J. Appl. Prob. **49,** 1134–1143 (2012)
Printed in England

UPPER DEVIATIONS FOR SPLIT TIMES OF BRANCHING PROCESSES

HAMED AMINI,* ** AND
MARC LELARGE,* *** *École Normale Supérieure and INRIA*

Abstract

Upper deviation results are obtained for the split time of a supercritical continuous-time Markov branching process. More precisely, we establish the existence of logarithmic limits for the likelihood that the split times of the process are greater than an identified value and determine an expression for the limiting quantity. We also give an estimation for the lower deviation probability of the split times, which shows that the scaling is completely different from the upper deviations.

Keywords: Continuous-time branching process; split time; large deviation

2010 Mathematics Subject Classification: Primary 60J80
Secondary 60F10

1. Introduction

We consider a one-dimensional continuous-time Markov branching process $\{Z(t); t \geq 0\}$ with infinitesimal generating function $f(s) - s$, where $f(s) = \sum_{i\geq 0} p_i s^i$ with $p_i \geq 0$ for all $i \geq 0$ and $\sum_{i\geq 0} p_i = 1$ (see, e.g. [3]). We recall the construction of the process $\{Z(t)\}$ and introduce some required notation. Let $\{\xi_i,\ i \geq 1\}$ be a sequence of independent and identically distributed (i.i.d.) random variables with generating function f, and let $\eta_i = \xi_i - 1$. We define $S_0^k = Z(0) = k$ for $k \geq 1$, and $S_n^k = k + \sum_{i=1}^n \eta_i$ for $n \geq 1$. Let $I := \inf\{n \in \mathbb{N};\ S_n^k = 0\}$. If $S_n^k \neq 0$ for all $n \in \mathbb{N}$ then $I = \infty$. Given the sequence $\{\xi_i\}$, let $\tau_1^k, \dots, \tau_I^k$ be mutually independent exponential random variables with means

$$\mathrm{E}[\tau_j^k \mid \{\xi_i\}] = \frac{1}{S_{j-1}^k}.$$

We define the sequence of split times by $T_0^k = 0$ and $T_n^k = \tau_1^k + \cdots + \tau_n^k$ for $1 \leq n \leq I$, and let

$$Z(t) = \begin{cases} S_{n-1}^k & \text{for } T_{n-1}^k \leq t < T_n^k,\ 1 \leq n \leq I, \\ 0 & \text{for } T_I^k \leq t. \end{cases}$$

On the event $I < \infty$, we employ the convention that $T_n^k = +\infty$ for any $n \geq I + 1$. This event corresponds to the extinction of the branching process. Let λ denote its probability, which is the smallest root in $[0, 1]$ to $s = f(s)$. We also define $\lambda_* = f'(\lambda)$.

We can now state our main theorem.

Received 3 January 2012; revision received 18 May 2012.
* Postal address: INRIA-ENS Project TREC, École Normale Supérieure, 45 rue d'Ulm, 75005 Paris, France.
** Email address: hamed.amini@ens.fr
*** Email address: marc.lelarge@ens.fr

https://doi.org/10.1239/jap/1354716662 Published online by Cambridge University Press

Theorem 1. *In the above setting with $1 < f'(1) < \infty$, we define $\xi_{\min} = \min\{i,\ p_i > 0\}$. For any $x > 0$ and $k \geq 1$, we have*

$$\lim_{n\to\infty} \frac{1}{\log n} \log \mathrm{P}\left(\infty > T_n^k \geq \left(x + \frac{1}{f'(1)-1}\right) \log n\right) = -x g(\xi_{\min}, k), \tag{1}$$

with

$$g(\xi_{\min}, k) = k\,\mathbf{1}(\xi_{\min} \geq 2) + k(1-p_1)\,\mathbf{1}(\xi_{\min} = 1) + (1-\lambda_*)\,\mathbf{1}(\xi_{\min} = 0).$$

Although split times have been studied in [2] (in the case where $\xi_{\min} \geq 2$) and very precise large deviations are known for supercritical Galton–Watson processes [4], [5], [8], [10], [11], to the best of our knowledge, (1) cannot be obtained directly from results in the literature and so we provide a complete proof in the next section.

To gain some insights about (1), we interpret the process $Z(t)$ as a population process describing living particles at time t. Note that if $\xi_{\min} \geq 2$ then T_n^k is exactly the first time when the population increased by an amount n, starting with an initial population of k. Thanks to [2], the growth rate of T_n^k is of the order $\log n/(f'(1)-1)$. Hence, the upper deviation for T_n^k analysed in (1) corresponds to an event where the population remains small for a large amount of time. Since each initial particle gives rise to a population which cannot become extinct, each of these subpopulations has to remain small and the linearity of g in k is then easy to understand. When $\xi_{\min} = 1$, note that some split times are not real split times for the population since a particle can replicate itself at a split. An easy change-of-time argument allows us to reduce this case to the previous case and introduces a factor $1 - p_1$. Finally, when $\xi_{\min} = 0$, the situation is radically different as extinction is possible. In this case, the function g does not depend on k. Indeed, the typical event for a population starting with k particles to remain small but positive for a long time is that $k-1$ of these particles become extinct while only one particle replicates itself for a very long time. This can be seen by the following argument giving an interpretation of λ^*.

Let $\mathcal{X}_p$ be a Galton–Watson tree with offspring distribution $p = (p_i)_{i\geq 0}$, i.e. a random tree where the root and all successive descendants have a random number of children independent from the rest and with distribution p. Let $\mathcal{X}_p^+ \subseteq \mathcal{X}_p$ be the set of particles of $\mathcal{X}_p$ that survive, i.e. have descendants in all future generations. Then $\mathcal{X}_p^+$ contains the root of the original branching process with probability $1-\lambda$, and is empty otherwise. We denote by ξ the random variable with distribution $\mathrm{P}(\xi = k) = p_k$. For $\eta \in [0, 1]$, we let ξ_η be the *thinning* of ξ obtained by taking ξ points and then randomly and independently keeping each of them with probability η, i.e.

$$\mathrm{P}(\xi_\eta = k) = \sum_{r=k}^{\infty} p_r \binom{r}{k} \eta^k (1-\eta)^{r-k}.$$

Note that the number of surviving children has the distribution $\xi_{1-\lambda}$. Let ξ^+ denote the offspring distribution in $\mathcal{X}_p^+$. Conditioning on a particle being in $\mathcal{X}_p^+$ is exactly the same as conditioning on at least one of its children surviving, so that

$$\mathrm{P}(\xi^+ = 1) = \mathrm{P}(\xi_{1-\lambda} = 1 \mid \xi_{1-\lambda} \geq 1) = \frac{\sum_{k\geq 0} k p_k \lambda (1-\lambda)^{k-1}}{\lambda} = f'(\lambda) = \lambda_*.$$

Hence, the factor $1 - \lambda_*$ in (1) is obtained by the same kind of change of time that led to the factor $1 - p_1$ in the case $\xi_{\min} = 1$.

Theorem 1 will be used to establish the asymptotic of the diameter in random graphs with exponential edge weights (see [6] for regular graphs corresponding to $p_r = 1$ for some $r \geq 2$), which is the subject of our forthcoming paper [1]. We think that it is of independent interest as it gives some insights into how slow the growth of a continuous-time branching process can be.

Remark 1. We describe here how to heuristically derive our main result using known properties of continuous-time Markov branching processes. It is well known (see, e.g. [3, Theorem 1, p. 111]) that $Z(t)\mathrm{e}^{-(f'(1)-1)t}$ has an almost-sure limit W. The limiting random variable W has an atom at 0 of size λ. Furthermore, the random variable W, conditioned to be positive, admits a continuous density on $\mathbb{R}^+$, and we define $W_\lambda \stackrel{\mathrm{D}}{=} (W \mid W > 0)$. The random variable W_λ is given by (see, e.g. [3, Theorem 3, p. 116], [9], and [12])

$$W_\lambda = \sum_{i=1}^{k} \tilde{W}_i \mathrm{e}^{-(f'(1)-1)E_i},$$

where the E_is are i.i.d. Exp(1), independent of the $\tilde{W}_i$ which are i.i.d. with Laplace transform $\tilde{\phi}(t) = \mathrm{E}[\mathrm{e}^{-t\tilde{W}}]$ whose inverse function is given by

$$\tilde{\phi}^{-1}(x) = (1-x)\exp\biggl(\int_1^x \biggl(\frac{f'(1)-1}{f(s)-s} + \frac{1}{1-s}\biggr)\,\mathrm{d}s\biggr), \qquad \lambda < x \leq 1. \tag{2}$$

Note that

$$\begin{aligned}
\mathrm{P}&\biggl(\infty > T_n^k \geq \biggl(x + \frac{1}{f'(1)-1}\biggr)\log n\biggr) \\
&= \mathrm{P}\biggl(Z\biggl(\biggl(x + \frac{1}{f'(1)-1}\biggr)\log n\biggr) < n,\ I < \infty\biggr) \\
&= \mathrm{P}\biggl(\exp\biggl(-(f'(1)-1)\biggl(x + \frac{1}{f'(1)-1}\biggr)\log n\biggr) \\
&\qquad \times Z\biggl(\biggl(x + \frac{1}{f'(1)-1}\biggr)\log n\biggr) < n^{-x(f'(1)-1)},\ I < \infty\biggr).
\end{aligned}$$

Since the left-hand term in the above probability converges to W as n goes to ∞, it is reasonable to analyze $\mathrm{P}(W_\lambda < n^{-x(f'(1)-1)})$. Note that

$$\psi(s) := \mathrm{E}[\mathrm{e}^{-sW_\lambda}] = (\mathrm{E}[\exp(-s\tilde{W}\mathrm{e}^{-(f'(1)-1)E})])^k,$$

where E is an Exp(1) random variable and $\psi(\cdot)$ is the Laplace transform of W_λ. Moreover,

$$\mathrm{E}[\exp(-s\tilde{W}\mathrm{e}^{-(f'(1)-1)E})] = \int_0^\infty \mathrm{e}^{-x}\tilde{\phi}(s\mathrm{e}^{-(f'(1)-1)x})\,\mathrm{d}x.$$

Combining (2) together with the Tauberian theorem [7, Section XIII.5], it is easy to infer that

$$\mathrm{P}(W_\lambda < n^{-x(f'(1)-1)}) \approx n^{-xg(\xi_{\min},k)}.$$

In the next section we present the proof of Theorem 1. For completeness, in Section 3 we give an estimation for the lower deviation probability of the split times which shows that the scaling is completely different from the upper deviations.

2. Proof of Theorem 1

Let us define $\alpha_n := \lfloor \log^3 n \rfloor$. In the sequel, we will use the following property of the exponential random variables, sometimes without mention. If Y is an exponential random variable of rate γ then, for any $\theta < \gamma$, we have $\mathrm{E}[\mathrm{e}^{\theta Y}] = \gamma/(\gamma - \theta)$.

We first assume that $p_0 = p_1 = 0$, so $\xi_{\min} \geq 2$. Thus, $\eta_i \geq 1$ for all i, so S_n^k is increasing in n and $I = \infty$. In this case, all the elements of the sequence $\{\tau_n^k\}$ are finite and this sequence has been studied in [2].

We now prove the upper bound in (1) for this specific case. Note that, since $S_{i-1}^k \geq k + i - 1$ for all $i \geq 1$, we have, for any $\theta < k + 1$,

$$\mathrm{E}[\exp(\theta \tau_i^k)] = \mathrm{E}\bigg[\frac{S_{i-1}^k}{S_{i-1}^k - \theta}\bigg] \leq \frac{k+i-1}{k+i-1-\theta}. \tag{3}$$

We have, for any $\varepsilon > 0$,

$$\begin{aligned} &\mathrm{P}\bigg(T_n^k \geq \bigg(x + \frac{1}{f'(1)-1}\bigg)\log n\bigg) \\ &\quad \leq \mathrm{P}\bigg(\sum_{i \leq \alpha_n} \tau_i^k \geq x(\varepsilon)\log n\bigg) + \mathrm{P}\bigg(\sum_{i=\alpha_n+1}^{n} \tau_i^k \geq \frac{1+\varepsilon}{f'(1)-1}\log n\bigg), \end{aligned} \tag{4}$$

where $x(\varepsilon) = x - \varepsilon/(f'(1) - 1)$.

For the first term, we have

$$\mathrm{P}\bigg(\sum_{i \leq \alpha_n} \tau_i^k \geq x(\varepsilon)\log n\bigg) = \int_0^{x(\varepsilon)\log n} \mathrm{P}\bigg(\sum_{2 \leq i \leq \alpha_n} \tau_i^k \geq x(\varepsilon)\log n - y\bigg) k\mathrm{e}^{-ky}\,\mathrm{d}y,$$

since τ_1^k is an exponential random variable with mean k independent of the τ_i^ks with $i \geq 2$. We need to bound the right-hand term and we proceed as follows:

$$\begin{aligned} \mathrm{P}\bigg(\sum_{2 \leq i \leq \alpha_n} \tau_i^k \geq x(\varepsilon)\log n - y\bigg) &\leq n^{-kx(\varepsilon)}\mathrm{e}^{ky}\,\mathrm{E}\bigg[\exp\bigg(k\sum_{2 \leq i \leq \alpha_n} \tau_i^k\bigg)\bigg] \quad \text{(by Chernoff's bound)} \\ &\leq n^{-kx(\varepsilon)}\mathrm{e}^{ky}\prod_{i=2}^{\alpha_n}\frac{k+i-1}{i-1} \quad \text{(by (3))} \\ &\leq n^{-kx(\varepsilon)}\mathrm{e}^{ky}\exp\bigg(\sum_{i=2}^{\alpha_n}\frac{k}{i-1}\bigg) \quad (\text{since } 1 + x \leq \mathrm{e}^x) \\ &< n^{-kx(\varepsilon)}\mathrm{e}^{ky}\alpha_n^k \log n \quad \text{for sufficiently large } n. \end{aligned}$$

Hence, we obtain (for sufficiently large n)

$$\mathrm{P}\bigg(\sum_{i \leq \alpha_n} \tau_i^k \geq x(\varepsilon)\log n\bigg) < kx(\varepsilon)n^{-kx(\varepsilon)}\alpha_n^k \log^2 n < n^{-kx(\varepsilon)}\alpha_n^{k+1}.$$

We now give an upper bound for the second term in (4). We first recall a basic result of probability: for any $\varepsilon > 0$, there is a constant $\gamma > 0$ such that, for large enough n, we have

$$\mathrm{P}\bigg(S_n^k \leq k + n\frac{f'(1)-1}{1+\varepsilon}\bigg) \leq \mathrm{e}^{-\gamma n}.$$

We define the event $\mathcal{E}_n = \{S_i^k \geq k + i(f'(1) - 1)/(1 + \varepsilon/3),\ \alpha_n \leq i \leq n\}$, so that (by the union bound) we have $\mathrm{P}(\mathcal{E}_n) \geq 1 - o(n^{-\log n})$. Using the fact that $\sqrt{\alpha_n} = o(\alpha_n)$, we have, for sufficiently large n,

$$\begin{aligned}\mathrm{E}\bigg[\exp\bigg(\sqrt{\alpha_n}\sum_{i=\alpha_n+1}^{n}\tau_i^k\bigg)\ \bigg|\ \mathcal{E}_n\bigg] &\leq \prod_{i=\alpha_n}^{n-1}\bigg(1 + \frac{\sqrt{\alpha_n}}{k + i(f'(1)-1)/(1+\varepsilon/3) - \sqrt{\alpha_n}}\bigg)\\ &\leq \prod_{i=\alpha_n}^{n-1}\bigg(1 + \frac{\sqrt{\alpha_n}(1+\varepsilon/2)}{i(f'(1)-1)}\bigg)\\ &\leq \exp\bigg(\frac{\sqrt{\alpha_n}(1+\varepsilon/2)}{f'(1)-1}\log n\bigg).\end{aligned}$$

Now we have (by Markov's inequality)

$$\begin{aligned}&\mathrm{P}\bigg(\sum_{i=\alpha_n+1}^{n}\tau_i^k \geq \frac{1+\varepsilon}{f'(1)-1}\log n\ \bigg|\ \mathcal{E}_n\bigg)\\ &\quad\leq \mathrm{E}\bigg[\exp\bigg(\sqrt{\alpha_n}\sum_{i=\alpha_n+1}^{n}\tau_i^k\bigg)\ \bigg|\ \mathcal{E}_n\bigg]\exp\bigg(\frac{-\sqrt{\alpha_n}(1+\varepsilon)}{f'(1)-1}\log n\bigg)\\ &\quad\leq \exp\bigg(-\frac{\sqrt{\alpha_n}\varepsilon}{2(f'(1)-1)}\log n\bigg)\\ &\quad\leq o(n^{-\log n}).\end{aligned}$$

Hence, we obtain

$$\mathrm{P}\bigg(\sum_{i=\alpha_n+1}^{n}\tau_i^k \geq \frac{1+\varepsilon}{f'(1)-1}\log n\bigg) \leq 1 - \mathrm{P}(\mathcal{E}_n) + o(n^{-\log n}) = o(n^{-\log n}). \tag{5}$$

Note that in order to get (5), we only used the fact that $p_0 = p_1 = 0$ to ensure that $T_n^k < \infty$ for all n. In particular, the argument is still valid if $p_0 = 0$.

To summarize in the case $p_0 = p_1 = 0$, we obtain, for any $\varepsilon > 0$ and with $x(\varepsilon) = x - \varepsilon/(f'(1)-1)$,

$$\mathrm{P}\bigg(T_n^k \geq \bigg(x + \frac{1}{f'(1)-1}\bigg)\log n\bigg) \leq n^{-kx(\varepsilon)}\alpha_n^{k+1} + o(n^{-\log n}),$$

and the upper bound for (1) follows.

We now prove a lower bound for (1). We start with (for any $\varepsilon > 0$)

$$\mathrm{P}\bigg(T_n^k \geq \bigg(x + \frac{1}{f'(1)-1}\bigg)\log n\bigg) \geq \mathrm{P}(\tau_1^k \geq \tilde{x}(\varepsilon)\log n)\,\mathrm{P}\bigg(\sum_{i=2}^{n}\tau_i^k \geq \frac{1-\varepsilon}{f'(1)-1}\log n\bigg),$$

where $\tilde{x}(\varepsilon) = x + \varepsilon/(f'(1)-1)$. Since $\mathrm{P}(\tau_1^k \geq \tilde{x}(\varepsilon)\log n) = n^{-k\tilde{x}(\varepsilon)}$, we need to show that

$$\liminf_{n\to\infty}\mathrm{P}\bigg(\sum_{i=2}^{n}\tau_i^k \geq \frac{1-\varepsilon}{f'(1)-1}\log n\bigg) > 0.$$

This follows from the almost-sure convergence of $(1/\log n)\sum_{i=2}^{n}\tau_i^k$ to $(f'(1)-1)^{-1}$ (see, e.g. Corollary 1 in Section 2 of [2]).

We now consider the case where $p_1 > 0$ while $p_0 = 0$. We start with the upper bound and decomposition (4). Note that $f'(1) > 1$ implies that $p_1 < 1$. Let the $\tilde{\tau}_i^k$ be the *real* split times of the process $Z(t)$, i.e. the times where $Z(t)$ increases. Let N be a geometric random variable with parameter p_1, i.e. $\mathrm{P}(N = j) = p_1^j(1-p_1)$, independent of everything else. Then $\tilde{\tau}_1^k$ has the same law as a sum of N mutually independent exponential random variables with mean k^{-1}. Hence, it is distributed as an exponential random variable with mean $1/k(1-p_1)$. For the first term, note that $\tilde{\tau}_i^k \geq \tau_i^k$, so we obtain

$$\mathrm{P}\biggl(\sum_{i\leq\alpha_n}\tau_i^k \geq x(\varepsilon)\log n\biggr) \leq \mathrm{P}\biggl(\sum_{i\leq\alpha_n}\tilde{\tau}_i^k \geq x(\varepsilon)\log n\biggr).$$

It was shown in [3, Section III.9] that the branching process $\tilde{Z}(t)$ associated to the split times $\tilde{\tau}_i^k$ is still a Markov branching process with the same infinitesimal generating function but with $\tilde{p}_1 = 0$ and with

$$\mathrm{E}[\tilde{\tau}_j^k \mid \{\tilde{\xi}_i\}] = \frac{1}{(1-p_1)\tilde{S}_{j-1}^k},$$

where $\{\tilde{\xi}_i\}$ is a sequence of i.i.d. random variables with generating function

$$\tilde{f}(s) = \frac{f(s) - p_1 s}{1-p_1} = \sum_{j=2}^{\infty}\tilde{p}_j s^j,$$

and with $\tilde{S}_j^k = k + \sum_{i=1}^{j}\tilde{\xi}_i - j$.

Thanks to previous analysis, we therefore obtain

$$\mathrm{P}\biggl(\sum_{i\leq\alpha_n}\tilde{\tau}_i^k \geq x(\varepsilon)\log n\biggr) \leq n^{-k(1-p_1)x(\varepsilon)}\alpha_n^{(1-p_1)k+1}.$$

Since (5) is still valid, the upper bound follows from (4).

The lower bound also follows from a simple adaptation of the argument above:

$$\mathrm{P}\biggl(T_n^k \geq \biggl(x + \frac{1}{f'(1)-1}\biggr)\log n\biggr) \geq \mathrm{P}\biggl(\tilde{\tau}_1^k \geq \tilde{x}(\varepsilon)\log n,\ \sum_{i=N}^{n}\tau_i^k \geq \frac{1-\varepsilon}{f'(1)-1}\log n\biggr).$$

At time $\tilde{\tau}_1^k$ we have $Z(\tilde{\tau}_1^k) = k + \tilde{\eta}$, where the random variable $\tilde{\eta}$ is distributed as η_1 conditioned on being greater than 1. Let j be such that $\mathrm{P}(\tilde{\eta} \leq j) \geq \frac{1}{2}$. The process $\{Z(t);\, t \geq \tilde{\tau}_1^k\}$ has the same law as the original process starting with $k + \tilde{\eta}$ particles and is independent of $\tilde{\tau}_1^k$. Since τ_i^k is stochastically decreasing in k, we have

$$\begin{aligned}
&\mathrm{P}\biggl(T_n^k \geq \biggl(x + \frac{1}{f'(1)-1}\biggr)\log n\biggr)\\
&\quad\geq \mathrm{P}(\tilde{\tau}_1^k \geq \tilde{x}(\varepsilon)\log n,\ N \leq \alpha_n)\,\mathrm{P}\biggl(\tilde{\eta} \leq j,\ \sum_{i=1}^{n-\alpha_n}\tau_i^{k+j} \geq \frac{1-\varepsilon}{f'(1)-1}\log n\biggr)\\
&\quad\geq (n^{-k(1-p_1)\tilde{x}(\varepsilon)} - p_1^{\alpha_n})\frac{1}{2}\,\mathrm{P}\biggl(\sum_{i=1}^{n-\alpha_n}\tau_i^{k+j} \geq \frac{1-\varepsilon}{f'(1)-1}\log n\biggr).
\end{aligned}$$

Since we still have the almost-sure convergence of $(1/\log n)\sum_{i=1}^{n-\alpha_n}\tau_i^{k+j}$ to $(f'(1)-1)^{-1}$, we obtain the lower bound.

We now consider the case where $p_0>0$, so that the probability of extinction $\mathrm{P}(I<\infty)=\lambda$ is positive and strictly less than 1 (because $f'(1)>1$). Following [3], we define

$$\tilde{Z}(t)=\begin{cases}0 & \text{if } I<\infty,\\ \text{the number of particles among } Z_t \text{ which have} & \\ \text{an infinite line of descent} & \text{if } I=\infty.\end{cases}$$

We have $\tilde{Z}(0)=\mathrm{Bin}(k,1-\lambda)$ and, for $1\le j\le k$,

$$\mathrm{P}(\tilde{Z}(0)=j\mid I=\infty)=\binom{k}{j}\frac{(1-\lambda)^j\lambda^{k-j}}{1-\lambda^k}. \tag{6}$$

By Theorem I.12.1 of [3], the process $\{\tilde{Z}(t);t\ge 0\}$ conditioned on the event $I=\infty$ is a Markov branching process with infinitesimal generating function $\tilde{f}(s)-s$, where

$$\tilde{f}(s)=\sum_i\tilde{p}_i s^i=\frac{f((1-\lambda)s+\lambda)-\lambda}{1-\lambda}\quad\text{for } 0\le s\le 1.$$

Clearly, $\tilde{f}(0)=\tilde{p}_0=0$, so this process survives and we define the corresponding split times $\tilde{T}_n=\tilde{\tau}_1+\cdots+\tilde{\tau}_n$ for all $n\ge 1$ as we did for the original process $Z(t)$ (but now with a random number of initial particles given by (6)). Note that we have $\tilde{f}'(1)=f'(1)$ and $\tilde{p}_1=f'(\lambda)=\lambda_*\in(0,1)$.

On the event $I=\infty$, we clearly have $\tilde{\tau}_n\ge\tau_n$ for all $n\ge 1$; hence, thanks to the previous analysis, we have, for any $1\le j\le k$,

$$\limsup_{n\to\infty}\frac{1}{\log n}\log\mathrm{P}\biggl(T_n^k\ge\biggl(x+\frac{1}{f'(1)-1}\biggr)\log n\biggm| I=\infty,\ \tilde{Z}(0)=j\biggr)\le -xj(1-\lambda_*),$$

and the upper bound follows in the case $I=\infty$. We now consider the case $I<\infty$. First note that $T_n^k\le T_n^1$; hence, we need to consider only the case $k=1$. It follows from Theorem I.12.3 of [3] that, conditioned on the event $I<\infty$, the branching process has the same law as a Markov branching process $\bar{Z}(t)$ with infinitesimal generating function $\bar{f}(s)-s$, where $\bar{f}(s)=\lambda^{-1}f(\lambda)$. The total progeny is finite and $\bar{f}'(1)=f'(\lambda)=\lambda_*<1$. Moreover, we have

$$\begin{aligned}\mathrm{P}\biggl(T_n^1\ge\biggl(x+\frac{1}{f'(1)-1}\biggr)\log n\biggm| I<\infty\biggr)&\le\mathrm{P}\biggl(\bar{Z}\biggl(\biggl(x+\frac{1}{f'(1)-1}\biggr)\log n\biggr)>0\biggr)\\&\le\mathrm{E}\biggl[\bar{Z}\biggl(\biggl(x+\frac{1}{f'(1)-1}\biggr)\log n\biggr)\biggr]\\&=n^{-x(1-\lambda_*)-(1-\lambda_*)/(f'(1)-1)},\end{aligned}$$

and the upper bound follows.

We now derive a lower bound. First note that, using the Markov property, given $Z(t)=j$ for $j\ge 1$, the random variable $\tilde{Z}(t)$ is distributed as a binomial random variable with j trials and probability of 'success' $1-\lambda$. Let $\{X_i,\ i\ge 1\}$ be a sequence of independent Bernoulli random variables with $\mathrm{E}[X_1] = 1 - \lambda$. To ease notation, let $x_n = (x + 1/(f'(1) - 1)) \log n$ and, for $\varepsilon > 0$, let $\mathcal{E}$ be the event $\mathcal{E} = \{\sum_{i=1}^{m} X_i \geq (1 - \lambda - \varepsilon)m$ for all $m \geq 1\}$. We have

$$\begin{aligned} \mathrm{P}(\infty > T_n^k \geq x_n) &= \mathrm{P}(Z(x_n) \leq n, I = \infty \mid Z(0) = k) \\ &\geq \mathrm{P}\biggl(Z(x_n) \leq n,\ \tilde{Z}(x_n) = \sum_{i=1}^{Z(x_n)} X_i,\ I = \infty, \mathcal{E} \biggm| Z(0) = k\biggr). \end{aligned}$$

On the event $\mathcal{E}$, we have $\sum_{i=1}^{Z(x_n)} X_i \geq (1 - \lambda - \varepsilon) Z(x_n)$. Thus,

$$\begin{aligned} &\mathrm{P}(\infty > T_n^k \geq x_n) \\ &\quad\geq \mathrm{P}\biggl(\tilde{Z}(x_n) \leq \frac{n}{1 - \lambda - \varepsilon},\ I = \infty, \mathcal{E} \biggm| Z(0) = k\biggr) \\ &\quad\geq \mathrm{P}\biggl(\tilde{Z}(x_n) \leq \frac{n}{1 - \lambda - \varepsilon} \biggm| \tilde{Z}(0) = 1\biggr) \mathrm{P}(\tilde{Z}(0) = 1 \mid Z(0) = k)\, \mathrm{P}(\mathcal{E} \mid Z(0) = k). \end{aligned}$$

We have $\mathrm{P}(\tilde{Z}(0) = 1 \mid Z(0) = k)\, \mathrm{P}(\mathcal{E} \mid Z(0) = k) > 0$, and the process $\{\tilde{Z}(t); t \geq 0\}$ conditioned on the event $\tilde{Z}(0) = 1$ is a Markov branching process with infinitesimal generating function $\tilde{f}(s) - s$ with $\tilde{p}_1 = \lambda_*$. Hence, the lower bound follows from the previous analysis.

3. On the lower deviation probability

In this section we consider the lower deviation probability for the split time of a branching process.

Proposition 1. *In the setting outlined in the introduction with* $1 < f'(1) < \infty$*, for any* $x > 0$*,* $k \geq 1$*, and a sufficiently large constant* C*, we have*

$$\mathrm{P}\biggl(T_n^k < \biggl(\frac{1 - x}{f'(1) - 1}\biggr) \log n\biggr) = o(n^C \mathrm{e}^{-n^x}).$$

The proof relies on the following lemma.

Lemma 1. *Let* $X_1, \ldots, X_t$ *be a random process adapted to a filtration* $\mathcal{F}_0 = \sigma[\varnothing], \mathcal{F}_1, \ldots, \mathcal{F}_t$*, and let* $\mu_i = \mathrm{E}\, X_i$*,* $\Sigma_i = X_1 + \cdots + X_i$*, and* $\Lambda_i = \mu_1 + \cdots + \mu_i$*. Let* $Y_i \sim \mathrm{Exp}(\Sigma_i)$ *and* $Z_i \sim \mathrm{Exp}(\Lambda_i)$*, where all exponential variables are independent. Then we have*

$$Y_1 + \cdots + Y_t \geq_{\mathrm{st}} Z_1 + \cdots + Z_t.$$

Proof. By Jensen's inequality, it is easy to see that, for a positive random variable X, we have

$$\mathrm{Exp}(X) \geq_{\mathrm{st}} \mathrm{Exp}(\mathrm{E}\, X).$$

Then, by induction, it suffices to prove that, for a pair of random variables X_1 and X_2, we have $Y_1 + Y_2 \geq_{\mathrm{st}} Z_1 + Z_2$. We have

$$\begin{aligned} \mathrm{P}(Y_1 + Y_2 > s) &= \mathrm{E}_{X_1}[\mathrm{P}(Y_1 + Y_2 > s \mid X_1)] \\ &\geq \mathrm{E}_{X_1}[\mathrm{P}(\mathrm{Exp}(X_1) + \mathrm{Exp}(X_1 + \mu_2) > s)] \\ &\geq \mathrm{P}(Z_1 + Z_2 > s). \end{aligned}$$

https://doi.org/10.1239/jap/1354716662 Published online by Cambridge University Press

We infer, by Lemma 1,

$$T_n^k \geq_{\mathrm{st}} \sum_{i=1}^{n} \mathrm{Exp}(k + (f'(1) - 1)i),$$

where all exponential variables are independent. Thus, we have

$$\begin{aligned}
\mathrm{P}(T_n^k \leq t) &\leq \int_{\sum x_i \leq t} \exp\biggl(-\sum_{i=1}^{n}((f'(1)-1)i + k)x_i\biggr)\,\mathrm{d}x_1 \cdots \mathrm{d}x_n \prod_{i=1}^{n}((f'(1)-1)i + k) \\
&= \int_{0 \leq y_1 \leq \cdots \leq y_n \leq t} \exp\biggl(-(f'(1)-1)\sum_{i=1}^{n} y_i\biggr)\mathrm{e}^{-ky_n}\,\mathrm{d}y_1 \cdots dy_n \\
&\quad \times \prod_{i=1}^{n}((f'(1)-1)i + k),
\end{aligned}$$

where $y_k = \sum_{i=0}^{k-1} x_{n-i}$. Letting y play the role of y_n, and accounting for all permutations over $y_1, \ldots, y_{n-1}$ (giving each such variable the range $[0, y]$), we obtain

$$\begin{aligned}
\mathrm{P}(T_n^k \leq t) &\leq (f'(1)-1)^n \frac{\prod_{i=1}^{n}(i + k/(f'(1)-1))}{(n-1)!} \\
&\quad \times \int_0^t \mathrm{e}^{-(f'(1)-1+k)y}\biggl(\int_{[0,y]^{n-1}} \exp\biggl(-(f'(1)-1)\sum_{i=1}^{n-1} y_i\biggr)\mathrm{d}y_1 \cdots \mathrm{d}y_{n-1}\biggr)\mathrm{d}y \\
&= n\frac{\prod_{i=1}^{n}(i + k/(f'(1)-1))}{n!}(f'(1)-1) \\
&\quad \times \int_0^t \mathrm{e}^{-(f'(1)-1+k)y}\biggl(\prod_{i=1}^{n-1}\int_0^y (f'(1)-1)\mathrm{e}^{-(f'(1)-1)y_i}\,\mathrm{d}y_i\biggr)\mathrm{d}y \\
&= n\prod_{i=1}^{n}\biggl(1 + \frac{k}{(f'(1)-1)i}\biggr)(f'(1)-1) \\
&\quad \times \int_0^t \mathrm{e}^{-(f'(1)-1+k)y}(1 - \mathrm{e}^{-(f'(1)-1)y})^{n-1}\,\mathrm{d}y \\
&\leq cn^{k/(f'(1)-1)+1}(f'(1)-1)\int_0^t \mathrm{e}^{-(f'(1)-1+k)y}(1 - \mathrm{e}^{-(f'(1)-1)y})^{n-1}\,\mathrm{d}y,
\end{aligned}$$

where $c > 0$ is an absolute constant. Now we use the fact that

$$(1 - \mathrm{e}^{-(f'(1)-1)y})^{n-1} \leq \mathrm{e}^{-n^x} \quad \text{for all } 0 \leq y \leq \frac{1-x}{f'(1)-1}\log n =: t_n.$$

We infer that (for $C > k/(f'(1)-1) + 1$)

$$\mathrm{P}\biggl(T_n^k \leq \biggl(\frac{1-x}{f'(1)-1}\biggr)\log n\biggr) \leq c(f'(1)-1)n^{k/(f'(1)-1)+1}\int_0^{t_n} \mathrm{e}^{-n^x}\,\mathrm{d}y = o(n^C \mathrm{e}^{-n^x}).$$

Acknowledgement

The authors thank an anonymous referee for the suggestions which improved the presentation of the paper.

References

[1] Amini, H. and Lelarge, M. (2012). The diameter of weighted random graphs. In preparation.

[2] Athreya, K. B. and Karlin, S. (1967). Limit theorems for the split times of branching processes. *J. Math. Mech.* **17,** 257–277.

[3] Athreya, K. B. and Ney, P. E. (2004). *Branching Processes*. Dover Publications, Mineola, NY.

[4] Biggins, J. D. and Bingham, N. H. (1993). Large deviations in the supercritical branching process. *Adv. Appl. Prob.* **25,** 757–772.

[5] Bingham, N. H. (1988). On the limit of a supercritical branching process. In *A Celebration of Applied Probability* (J. Appl. Prob. Spec. Vol. **25A**), ed. J. Gani, Applied Probability Trust, Sheffield, pp. 215–228.

[6] Ding, J., Kim, J. H., Lubetzky, E. and Peres, Y. (2010). Diameters in supercritical random graphs via first passage percolation. *Combinatorics Prob. Comput.* **19,** 729–751.

[7] Feller, W. (1991). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.

[8] Fleischmann, K. and Wachtel, V. (2007). Lower deviation probabilities for supercritical Galton–Watson processes. *Ann. Inst. H. Poincaré Prob. Statist.* **43,** 233–255.

[9] Harris, T. E. (1951). Some mathematical models for branching processes. In *Proc. 2nd Berkeley Symp. Mathematical Statistics and Probability*, University of California Press, Berkeley, pp. 305–328.

[10] Liu, Q. (1999). Asymptotic properties of supercritical age-dependent branching processes and homogeneous branching random walks. *Stoch. Process. Appl.* **82,** 61–87.

[11] Ney, P. E. and Vidyashankar, A. N. (2004). Local limit theory and large deviations for supercritical branching processes. *Ann. Prob.* **14,** 1135–1166.

[12] Sevast'yanov, B. A. (1951). The theory of branching random processes. *Usephi Mat. Nauk.* **6,** 47–99 (in Russian).