mathbf{E}\left[\left(\frac{\hat{Z}^{(n)}(-\infty,\sqrt{n}x]}{n}\right)^r\right] = \frac{1}{n^r}\mathbf{E}\left[\sum_{u^1,\cdots,u^r\in\hat{G}_n}\mathbf{1}_{\left\{V(u^1)\le\sqrt{n}x,\cdots,V(u^r)\le\sqrt{n}x\right\}}\right]
$$

$$
= \frac{1}{n^r}\mathbf{E}_{\mathbf{Q}^{[r]}}\left[\mathbf{1}_{\left\{\xi_n^1\le\sqrt{n}x,\cdots,\xi_n^r\le\sqrt{n}x\right\}}\prod_{w\in skel^{(r)}(n)\setminus\{\varnothing\}}m_{D\left(\overleftarrow{w}\right)}(|w|)\right]. \quad (4.13)
$$

To compute the right side of (4.13), we should express the product in the expectation according to the marks the particles carried. We denote by $T^{[r]}$ the split time at which r marks are carried by different children. Let $\{T^{[r]}>n\}$ denote the event that all r marks are carried by the same particle up to time n, that is r marks do not split, and $\{T^{[r]}\le n\}$ denote the event that the r marks have been carried by different particles at some times $0\le j_1,\cdots,j_i\le n$. For the latter case, more specifically, we let $T_1^{[r]}(1,r-1)$ denote the first split time at which one mark and the rest $r-1$ marks are carried by two different particles, $T_1^{[r]}(1,1,r-2)$ denote the first split time at which two of the r marks are carried by two different children and the rest $r-2$ marks are carried by a different particle, $T_2^{[r]}\big(3,(1,r-4)\big)$ denote the second split time at which one of the $r-3$ marks and the rest $r-4$ marks are carried by two different particles. More generally, for $1\le i\le r-1$, we denote by $T_i^{[r]}\Big(r_1,\big(r_2\cdots,(r_i,r-r_1-\cdots-r_i)\cdots\big)\Big)$ the i-th split time at which r_i of the $r-r_1-\cdots-r_{i-1}$ marks and the remaining $r-r_1-\cdots-r_i$ marks are carried by two different children.

- On the event that all r marks do not split up to time n, we have

$$
\frac{1}{n^r}\mathbf{E}_{\mathbf{Q}^{[r]}}\left[\mathbf{1}_{\left\{\xi_n^1\le\sqrt{n}x,\cdots,\xi_n^r\le\sqrt{n}x\right\}}\prod_{w\in skel^{(r)}(n)\setminus\{\varnothing\}}m_{D\left(\overleftarrow{w}\right)}(|w|)\mathbf{1}_{\left\{T^{[r]}>n\right\}}\right]
$$

$$
= \frac{1}{n^r}\mathbf{E}_{\mathbf{Q}^{[r]}}\left[\mathbf{1}_{\left\{\xi_n^1\le\sqrt{n}x\right\}}\prod_{k=1}^n m_r(k)\right]\mathbf{Q}^{[r]}(T^{[r]}>n).
$$

Under $\mathbf{Q}^{[r]}$, the particle which carries r marks branches at time j into l particles with probability $l^r p_l(j,n)/m_r(j)$. At such a branching event, the r marks follow the same particle with probability $1/l^{r-1}$. Thus, we have

$$
\mathbf{Q}^{[r]}(T^{[r]}>n) = \prod_{k=1}^n \frac{m_1(k)}{m_r(k)}.
$$

It follows that

$$
\frac{1}{n^r}\mathbf{E}_{\mathbf{Q}^{[r]}}\left[\mathbf{1}_{\left\{\xi_n^1\le\sqrt{n}x,\cdots,\xi_n^r\le\sqrt{n}x\right\}}\prod_{w\in skel^{(r)}(n)\setminus\{\varnothing\}}m_{D\left(\overleftarrow{w}\right)}(|w|)\mathbf{1}_{\left\{T^{[r]}>n\right\}}\right]
$$

$$
= \frac{1}{n^r}\mathbf{Q}^{[r]}(\xi_n^1\le\sqrt{n}x)\prod_{k=1}^n m_1(k)
$$

$$
= \frac{1}{n^r}\mathbf{Q}^{[r]}(\xi_n^1\le\sqrt{n}x)\prod_{k=1}^n \frac{1-f_{(n-k)}(0)}{1-f_{(n-k+1)}(0)}
$$

$$
= \frac{1}{n^r}\mathbf{Q}^{[r]}(\xi_n^1\le\sqrt{n}x)\frac{1}{1-f_{(n)}(0)}.
$$

Since $r \geq 3$, we obtain, by taking limit as $n \to \infty$,

$$\lim_{n\to\infty} \frac{1}{n^r} \mathbf{Q}^{[r]}\big(\xi_n^1 \leq \sqrt{n}x\big) \frac{1}{1 - f_{(n)}(0)} = 0.$$

The value on the right side of (4.13) converges to zero as $n \to \infty$ on the event that r marks are carried by the same particle up to time n.

- On the event that r marks are carried by r different children at the first split time $T_1^{[r]}(1, \cdots, 1)$, we have

$$\sum_{j=1}^n \frac{1}{n^r} \mathbf{E}_{\mathbf{Q}^{[r]}} \bigg[\mathbf{1}_{\{\xi_n^1 \leq \sqrt{n}x, \cdots, \xi_n^r \leq \sqrt{n}x\}} \prod_{w \in skel^{(r)}(n)\backslash\{\varnothing\}} m_{D(\overleftarrow{w})}\big(|w|\big) \mathbf{1}_{\{T_1^{[r]}(1,\cdots,1)=j\}} \bigg]$$

$$= \sum_{j=1}^n \frac{1}{n^r} \mathbf{E}_{\mathbf{Q}^{[r]}} \bigg[\mathbf{1}_{\{\xi_{j,n}^1 \leq \sqrt{n}x, \cdots, \xi_{j,n}^r \leq \sqrt{n}x\}} \prod_{k=1}^j m_r(k) \prod_{k=j+1}^n \big(m_1(k)\big)^r \bigg]$$

$$\times \mathbf{Q}^{[r]}\big(T_1^{[r]}(1,\cdots,1)=j\big),$$

where $\xi_{j,n}^1, \cdots, \xi_{j,n}^r$ are r random walks having the same trajectory before time j and splitting into r different trajectories at time j. Under $\mathbf{Q}^{[r]}$, the particle which carries r marks branches at time j into l particles with probability $l^r p_l(j,n)/m_r(j)$. At such a branching event, the r marks follow r different particles with probability $\frac{l\cdots(l-r+1)}{l^r}$. Thus, we have

$$\mathbf{Q}^{[r]}\big(T_1^{[r]}(1,\cdots,1)=j\big) = \prod_{k=1}^{j-1} \frac{m_1(k)}{m_r(k)} \frac{\hat{f}_{j,n}^{(r)}(1)}{m_r(j)}.$$

Then,

$$\sum_{j=1}^n \frac{1}{n^r} \mathbf{E}_{\mathbf{Q}^{[r]}} \bigg[\mathbf{1}_{\{\xi_n^1 \leq \sqrt{n}x, \cdots, \xi_n^r \leq \sqrt{n}x\}} \prod_{w \in skel^{(r)}(n)\backslash\{\varnothing\}} m_{D(\overleftarrow{w})}\big(|w|\big) \mathbf{1}_{\{T_1^{[r]}(1,\cdots,1)=j\}} \bigg]$$

$$= \sum_{j=1}^n \frac{1}{n^r} \mathbf{Q}^{[r]}\big(\xi_{j,n}^1 \leq \sqrt{n}x, \cdots, \xi_{j,n}^r \leq \sqrt{n}x\big) \prod_{k=j+1}^n \big(m_1(k)\big)^r \prod_{k=1}^{j-1} m_1(k)\hat{f}_{j,n}^{(r)}(1)$$

$$= \sum_{j=1}^n \frac{1}{n^r} \mathbf{Q}^{[r]}\big(\xi_{j,n}^1 \leq \sqrt{n}x, \cdots, \xi_{j,n}^r \leq \sqrt{n}x\big) \frac{f^{(r)}(1)}{1 - f_{(n)}(0)},$$

where the last equality holds from Lemma 2.1 and (4.1). Taking limit as $n \to \infty$, we get

$$\lim_{n\to\infty} \sum_{j=1}^n \frac{1}{n^r} \mathbf{Q}^{[r]}\big(\xi_{j,n}^1 \leq \sqrt{n}x, \cdots, \xi_{j,n}^r \leq \sqrt{n}x\big) \frac{f^{(r)}(1)}{1 - f_{(n)}(0)} = 0.$$

That is the term on the right side of (4.13) vanishes as $n \to \infty$ on the event that r marks are carried by r different children at the first split time $T_1^{[r]}(1, \cdots, 1)$.

- On the event that r marks are carried by $r-1$ different children at the first split time $T_1^{[r]}(1,\cdots,1,2)$, we have

$$\sum_{j_1=1}^{n} \frac{1}{n^r} \mathbf{E}_{\mathbf{Q}^{[r]}} \left[\mathbf{1}_{\left\{ \xi_n^1 \le \sqrt{n}x, \cdots, \xi_n^r \le \sqrt{n}x \right\}} \prod_{w \in skel^{(r)}(n) \setminus \{\varnothing\}} m_{D(\overleftarrow{w})}(|w|) \mathbf{1}_{\left\{ T_1^{[r]}(1,\cdots,1,2) = j_1 \right\}} \right]$$

$$= \sum_{j_1=1}^{n} \sum_{j_2=j_1+1}^{n} \frac{1}{n^r} \mathbf{Q}^{[r]} \left(\xi_{j_1,n}^1 \le \sqrt{n}x, \cdots, \xi_{j_1,n}^{r-2} \le \sqrt{n}x, \xi_{j_1,j_2,n}^{r-1} \le \sqrt{n}x, \xi_{j_1,j_2,n}^r \le \sqrt{n}x \right)$$

$$\times \prod_{k=1}^{j_1} m_r(k) \prod_{k=j_1+1}^{n} (m_1(k))^{r-2} \prod_{k=j_1+1}^{j_2} m_2(k) \prod_{k=j_2+1}^{n} (m_1(k))^2$$

$$\times \mathbf{Q}^{[r]} \left(T_1^{[r]}(1,\cdots,1,2) = j_1, T_2^{[r]}(1,\cdots,1,(1,1)) = j_2 \right)$$

$$+ \sum_{j_1=1}^{n} \frac{1}{n^r} \mathbf{Q}^{[r]} \left(\xi_{j_1,n}^1 \le \sqrt{n}x, \cdots, \xi_{j_1,n}^{r-1} \le \sqrt{n}x \right)$$

$$\times \prod_{k=1}^{j_1} m_r(k) \prod_{k=j_1+1}^{n} (m_1(k))^{r-2} \prod_{k=j_1+1}^{n} m_2(k)$$

$$\times \mathbf{Q}^{[r]} \left(T_1^{[r]}(1,\cdots,1,2) = j_1, T_2^{[r]}(1,\cdots,1,(1,1)) > n \right),$$

where $\xi_{j_1,j_2,n}^{r-1}, \xi_{j_1,j_2,n}^r$ are two random walks having the same trajectory as $\xi_{j_1,n}^1$ before time j_1 and splitting into two different trajectories at time j_2. Under $\mathbf{Q}^{[r]}$, the particle which carries r marks branches at time j_1 into l particles with probability $l^r p_l(j_1, n)/m_r(j_1)$. At such a branching event, the r marks follow $r-1$ different particles with probability $\binom{r}{2} \frac{l \cdots (l-r+2)}{l^r}$. As a result, we can get

$$\mathbf{Q}^{[r]} \left(T_1^{[r]}(1,\cdots,1,2) = j_1, T_2^{[r]}(1,\cdots,1,(1,1)) = j_2 \right)$$

$$= \prod_{k=1}^{j_1-1} \frac{m_1(k)}{m_r(k)} \frac{\binom{r}{2} \hat{f}_{j_1,n}^{(r-1)}(1)}{m_r(j_1)} \prod_{k=j_1+1}^{j_2-1} \frac{m_1(k)}{m_2(k)} \frac{\hat{f}_{j_2,n}^{(2)}(1)}{m_2(j_2)},$$

and

$$\mathbf{Q}^{[r]} \left(T_1^{[r]}(1,\cdots,1,2) = j_1, T_2^{[r]}(1,\cdots,1,(1,1)) > n \right)$$

$$= \prod_{k=1}^{j_1-1} \frac{m_1(k)}{m_r(k)} \frac{\binom{r}{2} \hat{f}_{j_1,n}^{(r-1)}(1)}{m_r(j_1)} \prod_{k=j_1+1}^{n} \frac{m_1(k)}{m_2(k)}.$$

Then,

$$\sum_{j_1=1}^{n} \frac{1}{n^r} \mathbf{E}_{\mathbf{Q}^{[r]}} \left[\mathbf{1}_{\left\{ \xi_n^1 \le \sqrt{n}x, \cdots, \xi_n^r \le \sqrt{n}x \right\}} \prod_{w \in skel^{(r)}(n) \setminus \{\varnothing\}} m_{D(\overleftarrow{w})}(|w|) \mathbf{1}_{\left\{ T_1^{[r]}(1,\cdots,1,2) = j_1 \right\}} \right]$$

$$= \sum_{j_1=1}^{n} \sum_{j_2=j_1+1}^{n} \frac{1}{n^r} \mathbf{Q}^{[r]} \left(\xi_{j_1,n}^1 \le \sqrt{n}x, \cdots, \xi_{j_1,n}^{r-2} \le \sqrt{n}x, \xi_{j_1,j_2,n}^{r-1} \le \sqrt{n}x, \xi_{j_1,j_2,n}^r \le \sqrt{n}x \right)$$

$$\times \frac{\binom{r}{2} f^{(r-1)}(1)\sigma^2}{1-f_{(n)}(0)} + \sum_{j_1=1}^{n} \frac{1}{n^r} \mathbf{Q}^{[r]} \left(\xi_{j_1,n}^1 \le \sqrt{n}x, \cdots, \xi_{j_1,n}^{r-1} \le \sqrt{n}x \right) \frac{\binom{r}{2} f^{(r-1)}(1)}{1-f_{(n)}(0)}, \qquad (4.14)$$

where the last equality holds from Lemma 2.1 and (4.1).

– When $r = 3$, taking limit as $n \to \infty$, we get

$$\lim_{n \to \infty} \sum_{j_1=1}^{n} \frac{1}{n^r} \mathbf{Q}^{[r]}\left(\xi_{j_1,n}^{1} \le \sqrt{n}x, \cdots, \xi_{j_1,n}^{r-1} \le \sqrt{n}x \right) \frac{\binom{r}{2} f^{(r-1)}(1)}{1 - f_{(n)}(0)} = 0$$

and

$$\lim_{n \to \infty} \sum_{j_1=1}^{n} \sum_{j_2=j_1+1}^{n} \frac{\binom{r}{2} f^{(r-1)}(1)\sigma^2}{n^r \left(1 - f_{(n)}(0)\right)}$$

$$\times \mathbf{Q}^{[r]}\left(\xi_{j_1,n}^{1} \le \sqrt{n}x, \cdots, \xi_{j_1,n}^{r-2} \le \sqrt{n}x, \xi_{j_1,n}^{r-1} \le \sqrt{n}x, \xi_{j_1,j_2,n}^{r} \le \sqrt{n}x \right)$$

$$= \frac{3\sigma^6}{2} \lim_{n \to \infty} \sum_{j_1=1}^{n} \sum_{j_2=j_1+1}^{n} \frac{1}{n^2}$$

$$\times \mathbf{Q}^{[3]}\left(\xi_{j_1,n}^{1} \le \sqrt{n}x, \xi_{j_1,j_2,n}^{2} \le \sqrt{n}x, \xi_{j_1,j_2,n}^{3} \le \sqrt{n}x \right)$$

$$= \frac{3\sigma^6}{2} \lim_{n \to \infty} \sum_{j_1=1}^{n} \sum_{j_2=j_1+1}^{n} \frac{1}{n^2}$$

$$\times \mathbf{Q}^{[3]}\left[\mathbf{Q}_{\xi_{j_1,j_1}^{1}}^{[3]}\left(\xi_{j_1,n}^{1} \le \sqrt{n}x \right) \mathbf{Q}_{\xi_{j_1,j_1}^{1}}^{[3]}\left(\xi_{j_1,j_2,n}^{2} \le \sqrt{n}x, \xi_{j_1,j_2,n}^{3} \le \sqrt{n}x \right) \right]$$

$$= \frac{3\sigma^6}{2} \int_0^1 \mathbf{E}\left[\Phi_{1-t_1}\left(x - B_{t_1}^1 \right) \int_{t_1}^1 \mathbf{E}_{B_{t_1}^1}\left[\Phi_{1-t_2}^2\left(x - \left(B_{t_2}^2 - B_{t_1}^2\right) - B_{t_1}^1 \right) \right] dt_2 \right] dt_1,$$

where B^1 and B^2 are two Brownian motions such that $B_s^1 = B_s^2$ for all $s \le t_1$ and $\left\{ B_{s+t_1}^1 - B_{t_1}^1 \right\}_{s \ge 0}$ and $\left\{ B_{s+t_1}^2 - B_{t_1}^2 \right\}_{s \ge 0}$ are independent, the last equality holds by the similar calculations as (4.10).

So, when $r = 3$, we obtain

$$\mu_3(x) = \sum_{i=1}^{2} \binom{3}{i} \int_0^1 \mathbf{E}\left[\mu_i^t(x - B_t) \mu_{3-i}^t(x - B_t) \right] dt.$$

– When $r \ge 4$, it is easy to show that both terms on the right side of (4.14) tend to zero as $n \to \infty$.

• Repeating the calculations, for any $r \ge 4$, only the case that r marks follow two different particles at each split time contributes to the limit of the right side of (4.13). Indeed, under $\mathbf{Q}^{[r]}$, the particle which carries r marks branches at time j_1 into l particles with probability $l^r p_l(j_1, n)/m_r(j_1)$. At such a branching event, the r marks follow three different particles carrying i_1, i_2 and $r - i_1 - i_2$ marks respectively, with probability $\binom{r}{i_1}\binom{r-i_1}{i_2}\frac{l(l-1)(l-2)}{l^r}$. Then

$$\mathbf{Q}^{[r]}\left(T_1^{[r]}(i_1, i_2, r - i_1 - i_2) = j_1 \right) = \prod_{k=1}^{j_1-1} \frac{m_1(k)}{m_r(k)} \frac{\binom{r}{i_1}\binom{r-i_1}{i_2} \hat{f}_{j_1,n}^{(3)}(1)}{m_r(j_1)}.$$

Thus, we have

$$
\lim_{n\to\infty} \frac{1}{n^r} \mathbf{E}_{\mathbf{Q}^{[r]}}\left[\mathbf{1}_{\left\{\xi_n^1 \le \sqrt{n}x,\cdots,\xi_n^r \le \sqrt{n}x\right\}} \prod_{w\in skel^{(r)}(n)\backslash\{\varnothing\}} m_{D\left(\overleftarrow{w}\right)}\left(|w|\right) \mathbf{1}_{\left\{T_1^{[r]}\left(i_1,i_2,r-i_1-i_2\right)=j_1\right\}} \right]
$$

$$
= \lim_{n\to\infty} \frac{1}{n^r} \sum_{j_1=1}^{n} \mathbf{E}_{\mathbf{Q}^{[r]}}\left[\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^1 \le \sqrt{n}x-\xi_{j_1},\cdots,\tilde{\xi}_{n-j_1}^r \le \sqrt{n}x-\xi_{j_1}\right\}} \prod_{w\in skel^{(r)}(n-j_1)\backslash\{w_{j_1}\}} m_{D\left(\overleftarrow{w}\right)}\left(|w|\right) \right]
$$

$$
\times \prod_{k=1}^{j_1} m_r(k)\mathbf{Q}^{[r]}\left(T_1^{[r]}\left(i_1,i_2,r-i_1-i_2\right)=j_1\right)
$$

$$
= \lim_{n\to\infty} \frac{1}{n^2} \sum_{j_1=1}^{n} \mathbf{E}_{\mathbf{Q}^{[r]}}\left[\frac{1}{n^{i_1-1}}\mathbf{E}_{\mathbf{Q}^{[i_1]}}\left(\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^{k_1} \le \sqrt{n}x-\xi_{j_1},\cdots,\tilde{\xi}_{n-j_1}^{k_{i_1}} \le \sqrt{n}x-\xi_{j_1}\right\}} \right.\right.
$$

$$
\times \prod_{w\in skel^{(i_1)}(n-j_1)\backslash\{w_{j_1}\}} m_{D\left(\overleftarrow{w}\right)}\left(|w|\right) \mid \sigma\left(w_{j_1},\xi_{j_1}\right)\Bigg)
$$

$$
\times \frac{1}{n^{i_2-1}}\mathbf{E}_{\mathbf{Q}^{[i_2]}}\left(\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^{k_{i_1}+1} \le \sqrt{n}x-\xi_{j_1},\cdots,\tilde{\xi}_{n-j_1}^{k_{i_2}} \le \sqrt{n}x-\xi_{j_1}\right\}} \right.
$$

$$
\times \prod_{w\in skel^{(i_2)}(n-j_1)\backslash\{w_{j_1}\}} m_{D\left(\overleftarrow{w}\right)}\left(|w|\right) \mid \sigma\left(w_{j_1},\xi_{j_1}\right)\Bigg)
$$

$$
\times \frac{1}{n^{r-i_1-i_2-1}}\mathbf{E}_{\mathbf{Q}^{[r-i_1-i_2]}}\left(\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^{k_{i_2}+1} \le \sqrt{n}x-\xi_{j_1},\cdots,\tilde{\xi}_{n-j_1}^{k_r} \le \sqrt{n}x-\xi_{j_1}\right\}} \right.
$$

$$
\times \prod_{w\in skel^{(r-i_1-i_2)}(n-j_1)\backslash\{w_{j_1}\}} m_{D\left(\overleftarrow{w}\right)}\left(|w|\right) \mid \sigma\left(w_{j_1},\xi_{j_1}\right)\Bigg)\Bigg]
$$

$$
\times \frac{1}{n}\prod_{k=1}^{j_1-1} m_1(k)\binom{r}{i_1}\binom{r-i_1}{i_2}\hat{f}_{j_1,n}^{(3)}(1), \tag{4.15}
$$

where $skel^{(r)}(n-j_1)\backslash\{w_{j_1}\}$ is the set of all descendants of w_{j_1} up to time n that have carried at least one mark, $\tilde{\xi}_{n-j_1}^i := \xi_n^i - \xi_{j_1}$ and $\{k_1,\cdots,k_r\}$ is a permutation of $\{1,\cdots,r\}$.

The value on the right side of (4.13) converges to zero as $n\to\infty$ on the event that the r marks follow three different particles carrying i_1, i_2 and $r-i_1-i_2$ marks respectively, since one would obtain at least one extra $\frac{1}{n}$ term which makes (4.15) tend to zero.

Hence, we can only consider the case that r marks follow two different particles at each split time. At the first split time, the r marks follow two different particles carrying i and $r-i$ marks respectively, with probability $\binom{r}{i}\frac{l(l-1)}{l^r}$. So, we have

$$
\mathbf{Q}^{[r]}\left(T_1^{[r]}\left(i,r-i\right)=j_1\right) = \prod_{k=1}^{j_1-1} \frac{m_1(k)}{m_r(k)} \frac{\binom{r}{i}\hat{f}_{j_1,n}^{(2)}(1)}{m_r\left(j_1\right)}.
$$

Note that $\left\{T_1^{[r]}(i, r-i) = j_1\right\}$ is the same event as $\left\{T_1^{[r]}(r-i, i) = j_1\right\}$. It follows from the above argument that

$$
\begin{aligned}
\mu_r(x) &= \lim_{n\to\infty} \frac{1}{n^r} \mathbf{E}_{\mathbf{Q}^{[r]}} \left[\mathbf{1}_{\left\{\xi_n^1 \leq \sqrt{n}x, \cdots, \xi_n^r \leq \sqrt{n}x\right\}} \prod_{w\in skel^{(r)}(n)\setminus\{\varnothing\}} m_{D(\overleftarrow{w})}(|w|) \right] \\
&= \lim_{n\to\infty} \frac{1}{n^r} \sum_{i=1}^{r-1} \frac{1}{2} \sum_{j_1=1}^{n} \mathbf{E}_{\mathbf{Q}^{[r]}} \left[\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^1 \leq \sqrt{n}x-\xi_{j_1}, \cdots, \tilde{\xi}_{n-j_1}^r \leq \sqrt{n}x-\xi_{j_1}\right\}} \right. \\
&\quad \times \prod_{w\in skel^{(r)}(n-j_1)\setminus\{w_{j_1}\}} m_{D(\overleftarrow{w})}(|w|) \left] \prod_{k=1}^{j_1} m_r(k) \mathbf{Q}^{[r]}\left(T_1^{[r]}(i, r-i) = j_1\right) \right. \\
&= \lim_{n\to\infty} \sum_{i=1}^{r-1} \binom{r}{i} \frac{1}{n} \sum_{j_1=1}^{n} \mathbf{E}_{\mathbf{Q}^{[r]}} \left[\frac{1}{n^{i-1}} \mathbf{E}_{\mathbf{Q}^{[i]}} \left(\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^{k_1} \leq \sqrt{n}x-\xi_{j_1}, \cdots, \tilde{\xi}_{n-j_1}^{k_i} \leq \sqrt{n}x-\xi_{j_1}\right\}} \right.\right. \\
&\quad \times \prod_{w\in skel^{(i)}(n-j_1)\setminus\{w_{j_1}\}} m_{D(\overleftarrow{w})}(|w|) \mid \sigma(w_{j_1}, \xi_{j_1}) \Big) \\
&\quad \times \frac{1}{n^{r-i-1}} \mathbf{E}_{\mathbf{Q}^{[r-i]}} \left(\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^{k_{i+1}} \leq \sqrt{n}x-\xi_{j_1}, \cdots, \tilde{\xi}_{n-j_1}^{k_r} \leq \sqrt{n}x-\xi_{j_1}\right\}} \right. \\
&\quad \times \prod_{w\in skel^{(r-i)}(n-j_1)\setminus\{w_{j_1}\}} m_{D(\overleftarrow{w})}(|w|) \mid \sigma(w_{j_1}, \xi_{j_1}) \Big) \right] \frac{1}{n} \prod_{k=1}^{j_1-1} m_1(k) \frac{\hat{f}_{j_1,n}^{(2)}(1)}{2} \\
&= \sum_{i=1}^{r-1} \binom{r}{i} \lim_{n\to\infty} \frac{1}{n} \sum_{j_1=1}^{n} \mathbf{E}_{\mathbf{Q}^{[r]}} \left[\frac{1}{n^{i-1}} \mathbf{E}_{\mathbf{Q}^{[i]}} \left(\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^{k_1} \leq \sqrt{n}x-\xi_{j_1}, \cdots, \tilde{\xi}_{n-j_1}^{k_i} \leq \sqrt{n}x-\xi_{j_1}\right\}} \right.\right. \\
&\quad \times \prod_{w\in skel^{(i)}(n-j_1)\setminus\{w_{j_1}\}} m_{D(\overleftarrow{w})}(|w|) \mid \sigma(w_{j_1}, \xi_{j_1}) \Big) \\
&\quad \times \frac{1}{n^{r-i-1}} \mathbf{E}_{\mathbf{Q}^{[r-i]}} \left(\mathbf{1}_{\left\{\tilde{\xi}_{n-j_1}^{k_{i+1}} \leq \sqrt{n}x-\xi_{j_1}, \cdots, \tilde{\xi}_{n-j_1}^{k_r} \leq \sqrt{n}x-\xi_{j_1}\right\}} \right. \\
&\quad \times \prod_{w\in skel^{(r-i)}(n-j_1)\setminus\{w_{j_1}\}} m_{D(\overleftarrow{w})}(|w|) \mid \sigma(w_{j_1}, \xi_{j_1}) \Big) \right] \frac{\left(1 - f_{(n-j_1)}(0)\right)^2 \sigma^2}{2n\left(1 - f_{(n)}(0)\right)},
\end{aligned}
$$

where the sum "$\sum_{i=1}^{r-1}$" overcounting the case $i = 1, \cdots, [r/2]$, we multiply by $\frac{1}{2}$ in the computation. By similar calculations as above, the term $\left(1 - f_{(n-j_1)}(0)\right)^2$ can be dealt with $\prod_{w\in skel^{(i)}(n-j_1)\setminus\{w_{j_1}\}} m_{D(\overleftarrow{w})}(|w|)$ and $\prod_{w\in skel^{(r-i)}(n-j_1)\setminus\{w_{j_1}\}} m_{D(\overleftarrow{w})}(|w|)$, respectively, since particles in these two terms start from time j_1. Then, we obtain

$$
\mu_r(x) = \sum_{i=1}^{r-1} \binom{r}{i} \int_0^1 \mathbf{E}\left[\mu_i^t(x - B_t)\mu_{r-i}^t(x - B_t)\right] dt.
$$

Therefore, we complete the proof of Corollary 4.6.   □

4.2. *Proof of Theorem 1.1.*

Proof of Theorem 1.1: Note that

$$\limsup_{r \to \infty} \mu_{2r}^{1/2r}(x)/2r < \infty,$$

in fact, $\mu_{2r}(x) \leq \mu_{2r}(\infty)$ and $\mu_{2r}^{1/2r}(\infty)/2r < \infty$ by Remark 1.2. And then by Theorem 3.3.12 of Durrett (2010) and Corollary 4.2, 4.4 and 4.6, $\mathcal{L}\left(\frac{Z^{(n)}(-\infty, \sqrt{n}x]}{n} \mid Z_n > 0\right)$ converges weakly to the unique distribution with these moments. □

Acknowledgements

References

Aïdékon, E. Convergence in law of the minimum of a branching random walk. *Ann. Probab.*, **41** (3A), 1362–1426 (2013). MR3098680.

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.

Biggins, J. D. The first- and last-birth problems for a multitype age-dependent branching process. *Advances in Appl. Probability*, **8** (3), 446–459 (1976). MR420890.

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.

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

Fleischmann, K. and Siegmund-Schultze, R. An invariance principle for reduced family trees of critical spatially homogeneous branching processes. *Serdica*, **4** (2-3), 111–134 (1978). MR541814.

Gao, Z., Liu, Q., and Wang, H. Central limit theorems for a branching random walk with a random environment in time. *Acta Math. Sci. Ser. B (Engl. Ed.)*, **34** (2), 501–512 (2014). MR3174096.

Geiger, J. A new proof of Yaglom's exponential limit law. In *Mathematics and computer science (Versailles, 2000)*, Trends Math., pp. 245–249. Birkhäuser, Basel (2000). MR1798303.

Hammersley, J. M. Postulates for subadditive processes. *Ann. Probab.*, **2**, 652–680 (1974). MR370721.

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.

Hong, W. and Yao, D. Conditional central limit theorem for subcritical branching random walk. *ALEA Lat. Am. J. Probab. Math. Stat.*, **20** (2), 1411–1432 (2023). MR4683380.

Hu, Y. and Shi, Z. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.*, **37** (2), 742–789 (2009). MR2510023.

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

Kesten, H. Branching random walk with a critical branching part. *J. Theoret. Probab.*, **8** (4), 921–962 (1995). MR1353560.

Kesten, H., Ney, P., and Spitzer, F. The Galton-Watson process with mean one and finite variance. *Teor. Verojatnost. i Primenen.*, **11**, 579–611 (1966). MR207052.

Kingman, J. F. C. The first birth problem for an age-dependent branching process. *Ann. Probab.*, **3** (5), 790–801 (1975). MR400438.

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

Lalley, S. P. and Shao, Y. On the maximal displacement of critical branching random walk. *Probab. Theory Related Fields*, **162** (1-2), 71–96 (2015). MR3350041.

Lyons, R., Pemantle, R., and Peres, Y. Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.*, **23** (3), 1125–1138 (1995). MR1349164.

Ren, Y.-X., Song, R., and Sun, Z. A 2-spine decomposition of the critical Galton-Watson tree and a probabilistic proof of Yaglom's theorem. *Electron. Commun. Probab.*, **23**, Paper No. 42, 12 (2018). MR3841403.

Sakhanenko, A. I. Estimates in the invariance principle in terms of truncated power moments. *Sibirsk. Mat. Zh.*, **47** (6), 1355–1371 (2006). MR2302850.

Stam, A. J. On a conjecture by Harris. *Z. Wahrsch. 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.