Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2014, Vol. 50, No. 2, 602–627
DOI: 10.1214/12-AIHP526

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Conditional limit theorems for intermediately subcritical branching processes in random environment[1]

V. I. Afanasyev[a], Ch. Böinghoff[b], G. Kersting[b] and V. A. Vatutin[a]

[a]*Department of Discrete Mathematics, Steklov Mathematical Institute, 8 Gubkin Street, 119 991 Moscow, Russia.*
E-mail: viafan@mail.ru; vatutin@mi.ras.ru
[b]*Fachbereich Mathematik, Universität Frankfurt, Fach 187, D-60054 Frankfurt am Main, Germany.*
E-mail: boeinghoff@math.uni-frankfurt.de; kersting@math.uni-frankfurt.de

Received 13 January 2012; revised 18 September 2012; accepted 24 September 2012

Abstract. For a branching process in random environment it is assumed that the offspring distribution of the individuals varies in a random fashion, independently from one generation to the other. For the subcritical regime a kind of phase transition appears. In this paper we study the intermediately subcritical case, which constitutes the borderline within this phase transition. We study the asymptotic behavior of the survival probability. Next the size of the population and the shape of the random environment conditioned on non-extinction is examined. Finally we show that conditioned on non-extinction periods of small and large population sizes alternate. This kind of 'bottleneck' behavior appears under the annealed approach only in the intermediately subcritical case.

Résumé. Nous considérons un processus de branchement dans un environnement aléatoire dont la distribution des enfants des individus varie aléatoirement de façon indépendante d'une génération à l'autre. Dans le régime sous critique, une transition de phase apparaît. Cet article est consacré à l'étude de la région proche de la transition. Nous étudions le comportement asymptotique de la probabilité de survie ainsi que la taille de la population et la forme de l'environnement aléatoire sous la condition de non-extinction. Nous montrons finalement que conditionnée à la non-extinction, la population alterne des périodes de petite et de grande taille. Ce type de comportement apparaît sous la mesure moyennée uniquement dans ce régime sous critique proche de la transition.

MSC: Primary 60J80; secondary 60K37; 60G50; 60F17

Keywords: Branching process; Random environment; Random walk; Change of measure; Survival probability; Functional limit theorem; Tree

1. Introduction and main results

Branching processes in random environment (BPRE), which have been introduced in [7,28], are a discrete time model for the development of a (discrete) population. You can think of a population of plants having a one-year life-cycle. In each year, the environment varies in a random fashion, independently from one generation to the other. Given the environment, all individuals reproduce independently according to the same mechanism.

More precisely, let Δ be the space of all probability measures on $\mathbb{N}_0$. Equipped with the total variation metric, Δ is a Polish space. Let Q be a random variable taking values in Δ. Then an infinite sequence $\Pi = (Q_1, Q_2, \ldots)$ of i.i.d. copies of Q is called a *random environment* and Q_n is the (random) offspring distribution of an individual at generation $n-1$. Let us denote by Z_n the number of individuals in generation n. Then Z_n is the sum of Z_{n-1} independent random variables, each of which has distribution Q_n. A sequence of $\mathbb{N}_0$-valued random variables $Z_0, Z_1, \ldots$

[1]This paper is a part of the research project 'Branching processes and random walks in random environment' supported by the German Research Foundation (DFG) and the Russian Foundation of Basic Research (RFBF, Grant DFG-RFBR 08-01-91954).

is called a *branching process in the random environment* $\varPi$, if Z_0 is independent of $\varPi$ and, given $\varPi$ the process $Z = (Z_0, Z_1, \ldots)$ is a Markov chain with

$$\mathcal{L}\big(Z_n | Z_{n-1} = z, \varPi = (q_1, q_2, \ldots)\big) = q_n^{*z} \tag{1.1}$$

for every $n \in \mathbb{N} = \{1, 2, \ldots\}$, $z \in \mathbb{N}_0$ and $q_1, q_2, \ldots \in \Delta$, where q^{*z} is the z-fold convolution of the measure q. By $\mathbb{P}$, we will denote the corresponding probability measure on the underlying probability space. For convenience, we assume that the process starts with a single founding ancestor, i.e. $Z_0 = 1$ and we exclude the trivial case $\mathbb{P}(Q(0) = 1) = 1$ throughout this paper. (We shorten $Q(\{y\}), q(\{y\})$ to $Q(y), q(y)$.)

As it turns out, the fine structure of the offspring distributions is only of little importance and the asymptotic behavior of the BPRE is mainly determined by the properties of $\log m(Q)$, where

$$m(q) = \sum_{y=0}^{\infty} y q(y)$$

is the mean of the offspring distribution $q \in \Delta$. More precisely, to describe the asymptotic behavior of the BPRE, we have to examine the so-called *associated random walk* $S = (S_n)_{n \geq 0}$ which is defined by $S_0 = 0$ and the increments $X_n = S_n - S_{n-1}$, $n \geq 1$ with

$$X = \log m(Q) \quad \text{resp. } X_n = \log m(Q_n).$$

Note that the expectation of Z_n, conditioned on the environment, can be expressed in terms of S_n, i.e. using (1.1) and the assumption $Z_0 = 1$ a.s.,

$$\mathbb{E}[Z_n | \varPi] = \prod_{k=1}^{n} m(Q_k) = \exp(S_n) \quad \mathbb{P}\text{-a.s.}$$

and, averaging over the environment,

$$\mathbb{E}[Z_n] = \mathbb{E}\big[m(Q)\big]^n, \tag{1.2}$$

where we assume that the expectation is finite. Using a first moment estimate yields $\mathbb{P}(Z_n > 0 | \varPi) = \min_{0 \leq k \leq n} \mathbb{P}(Z_k > 0 | \varPi) \leq \min_{0 \leq k \leq n} \mathbb{E}[Z_k | \varPi]$ and thus

$$\mathbb{P}(Z_n > 0 | \varPi) \leq \exp\Big(\min_{0 \leq k \leq n} S_k\Big) \quad \mathbb{P}\text{-a.s.} \tag{1.3}$$

This is a classical upper estimate for the survival probability. If S is an oscillating random walk, the branching process is called *critical* (see [4]) resp. *subcritical* if S drifts to $-\infty$. From the strong law of large numbers, it results that the conditional non-extinction probability at n decays in this case at an exponential rate for almost every environment as $n \to \infty$.

As was observed by Afanasyev [1] and later independently by Dekking [18] there are three possibilities for the asymptotic behavior of subcritical branching processes. The asymptotic behavior changes with the sign of $\mathbb{E}[X\mathrm{e}^X]$, which we assume to be finite. A BPRE is *weakly* subcritical if $\mathbb{E}[X\mathrm{e}^X] > 0$, *intermediately* subcritical if $\mathbb{E}[X\mathrm{e}^X] = 0$ and *strongly* subcritical if $\mathbb{E}[X\mathrm{e}^X] < 0$.

The present article is a part of several publications having started with [3–5], in which we try to develop characteristic properties of the different cases. The phase transition already becomes visible when looking at the asymptotic survival probability, i.e. there are positive constants θ_1, θ_2 such that $\mathbb{P}(Z_n > 0) \sim \theta_1 \mathbb{P}(\min(S_1, \ldots, S_n) \geq 0)$ in the critical case and in the weakly subcritical case (see [3,4]), whereas $\mathbb{P}(Z_n > 0) \sim \theta_2 \mathbb{E}[m(Q)]^n$ in the strongly subcritical case (see [5]). Characteristics of the different regimes become more evident in the typical magnitude of Z_k, $k \leq n$, conditioned on $\{Z_n > 0\}$. In the weakly subcritical case Z_k is very large, unless k is close to 0 or n [3]. Contrarily in the strongly subcritical case Z_k stays small for all $0 \leq k \leq n$, see [5]. The phase transition between weak and strong subcriticality can also be observed on the level of large deviations, see [8,13,24,25]. For a more detailed comparative discussion we refer the reader to [12].

Here, our goal is to describe the intermediate case. It is located at the borderline between the weakly and strongly subcritical cases. The passage corresponds to the phase transition in the model, thus a particular rich behavior can be expected for the intermediate case. This is reflected in our results below. In particular we shall observe a kind of bottleneck phenomenon, i.e. that conditioned on survival, there are times when the branching process is small (*bottlenecks*), yet very large inbetween. This phenomenon does not occur elsewhere under the annealed approach. In the critical regime under the quenched approach, the branching process conditioned on survival also exhibits periods of large population sizes alternating with bottlenecks (see [31,32] and [33]).

Assumption A1. *The process Z is intermediately subcritical*, *i.e.*

$$\mathbb{E}\big[X\mathrm{e}^X\big] = 0.$$

The assumption suggests to change from $\mathbb{P}$ to a measure $\mathbf{P}$: For every $n \in \mathbb{N}$ and every bounded, measurable function $\varphi : \Delta^n \times \mathbb{N}_0^{n+1} \to \mathbb{R}$, $\mathbf{P}$ is given by its expectation

$$\mathbf{E}\big[\varphi(Q_1, \dots, Q_n, Z_0, \dots, Z_n)\big] = \gamma^{-n}\mathbb{E}\big[\varphi(Q_1, \dots, Q_n, Z_0, \dots, Z_n)\mathrm{e}^{S_n - S_0}\big],$$

with

$$\gamma = \mathbb{E}\big[\mathrm{e}^X\big].$$

(We include S_0 in the above expression, because later on we shall also consider cases where $S_0 \neq 0$.) From (1.2) we obtain

$$\mathbb{E}[Z_n] = \gamma^n. \tag{1.4}$$

The assumption $\mathbb{E}[X\mathrm{e}^X] = 0$ translates into

$$\mathbf{E}[X] = 0.$$

Thus S becomes a recurrent random walk under $\mathbf{P}$.

As to the regularity of the distribution of X we make the following assumptions.

Assumption A2. *The distribution of X has finite variance with respect to* $\mathbf{P}$ *or* (*more generally*) *belongs to the domain of attraction of some stable law with index $\alpha \in (1, 2]$*. *It is non-lattice*.

Since $\mathbf{E}[X] = 0$ this means that there is an increasing sequence of positive numbers

$$a_n = n^{1/\alpha}\ell_n$$

with a slowly varying sequence $\ell_1, \ell_2, \dots$ such that for $n \to \infty$

$$\mathbf{P}\left(\frac{1}{a_n}S_n \in \mathrm{d}x\right) \to s(x)\,\mathrm{d}x$$

weakly, where $s(x)$ denotes the density of the limiting stable law. In the case of finite variance $\sigma^2 = \mathbf{E}[X^2] < \infty$, the slowly varying sequence is constant, i.e. $\ell_n = \sigma$.

Note that due to the change of measure X^- always has finite variance and an infinite variance may only arise from X^+. In case of $\alpha < 2$ this is the so-called spectrally positive case ([11], Section 8.2.9).

Our last assumption on the environment concerns the standardized truncated second moment of Q,

$$\zeta(a) = \frac{1}{m(Q)^2}\sum_{y=a}^{\infty} y^2 Q(y), \quad a \in \mathbb{N}.$$

Assumption A3. *For some* $\varepsilon > 0$ *and some* $a \in \mathbb{N}$

$$\mathbf{E}\big[\big(\log^+ \zeta(a)\big)^{\alpha+\varepsilon}\big] < \infty,$$

where $\log^+ x = \log(x \vee 1)$.

See [4] for examples where this assumption is fulfilled for any $\alpha \in (1, 2]$. For binary branching processes in random environment (where individuals have either two children or none) $\zeta(3) = 0$, and for cases where Q is a.s. a Poisson distribution $\zeta(2) \le 2$ or a.s. a geometric distribution $\zeta(2) \le 4$.

The following theorem has been obtained under quite stronger assumptions in [1,21,30]. Let

$$\tau_n = \min\{k \le n | S_k \le S_0, S_1, \dots, S_n\}$$

be the moment, when S_k takes its minimum within S_0 to S_n for the first time.

Theorem 1.1. *Under Assumptions A1 to A3, there is a constant* $0 < \theta < \infty$ *such that as* $n \to \infty$

$$\mathbb{P}(Z_n > 0) \sim \theta \gamma^n \mathbf{P}(\tau_n = n).$$

In this form the result holds in the strongly subcritical case too [22], however it differs from the corresponding result in the weakly subcritical case [3]. Along the way of proving the subsequent results we also obtain a proof of the above theorem (see the end of Lemma 3.3). Since $\mathbf{P}(\tau_n = n) \sim 1/b_n$ with

$$b_n = n^{1-\alpha^{-1}} \ell'_n$$

for some slowly varying sequence (ℓ'_n) (see Lemma 2.2 below), it follows

$$\mathbb{P}(Z_n > 0) \sim \theta \frac{\gamma^n}{b_n}.$$

If $\sigma^2 < \infty$, then l'_n is constant (see [11], Theorems 8.9.12/8.9.13). The next theorem deals with the branching process conditioned on survival at time n.

Theorem 1.2. *Under Assumptions A1 to A3 the distribution of* Z_n *conditioned on the event* $Z_n > 0$ *converges weakly to a probability distribution on* $\mathbb{N}$. *Also for every* $\beta < 1$ *the sequence* $\mathbb{E}[Z_n^\beta | Z_n > 0]$ *is bounded.*

For $\beta = 1$ this statement is no longer true, since $\mathbb{E}[Z_n] = \gamma^n$ from (1.4) and consequently $\mathbb{E}[Z_n | Z_n > 0] \to \infty$ for $n \to \infty$.

The next theorem captures the typical appearance of the random environment, when conditioned on survival. Let S^n be the stochastic process with paths in the Skorohod space $D[0, 1]$ of càdlàg functions on $[0, 1]$ given by

$$S_t^n = S_{nt}, \quad 0 \le t \le 1.$$

We agree on the convention $S_{nt} = S_{\lfloor nt \rfloor}$, which we use correspondingly for Z_{nt}, τ_{nt}.

Also let L^* denote a process, which can be understood as a Lévy-process on $[0, 1]$ conditioned to attain its minimum at time $t = 1$. Formally we will define it in Section 2 in such a way that $(L_1^* - L_{(1-t)-}^*)_{0 \le t \le 1}$ is a Lévy-meander, as introduced in [16]. If $\mathbf{E}[X^2] < \infty$, this is the dual process of a Brownian meander.

Theorem 1.3. *Assume Assumptions A1 to A3. Then, as* $n \to \infty$, *the distribution of* $n - \tau_n$ *conditioned on the event* $Z_n > 0$ *converges to a probability distribution* p *on* $\mathbb{N}_0$ *and*

$$\left(\frac{1}{a_n} S^n \middle| Z_n > 0\right) \xrightarrow{d} L^*$$

in the Skorohod space $D[0,1]$. *Also both quantities are asymptotically independent, namely for every bounded continuous* $\varphi : D[0,1] \to \mathbb{R}$ *and every* $B \subset \mathbb{N}_0$

$$\mathbb{E}\left[\varphi\left(\frac{1}{a_n} S^n\right); n-\tau_n \in B \middle| Z_n > 0\right] \to \mathbf{E}\left[\varphi\left(L^*\right)\right] p(B).$$

In the strongly subcritical case, $n-\tau_n$ also converges to a probability distribution on $\mathbb{N}_0$ (consequence of [5], Theorem 1.3). However, this statement is not true for critical (see [4], Theorem 1.4) or weakly subcritical BPREs (see [3], Theorem 1.1 and its proof). The limit L^* only appears in the intermediate case.

The last theorem characterizes the typical behavior of Z, conditioned on survival. For a partial result see Theorem 1 in [2]. Recall that τ_{nt} is the moment when $S_0, \dots, S_{nt}$ takes its minimum.

Theorem 1.4. *Let* $0 < t_1 < \cdots < t_r < 1$. *For* $i = 1, \dots, r$ *let*

$$\mu(i) = \min\left\{ j \le i : \inf_{t \le t_j} L^*_t = \inf_{t \le t_i} L^*_t \right\}.$$

Then under Assumptions A1 to A3 there are i.i.d. random variables $V_1, \dots, V_r$ *with values in* $\mathbb{N}$ *and independent of* L^* *such that*

$$\left((Z_{\tau_{nt_1}}, \dots, Z_{\tau_{nt_r}}) | Z_n > 0\right) \xrightarrow{d} (V_{\mu(1)}, \dots, V_{\mu(r)})$$

as $n \to \infty$. *Also there are i.i.d. strictly positive random variables* $W_1, \dots, W_r$ *independent of* L^* *such that*

$$\left(\left(\frac{Z_{nt_1}}{\mathrm{e}^{S_{nt_1} - S_{\tau_{nt_1}}}}, \dots, \frac{Z_{nt_r}}{\mathrm{e}^{S_{nt_r} - S_{\tau_{nt_r}}}}\right) \middle| Z_n > 0\right) \xrightarrow{d} (W_{\mu(1)}, \dots, W_{\mu(r)})$$

as $n \to \infty$.

Note that this theorem cannot be generalized to a functional limit result in Skorohod space. The limiting process would consist of paths which are constant within excursions between descending ladder points of the process $L^\star$ and change independently from one excursion to the next. However, such a process is not càdlàg at these ladder points.

The content of Theorems 1.2 to 1.4 may be understood as follows: As in Theorem 1.2 one expects that the population size is small at time n, conditioned on $\{Z_n > 0\}$. This requires that S_n is not much larger than $\min(S_0, \dots, S_n)$, because otherwise the population would grow again at the end. Indeed this is confirmed by Theorem 1.3. Similarly one expects that the population size is small in decreasing ladder points before n, which is stated in the first part of Theorem 1.4. On the other hand within upward excursions between such points of minimum, the population development is unaffected of the condition $\{Z_n > 0\}$.

More precisely for $r = 1$ and $t_1 = t$ Theorem 1.4 says the following: At time τ_{nt} the population consists only of few individuals, whereas at time nt it is large, namely of order $\mathrm{e}^{S_{nt} - S_{\tau_{nt}}}$-many individuals, which for every $\varepsilon > 0$ is bigger than $\mathrm{e}^{\delta a_n}$ with probability $1 - \varepsilon$, if $\delta > 0$ is small enough. Thus the minimum of the random walk at time τ_{nt} acts as a bottleneck for the population, whereas afterwards the increasing random walk generates an environment, which is favorable for growth.

Moreover: In case of $r = 2$ either $\tau_{nt_1} < \tau_{nt_2}$ or $\tau_{nt_1} = \tau_{nt_2}$, which for the limiting process L^* means $\mu(2) = 2$ or $\mu(2) = 1$. The theorem says that in the first situation of two bottlenecks the population sizes $Z_{\tau_{nt_1}}$ and $Z_{\tau_{nt_2}}$ are asymptotically independent, as well as the sizes Z_{nt_1} and Z_{nt_2}. In the second situation of one common bottleneck certainly $Z_{\tau_{nt_1}}$ and $Z_{\tau_{nt_2}}$ are equal. Interestingly this is asymptotically true as well for $Z_{nt_1}/\mathrm{e}^{S_{nt_1} - S_{\tau_{nt_1}}}$ and $Z_{nt_2}/\mathrm{e}^{S_{nt_2} - S_{\tau_{nt_2}}}$. Here a law of large numbers is at work, in a similar fashion as for supercritical Galton–Watson processes.

Theorems 1.3 and 1.4 may lead to the conjecture that $(\frac{1}{a_n} \log Z_{nt})_{0 \le t \le 1}$ converges to a Lévy-process, conditioned to take its minimum at the end and reflected at zero. For the finite dimensional distributions this follows from the theorems together with path properties of Lévy-processes. For linear fractional offspring distributions this was already obtained in [2], Theorem 2. The proof of tightness is somewhat involved and given in a separate paper (see [14]).

The proofs rest largely on the fact that the event $Z_n > 0$ asymptotically entails that τ_n takes a value close to n, as stated in Theorem 1.3. Thus it is our strategy to replace the conditioning event $Z_n > 0$ by events $\tau_n = n - m$, which are easier to handle. Here we can build on some random walk theory. For the proof of the last theorem we also make use of constructions of *trees with stem* going back to Lyons, Perez and Pemantle [26] and Geiger [20] for Galton–Watson processes. They establish a connection between branching processes conditioned to survive and branching processes with immigration.

The paper is organized as follows: In Section 2 we compile and prove several results on random walks. In Section 3 the proofs of the first three theorems are given. Section 4 deals with trees with stem and Section 5 contains the proof of our last theorem.

2. Results on random walks

In this section we assemble several auxiliary results on the random walk S. We allow for an arbitrary initial value $S_0 = x$. Then we write $\mathbf{P}_x(\cdot)$ and $\mathbf{E}_x[\cdot]$ for the corresponding probabilities and expectations. Thus $\mathbf{P} = \mathbf{P}_0$.

2.1. *Some asymptotic results*

Let us introduce for $n \ge 1$

$$L_n = \min(S_1, \dots, S_n), \qquad M_n = \max(S_1, \dots, S_n)$$

and as above for $n \ge 0$

$$\tau_n = \min\bigl\{k \le n\colon\ S_k = \min(0, L_n)\bigr\}.$$

There is a connection between M_n and τ_n, set up by the dual random walk

$$\hat{S}_k = S_n - S_{n-k}, \quad 0 \le k \le n.$$

Namely $\{\tau_n = n\} = \{\hat{M}_n < 0\}$ with $\hat{M}_n = \max(\hat{S}_1, \dots, \hat{S}_n)$ and consequently

$$\mathbf{P}(\tau_n = n) = \mathbf{P}(M_n < 0).$$

In particular $\mathbf{P}(\tau_n = n)$ is decreasing.

Next define the renewal functions $u : \mathbb{R} \to \mathbb{R}$ and $v : \mathbb{R} \to \mathbb{R}$ by

$$u(x) = 1 + \sum_{k=1}^{\infty} \mathbf{P}(-S_k \le x, M_k < 0), \quad x \ge 0,$$

$$v(x) = 1 + \sum_{k=1}^{\infty} \mathbf{P}(-S_k > x, L_k > 0), \quad x < 0,$$

$$v(0) = \mathbf{E}\bigl[v(X); X < 0\bigr],$$

and 0 elsewhere. In particular $u(0) = 1$. It is well-known that $0 < v(0) \le 1$, for details we refer to [15], Appendix B, and [32]. (Our function $v(x)$ coincides with the function $v(x)$ in [3] up to a constant.) Also $u(x)$ and $v(-x)$ are of order $\mathrm{O}(x)$ for $x \to \infty$.

Lemma 2.1. *Under Assumption A2, for every $r > 0$ there exists a $\kappa > 0$ such that*

$$\mathbf{E}\bigl[\mathrm{e}^{-rS_n}; L_n \ge 0\bigr] \sim \kappa n^{-1} a_n^{-1}$$

as $n \to \infty$.

For the proof we refer to Proposition 2.1 in [3].

Lemma 2.2. *Under Assumption A2 there are real numbers*

$$b_n = n^{1-\alpha^{-1}} \ell'_n, \quad n \geq 1$$

with a sequence (ℓ'_n) *slowly varying at infinity such that for every* $x \geq 0$

$$\mathbf{P}(M_n < x) \sim v(-x) b_n^{-1}$$

as $n \to \infty$. *Also there is a constant* $c > 0$ *such that for all* $x \geq 0$

$$\mathbf{P}(M_n < x) = \mathbf{P}_{-x}(M_n < 0) \leq c v(-x) b_n^{-1}.$$

Proof. The corresponding statements for $\mathbf{P}(M_n \leq x)$ are well-known. Indeed the first one is contained in Theorem 8.9.12 in [11], where ρ now is equal to $1 - \alpha^{-1}$, since we are in the spectrally positive case (note that the proof therein works for all $x \geq 0$ and not only, as stated, for the continuity points of v).

For $x > 0$ this proof completely translates to $\mathbf{P}(M_n < x)$. Therefrom the case $x = 0$ can be treated as follows:

$$\begin{aligned}\mathbf{P}(M_n < 0) &= \mathbf{E}\big[\mathbf{P}_{X_1}(M_{n-1} < 0); X_1 < 0\big] \\ &= b_{n-1}\mathbf{E}\left[\frac{\mathbf{P}_{X_1}(M_{n-1} < 0)}{b_{n-1}}; X_1 < 0\right].\end{aligned}$$

As for every $x > 0$, $\mathbb{P}_{X_1}(M_n < 0) \leq \mathbb{P}_{X_1}(M_n < x)$, applying the bound for $\mathbb{P}_{X_1}(M_n < x)$ yields that $b_{n-1}^{-1}\mathbb{P}_{X_1}(M_n < 0)$ is bounded by an integrable function. Thus from dominated convergence and from $b_n \sim b_{n-1}$ we get

$$\mathbf{P}(M_n < 0) \sim b_n \mathbf{E}\big[v(X_1); X_1 < 0\big].$$

Now from Eq. (2.1) below the conditional expectation in the right-hand side is equal to $v(0)$, as defined above, which gives the claim.

The second statement is obtained just as in Lemma 2.1 in [4]. □

2.2. *The probability measures* $\mathbf{P}^+$ *and* $\mathbf{P}^-$

The fundamental properties of u, v are the identities

$$\begin{aligned}\mathbf{E}\big[u(x+X); X+x \geq 0\big] &= u(x), \quad x \geq 0, \\ \mathbf{E}\big[v(x+X); X+x < 0\big] &= v(x), \quad x \leq 0,\end{aligned} \tag{2.1}$$

which hold for every oscillating random walk (see e.g. [32]). It follows that u and v give rise to further probability measures $\mathbf{P}^+$ and $\mathbf{P}^-$. The construction procedure is standard and explained for $\mathbf{P}^+$ in detail in [4,10]. We shortly summarize it below.

Consider the filtration $\mathcal{F} = (\mathcal{F}_n)_{n \geq 0}$, where $\mathcal{F}_n = \sigma(Q_1, \ldots, Q_n, Z_0, \ldots, Z_n)$. Thus S is adapted to $\mathcal{F}$ and X_{n+1} (as well as Q_{n+1}) is independent ot $\mathcal{F}_n$ for all $n \geq 0$. Then for every bounded, $\mathcal{F}_n$-measurable random variable R_n

$$\begin{aligned}\mathbf{E}_x^+[R_n] &= \frac{1}{u(x)}\mathbf{E}_x\big[R_n u(S_n); L_n \geq 0\big], \quad x \geq 0, \\ \mathbf{E}_x^-[R_n] &= \frac{1}{v(x)}\mathbf{E}_x\big[R_n v(S_n); M_n < 0\big], \quad x \leq 0.\end{aligned}$$

These are Doob's transforms from the theory of Markov chains. Shortly speaking $\mathbf{P}_x^+$ and $\mathbf{P}_x^-$ correspond to conditioning the random walk S not to enter $(-\infty, 0)$ and $[0, \infty)$ respectively.

The following lemma is taken from [4,10].

Lemma 2.3. *Assume Assumption A2 and let $U_1, U_2, \dots$ be a sequence of uniformly bounded random variables, adapted to the filtration $\mathcal{F}$. If $U_n \to U_\infty$ $\mathbf{P}^+$-a.s. for some limiting random variable U_∞, then as $n \to \infty$*

$$\mathbf{E}[U_n | L_n \geq 0] \to \mathbf{E}^+[U_\infty].$$

Similarly, if $U_n \to U_\infty$ $\mathbf{P}^-$-a.s., then as $n \to \infty$

$$\mathbf{E}[U_n | M_n < 0] \to \mathbf{E}^-[U_\infty].$$

The first part coincides with Lemma 2.5 from [4]. The proof of the second part follows exactly the same lines using Lemma 2.2.

2.3. *Two functional limit results*

Because of Assumption A2 there exists a Lévy-process $L = (L_t)_{t\geq 0}$ such that the processes $S^n = (\frac{1}{a_n} S_{nt})_{0\leq t\leq 1}$ converge in distribution to L in the Skorohod space $D[0,1]$. Let $L^- = (L_t^-)_{0\leq t\leq 1}$ denote the corresponding non-positive Lévy-meander. This is the process $(L_t)_{0\leq t\leq 1}$, conditioned on the event $\sup_{t\leq 1} L_t \leq 0$ (see [9] and [16]).

Lemma 2.4. *Under Assumptions A1 and A2 for every $x \geq 0$ and $n \to \infty$*

$$\left(\frac{1}{a_n} S^n \middle| M_n < -x\right) \xrightarrow{d} L^-$$

in the Skorohod space $D[0,1]$.

The proof follows exactly the same arguments as the proof of Lemma 2.3 in [4], i.e. using the suitably adapted decomposition (2.10) therein and [19].

From L^- we obtain the process L^* as follows. Let $\Lambda : D[0,1] \to D[0,1]$ be the mapping $g \mapsto \hat{g}$ given by

$$\hat{g}(t) = g(1) - g(s-), \quad 0 \leq t \leq 1, \text{ with } s = 1 - t$$

and $g(0-) = 0$. Λ is a continuous mapping and $\Lambda^{-1} = \Lambda$. Note that Λ maps the subset $D^- = \{g \in D[0,1]: \sup_{t\geq\varepsilon} g(t) < 0 \text{ for all } \varepsilon > 0\}$ onto the set $D^* = \{g \in D[0,1]: \inf_{s\leq 1-\varepsilon} g(s) > g(1) \text{ for all } \varepsilon > 0\}$.

Now let

$$L^* = \Lambda(L^-).$$

Since $L^- \in D^-$ a.s. it follows that $L^* \in D^*$ a.s. This means that L^* takes its infimum at the end a.s. L^* may be viewed as the process $(L_t)_{0\leq t\leq 1}$, conditioned to attain its infimum at $t = 1$. This becomes clear from the following result.

Lemma 2.5. *Under Assumptions A1 and A2, for $n \to \infty$*

$$\left(\frac{1}{a_n} S^n \middle| \tau_n = n\right) \xrightarrow{d} L^*$$

in $D[0,1]$.

Proof. In Lemma 2.4, we may replace S^n by the process T^n given by $T_t^n = S_{t+1/n}^n$ for $t \leq 1 - \frac{1}{n}$ and $T_t^n = S_1^n$ for $1 - \frac{1}{n} < t \leq 1$. It follows that

$$\left(\frac{1}{a_n} \Lambda(T^n) \middle| M_n < 0\right) \xrightarrow{d} L^*.$$

Now $\Lambda(T^n)$ is obtained from S^n, if we just interchange the jumps in S^n from $X_1, \dots, X_n$ to $X_n \dots, X_1$. This corresponds to proceeding to the dual random walk, and it follows

$$\left(\frac{1}{a_n} S^n \middle| \tau_n = n\right) \xrightarrow{d} L^*.$$

This is the claim. □

We end this section by some remarks on the distribution of L_1^*. First L_1 has a stable distribution, thus it has a density with respect to Lebesgue measure and is unbounded from below. Since we are in the spectrally positive case, L has no negative jumps a.s. Therefore we may use fluctuation theory for the process $L^{\downarrow}$, which is the Lévy-process, conditioned to take values in $(-\infty, 0]$, see [9], Section VII.3. From Corollary 16 therein it follows that $L_1^{\downarrow}$ has a density and is unbounded from below, too. As stated in [16] (see also [15]) the distributions of $L_1^{\downarrow}$ and L_1^* are mutually absolutely continuous, therefore also the distribution ν of L_1^* has a density and is not concentrated on some compact interval.

2.4. *Further limit results*

Let $Q_j = Q_1$ for $j \le 0$.

Lemma 2.6. *Under Assumptions A1 and A2 for $m \ge 0, k \ge 1$, for $n \to \infty$ the distribution of*

$$\left((Q_{\tau_n+1}, \dots, Q_{\tau_n+k}), (Q_{\tau_n}, \dots, Q_{\tau_n-k+1}), \frac{(S_{\tau_n}, S_{n-m})}{a_n}\right)$$

converges weakly to a probability measure $\mu_k^+ \otimes \mu_k^- \otimes \mu$, where μ_k^+, μ_k^- are the distributions of $(Q_1, \dots, Q_k)$ under the probability measures $\mathbf{P}^+, \mathbf{P}^-$ and μ is a non-degenerate probability measure on $\mathbb{R}^2$, i.e. the measure is not a Dirac measure.

Proof. Let for $r \ge 0$

$$Q^+(r) = (Q_{r+1}, \dots, Q_{r+k}), \qquad Q^-(r) = (Q_r, \dots, Q_{r-k+1}).$$

Let $\phi_1, \phi_2 : \Delta^k \to \mathbb{R}$ be bounded functions and $\phi_3, \phi_4 : \mathbb{R} \to \mathbb{R}$ be bounded continuous functions. A decomposition with respect to τ_n yields

$$\begin{aligned}
&\mathbf{E}\left[\phi_1\big(Q^-(\tau_n)\big)\phi_2\big(Q^+(\tau_n)\big)\phi_3\left(\frac{S_{\tau_n}}{a_n}\right)\phi_4\left(\frac{S_{n-m} - S_{\tau_n}}{a_n}\right)\right] \\
&\quad = \sum_{r=0}^{n} \mathbf{E}\left[\phi_1\big(Q^-(r)\big)\phi_2\big(Q^+(r)\big)\phi_3\left(\frac{S_r}{a_n}\right)\phi_4\left(\frac{S_{n-m} - S_r}{a_n}\right); \tau_n = r\right].
\end{aligned} \tag{2.2}$$

Letting $L_{r,n} = \min(S_{r+1}, \dots, S_n) - S_r$ and using duality we get for $r > k$

$$\begin{aligned}
&\mathbf{E}\left[\phi_1\big(Q^-(r)\big)\phi_2\big(Q^+(r)\big)\phi_3\left(\frac{S_r}{a_n}\right)\phi_4\left(\frac{S_{n-m} - S_r}{a_n}\right); \tau_n = r\right] \\
&\quad = \mathbf{E}\left[\phi_1\big(Q^-(r)\big)\phi_3\left(\frac{S_r}{a_n}\right)\phi_2\big(Q^+(r)\big)\phi_4\left(\frac{S_{n-m} - S_r}{a_n}\right); \tau_r = r, L_{r,n} \ge 0\right] \\
&\quad = \mathbf{E}\left[\phi_1\big(Q^-(r)\big)\phi_3\left(\frac{S_r}{a_n}\right); \tau_r = r\right]\mathbf{E}\left[\phi_2\big(Q^+(0)\big)\phi_4\left(\frac{S_{n-m-r}}{a_n}\right); L_{n-r} \ge 0\right] \\
&\quad = \mathbf{E}\left[\phi_1\big(Q^+(0)\big)\phi_3\left(\frac{S_r}{a_n}\right); M_r < 0\right]\mathbf{E}\left[\phi_2\big(Q^+(0)\big)\phi_4\left(\frac{S_{n-m-r}}{a_n}\right); L_{n-r} \ge 0\right].
\end{aligned}$$

Moreover for $r > k$

$$\frac{\mathbf{E}[\phi_1(Q^+(0))\phi_3(S_r/a_n); M_r < 0]}{\mathbf{P}(M_r < 0)}$$
$$= \mathbf{E}\left[\phi_1\big(Q^+(0)\big)\mathbf{E}_{S_k}\left[\phi_3\left(\frac{S_{r-k}}{a_n}\right)\Big| M_{r-k} < 0\right]\frac{\mathbf{P}_{S_k}(M_{r-k} < 0)}{\mathbf{P}(M_r < 0)}; M_k < 0\right].$$

Therefore by Lemmas 2.2, 2.4 and dominated convergence, if $r_n \sim tn$ for some $0 < t < 1$, then $a_{r_n}/a_n \sim t^{1/\alpha}$ and

$$\frac{\mathbf{E}[\phi_1(Q^+(0))\phi_3(S_{r_n}/a_n); M_{r_n} < 0]}{\mathbf{P}(M_{r_n} < 0)}$$
$$\to \mathbf{E}\big[\phi_1\big(Q^+(0)\big)v(S_k); M_k < 0\big]\mathbf{E}\big[\phi_3\big(t^{1/\alpha}L_1^-\big)\big]$$
$$= \mathbf{E}^-\big[\phi_1\big(Q^+(0)\big)\big]\mathbf{E}\big[\phi_3\big(t^{1/\alpha}L_1^-\big)\big].$$

In much the same way, letting L^+ be the positive Lévy meander and using Lemma 2.3 from [4], it follows that

$$\frac{\mathbf{E}[\phi_2(Q^+(0))\phi_4(S_{n-m-r_n}/a_n); L_{n-r_n} \geq 0]}{\mathbf{P}(L_{n-r_n} \geq 0)}$$
$$\to \mathbf{E}^+\big[\phi_2\big(Q^+(0)\big)\big]\mathbf{E}\big[\phi_4\big((1-t)^{1/\alpha}L_1^+\big)\big].$$

Since $\mathbf{P}(M_{r_n} < 0)\mathbf{P}(L_{n-r_n} \geq 0) = \mathbf{P}(\tau_n = r_n)$, we obtain for $r_n \sim tn$ and $0 < t < 1$

$$\mathbf{E}\left[\phi_1\big(Q^-(r_n)\big)\phi_2\big(Q^+(r_n)\big)\phi_3\left(\frac{S_{r_n}}{a_n}\right)\phi_4\left(\frac{S_{n-m} - S_{r_n}}{a_n}\right)\Big|\tau_n = r_n\right]$$
$$\to \mathbf{E}^-\big[\phi_1(Q_1, \dots, Q_k)\big]\mathbf{E}^+\big[\phi_2(Q_1, \dots, Q_k)\big]\mathbf{E}\big[\phi_3\big(t^{1/\alpha}L_1^-\big)\big]\mathbf{E}\big[\phi_4\big((1-t)^{1/\alpha}L_1^+\big)\big].$$

Now in view of Assumption A2, the generalized arcsine law (see [11]) is valid for τ_n, i.e. τ_n/n is convergent in distribution to a Beta-distribution with a density, which we denote by $g(t)\,\mathrm{d}t$. Therefore it follows from (2.2) that

$$\mathbf{E}\left[\phi_1\big(Q^-(\tau_n)\big)\phi_2\big(Q^+(\tau_n)\big)\phi_3\left(\frac{S_{\tau_n}}{a_n}\right)\phi_4\left(\frac{S_{n-m} - S_{\tau_n}}{a_n}\right)\right]$$
$$\to \mathbf{E}^-\big[\phi_1(Q_1, \dots, Q_k)\big]\mathbf{E}^+\big[\phi_2(Q_1, \dots, Q_k)\big]$$
$$\times \int_0^1 \mathbf{E}\big[\phi_3\big(t^{1/\alpha}L_1^-\big)\big]\mathbf{E}\big[\phi_4\big((1-t)^{1/\alpha}L_1^+\big)\big]g(t)\,\mathrm{d}t.$$

This gives the claim. □

Next let $0 = t_0 < t_1 < \cdots < t_r < t_{r+1} = 1$ and for $1 \leq i \leq r$

$$\sigma_{i,n} = \min\{k\colon nt_{i-1} \leq k \leq nt_i, S_k \leq S_j \text{ for all } nt_{i-1} \leq j \leq nt_i\} \tag{2.3}$$

be the first moment, when S_k takes its minimum between nt_{i-1} and nt_i.

Lemma 2.7. *Let $m \geq 0$ and $k, r \geq 1$. Then under Assumptions A1, A2, given the event $\tau_{n-m} = n - m$, the random elements in Δ^{2k}*

$$Q^{(i)} = \big((Q_{\sigma_{i,n}+1}, \dots, Q_{\sigma_{i,n}+k}), (Q_{\sigma_{i,n}}, \dots, Q_{\sigma_{i,n}-k+1})\big), \quad i = 1, \dots, r,$$

are for $n \to \infty$ asymptotically independent with asymptotic distribution $\mu_k^+ \otimes \mu_k^-$. Also, given $\tau_{n-m} = n - m$, they are asymptotically independent from the random vector

$$\frac{1}{a_n}(S_{\sigma_{1,n}}, S_{nt_1}, \dots, S_{\sigma_{r,n}}, S_{nt_r}).$$

Proof. Recall from above that, given $\tau_n = n$, the distribution of $\frac{1}{a_n} S_n$ is weakly convergent to a probability measure ν on $\mathbb{R}^-$, the distribution of L_1^*, which possesses a density and is not concentrated on a compact interval.

Let

$$\sigma_{r+1,n} = \min\{k\colon\ nt_r \le k \le n-m,\ S_k \le S_j \text{ for all } nt_r \le j \le n-m\}$$

and for $i \le r$

$$U_i = \frac{1}{a_n}(S_{\sigma_{i,n}} - S_{nt_{i-1}}), \qquad V_i = \frac{1}{a_n}(S_{nt_i} - S_{nt_{i-1}}), \qquad V_{r+1} = \frac{1}{a_n}(S_{n-m} - S_{nt_r})$$

and $W_i = (U_i, V_i)$. Since (a_n) is regularly varying, from the last lemma and from our assumptions on independence it follows that the random variables $Q^{(1)}, \dots, Q^{(r)}, W_1, \dots, W_r, V_{r+1}$ are asymptotically independent. The event $\sigma_{r+1,n} = n-m$ is independent of $Q^{(1)}, \dots, Q^{(r)}, W_1, \dots, W_r$ and thus only affects V_{r+1}. Thus from Lemma 2.6

$$\big(Q^{(1)}, \dots, Q^{(r)}, W_1, \dots, W_r, V_{r+1} | \sigma_{r+1,n} = n-m\big) \xrightarrow{d} \big(\mu_k^+ \otimes \mu_k^-\big)^{\otimes r} \mu_1 \otimes \cdots \otimes \mu_r \otimes \nu,$$

where the probability measures μ_i also depend on $t_i - t_{i-1}$. If a Borel set $A \subset \mathbb{R}^{2r+1}$ satisfies $\mu_1 \otimes \cdots \otimes \mu_r \otimes \nu(A) > 0$ and $\mu_1 \otimes \cdots \otimes \mu_r \otimes \nu(\partial A) = 0$, it follows

$$\big(Q^{(1)}, \dots, Q^{(r)} | (W_1, \dots, W_r, V_{r+1}) \in A, \sigma_{r+1,n} = n-m\big) \xrightarrow{d} \big(\mu_k^+ \otimes \mu_k^-\big)^{\otimes r}.$$

We apply this result to A of the form $A = B \cap C$, where the Borel set B satisfies the same conditions as A, and

$$C = \left\{(u_1, v_1, \dots, u_r, v_r, v_{r+1})\colon\ u_j > \sum_{i=j}^{r+1} v_i \text{ for } j \le r\right\}.$$

Since ν is not concentrated on a compact set, $\mu_1 \otimes \cdots \otimes \mu_r \otimes \nu(C) > 0$, and because ν has a density, $\mu_1 \otimes \cdots \otimes \mu_r \otimes \nu(\partial C) = 0$. As

$$\begin{aligned}\{\tau_{n-m} = n-m\} &= \{S_{\sigma_{j,n}} > S_{n-m} \text{ for } j \le r, \sigma_{r+1,n} = n-m\} \\ &= \big\{(W_1, \dots, W_r, V_{r+1}) \in C, \sigma_{r+1,n} = n-m\big\}\end{aligned}$$

we obtain

$$\big(Q^{(1)}, \dots, Q^{(r)} | (W_1, \dots, W_r, V_{r+1}) \in B, \tau_{n-m} = n-m\big) \xrightarrow{d} \big(\mu_k^+ \otimes \mu_k^-\big)^{\otimes r}.$$

The choice $B = \mathbb{R}^{2r+1}$ gives the asymptotic distribution of $(Q^{(1)}, \dots, Q^{(r)})$. Since $(S_{\sigma_{1,n}}, S_{nt_1}, \dots, S_{\sigma_{r,n}}, S_{nt_r})$ is obtained from $(W_1, \dots, W_r, V_{r+1})$ by linear combinations, also the asymptotic independence follows. □

3. Proofs of Theorems 1.1 to 1.3

Define

$$\eta_k = \sum_{y=0}^{\infty} y(y-1) Q_k(y) \Big/ \left(\sum_{y=0}^{\infty} y Q_k(y)\right)^2, \quad k \ge 1.$$

Lemma 3.1. *Assume Assumptions A1 to A3. Then for all $x \ge 0$*

$$\sum_{k=0}^{\infty} \eta_{k+1} e^{-S_k} < \infty \quad \mathbf{P}_x^+\text{-a.s.}$$

and for all $x \le 0$

$$\sum_{k=1}^{\infty} \eta_k e^{S_k} < \infty \quad \mathbf{P}_x^- \text{-a.s.}$$

The proof of the first statement can be found in [4] (see Lemma 2.7 therein under conditions B1 and B2), the second one is proven in just the same way.

Lemma 3.2. *Under Assumptions A1 to A3, there is a non-vanishing finite measure* p' *on* $\mathbb{N}_0$ *with* $p'(0) = 0$ *such that the following holds: Let* Y_n *be uniformly bounded random variables of the form* $Y_n = \varphi_n(Q_1, \dots, Q_{n-r_n})$ *with natural numbers* $r_n \to \infty$ *and let* ℓ *be a real number such that for every* $m \in \mathbb{N}_0$

$$\mathbf{E}[Y_n | \tau_{n-m} = n - m] \to \ell$$

as $n \to \infty$. *Also let* $\psi : \mathbb{N}_0 \to \mathbb{R}$ *be a bounded function with* $\psi(0) = 0$. *Then*

$$\mathbf{E}\big[Y_n \psi(Z_n) e^{-S_n} | \tau_n = n\big] \to \ell \int \psi \, dp'$$

as $n \to \infty$.

Proof. Let $f_n(s) = \sum_{k \ge 0} s^k Q_n(k)$, $0 \le s \le 1$, be the (random) generating function of Q_n, $n \ge 1$, and denote

$$f_{j,k} = \begin{cases} f_{j+1} \circ f_{j+2} \circ \cdots \circ f_k & \text{for } 0 \le j < k, \\ \text{id} & \text{for } j = k, \\ f_j \circ f_{j-1} \circ \cdots \circ f_{k+1} & \text{for } 0 \le k < j. \end{cases}$$

For $0 \le k < n$ set

$$L_{k,n} = \min(S_{k+1}, \dots, S_n) - S_k \quad \text{and} \quad L_{n,n} = 0.$$

First we look at the case $\psi(z) = 1 - s^z$ with $0 \le s < 1$ (with $0^0 = 1$). We decompose the expectation according to the value of τ_{n-m} for some fixed $m \in \mathbb{N}_0$. For convenience we assume $0 \le Y_n \le 1$. Then for $l > m$ because of $\mathbf{E}[Z_n | \Pi] = e^{S_n}$ a.s. and $1 - s^z \le z$

$$\begin{aligned} &\mathbf{E}\big[Y_n\big(1 - s^{Z_n}\big)e^{-S_n}; \tau_{n-m} < n - l, \tau_n = n\big] \\ &\quad \le \mathbf{E}\big[Z_n e^{-S_n}; \tau_{n-m} < n - l, \tau_n = n\big] = \mathbf{P}(\tau_{n-m} < n - l, \tau_n = n). \end{aligned}$$

From duality

$$\begin{aligned} \mathbf{P}(\tau_{n-m} < n - l, \tau_n = n) &= \mathbf{P}\Big(M_n < 0, \max_{m \le j \le n} S_j < \max_{l < j \le n} S_j < 0\Big) \\ &\le \mathbf{P}(S_k \ge S_m \text{ for some } l < k \le n, M_n < 0) \end{aligned}$$

and in view of Lemma 2.3

$$\begin{aligned} &\mathbf{P}(S_k \ge S_m \text{ for some } l < k \le n, M_n < 0) \\ &\quad \sim \mathbf{P}^-(S_k \ge S_m \text{ for some } k > l)\mathbf{P}(M_n < 0). \end{aligned}$$

Since $S_k \to -\infty$ $\mathbf{P}^-$-a.s. (see Lemma 2.6 in [4]), we obtain that for given $\varepsilon > 0$ and $m \in \mathbb{N}$ the estimate $\mathbf{P}^-(S_k \ge S_m \text{ for some } k > l) < \varepsilon$ is valid, if l is chosen large enough. Altogether this implies that for l sufficiently large

$$\begin{aligned} &\mathbf{E}\big[Y_n\big(1 - s^{Z_n}\big)e^{-S_n}; \tau_n = n\big] \\ &\quad = \mathbf{E}\big[Y_n\big(1 - s^{Z_n}\big)e^{-S_n}; \tau_{n-m} \ge n - l, \tau_n = n\big] + \chi_1 \end{aligned}$$

where $|\chi_1| \le \varepsilon \mathbf{P}(M_n < 0) = \varepsilon \mathbf{P}(\tau_n = n)$.

Next from the branching property

$$\begin{aligned}&\mathbf{E}\big[Y_n\big(1-s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_{n-m} \ge n-l, \tau_n = n\big]\\&\quad = \mathbf{E}\big[Y_n\big(1-f_{0,n}(s)\big)\mathrm{e}^{-S_n}; \tau_{n-m} \ge n-l, \tau_n = n\big].\end{aligned}$$

By means of duality

$$\begin{aligned}&\big|\mathbf{E}\big[Y_n\big(1-f_{0,n}(s)\big)\mathrm{e}^{-S_n}; \tau_{n-m} \ge n-l, \tau_n = n\big]\\&\qquad - \mathbf{E}\big[Y_n\big(1-f_{n-m,n}(s)\big)\mathrm{e}^{-(S_n-S_{n-m})}; \tau_{n-m} \ge n-l, \tau_n = n\big]\big|\\&\quad \le \mathbf{E}\big[\big|\big(1-f_{0,n}(s)\big)\mathrm{e}^{-S_n} - \big(1-f_{n-m,n}(s)\big)\mathrm{e}^{-(S_n-S_{n-m})}\big|; \tau_n = n\big]\\&\quad = \mathbf{E}\big[\big|\big(1-f_{n,0}(s)\big)\mathrm{e}^{-S_n} - \big(1-f_{m,0}(s)\big)\mathrm{e}^{-S_m}\big|; M_n < 0\big].\end{aligned}$$

Now $U_n(s) = (1-f_{n,0}(s))\mathrm{e}^{-S_n}$ is decreasing in n (see Lemma 2.3 in [21]) with limit $U_\infty(s)$, and for given $\varepsilon > 0$ we obtain from Lemma 2.3 for n large enough

$$\begin{aligned}&\big|\mathbf{E}\big[Y_n\big(1-f_{0,n}(s)\big)\mathrm{e}^{-S_n}; \tau_{n-m} \ge n-l, \tau_n = n\big]\\&\qquad - \mathbf{E}\big[Y_n\big(1-f_{n-m,n}(s)\big)\mathrm{e}^{-(S_n-S_{n-m})}; \tau_{n-m} \ge n-l, \tau_n = n\big]\big|\\&\quad \le 2\mathbf{E}^-\big[U_m(s) - U_\infty(s)\big]\mathbf{P}(\tau_n = n) \le \varepsilon \mathbf{P}(\tau_n = n),\end{aligned}$$

if only m is chosen large enough. Now $\{\tau_{n-m} \ge n-l, \tau_n = n\}$ may be decomposed as $\bigcup_{j=m}^{l}(\{\tau_{n-j} = n-j\} \cup \{L_{n-j,n-m} \ge 0, \tau_n = n\})$ and for large n by $Y_n = \varphi_n(Q_1, \ldots, Q_{n-r_n})$

$$\begin{aligned}&\mathbf{E}\big[Y_n\big(1-f_{n-m,n}(s)\big)\mathrm{e}^{-(S_n-S_{n-m})}; \tau_{n-j} = n-j, L_{n-j,n-m} \ge 0, \tau_n = n\big]\\&\quad = \mathbf{E}[Y_n; \tau_{n-j} = n-j]\mathbf{E}\big[\big(1-f_{j-m,j}(s)\big)\mathrm{e}^{-(S_j-S_{j-m})}; L_{j-m} \ge 0, \tau_j = j\big].\end{aligned}$$

By assumption $\mathbf{E}[Y_n; \tau_{n-j} = n-j] \sim \ell \mathbf{P}(\tau_n = n)$. Putting pieces together we obtain

$$\begin{aligned}&\mathbf{E}\big[Y_n\big(1-s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_n = n\big]\\&\quad = \mathbf{E}\big[Y_n\big(1-s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_{n-m} \ge n-l, \tau_n = n\big] + \chi_1\\&\quad = \ell \mathbf{P}(\tau_n = n)\sum_{j=m}^{l}\mathbf{E}\big[\big(1-f_{j-m,j}(s)\big)\mathrm{e}^{-(S_j-S_{j-m})}; L_{j-m} \ge 0, \tau_j = j\big] + \chi_2\end{aligned}$$

where $|\chi_2| \le 3\varepsilon \mathbf{P}(\tau_n = n)$. In particular we may apply this formula for $Y_n = 1$, to obtain for large n

$$\big|\mathbf{E}\big[Y_n\big(1-s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_n = n\big] - \ell \mathbf{E}\big[\big(1-s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_n = n\big]\big| \le 6\varepsilon \mathbf{P}(\tau_n = n)$$

and our computations boil down to the formula

$$\mathbf{E}\big[Y_n\big(1-s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_n = n\big] \sim \ell \mathbf{E}\big[\big(1-s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_n = n\big].$$

The right-hand side may be written as $\ell \mathbf{E}[(1-f_{n,0}(s))\mathrm{e}^{-S_n}; M_n < 0]$ and another application of Lemma 2.3 gives altogether

$$\mathbf{E}\big[Y_n\big(1-s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_n = n\big] \sim \ell \mathbf{E}^-\big[U_\infty(s)\big]\mathbf{P}(\tau_n = n).$$

In view of $s^z 1_{z>0} = (1-0^z) - (1-s^z)$ this implies

$$\mathbf{E}\big[Y_n s^{Z_n}\mathrm{e}^{-S_n}; Z_n > 0, \tau_n = n\big] \sim \ell h(s)\mathbf{P}(\tau_n = n) \tag{3.1}$$

with $h(s) = \mathbf{E}^-[U_\infty(0) - U_\infty(s)]$.

Now we show that $h(1) = \mathbf{E}^-[U_\infty(0)] > 0$. This follows from an estimate due to Agresti (see [6] and the proof of Proposition 3.1 in [4]), which in our case reads

$$\big(1 - f_{k,0}(s)\big)\mathrm{e}^{-S_k} \ge \left(\frac{1}{1-s} + \sum_{i=1}^{k} \eta_i \mathrm{e}^{S_i}\right)^{-1}.$$

Letting $k \to \infty$ Lemma 3.1 implies $U_\infty(s) > 0$ $\mathbf{E}^-$-a.s. and thus $h(s) > 0$ for all $s < 1$. For $s = 0$ it follows that $h(1) = \mathbf{E}^-[U_\infty(0)] > 0$.

Also from $\mathbf{E}[Z_n \mathrm{e}^{-S_n}; \tau_n = n] = \mathbf{E}[\mathbf{E}[Z_n|\Pi]\mathrm{e}^{-S_n}; \tau_n = n] = \mathbf{P}(\tau_n = n)$ and from $1 - s^z \le z(1-s)$ we get

$$\begin{aligned}\mathbf{E}\big[U_n(s); M_n < 0\big] &= \mathbf{E}\big[\big(1 - s^{Z_n}\big)\mathrm{e}^{-S_n}; \tau_n = n\big] \\ &\le (1-s)\mathbf{E}\big[Z_n \mathrm{e}^{-S_n}; \tau_n = n\big] = (1-s)\mathbf{P}(\tau_n = n)\end{aligned}$$

which for $n \to \infty$ implies $h(1) - h(s) = \mathbf{E}^-[U_\infty(s)] \le 1 - s$. Therefore h is continuous at $s = 1$. Our claim follows now from (3.1) and the continuity theorem for generating functions. □

Lemma 3.3. *Let Y_n fulfil the same conditions as in Lemma 3.2. Then under Assumptions A1 to A3 there is a non-vanishing finite measure p'' on $\mathbb{N} \times \mathbb{N}_0$ such that for every bounded $\psi : \mathbb{N} \times \mathbb{N}_0 \to \mathbb{R}$*

$$\frac{\mathbb{E}[Y_n \psi(Z_n, n - \tau_n); Z_n > 0]}{\gamma^n \mathbf{P}(\tau_n = n)} \to \ell \int \psi \, \mathrm{d}p''$$

as $n \to \infty$.

Proof. We have for fixed $j \in \mathbb{N}_0$ and every $n \in \mathbb{N}$ with $n > j$

$$\begin{aligned}&\gamma^{-n}\mathbb{E}\big[Y_n \psi(Z_n, n - \tau_n); Z_n > 0, \tau_n = n - j\big] \\ &\quad = \mathbf{E}\big[Y_n \psi_j(Z_{n-j})\mathrm{e}^{-S_{n-j}}; \tau_{n-j} = n - j\big]\end{aligned}$$

with $\psi_j(z) = \mathbf{E}[\psi(Z_j, j)\mathrm{e}^{-S_j}; Z_j > 0, L_j \ge 0 | Z_0 = z]$ for $z > 0$ and $\psi_j(0) = 0$. Also there is a finite measure p'_j such that $\int \psi_j \, \mathrm{d}p' = \int \psi(\cdot, j) \, \mathrm{d}p'_j$. From the preceding lemma

$$\frac{\mathbb{E}[Y_n \psi(Z_n, n - \tau_n); Z_n > 0, \tau_n = n - j]}{\gamma^n \mathbf{P}(\tau_n = n)} \to \ell \int \psi(\cdot, j) \, \mathrm{d}p'_j.$$

In particular p'_0 is non-vanishing. Thus it remains to show that for given $\varepsilon > 0$ there is a natural number k such that

$$\gamma^{-n}\mathbb{E}\big[Y_n \psi(Z_n, n - \tau_n); Z_n > 0, \tau_n \le n - k\big] \le \varepsilon \mathbf{P}(\tau_n = n)$$

for large n. Without loss of generality $0 \le Y_n \le 1$ and $0 \le \psi \le 1$. Then, using the inequality $\mathbb{P}(Z_i > 0|\Pi) \le \mathrm{e}^{S_i}$, we have

$$\begin{aligned}\gamma^{-n}\mathbb{E}\big[Y_n \psi(Z_n, n - \tau_n); Z_n > 0, \tau_n \le n - k\big] &\le \mathbf{E}\big[\mathrm{e}^{-S_n}; Z_n > 0, \tau_n \le n - k\big] \\ &\le \sum_{i=0}^{n-k} \mathbf{E}\big[\mathrm{e}^{-S_n}; Z_i > 0, \tau_i = i, L_{i,n} \ge 0\big] \\ &\le \sum_{i=0}^{n-k} \mathbf{E}\big[\mathrm{e}^{S_i - S_n}; \tau_i = i, L_{i,n} \ge 0\big] \\ &= \sum_{i=0}^{n-k} \mathbf{P}(\tau_i = i)\mathbf{E}\big[\mathrm{e}^{-S_{n-i}}; L_{n-i} \ge 0\big].\end{aligned}$$

From Lemmas 2.1, 2.2 both $\mathbf{P}(\tau_n = n)$ and $\mathbf{E}[\mathrm{e}^{-S_n}; L_n \geq 0]$ are regularly varying with negative indices. Therefore for large n

$$\begin{aligned}
&\gamma^{-n}\mathbb{E}\big[Y_n\psi(Z_n, n-\tau_n); Z_n > 0, \tau_n \leq n-k\big] \\
&\quad \leq \mathbf{E}\big[\mathrm{e}^{-S_{n/3}}; L_{n/3} \geq 0\big] \sum_{i\leq n/2} \mathbf{P}(\tau_i = i) \\
&\qquad + \mathbf{P}(\tau_{n/3} = n/3) \sum_{k\leq j\leq n/2} \mathbf{E}\big[\mathrm{e}^{-S_j}; L_j \geq 0\big],
\end{aligned}$$

where we used th fact that $\mathbb{P}(\tau_n = n) = \mathbb{P}(M_n < 0)$ is non-increasing in n. Also $\mathbf{E}[\mathrm{e}^{-S_n}; L_n \geq 0] = \mathrm{o}(\frac{1}{n})$ and $\sum_{i\leq n}\mathbf{P}(\tau_i = i) = \mathrm{O}(n\mathbf{P}(\tau_n = n))$ and $\sum_{j\geq 1}\mathbf{E}[\mathrm{e}^{-S_j}; L_j \geq 0] < \infty$. Therefore for every $\varepsilon > 0$ the right-hand side of the inequality above is bounded by $\varepsilon\mathbf{P}(\tau_n = n)$, if k is large enough. This gives the claim. □

Choosing $Y_n = 1$ and $\psi = 1_{\mathbb{N}\times\mathbb{N}_0}$, we obtain Theorem 1.1.

Proof of Theorem 1.2. In view of Theorem 1.1, the first part is a special case of Lemma 3.3 with $Y_n = 1$ and $\psi(Z_n, n-\tau_n) = 1 - s^{Z_n}$. For the second part we use Hölder's inequality (with $1/p = \beta$, $1/q = 1-\beta$) and (1.3)

$$\begin{aligned}
\gamma^{-n}\mathbb{E}\big[Z_n^\beta\big] &= \mathbf{E}\big[\mathbf{E}\big[Z_n^\beta 1_{Z_n>0}|\varPi\big]\mathrm{e}^{-S_n}\big] \\
&\leq \mathbf{E}\big[\mathbf{E}[Z_n|\varPi]^\beta\mathbf{P}(Z_n > 0|\varPi)^{1-\beta}\mathrm{e}^{-S_n}\big] \leq \mathbf{E}\big[\mathrm{e}^{(1-\beta)(L_n-S_n)}\big].
\end{aligned}$$

Again we decompose with τ_n and obtain

$$\begin{aligned}
\gamma^{-n}\mathbb{E}\big[Z_n^\beta\big] &\leq \sum_{i=0}^{n}\mathbf{E}\big[\mathrm{e}^{(1-\beta)(L_n-S_n)}; \tau_i = i, L_{i,n} \geq 0\big] \\
&= \sum_{i=0}^{n}\mathbf{P}(\tau_i = i)\mathbf{E}\big[\mathrm{e}^{-(1-\beta)S_{n-i}}; L_{n-i} \geq 0\big].
\end{aligned}$$

As above we show by means of Lemma 2.1 with $r = 1-\beta$ that this quantity is of order $\mathbf{P}(\tau_n = n)$, and the claim follows. □

Proof of Theorem 1.3. Again the first part is a special case of Lemma 3.3. Next let $\varphi : D[0,1] \to \mathbb{R}$ be bounded and continuous. We apply Lemma 3.3 to $Y_n = \varphi(\frac{1}{a_n}\bar{S}^n)$, where $\bar{S}_t^n = S_{nt\wedge r_n}$ with natural numbers $r_n \to \infty$. If $n - r_n = \mathrm{o}(n)$, then it follows from Lemma 2.5 and standard arguments that $\mathbf{E}[Y_n|\tau_{n-m} = n-m] \to \mathbf{E}[\varphi(L^*)]$. Lemma 3.3 yields

$$\mathbb{E}\left[\varphi\left(\frac{1}{a_n}\bar{S}^n\right)\Big|Z_n > 0\right] \to \mathbf{E}\big[\varphi\big(L^*\big)\big].$$

Thus $(\frac{1}{a_n}\bar{S}^n | Z_n > 0) \stackrel{d}{\to} L^*$. Also conditional asymptotic independence follows from Lemma 3.3. Finally for fixed r

$$\begin{aligned}
&\mathbb{P}\big(|X_{n-r+1}| + \cdots + |X_n| \geq \sqrt{a_n}; Z_n > 0\big) \\
&\quad \leq \mathbb{P}(Z_{n-r} > 0)\mathbb{P}\big(|X_1| + \cdots + |X_r| \geq \sqrt{a_n}\big) = \mathrm{o}\big(\mathbb{P}(Z_n > 0)\big).
\end{aligned}$$

This holds true also, if $r = r_n \to \infty$ sufficiently slow. It follows

$$\gamma^{-n}\mathbb{P}\left(\frac{1}{a_n}\sup\big|S^n - \bar{S}^n\big| \geq \varepsilon\Big|Z_n > 0\right) \to 0$$

for all $\varepsilon > 0$. Therefore $(\frac{1}{a_n}(S^n - \bar{S}^n)|Z_n > 0) \xrightarrow{d} 0$ in $D[0,1]$ and consequently $(\frac{1}{a_n} S^n | Z_n > 0) \xrightarrow{d} L^*$. This finishes the proof. □

4. Trees with stem

For every $n = 0, 1, \ldots, \infty$ let $\mathcal{T}_n$ be the set of all ordered rooted trees of height exactly n. For a precise definition we refer to the coding of ordered trees and their nodes given by Neveu [27]. Then $\mathcal{T}_{\geq n} = \mathcal{T}_n \cup \mathcal{T}_{n+1} \cup \cdots \cup \mathcal{T}_\infty$ is the set of ordered rooted trees of at least height n. With $[\,]_n : \mathcal{T}_{\geq n} \to \mathcal{T}_n$ we denote the operation of pruning a tree $t \in \mathcal{T}_{\geq n}$ to a tree $[t]_n \in \mathcal{T}_n$ of height exactly n by eliminating all nodes of larger height.

For $n = 0, 1, \ldots, \infty$ a tree with a stem of height n, shortly a *trest* of height n, is a pair

$$\mathsf{t} = (t, k_0 k_1 \ldots k_n),$$

where $t \in \mathcal{T}_{\geq n}$ and $k_0, \ldots, k_n$ are nodes in t such that k_0 is the root (founding ancestor) and k_i is an offspring of k_{i-1}. Thus k_i belongs to generation i. We call $k_0 \ldots k_n$ the stem within t, it is determined by k_n. $\mathcal{T}'_n$ denotes the set of all trests of height n. Such a construction is in the spirit of spinal decompositions as in [17,23] and others.

A trest $\mathsf{t} = (t, k_0 k_1 \ldots k_n)$ of height n can also be pruned at height $m \leq n$ to obtain the trest of height m

$$[\mathsf{t}]_m = \big([t]_m, k_0 \ldots k_m\big).$$

To every tree $t \in \mathcal{T}_{\geq n}$ there belongs a unique trest

$$\langle t \rangle_n = \big([t]_n, k_0(t) \ldots k_n(t)\big)$$

of height n, where $k_0(t) \ldots k_n(t)$ is the *leftmost* stem, which can be fitted into $[t]_n$. Notice that this stem is uniquely determined, since t is ordered and of at least height n.

Now let $\pi = (q_1, q_2, \ldots)$ be a fixed environment. Define the distribution $\tilde{q}_i$ by its weights

$$\tilde{q}_i(y) = \frac{1}{m(q_i)} y q_i(y), \quad y = 0, 1, \ldots$$

Then a corresponding *LPP-trest* (Lyons–Pemantle–Peres trest) is the random trest $\tilde{\mathsf{T}} = (\tilde{T}, \tilde{K}_0 \tilde{K}_1 \ldots)$ with values in $\mathcal{T}'_\infty$ satisfying the following properties:

Given $\Pi = (q_1, q_2, \ldots)$ a.s.

- the offspring numbers of all individuals are independent random variables,
- the offspring number of $\tilde{K}_{i-1}$ has distribution $\tilde{q}_i$ and the offspring number of any other individual in generation $i - 1$ has distribution q_i, and
- the node $\tilde{K}_i$ is uniformly distributed among all children of $\tilde{K}_{i-1}$, given the offspring number of $\tilde{K}_{i-1}$ and given all other random quantities.

Shortly speaking: From the infinite stem individuals grow according to a size biased distribution, and from the other individuals ordinary branching trees arise to the right and left of the stem. Such type of trests have been considered by Lyons, Peres and Pemantle [26] in the Galton–Watson case.

Let $\tilde{Z}_n$ be the population size of the LPP-trest in generation n.

Lemma 4.1. *Under Assumptions A1 to A3, as $n \to \infty$*

$$e^{-S_n} \tilde{Z}_n \to W^+ \quad \mathbf{P}^+\text{-}a.s.$$

with some random variable W^+ fulfilling $W^+ > 0$ $\mathbf{P}^+$-a.s.

Proof. We use the representation

$$\tilde{Z}_n = 1 + \sum_{i=0}^{n-1} \tilde{Z}_n^i$$

where $\tilde{Z}_n^i$ is the number of individuals in generation n other than $\tilde{K}_n$, which descent from $\tilde{K}_i$ but not from $\tilde{K}_{i+1}$. Thus $\mathbf{E}[\tilde{Z}_{i+1}^i | \Pi] = \sum_y y \tilde{Q}_{i+1}(y) - 1 = \mathrm{e}^{X_{i+1}} \eta_{i+1}$ and a.s.

$$\mathbf{E}\big[\tilde{Z}_n^i | \Pi\big] = \mathrm{e}^{S_n - S_{i+1}} \mathbf{E}\big[\tilde{Z}_{i+1}^i | \Pi\big] = \eta_{i+1} \mathrm{e}^{S_n - S_i}. \tag{4.1}$$

Now given the environment, $\mathrm{e}^{-S_n} \sum_{i=k}^{n-1} \tilde{Z}_n^i$ is for $n > k$ a non-negative submartingale. Therefore Doob's inequality implies that for every $\varepsilon \in (0, 1)$

$$\mathbf{P}\left(\max_{k<m\le n} \mathrm{e}^{-S_m} \sum_{i=k}^{m-1} \tilde{Z}_m^i \ge \varepsilon \middle| \Pi \right) \le \frac{1}{\varepsilon} \sum_{i=k}^{n-1} \mathrm{e}^{-S_n} \mathbf{E}\big[\tilde{Z}_n^i | \Pi\big] \le \frac{1}{\varepsilon} \sum_{i \ge k} \eta_{i+1} \mathrm{e}^{-S_i}$$

and

$$\mathbf{P}^+\left(\sup_{m>k} \mathrm{e}^{-S_m} \sum_{i=k}^{m-1} \tilde{Z}_m^i \ge \varepsilon \right) \le \frac{1}{\varepsilon} \mathbf{E}^+\left[1 \wedge \sum_{i \ge k} \eta_{i+1} \mathrm{e}^{-S_i} \right].$$

From Lemma 3.1 it follows that

$$\mathbf{P}^+\left(\sup_{m>k} \mathrm{e}^{-S_m} \sum_{i=k}^{m-1} \tilde{Z}_m^i \ge \varepsilon \right) \le \varepsilon,$$

if k is chosen large enough. Also $\mathrm{e}^{-S_n} \tilde{Z}_n^i$ is for $n \ge i+1$ and a fixed environment a non-negative martingale, such that for $n \to \infty$

$$\mathrm{e}^{-S_n} \tilde{Z}_n^i \to W^i \quad \mathbf{P}^+\text{-a.s.}$$

These facts together with $S_n \to \infty$ $\mathbf{P}^+$-a.s. (see [4], proof of Lemma 2.6) imply that

$$\mathrm{e}^{-S_n} \tilde{Z}_n \to W^+ \quad \mathbf{P}^+\text{-a.s.}$$

for some random variable W^+. Also $W^+ \ge \sum_{i \ge 0} W^i$ $\mathbf{P}^+$-a.s.

Thus it remains to show that $\sum_{i \ge 0} W^i > 0$ $\mathbf{P}^+$-a.s. Given Π, the random variables W^i are independent, since they arise from independent branching processes in the LPP-trest. In view of the second Borel–Cantelli Lemma it is thus sufficient to prove

$$\sum_{i \ge 0} \mathbf{P}^+\big(W^i > 0 | \Pi\big) = \infty \quad \mathbf{P}^+\text{-a.s.}$$

Now we use the formula

$$\mathbf{P}^+\big(W^i > 0 | \Pi\big) \ge \left(\sum_{j=i}^{\infty} \eta_{j+1} \mathrm{e}^{-(S_j - S_i)} \right)^{-1},$$

which is taken from the proof of Proposition 3.1 in [4] (a few lines after (3.7) therein). Because of Lemma 3.1 above the right-hand side is strictly positive $\mathbf{P}^+$-a.s. Moreover there are random times $0 = \nu(0) < \nu(1) < \cdots$ such that

$$\left(\sum_{j=\nu(k)}^{\infty} \eta_{j+1} \mathrm{e}^{-(S_j - S_{\nu(k)})} \right)^{-1}, \quad k = 0, 1, \ldots$$

is a stationary sequence of random variables, which is a consequence of Tanaka's decomposition, see [29] and Lemma 2.6 in [4]. From Birkhoff's ergodic theorem it follows that

$$\frac{1}{n}\sum_{k=1}^{n}\left(\sum_{j=\nu(k)}^{\infty}\eta_{j+1}\mathrm{e}^{-(S_j-S_{\nu(k)})}\right)^{-1}$$

has a strictly positive limit $\mathbf{P}^+$-a.s. This implies our claim. □

We use the LPP-tree to approximate conditioned BPRE. Let us denote by T a branching tree in random environment Π. This is nothing else than the ordered rooted tree belonging to a BPRE in environment Π. Again let Z_n denote its number of individuals in generation n.

Theorem 4.2. *Assume Assumptions A1 to A3. Let $0 \le r_n < n$ be a sequence of natural numbers with $r_n \to \infty$. Let Y_n be uniformly bounded random variables of the form $Y_n = \varphi(Q_1, \dots, Q_{n-r_n})$ and let $B_n \subset \mathcal{T}'_{n-r_n}$, $n \ge 1$. If for some $\ell \ge 0$*

$$\mathbf{E}\big[Y_n; [\tilde{\mathsf{T}}]_{n-r_n} \in B_n | \tau_{n-m} = n-m\big] \to \ell$$

for all $m \ge 0$, then

$$\mathbb{E}\big[Y_n; \big[\langle T\rangle_n\big]_{n-r_n} \in B_n | Z_n > 0\big] \to \ell.$$

B_n may be random, depending only on the environment Π.

For the proof we use the following theorem due to J. Geiger (see [20]). Let $\pi = (q_1, q_2, \dots)$ be a fixed environment, let $\mathbf{P}_\pi(\cdot)$ be the corresponding probabilities and let

$$\mathsf{T}_{n,\pi} = (T_n, K_0 \dots K_n)$$

denote a random trest of height n and let for $i = 1, \dots, n$

$$\begin{aligned}
T'_i &= \text{subtree within } T_n \text{ right to the stem with root } K_{i-1},\\
T''_i &= \text{subtree within } T_n \text{ left to the stem with root } K_{i-1},\\
R_i &= \text{size of the first generation of } T'_i,\\
L_i &= \text{size of the first generation of } T''_i.
\end{aligned}$$

For $\mathsf{T}_{n,\pi}$ the following properties are required:

- $\mathbf{P}_\pi(R_i = r, L_i = l) = q_i(r+l+1)\frac{\mathbf{P}_\pi(Z_n>0|Z_i=1)\mathbf{P}_\pi(Z_n=0|Z_i=1)^l}{\mathbf{P}_\pi(Z_n>0|Z_{i-1}=1)}$.
- T'_i, if decomposed at its first generation, consists of R_i subtrees τ'_{ij}, $j = 1, \dots, R_i$, which are branching trees within the fixed environment $(q_{i+1}, q_{i+2}, \dots)$.
- Similarly T''_i consists of L_i subtrees τ''_{ij}, which are branching trees within the fixed environment $(q_{i+1}, q_{i+2}, \dots)$ conditioned to be extinct before generation $n - i$.
- All pairs (R_i, L_i) and all subtrees τ'_{ij}, τ''_{ij} are independent.

These properties determine the distribution of $\mathsf{T}_{n,\pi}$ up to possible offspring of K_n and thus the distribution of $\langle \mathsf{T}_{n,\pi}\rangle_n$.

Theorem 4.3. *For almost all π the conditional distribution of $\langle T\rangle_n$ given $\Pi = \pi$, $Z_n > 0$ is equal to the distribution of $\langle \mathsf{T}_{n,\pi}\rangle_n$.*

Geiger proved this result for a fixed environment $q_1 = q_2 = \cdots$ i.e. in the Galton–Watson case, see Proposition 2.1 in [20]. His proof carries over straightforward to a varying environment.

Proof of Theorem 4.2. For the trest $\tilde{\mathsf{T}}$ we introduce the notations $\tilde{T}_i', \tilde{T}_i'', \tilde{R}_i, \tilde{L}_i, \tilde{\tau}_{ij}', \tilde{\tau}_{ij}''$. They have the same meaning as above $T_i', T_i'', R_i, L_i, \tau_{ij}', \tau_{ij}''$ for the trest $\mathsf{T}_{n,\pi}$. From the construction of $\tilde{\mathsf{T}}$

$$\mathbf{P}_\pi(\tilde{R}_i = r, \tilde{L}_i = l) = q_i(r + l + 1)\mathrm{e}^{-X_i}.$$

$\tilde{\tau}_{ij}'$ and τ_{ij}' are equal in distribution, whereas $\tilde{\tau}_{ij}''$ is no longer conditioned to be extinct in generation $n - i$, as this is the case for τ_{ij}''.

In order to compare both trests we will couple them. We first consider the branching process in a fixed environment $\pi = (q_1, q_2, \ldots)$ and again write the corresponding probabilities as $\mathbf{P}_\pi(\cdot)$. To begin with we estimate the total variation distance between the distributions of (R_i, L_i) and $(\tilde{R}_i, \tilde{L}_i)$. Note that

$$\begin{aligned}\mathbf{P}_\pi(Z_n > 0 | Z_{i-1} = 1) &= \sum_{j\ge 1} \mathbf{P}_\pi(Z_n > 0 | Z_i = j)\mathbf{P}_\pi(Z_i = j | Z_{i-1} = 1)\\ &\le \sum_{j\ge 1} j\mathbf{P}_\pi(Z_n > 0 | Z_i = 1)\mathbf{P}_\pi(Z_i = j | Z_{i-1} = 1)\\ &= \mathrm{e}^{X_i}\mathbf{P}_\pi(Z_n > 0 | Z_i = 1)\end{aligned}$$

such that for $r, l, m \ge 0$ and $i \le n - m$

$$\begin{aligned}&\mathbf{P}_\pi(\tilde{R}_i = r, \tilde{L}_i = l) - \mathbf{P}_\pi(R_i = r, L_i = l)\\ &\quad\le q_i(r + l + 1)\mathrm{e}^{-X_i}\big(1 - \mathbf{P}_\pi(Z_n = 0 | Z_i = 1)^l\big)\\ &\quad\le q_i(r + l + 1)\mathrm{e}^{-X_i} l\big(1 - \mathbf{P}_\pi(Z_n = 0 | Z_i = 1)\big)\\ &\quad\le l q_i(r + l + 1)\mathrm{e}^{-X_i}\mathbf{P}_\pi(Z_{n-m} > 0 | Z_i = 1)\\ &\quad\le l q_i(r + l + 1)\mathrm{e}^{-X_i}\mathrm{e}^{S_{n-m} - S_i}.\end{aligned}$$

Since the right-hand side is always non-negative, we may estimate the total variation distance as

$$\begin{aligned}&\frac{1}{2}\sum_{r,l\ge 0}\big|\mathbf{P}_\pi(\tilde{R}_i = r, \tilde{L}_i = l) - \mathbf{P}_\pi(R_i = r, L_i = l)\big|\\ &\quad= \sum_{r,l\ge 0}\big(\mathbf{P}_\pi(\tilde{R}_i = r, \tilde{L}_i = l) - \mathbf{P}_\pi(R_i = r, L_i = l)\big)^+\\ &\quad\le \mathrm{e}^{-X_i}\mathrm{e}^{S_{n-m} - S_i}\sum_{r,l\ge 0} l q_i(r + l + 1)\\ &\quad= \mathrm{e}^{-X_i}\mathrm{e}^{S_{n-m} - S_i}\frac{1}{2}\sum_{y=1}^{\infty} y(y-1)q_i(y) = \frac{1}{2}\eta_i\mathrm{e}^{S_{n-m} - S_{i-1}}.\end{aligned}$$

Similarly we estimate the total variation distance between the distributions of τ_{ij}'' and $\tilde{\tau}_{ij}''$. The second distribution is equal to the first distribution conditioned to be extinct in generation $n - i$. This event can be expressed as $\{\tau_{ij}'' \in B_i\}$ with the set B_i of trees of height less than $n - i$, thus for some tree t

$$\begin{aligned}\mathbf{P}_\pi\big(\tau_{ij}'' = t\big) - \mathbf{P}_\pi\big(\tilde{\tau}_{ij}'' = t\big) &= \mathbf{P}_\pi\big(\tau_{ij}'' = t\big) - \mathbf{P}_\pi\big(\tau_{ij}'' = t | \tau_{ij}'' \in B_i\big)\\ &\le \mathbf{P}_\pi\big(\tau_{ij}'' = t\big)1_{B_i^c}(t).\end{aligned}$$

Again, since the right-hand side is non-negative for $i \le n-m$

$$\frac{1}{2}\sum_t \bigl|\mathbf{P}_\pi\bigl(\tau''_{ij}=t\bigr)-\mathbf{P}_\pi\bigl(\tilde\tau''_{ij}=t\bigr)\bigr| \\ \le \mathbf{P}_\pi\bigl(\tau''_{ij}\in B_i^c\bigr)=\mathbf{P}_\pi(Z_n>0|Z_i=1)\le \mathbf{P}_\pi(Z_{n-m}>0|Z_i=1)\le \mathrm{e}^{S_{n-m}-S_i}.$$

Now we consider the following construction: Take couplings of the pairs (R_i, L_i), $(\tilde R_i, \tilde L_i)$ and of τ''_{ij} and $\tilde\tau''_{ij}$. Also let $\tau'_{ij}=\tilde\tau'_{ij}$. Put these components together to obtain (T'_i, T''_i) and $(\tilde T'_i, \tilde T''_i)$. If the couplings are all independent of each other, then the resulting trests have the required distributional properties. We denote the resulting probabilities again by $\mathbf{P}_\pi$. Thus

$$\mathbf{P}_\pi\bigl(\bigl(T'_i, T''_i\bigr)\ne\bigl(\tilde T'_i, \tilde T''_i\bigr)\bigr) \\ \le \mathbf{P}_\pi\bigl((R_i,L_i)\ne(\tilde R_i,\tilde L_i)\bigr)+\sum_{r,l\ge 0}\sum_{j=1}^{l}\mathbf{P}_\pi(\tilde R_i=r,\tilde L_i=l)\mathbf{P}_\pi\bigl(\tau''_{ij}\ne\tilde\tau''_{ij}\bigr).$$

For optimal couplings we may use the above estimates on the total variation distance and obtain for $i\le n-m$

$$\mathbf{P}_\pi\bigl(\bigl(T'_i, T''_i\bigr)\ne\bigl(\tilde T'_i, \tilde T''_i\bigr)\bigr)\le\frac{1}{2}\eta_i\mathrm{e}^{S_{n-m}-S_{i-1}}+\sum_{r,l\ge 0} lq_i(r+l+1)\mathrm{e}^{-X_i}\mathrm{e}^{S_{n-m}-S_i} \\ =\eta_i\mathrm{e}^{S_{n-m}-S_{i-1}}.$$

Altogether using Theorem 4.3 and the assumption that B_n depends only on Π, it follows for $m<r_n$

$$\bigl|\mathbf{P}_\pi\bigl(\bigl[\langle T\rangle_n\bigr]_{n-r_n}\in B_n|Z_n>0\bigr)-\mathbf{P}_\pi\bigl([\tilde{\mathsf{T}}]_{n-r_n}\in B_n\bigr)\bigr|\le 1\wedge\sum_{i=1}^{n-r_n}\eta_i\mathrm{e}^{S_{n-m}-S_{i-1}}.$$

Now from duality and from Lemmas 2.3, 3.1

$$\mathbf{E}\Biggl[1\wedge\sum_{i=1}^{n-r_n}\eta_i\mathrm{e}^{S_{n-m}-S_{i-1}}\Biggm|\tau_{n-m}=n-m\Biggr] \\ =\mathbf{E}\Biggl[1\wedge\sum_{i=r_n-m}^{n-m}\eta_i\mathrm{e}^{S_i}\Biggm|M_{n-m}<0\Biggr]\to 0.$$

According to our assumptions $\mathbf{E}[Y_n\mathbf{P}_\Pi([\tilde{\mathsf{T}}]_{n-r_n}\in B_n)|\tau_{n-m}=n-m]$ converges to ℓ. Our estimates thus imply that

$$\mathbf{E}\bigl[Y_n\mathbf{P}_\Pi\bigl(\bigl[\langle T\rangle_n\bigr]_{n-r_n}\in B_n|Z_n>0\bigr)\bigm|\tau_{n-m}=n-m)\bigr]\to\ell.$$

Thus we may apply Lemma 3.3 with $Y_n\mathbf{P}_\Pi(([\langle T\rangle_n]_{n-r_n}\in B_n|Z_n>0)$ instead of Y_n, $\psi=1$ to obtain

$$\frac{\mathbf{E}[Y_n\mathbf{P}_\Pi([\langle T\rangle_n]_{n-r_n}\in B_n|Z_n>0);Z_n>0]}{\gamma^n\mathbf{P}(\tau_n=n)}\to\ell p''(\mathbb{N}\times\mathbb{N}_0).$$

Also from Lemma 3.3 with $Y_n=1$ and $\psi=1$

$$\mathbb{P}(Z_n>0)\sim\gamma^n\mathbf{P}(\tau_n=n)p''(\mathbb{N}\times\mathbb{N}_0),$$

thus

$$\mathbf{E}\bigl[Y_n\mathbf{P}\bigl(\bigl[\langle T\rangle_n\bigr]_{n-r_n}\in B_n|\Pi,Z_n>0\bigr)\bigm|Z_n>0)\bigr]\to\ell.$$

Now

$$\begin{aligned}
&\mathbf{E}\big[Y_n\mathbf{P}\big(\big[\langle T\rangle_n\big]_{n-r_n}\in B_n\,|\,\Pi, Z_n>0\big); Z_n>0\big]\\
&\quad=\mathbf{E}\bigg[Y_n\frac{\mathbf{P}([\langle T\rangle_n]_{n-r_n}\in B_n, Z_n>0|\Pi)}{\mathbf{P}(Z_n>0|\Pi)}; Z_n>0\bigg]\\
&\quad=\mathbf{E}\big[Y_n\mathbf{P}\big(\big[\langle T\rangle_n\big]_{n-r_n}\in B_n, Z_n>0|\Pi\big)\big]\\
&\quad=\mathbf{E}\big[\mathbf{E}\big[Y_n; \big[\langle T\rangle_n\big]_{n-r_n}\in B_n, Z_n>0|\Pi\big]\big]\\
&\quad=\mathbf{E}\big[Y_n; \big[\langle T\rangle_n\big]_{n-r_n}\in B_n, Z_n>0\big].
\end{aligned}$$

This gives the claim of Theorem 4.2. □

5. Proof of Theorem 1.4

Let again $\tilde{\mathsf{T}}$ denote the LPP-trest. Recall that $\tilde{Z}^i_j$ is for $i<j$ the number of the individuals in generation j other than $\tilde{K}_j$, which descent from $\tilde{K}_i$ but not from $\tilde{K}_{i+1}$. For convenience we put $\tilde{Z}^i_j=0$ for $i\ge j$.

Lemma 5.1. *Let $0<t<1$. Then for every $\varepsilon>0$ there is a natural number a such that for any natural numbers m and $\varsigma\in[\tau_{nt}, nt]$*

$$\mathbf{P}\bigg(\sum_{i:|i-\tau_{nt}|\ge a}\frac{\tilde{Z}^i_\varsigma}{\mathrm{e}^{S_\varsigma-S_{\tau_{nt}}}}\ge\varepsilon\,\bigg|\,\tau_{n-m}=n-m\bigg)\le\varepsilon,$$

if n is sufficiently large (depending on ε, a and m). ς may be random, depending only on the random environment Π.

Proof. For $0<\varepsilon\le 1$ from Markov inequality and (4.1)

$$\begin{aligned}
&\varepsilon\mathbf{P}\bigg(\sum_{|i-\tau_{nt}|\ge a}\frac{\tilde{Z}^i_\varsigma}{\mathrm{e}^{S_\varsigma-S_{\tau_{nt}}}}\ge\varepsilon; \tau_{n-m}=n-m\bigg)\\
&\quad\le\mathbf{E}\bigg[1\wedge\sum_{i\le\varsigma,|i-\tau_{nt}|\ge a}\eta_{i+1}\mathrm{e}^{S_{\tau_{nt}}-S_i}; \tau_{n-m}=n-m\bigg].
\end{aligned}$$

Next we decompose with the value of τ_{nt} to obtain for $m\le(1-t)n$

$$\begin{aligned}
&\varepsilon\mathbf{P}\bigg(\sum_{|i-\tau_{nt}|\ge a}\frac{\tilde{Z}^i_\varsigma}{\mathrm{e}^{S_\varsigma-S_{\tau_{nt}}}}\ge\varepsilon; \tau_{n-m}=n-m\bigg)\\
&\quad\le\sum_{j\le nt}\mathbf{E}\bigg[1\wedge\sum_{i\le\varsigma,|i-j|\ge a}\eta_{i+1}\mathrm{e}^{S_j-S_i}; \tau_j=j, L_{j,nt}\ge 0\bigg]\\
&\qquad\times\mathbf{P}\big(\tau_{(1-t)n-m}=\lfloor(1-t)n\rfloor-m\big).
\end{aligned}$$

We split the expectation:

$$\begin{aligned}
&\sum_{j\le nt}\mathbf{E}\bigg[1\wedge\sum_{i\le\varsigma,|i-j|\ge a}\eta_{i+1}\mathrm{e}^{S_j-S_i}; \tau_j=j, L_{j,nt}\ge 0\bigg]\\
&\quad=\sum_{j\le nt}\mathbf{E}\bigg[1\wedge\sum_{i=0}^{j-a}\eta_{i+1}\mathrm{e}^{S_j-S_i}; \tau_j=j\bigg]\mathbf{P}(L_{nt-j}\ge 0)+\sum_{j\le nt}\mathbf{P}(\tau_j=j)\mathbf{E}\bigg[1\wedge\sum_{i=j+a}^{\varsigma}\eta_{i+1}\mathrm{e}^{S_j-S_i}; L_{j,nt}\ge 0\bigg].
\end{aligned}$$

Duality yields

$$\begin{aligned}
&\sum_{j\le nt}\mathbf{E}\Bigg[1\wedge\sum_{i\le\varsigma,|i-j|\ge a}\eta_{i+1}\mathrm{e}^{S_j-S_i};\tau_j=j,L_{j,nt}\ge 0\Bigg]\\
&\quad\le\sum_{a\le j\le nt}\mathbf{E}\Bigg[1\wedge\sum_{i=a}^{j}\eta_i\mathrm{e}^{S_i};M_j<0\Bigg]\mathbf{P}(L_{nt-j}\ge 0)\\
&\qquad+\sum_{a\le k\le nt}\mathbf{P}\big(\tau_{nt-k}=\lfloor nt\rfloor-k\big)\mathbf{E}\Bigg[1\wedge\sum_{i=a}^{k}\eta_{i+1}\mathrm{e}^{-S_i};L_k\ge 0\Bigg].
\end{aligned}$$

From Lemmas 2.3, 3.1 we may choose a so large that

$$\mathbf{E}\Bigg[1\wedge\sum_{i=a}^{j}\eta_i\mathrm{e}^{S_i};M_j<0\Bigg]\le\delta\mathbf{P}(M_j<0)$$

$$\mathbf{E}\Bigg[1\wedge\sum_{i=a}^{k}\eta_{i+1}\mathrm{e}^{-S_i};L_k\ge 0\Bigg]\le\delta\mathbf{P}(L_k\ge 0)$$

for all $j,k>a$ and given $\delta>0$. It follows from duality

$$\begin{aligned}
&\sum_{j\le nt}\mathbf{E}\Bigg[1\wedge\sum_{i\le\varsigma,|i-j|\ge a}\eta_{i+1}\mathrm{e}^{S_j-S_i};\tau_j=j,L_{j,nt}\ge 0\Bigg]\\
&\quad\le\delta\sum_{a\le j\le nt}\mathbf{P}(\tau_j=j)\mathbf{P}(L_{nt-j}\ge 0)\\
&\qquad+\delta\sum_{a\le k\le nt}\mathbf{P}\big(\tau_{nt-k}=\lfloor nt\rfloor-k\big)\mathbf{P}(L_k\ge 0)\\
&\quad\le 2\delta
\end{aligned}$$

and

$$\begin{aligned}
&\mathbf{P}\Bigg(\sum_{|i-\tau_{nt}|\ge a}\frac{\tilde{Z}^i_\varsigma}{\mathrm{e}^{S_\varsigma-S_{\tau_{nt}}}}\ge\varepsilon;\tau_n=n\Bigg)\\
&\quad\le\frac{2\delta}{\varepsilon}\mathbf{P}\big(\tau_{(1-t)n-m}=\big\lfloor(1-t)n\big\rfloor-m\big).
\end{aligned}$$

Since $\mathbf{P}(\tau_n=n)$ is regularly varying, the right-hand side is bounded by the term $\varepsilon\mathbf{P}(\tau_n=n)$, if δ is chosen small enough. This gives the claim. □

We now come to the proof of the first part of Theorem 1.4. Let $\sigma_{i,n}$ as in (2.3) and define $\mu_n(i)$ as the smallest natural number j between 1 and i such that $\tau_{nt_i}=\sigma_{j,n}$,

$$\mu_n(i)=\min\{j\le i:\ \tau_{nt_i}=\sigma_{j,n}\}. \tag{5.1}$$

Again let $\tilde{Z}_j$ be the number of individuals in generation j of the LPP-trest $\tilde{\mathsf{T}}$, thus

$$\tilde{Z}_j=1+\sum_{k=0}^{j-1}\tilde{Z}^k_j.$$

Therefore, given $\varepsilon > 0$ in view of the preceding lemma with $\varsigma = \tau_{nt}$ there is a natural number a such that given $\tau_{n-m} = n - m$ the probability is at least $1 - \varepsilon$ that the event

$$\tilde{Z}_{\tau_{nt_i}} = 1 + \sum_{|k-\tau_{nt_i}|\le a} \tilde{Z}^k_{\tau_{nt_i}} = 1 + \sum_{k=\sigma_{\mu_n(i),n}-a}^{\sigma_{\mu_n(i),n}} \tilde{Z}^k_{\sigma_{\mu_n(i),n}}$$

holds for all $i = 1, \dots, r$. Now note that given the environment Π the distribution of

$$1 + \sum_{k=\sigma_{j,n}-a}^{\sigma_{j,n}} \tilde{Z}^k_{\sigma_{j,n}}$$

only depends on $(Q_{\sigma_{j,n}-a}, \dots, Q_{\sigma_{j,n}})$. Lemma 2.7 says that given $\tau_{n-m} = n - m$ these random vectors are asymptotically i.i.d. Also this lemma gives asymptotic independence of these random vectors from

$$\frac{1}{a_n}(S_{\sigma_{1,n}}, S_{nt_1}, \dots, S_{\sigma_{r,n}}, S_{nt_r}),$$

which in turn determines $\mu_n(1), \dots, \mu_n(r)$. Finally in view of Lemma 2.5 $\mu_n(1), \dots, \mu_n(r)$ converges in distribution to $\mu(1), \dots, \mu(r)$.

These observations hold for every $\varepsilon > 0$. Therefore we may summarize our discussion as follows: For all $m \ge 1$

$$\big((\tilde{Z}_{\tau_{nt_1}}, \dots, \tilde{Z}_{\tau_{nt_r}})|\tau_{n-m} = n - m\big) \xrightarrow{d} (V_{\mu(1)}, \dots, V_{\mu(r)}),$$

where the right-hand term has just the properties as given in Theorem 1.4. Now Theorem 4.2 gives the claim.

The proof of the second part of Theorem 1.4 is prepared by the following lemma. Let for fixed a

$$\hat{Z}_{a,k} = \sum_{i:|i-\tau_{nt}|\le a} \tilde{Z}^i_k$$

and

$$\alpha_{a,n} = \mathrm{e}^{S_{\tau_{nt}} - S_{nt}} \hat{Z}_{a,nt}, \qquad \beta_{a,n} = \mathrm{e}^{S_{\tau_{nt}} - S_{\tau_{nt}+a}} \hat{Z}_{a,\tau_{nt}+a}.$$

Lemma 5.2. *Let $m \ge 1$, $\varepsilon > 0$ and $0 < t < 1$. Then, if a is sufficiently large*

$$\limsup_{n\to\infty} \mathbf{P}\big(|\alpha_{a,n} - \beta_{a,n}| > \varepsilon \big| \tau_{n-m} = n - m\big) \le \varepsilon.$$

Proof. Because of Markov inequality and (4.1)

$$\begin{aligned}
&\mathbf{P}(\beta_{a,n} > d|\tau_{n-m} = n - m) \\
&\quad \le \mathbf{P}\big(\mathrm{e}^{S_{\tau_{nt}} - S_{\tau_{nt}+a}} \mathbf{E}[\hat{Z}_{a,\tau_{nt}+a}|\Pi] > \sqrt{d}\big|\tau_{n-m} = n - m\big) + \frac{1}{\sqrt{d}} \\
&\quad \le \mathbf{P}\Bigg(\sum_{i:|i-\tau_{nt}|\le a} \eta_{i+1} \mathrm{e}^{S_{\tau_{nt}} - S_i} > \sqrt{d}\Bigg|\tau_{n-m} = n - m\Bigg) + \frac{1}{\sqrt{d}}.
\end{aligned}$$

From Lemma 2.7 (with $r = 1$, thus $\sigma_{1,n} = \tau_{nt}$) it follows that the sum converges in distribution for $n \to \infty$ and

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbf{P}(\beta_{a,n} > d|\tau_{n-m} = n - m) \\
&\quad \le \mathbf{P}^-\Bigg(\sum_{i\ge 1} \eta_i \mathrm{e}^{S_i} \ge \sqrt{d}\Bigg) + \mathbf{P}^+\Bigg(\sum_{i\ge 0} \eta_{i+1} \mathrm{e}^{-S_i} \ge \sqrt{d}\Bigg) + \frac{1}{\sqrt{d}}.
\end{aligned}$$

Therefore from Lemma 3.1 it results that there is a $d < \infty$ such that for all $a > 0$

$$\limsup_{n\to\infty} \mathbf{P}(\beta_{a,n} > d | \tau_{n-m} = n-m) < \varepsilon/2.$$

Moreover from Lemma 2.5 $t - \frac{1}{n}\tau_{nt}$ converges in distribution to a strictly positive random variable, thus $\mathbf{P}(\tau_{nt} + a \geq nt | \tau_{n-m} = n-m) \to 0$ for $n \to \infty$. Therefore for n large enough,

$$\mathbf{P}\big(|\beta_{a,n} - \alpha_{a,n}| > \varepsilon \big| \tau_{n-m} = n-m\big) \tag{5.2}$$

$$\leq \frac{\varepsilon}{2} + \mathbf{P}\big(|\alpha_{a,n} - \beta_{a,n}| > \varepsilon, \beta_{a,n} \leq d, \tau_{nt} + a \leq nt \big| \tau_{n-m} = n-m\big). \tag{5.3}$$

Now, given Π, $\hat{Z}_{a,\tau_{nt}+a}$ and $\tau_{nt} + a \leq nt$, the process $\hat{Z}_{a,k}$, $k \geq \tau_{nt} + a$ is a branching process in varying environment. Therefore $\mathbf{E}[\alpha_{a,n} | \Pi, \hat{Z}_{a,\tau_{nt}+a}] = \beta_{a,n}$ a.s. Also the branching property yields

$$\frac{\mathbf{Var}(Z_n | Z_0 = z, \Pi)}{\mathbf{E}[Z_n | Z_0 = 1, \Pi]^2} = z\left(\mathrm{e}^{-S_n} + \sum_{i=0}^{n-1} \eta_{i+1} \mathrm{e}^{-S_i} - 1\right), \tag{5.4}$$

therefore on $\tau_{nt} + a \leq nt$

$$\begin{aligned}
\varepsilon^2 \mathbf{P}\big(|\beta_{a,n} - \alpha_{a,n}| > \varepsilon \big| \Pi, \hat{Z}_{a,\tau_{nt}+a}\big) &\leq \mathbf{E}\big[(\beta_{a,n} - \alpha_{a,n})^2 | \Pi, \hat{Z}_{a,\tau_{nt}+a}\big] \\
&\leq \hat{Z}_{a,\tau_{nt}+a} \mathrm{e}^{2(S_{\tau_{nt}} - S_{\tau_{nt}+a})} \left(\mathrm{e}^{-(S_{nt} - S_{\tau_{nt}+a})} + \sum_{i=\tau_{nt}+a}^{\lfloor nt \rfloor} \eta_{i+1} \mathrm{e}^{-(S_i - S_{\tau_{nt}+a})}\right) \\
&= \beta_{a,n} \left(\mathrm{e}^{-(S_{nt} - S_{\tau_{nt}})} + \sum_{i=\tau_{nt}+a}^{\lfloor nt \rfloor} \eta_{i+1} \mathrm{e}^{-(S_i - S_{\tau_{nt}})}\right).
\end{aligned}$$

Inserting this estimate into (5.2), we obtain

$$\begin{aligned}
&\mathbf{P}\big(|\beta_{a,n} - \alpha_{a,n}| > \varepsilon; \tau_{n-m} = n-m\big) \\
&\quad \leq \frac{\varepsilon}{2} \mathbf{P}(\tau_{n-m} = n-m) \\
&\qquad + \frac{d}{\varepsilon^2} \mathbf{E}\left[1 \wedge \left(\mathrm{e}^{-(S_{nt} - S_{\tau_{nt}})} + \sum_{i=\tau_{nt}+a}^{\lfloor nt \rfloor} \eta_{i+1} \mathrm{e}^{-(S_i - S_{\tau_{nt}})}\right); \tau_{nt} + a \leq nt, \tau_{n-m} = n-m\right] \\
&\quad \leq \frac{\varepsilon}{2} \mathbf{P}(\tau_{n-m} = n-m) \\
&\qquad + \frac{d}{\varepsilon^2} \sum_{j \leq nt-a} \mathbf{P}(\tau_{nt} = j) \mathbf{E}\left[1 \wedge \left(\mathrm{e}^{-S_{nt-j}} + \sum_{i=a}^{\lfloor nt \rfloor - j} \eta_{i+1} \mathrm{e}^{-S_i}\right); L_{nt-j} \geq 0\right] \mathbf{P}\big(\tau_{n(1-t)-m} = \lfloor n(1-t) \rfloor - m\big).
\end{aligned}$$

From Lemmas 2.1, 2.3, 3.1 together with the fact that $\mathbf{P}(\tau_n = n)$ is regularly varying our claim follows for a sufficiently large. □

We are now ready to finish the proof of Theorem 1.4. We first treat the case $r = 1$. From $\tilde{Z}_{nt} = 1 + \hat{Z}_{a,nt} + \sum_{i:|i-\tau_{nt}|>a} \tilde{Z}^i_{nt}$

$$\begin{aligned}
&\mathbf{P}\big(\big|\mathrm{e}^{S_{\tau_{nt}} - S_{nt}} \tilde{Z}_{nt} - \beta_{a,n}\big| \geq 3\varepsilon \big| \tau_{n-m} = n-m\big) \\
&\quad \leq \mathbf{P}\big(\mathrm{e}^{S_{\tau_{nt}} - S_{nt}} \geq \varepsilon | \tau_{n-m} = n-m\big)
\end{aligned}$$

$$+\mathbf{P}\big(|\alpha_{a,n}-\beta_{a,n}|\geq\varepsilon\big|\tau_{n-m}=n-m\big)$$
$$+\mathbf{P}\bigg(\mathrm{e}^{S_{\tau_{nt}}-S_{nt}}\sum_{i:|i-\tau_{nt}|>a}\tilde{Z}^i_{nt}\geq\varepsilon\bigg|\tau_{n-m}=n-m\bigg).$$

From Lemma 2.5 it results that

$$\mathbf{P}\big(\mathrm{e}^{S_{\tau_{nt}}-S_{nt}}\geq\varepsilon|\tau_{n-m}=n-m\big)=\mathbf{P}\bigg(\frac{S_{\tau_{nt}}-S_{nt}}{a_n}\geq\frac{\log\varepsilon}{a_n}\bigg|\tau_{n-m}=n-m\bigg)\to 0.$$

Together with Lemmas 5.1, 5.2 it follows that for all $\varepsilon>0$ there is a natural number a such that

$$\mathbf{P}\big(\big|\mathrm{e}^{S_{\tau_{nt}}-S_{nt}}\tilde{Z}_{nt}-\beta_{a,n}\big|\geq 3\varepsilon|\tau_{n-m}=n-m\big)\leq 3\varepsilon$$

for large n.

Moreover from Lemma 2.7 we see that $\beta_{a,n}$, conditioned on $\tau_{n-m}=n-m$, converges in distribution for every a. This implies that $\mathrm{e}^{S_{\tau_{nt}}-S_{nt}}\tilde{Z}_{nt}$ conditioned on $\tau_{n-m}=n-m$ converges in distribution. Moreover from Lemma 4.1 there is a $\delta>0$ such that

$$\mathbf{P}^+\bigg(\mathrm{e}^{-S_a}\sum_{1\leq i\leq a}\tilde{Z}^i_a<\delta\bigg)<\varepsilon,$$

if a is sufficiently large. Then from Lemma 2.7

$$\mathbf{P}(\beta_{a,n}<\delta|\tau_{n-m}=n-m)<\varepsilon,$$

if n is sufficiently large. Therefore the limiting distribution of $\mathrm{e}^{S_{\tau_{nt}}-S_{nt}}\tilde{Z}_{nt}$ conditioned on $\tau_{n-m}=n-m$ has no atom in zero. An application of Theorem 4.2 now gives the claim for $r=1$.

Finally for $r>1$ we let

$$\beta_{a,n,i}=\mathrm{e}^{S_{\sigma_{i,n}}-S_{\sigma_{i,n}+a}}\hat{Z}_{a,\sigma_{i,n}+a},\quad i=1,\dots,r.$$

From (5.1) and our considerations above we know that for every $i\leq r$

$$\mathbf{P}\big(\big|\mathrm{e}^{S_{\tau_{nt_i}}-S_{nt_i}}\tilde{Z}_{nt_i}-\beta_{a,n,\mu_n(i)}\big|\geq\varepsilon\text{ for some }i\leq r|\tau_{n-m}=n-m\big)\leq\frac{\varepsilon}{r}$$

and the rest of the theorem follows by means of Lemma 2.7 and Theorem 4.2.

References

[1] V. I. Afanasyev. Limit theorems for a conditional random walk and some applications. Diss. Cand. Sci., MSU, Moscow, 1980.

[2] V. I. Afanasyev. Limit theorems for an intermediately subcritical and a strongly subcritical branching process in a random environment. *Discrete Math. Appl.* **11** (2001) 105–131.

[3] V. I. Afanasyev, Ch. Böinghoff, G. Kersting and V. A. Vatutin. Limit theorems for weakly subcritical branching processes in random environment. *J. Theoret. Probab.* **25** (2012) 703–732. MR2956209

[4] V. I. Afanasyev, J. Geiger, G. Kersting and V. A. Vatutin. Criticality for branching processes in random environment. *Ann. Probab.* **33** (2005) 645–673. MR2123206

[5] V. I. Afanasyev, J. Geiger, G. Kersting and V. A. Vatutin. Functional limit theorems for strongly subcritical branching processes in random environment. *Stochastic Process. Appl.* **115** (2005) 1658–1676. MR2165338

[6] A. Agresti. On the extinction times of varying and random environment branching processes. *J. Appl. Probab.* **12** (1975) 39–46. MR0365733

[7] K. B. Athreya and S. Karlin. On branching processes with random environments: I, II. *Ann. Math. Stat.* **42** (1971) 1499–1520, 1843–1858. MR0298781, MR0298780

[8] V. Bansaye and Ch. Böinghoff. Upper large deviations for branching processes in random environment with heavy tails. *Electron. J. Probab.* **16** (2011) 1900–1933. MR2851050

[9] J. Bertoin. *Lévy Processes*. Cambridge Univ. Press, Cambridge, 1996. MR1406564

[10] J. Bertoin and R. A. Doney. On conditioning a random walk to stay non-negative. *Ann. Probab.* **22** (1994) 2152–2167. MR1331218

[11] N. H. Bingham, C. M. Goldie and J. L. Teugels. *Regular Variation*. Cambridge Univ. Press, Cambridge, 1987. MR0898871

[12] M. Birkner, J. Geiger and G. Kersting. Branching processes in random environment – A view on critical and subcritical cases. In *Proceedings of the DFG-Schwerpunktprogramm Interacting Stochastic Systems of High Complexity* 269–291. Springer, Berlin, 2005. MR2118578

[13] Ch. Böinghoff and G. Kersting. Upper large deviations of branching processes in a random environment – Offspring distributions with geometrically bounded tails. *Stochastic Process. Appl.* **120** (2010) 2064–2077. MR2673988

[14] Ch. Böinghoff and G. Kersting. Simulations and a conditional limit theorem for intermediately subcritcal branching processes in random environment. Preprint, 2012. Available at arXiv:1209.1274.

[15] F. Caravenna and L. Chaumont. Invariance principles for random walks conditioned to stay positive. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** (2008) 170–190. MR2451576

[16] L. Chaumont. Excursion normalisée, méandre et pont pour les processus de Lévy stables. *Bull. Sci. Math.* **121** (1997) 377–403. MR1465814

[17] B. Chauvin, A. Rouault and A. Wakolbinger. Growing conditioned trees. *Stochastic Process. Appl.* **39** (1991) 117–130. MR1135089

[18] F. M. Dekking. On the survival probability of a branching process in a finite state i.i.d. environment. *Stochastic Process. Appl.* **27** (1988) 151–157.

[19] R. A. Doney. Conditional limit theorems for asymptotically stable random walks. *Z. Wahrsch. verw. Gebiete* **70** (1985) 351–360. MR0803677

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

[21] J. Geiger, G. Kersting and V. A. Vatutin. Limit theorems for subcritical branching processes in random environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **39** (2003) 593–620. MR1983172

[22] Y. Guivarc'h and Q. Liu. Propriétés asymptotiques des processus de branchement en environnement aléatoire. *C. R. Math. Acad. Sci. Paris Sér. I* **332** (2001) 339–344. MR1821473

[23] O. Kallenberg. Stability of critical cluster fields. *Math. Nachr.* **77** (1977) 7–43. MR0443078

[24] M. V. Kozlov. On large deviations of branching processes in a random environment: Geometric distribution of descendants. *Discrete Math. Appl.* **16** (2006) 155–174.

[25] M. V. Kozlov. On large deviations of strictly subcritical branching processes in a random environment with geometric distribution of progeny. *Theory Probab. Appl.* **54** (2010) 424–446.

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

[27] J. Neveu. Erasing a branching tree. *Adv. Appl. Probab.* **18** (1986) 101–108. MR0868511

[28] W. L. Smith and W. E. Wilkinson. On branching processes in random environments. *Ann. Math. Stat.* **40** (1969) 814–827. MR0246380

[29] H. Tanaka. Time reversal of random walks in one dimension. *Tokyo J. Math.* **12** (1989) 159–174. MR1001739

[30] V. A. Vatutin. A limit theorem for an intermediate subcritical branching process in a random environment. *Theory Probab. Appl.* **48** (2004) 481–492.

[31] V. A. Vatutin and E. E. Dyakonova. Galton–Watson branching processes in random environment. I: Limit theorems. *Theory Probab. Appl.* **48** (2004) 314–336.

[32] V. A. Vatutin and E. E. Dyakonova. Galton–Watson branching processes in random environment. II: Finite-dimensional distributions. *Theory Probab. Appl.* **49** (2005) 275–308.

[33] V. A. Vatutin and E. E. Dyakonova. Branching processes in random environment and the bottlenecks in the evolution of populations. *Theory Probab. Appl.* **51** (2007) 189–210.