



**HAL**  
open science

## On the cover time of $\lambda$ -biased walk on supercritical Galton–Watson trees

Tianyi Bai

► **To cite this version:**

Tianyi Bai. On the cover time of  $\lambda$ -biased walk on supercritical Galton–Watson trees. *Stochastic Processes and their Applications*, 2020, 130, pp.6863 - 6879. <10.1016/j.spa.2020.07.001>. <hal-03493482>

**HAL Id: hal-03493482**

**<https://hal.science/hal-03493482v1>**

Submitted on 17 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

# On the cover time of $\lambda$ -biased walk on supercritical Galton-Watson trees

Tianyi Bai\*

## Abstract

In this paper, we study the time required for a  $\lambda$ -biased ( $\lambda > 1$ ) walk to visit all the vertices of a supercritical Galton-Watson tree up to generation  $n$ . Inspired by the extremal landscape approach in [8] for the simple random walk on binary trees, we establish the scaling limit of the cover time in the biased setting.

## 1 Introduction

### 1.1 The model

A planar tree  $\mathbf{T}$  is a subset of  $\cup_{n \geq 0} \mathbb{N}_+^n$  such that

- The root  $\emptyset$  is in  $\mathbf{T}$ , where by convention,  $\mathbb{N}_+^0 = \{\emptyset\}$ .
- For every vertex  $x = (x_1, \dots, x_n) \in \mathbf{T}$ , its parent

$$\overleftarrow{x} = (x_1, \dots, x_{n-1}) \in \mathbf{T}.$$

- There exists an integer  $\nu_x(\mathbf{T}) \geq 0$  representing the number of children of  $x$ , i.e., for every  $j \in \mathbb{N}_+$ ,

$$(x_1, \dots, x_n, j) \in \mathbf{T} \text{ if and only if } 1 \leq j \leq \nu_x(\mathbf{T}).$$

For  $x \in \mathbf{T}$ , denote by  $|x| = n$  its height. For two vertices  $x, y \in \mathbf{T}$ , write  $x \preceq y$  if  $x$  is on the simple path from  $\emptyset$  to  $y$ . Denote by  $z = x \wedge y$  the common ancestor, i.e. the vertex with maximum height  $|z|$  such that  $z \preceq x, y$ . Denote

---

\*Laboratoire Analyse Géométrie et Applications, UMR 7539, CNRS, Université Paris 13 - Sorbonne Paris Cité, Université Paris 8, 99, Avenue Jean-Baptiste Clément, 93430 Villetaneuse. Email: bai@math.univ-paris13.fr

the tree  $\mathbf{T}$  chopped at height  $n$  by  $\mathbf{T}_n = \{x \in \mathbf{T}, |x| \leq n\}$  (in  $\mathbf{T}_n$ ,  $\nu_x = 0$  if  $|x| = n$ ), and the population in generation  $n$  by  $Z_n = \sum_{x \in \mathbf{T}_n} \mathbf{1}_{|x|=n}$ . Denote the subtree of  $\mathbf{T}$  rooted at  $x$  by  $\mathbf{T}^x = \{y \in \mathbf{T}, x \preceq y\}$ , and the population in the  $n$ -th generation of  $\mathbf{T}^x$  by  $Z_n^x = \sum_{y \in \mathbf{T}^x} \mathbf{1}_{|y|=n}$ . For convenience of further usage (local time related calculations), an artificial root  $\overleftarrow{\emptyset}$  with height  $-1$  is added to be the parent of  $\emptyset$ .

For any fixed probability distribution  $\mu = (\mu_n)_{n \in \mathbb{N}}$ , the Galton-Watson tree with offspring distribution  $\mu$  is a measure  $\mathbf{P}_{\text{GW}}$  on the set of planar trees, such that all the vertices have children distributed identically and independently as  $\mu$ , i.e.

$$\nu_x \stackrel{iid}{\sim} \mu, \forall x \neq \overleftarrow{\emptyset}.$$

The expectation under this probability measure is denoted by  $\mathbf{E}_{\text{GW}}$ .

Let  $m = \sum_{i=0}^{\infty} i\mu_i$  denote the average number of children for a vertex (except  $\overleftarrow{\emptyset}$ ), then there is a standard result on Galton-Watson trees which states that, if  $m > 1$  (so-called supercritical), then the tree extends to infinity with a positive probability. In other words, denote by  $\mathcal{S}$  the event that a Galton-Watson tree survives,  $\{\mathbf{T} : Z_n(\mathbf{T}) > 0, \forall n \geq 0\}$ , then  $\mathbf{P}_{\text{GW}}(\mathcal{S}) > 0$  for  $m > 1$ . To study the asymptotic behavior of the cover time, we study the first  $n$  generations of a supercritical Galton-Watson tree under the conditional probability measure  $\mathbf{P}_{\text{GW}}(\cdot | \mathcal{S})$ .

By the Kesten-Stigum theorem (cf. [18]),  $(\frac{Z_n}{m^n})_{n \geq 0}$  is a martingale, and that it converges to a nontrivial limit  $\mathbf{P}_{\text{GW}}(\cdot | \mathcal{S})$ -almost surely,

$$W := \lim_{n \rightarrow \infty} \frac{Z_n}{m^n} \in (0, \infty), \quad (1.1)$$

when  $\mathbf{E}_{\text{GW}}(Z_1 \log Z_1) = \sum_{k \geq 1} k \log k \mu_k < \infty$ .

Given a surviving tree  $\mathbf{T}$  and fix  $n > 0$ , let  $(X_n(t))_{t \geq 0}$  be a continuous time Markov jump process on  $\mathbf{T}_n$  starting at  $X_n(0) = \overleftarrow{\emptyset}$ , with transition rates  $p(\overleftarrow{\emptyset}, \emptyset) = 1$ ,

$$p(x, \overleftarrow{x}) = \frac{\lambda}{\lambda + \nu_x}, p(x, x^{(i)}) = \frac{1}{\lambda + \nu_x}, \forall x \in \mathbf{T}_n \setminus \{\overleftarrow{\emptyset}\}, 1 \leq i \leq \nu_x.$$

The probability measure for this random walk is denoted by  $\mathbf{P}_w(\cdot | \mathbf{T}_n)$ , and its corresponding expectation by  $\mathbf{E}_w(\cdot | \mathbf{T}_n)$ . In fact, it is more natural to call the discrete walk  $\lambda$ -biased random walk, and the results for both settings are the same, see Remark 1.2 (1).

We shall work under the following hypotheses denoted by  $(\mathbf{H})$ ,

$$\lambda > 1, m > 1, \sum_{k \geq 0} k^2 \mu_k < \infty.$$

**Conventions:**

- (1) The probability of a generic law independent of the constructions above is denoted by  $\mathbb{P}$ , with its expectation  $\mathbb{E}$ .
- (2) When an integer index is needed in the presence of a real value, we mean its integral part, for instance, we write  $Z_{\log n}$  for  $Z_{\lfloor \log n \rfloor}$ .
- (3) We write  $f \lesssim g$ , if there exists a constant  $C > 0$  such that  $f \leq Cg$ . This constant may depend on parameters  $m$  and  $\lambda$ .
- (4) All the  $O(\cdot)$ ,  $o(\cdot)$  notations are under the limit  $n \rightarrow \infty$ .
- (5) For convenience of the readers, notations used throughout the paper are gathered here:

$$\sigma_n = \sqrt{\frac{\lambda^{n+1} - 1}{\lambda - 1}}, \quad t_n^\mu = \sigma_n^2(\log Z_n + \mu), \quad s_n = \sum_{i=0}^n \frac{Z_i}{\lambda^i}.$$

## 1.2 Main results

The goal in this paper is to estimate the cover time defined by

$$T_n^{\text{cov}}(\mathbf{T}) = \inf \{t : \{X_n(s), 0 \leq s \leq t\} = \mathbf{T}_n\}.$$

The main result is

**Theorem 1.1.** *Under (H), for  $\mathbf{P}_{\text{GW}}(\cdot|\mathcal{S})$ -almost surely any tree  $\mathbf{T}$ , with  $x \in \mathbb{R}$  and  $n \rightarrow \infty$ , when  $\lambda > m$ ,*

$$\mathbf{P}_w \left( \frac{(\lambda - 1)T_n^{\text{cov}}}{2\lambda^{n+1} \sum_{i=0}^{\infty} \frac{Z_i}{\lambda^i}} - n \log m - \log W \leq x \middle| \mathbf{T}_n \right) \rightarrow e^{-e^{-x}};$$

when  $\lambda = m$ ,

$$\mathbf{P}_w \left( \frac{(m - 1)T_n^{\text{cov}}}{2m^{n+1} \sum_{i=0}^n Z_i} - n \log m - \log W \leq x \middle| \mathbf{T}_n \right) \rightarrow e^{-e^{-x}};$$

when  $1 < \lambda < m$ ,

$$\mathbf{P}_w \left( \frac{\left(\frac{m}{\lambda} - 1\right)(\lambda - 1)T_n^{\text{cov}}}{2Wm^{n+1}} - n \log m - \log W \leq x \middle| \mathbf{T}_n \right) \rightarrow e^{-e^{-x}}.$$

**Remark 1.2.** (1) *The the same result holds for the discrete time random walk with the same transition probabilities, since  $(X_n(t))_{0 \leq t \leq t_0}$  takes  $t_0 + O(\sqrt{t_0})$  steps, and the error is negligible for the cover time.*

- (2) In fact, the conditions in Lemma 2.1 are the only requirements for  $\mathbf{T}$ . Therefore, the result applies to other (random) trees satisfying these conditions, not necessarily the Galton-Watson trees.
- (3) The  $\lambda$ -biased case agrees with the case of the simple random walk (cf. eg. [10]) in first order in the number of excursions (round trips from  $\emptyset$ ) performed, for details see Remark 3.5.
- (4) If one define the random walk on  $\mathbf{T}$  instead of  $\mathbf{T}_n$ , the relation of  $\lambda$  and  $m$  discussed in Theorem 1.1 correspond to whether the walk is transient, positive recurrent, or null recurrent.

### 1.3 Related works

The cover time of a finite graph (by the simple random walk),  $T^{\text{cov}}(G)$ , is a fundamental object for a finite graph  $G = (V, E)$  (Section 2, Lovász [19]). For a simple graph  $G$  with  $|V| = n$ , a tight bound for its cover time was given in Feige [15], [16]

$$(1 - o(1))n \log n \leq T^{\text{cov}}(G) \leq 4n^3/27.$$

Bounds using hitting time were given in Matthews [20],

$$\max_{S \subseteq G} \min_{u, v \in S} H(u, v)(\log(\#E) - 1) \leq T^{\text{cov}}(G) \leq \max_{u, v \in G} H(u, v)(1 + \log n),$$

where  $H(u, v)$  is the expected time that the walk takes from  $u$  to  $v$ .

Up to the first order approximation, a general bound with Discrete Gaussian Free Field (DGFF) was given in Ding, Lee and Peres [12], then improved in Zhai [21],

$$\mathbb{P} \left( |T^{\text{cov}}(G) - \#EM^2| \geq \#E(\sqrt{sRM} + sR) \right) \leq Ce^{-cs},$$

where  $M = \mathbb{E}(\max_{x \in V} \eta_x)$ ,  $R = \max_{x, y \in V} R_{\text{eff}}(x, y)$ ,  $(\eta_x)_{x \in V}$  is a DGFF on  $G$  (centered Gaussian variables such that  $\text{Cov}(\eta_x, \eta_y) = R_{\text{eff}}(x, y)$ ), and  $R_{\text{eff}}$  is the effective resistance (cf. [3], consider each edge as a wire of electrical resistance 1, and take effective resistance in the physics sense, following Ohm's law).

Sharper results can be obtained if one restricts to trees. The first order estimate for the cover time of an  $m$ -ary trees was obtained in Aldous [2],

$$T_n^{\text{cov}} = (2 + o(1))n^2 \frac{m^{n+1}}{m-1} \log m.$$

It is showed in Andreoletti and Debs [4] that, the first  $R_n = (\gamma + o(1)) \log n$  generations are covered in  $n$  steps, by a recurrent Markovian random walk on the Galton-Watson tree, where  $\gamma$  is an explicit constant.

The case of simple random walk on binary trees received extensive studies recently, originally as a counterexample showing that at second order, the cover time is no longer determined by the DGFF (cf. [13]). Second order asymptotics with error  $O(\log \log^8 n)$  were given in Ding and Zeitouni [13], then refined to  $O(1)$  in Belius, Rosen and Zeitouni [6], and a scaling limit was established independently by Cortines, Loidor and Saglietti [8] and Dembo, Rosen and Zeitouni [10],

$$\mathbb{P} \left( \frac{T_n^{cov}}{2^{n+1}n} - n \log 2 + \log n \leq s \right) = \mathbb{E} \exp(-CZe^{-s}),$$

for some explicit constant  $C$  and distribution  $Z$ .

In the studies of the cover time, the continuous counterpart for trees is the two-dimensional torus. The first order estimate of its cover time was determined by Dembo, Peres, Rosen and Zeitouni [9], then the result was improved in Ding [11], Belius and Kistler [5], Abe [1], and most recently studied by Belius, Rosen and Zeitouni [7] to the extent that

$$\lim_{K \rightarrow \infty} \limsup_{\epsilon \rightarrow 0} \mathbb{P} \left( \left| \sqrt{T^{cov}(S^2)} - 2\sqrt{2} \left( \log \epsilon^{-1} - \frac{1}{4} \log \log \epsilon^{-1} \right) \right| > K \right) = 0,$$

where  $T_\epsilon^{cov}(S^2)$  is the time for the walk to intersect every ball of radius  $\epsilon$  on the 2-dimensional sphere.

## 1.4 Proof outline

The cover time is determined by the excursion time and local times defined as follows.

**Definition 1.3.** *Let  $\mathbf{T}$  be an infinite tree, and let  $(X_n(t))_{t \geq 0}$  be a random walk on  $\mathbf{T}_n$ .*

(1) *The excursion time is denoted by*

$$t_n^{cov} = \int_0^{T_n^{cov}} \mathbf{1}_{\{X_n(s) = \overleftarrow{\emptyset}\}} ds.$$

(2) *To establish the relation between  $t_n^{cov}$  and  $T_n^{cov}$ , let*

$$\tau_n(t) = \inf \left\{ r > 0 : \int_0^r \mathbf{1}_{\{X_n(s) = \overleftarrow{\emptyset}\}} ds \geq t \right\}.$$

(3) The (normalized) local time at  $x \in \mathbf{T}_n \setminus \{\overleftarrow{\emptyset}\}$  is

$$L_n^x(t) = \frac{1}{\pi_n(x)} \int_0^{\tau_n(t)} \mathbf{1}_{\{X_n(s)=x\}} ds,$$

where the normalization factor

$$\pi_n(x) = \frac{\lambda + \nu_x}{\lambda^{|x|+1}}, \quad x \in \mathbf{T}_n \setminus \{\overleftarrow{\emptyset}\},$$

is the stationary distribution of the  $\lambda$ -biased walk scaled at  $\pi_n(\overleftarrow{\emptyset}) = 1$ .

Intuitively, the walk  $(X_n(t))$  can be seen as independent samples of random walks that starts and ends at  $\overleftarrow{\emptyset}$ . Each of these trials is called an excursion. Then  $t_n^{cov}$  is the number of excursions performed to cover the tree  $\mathbf{T}_n$ ,  $\tau_n(t)$  is the actual time spent in  $t$  excursions, and  $L_n^x(t)$  encodes the status of  $x$  in  $t$  excursions. From these definitions, we have

$$\begin{aligned} \tau_n(t_n^{cov}) &\leq T_n^{cov} \leq \lim_{\epsilon \rightarrow 0^+} \tau_n(t_n^{cov} + \epsilon) \\ \tau_n(t) &= \sum_{x \in \mathbf{T}_n} \pi_n(x) L_n^x(t). \end{aligned} \tag{1.2}$$

In fact, it is not hard to determine  $L_t^x$  (Lemma 2.3) and  $\tau_n$  (Lemma 4.1), and the main part of the proof is to estimate  $t_n^{cov}$ . The key observation is that, when the tree is almost covered (at the first order estimate of the cover time), the non-visited vertices can be seen as independent. (This is inspired by the extremal landscape structure in [8], see Remark 3.5 (1) for details.) The scaling limit of the cover time is then established by characterizing the process afterwards.

The paper is organized as follows. In Section 2, we give the regularity conditions on trees and determine the distribution of local times  $L_n^x$ . In Section 3, we establish the scaling limit for  $t_n^{cov}$ . And in Section 4, we estimate  $T_n^{cov}$  by studying  $\tau_n$ , and finish the proof of Theorem 1.1.

## 2 Preliminaries

### 2.1 The trees

**Lemma 2.1.** *Let  $c, \epsilon > 0$ . Under **(H)**, for  $\mathbf{P}_{\text{GW}}(\cdot|\mathcal{S})$ -almost surely any tree  $\mathbf{T}$ , when  $n$  is large enough,*

$$\sum_{i=0}^n Z_i \lesssim Z_n, \quad (2.3)$$

$$|Z_n - m^n W| \leq m^{n/2} \log n, \quad (2.4)$$

$$\sum_{|x|=n-c \log n} (Z_n^x)^2 \leq Z_n^{1+\epsilon}. \quad (2.5)$$

*Proof.* By (1.1),  $\mathbf{P}_{\text{GW}}(\cdot|\mathcal{S})$ -almost surely,  $W \in (0, \infty)$ . Therefore, there exists a constant  $C > 1$  (depending on  $\mathbf{T}$ ) such that for all  $n \geq 0$ ,

$$\frac{1}{C} W m^n < Z_n < C W m^n.$$

(There are only finitely many  $n$  violating this relation with  $C = 2$ , take the maximum constant among these  $n$ .) One can then deduce (2.3).

Moreover, by Theorem 2, [17]:  $\mathbf{P}_{\text{GW}}(\cdot|\mathcal{S})$ -almost surely,

$$\limsup_{n \rightarrow \infty} \left| \frac{Z_n - m^n W}{\sqrt{Z_n \log n}} \right| = \sqrt{\frac{2 \text{Var} \mu}{m^2 - m}},$$

where we recall that  $\mu$  is the offspring distribution of  $\mathbf{P}_{\text{GW}}$ .

Therefore one can take  $n$  large enough such that

$$|Z_n - m^n W| \lesssim \sqrt{Z_n \log n} \leq \sqrt{C W m^n \log n},$$

and (2.4) follows.

By standard calculations,

$$\mathbf{E}_{\text{GW}} Z_n = m^n, \text{Var}_{\text{GW}}(Z_n) = m^{n-1} \frac{m^n - 1}{m - 1} \text{Var} \mu \lesssim m^{2n},$$

therefore, for  $n$  large enough,

$$\begin{aligned} \mathbf{P}_{\text{GW}} \left( \sum_{|x|=n-c \log n} (Z_n^x)^2 > m^{(1+\epsilon)n} \middle| \mathcal{S} \right) &\leq \frac{\mathbf{E}_{\text{GW}}(\sum_{|x|=n-c \log n} (Z_n^x)^2 | \mathcal{S})}{m^{(1+\epsilon)n}} \\ &\lesssim \frac{\mathbf{E}_{\text{GW}}(Z_{n-c \log n}) m^{2c \log n}}{m^{(1+\epsilon)n}} \\ &\lesssim m^{-\frac{\epsilon n}{2}}, \end{aligned}$$

and (2.5) follows from the union bound,

$$\mathbf{P}_{\text{GW}} \left( \exists n > N, \sum_{|x|=n-c \log n} (Z_n^x)^2 > m^{(1+\epsilon)n} \middle| \mathcal{S} \right) \lesssim \sum_{n>N} m^{-\frac{\epsilon n}{2}} \xrightarrow{N \rightarrow \infty} 0.$$

We remark that  $\sum_{|x|=n-r} (Z_n^x)^2$  is monotone increasing in  $r$ , therefore, the event (2.5) is decreasing in  $r$ , and once it is valid for  $c \log n$ , it is valid for all  $0 \leq r \leq c \log n$  simultaneously.  $\square$

## 2.2 The local times

**Definition 2.2.** Let  $\mathbf{T}$  be a tree.

(1) The effective resistance (for the  $\lambda$ -biased walk) between  $x, y \in \mathbf{T}$  is

$$R_{\text{eff}}(x, y) = \sum_{\substack{(z, \bar{z}) \\ \text{on the simple path} \\ \text{from } x \text{ to } y}} \lambda^{|z|}.$$

(2) The effective resistance between  $\overleftarrow{\emptyset}$  and any vertex at generation  $n$  is abbreviated as

$$\sigma_n^2 = \frac{\lambda^{n+1} - 1}{\lambda - 1}.$$

(3) For  $a, b > 0$ , let  $\text{PG}(a, b)$  be the distribution of  $\sum_{i=1}^P E_i$ , where  $P$  and  $E_i$  are independent random variables such that  $P \sim \text{Poisson}(a)$  has Poisson distribution of expected value  $a$ , and  $E_i \sim \text{Exp}(b)$  has exponential distribution of expected value  $\frac{1}{b}$ .

The following lemma characterizes local times.

**Lemma 2.3.** On any infinite tree  $\mathbf{T}$ , let  $x, y \in \mathbf{T}_n$  such that  $y \prec x$  (i.e.  $y$  is a strict ancestor of  $x$ ), let  $s, t > 0$ . Then under  $\mathbf{P}_w(\cdot | \mathbf{T}_n)$ ,

$$L_n^x(t) \sim \text{PG} \left( \frac{t}{\sigma_{|x|}^2}, \frac{1}{\sigma_{|x|}^2} \right),$$

$$(L_n^x(t) | L_n^y(t) = s) \sim \text{PG} \left( \frac{s}{\sigma_{|x|}^2 - \sigma_{|y|}^2}, \frac{1}{\sigma_{|x|}^2 - \sigma_{|y|}^2} \right).$$

*Proof.* By the memoryless property of the exponential distribution (if  $X \sim \text{Exp}(1)$ , then  $(X - c | X > c) \stackrel{d}{=} X$ ),  $L_n^x(t)$  is only affected by local times on the ray from  $\overleftarrow{\emptyset}$  to  $x$ , independent of movements on other branches or offspring of  $x$ .

By knowledge of reversible Markov chains (cf. eg. (3.24), p.69, [3], it can be easily checked that  $R_{\text{eff}}$  defined in Definition 2.2 do correspond to that of a reversible Markov chain, since there is only one simple path between two vertices),

$$\frac{1}{R_{\text{eff}}(x, y)} = \pi(x) \mathbf{P}_x(\tau_y < \tau_x^+ | \mathbf{T}_n),$$

where  $\mathbf{P}_x$  is the  $\lambda$ -biased walk starting at  $x$ ,  $\tau_y$  is the first hitting times of  $y$ , i.e.  $\tau_y = \inf\{t > 0, X_n(t) = y\}$ , and  $\tau_x^+$  is the first returning time to  $x$ ,  $\tau_x^+ = \inf\{t > 0, X_n(t) = x, X_n(s) \neq x, 0 \leq s \leq t\}$ . In particular,

$$\mathbf{P}_x(\tau_{\overleftarrow{\emptyset}} < \tau_x^+ | \mathbf{T}_n) = \frac{1}{\sigma_{|x|}^2 \pi_n(x)}.$$

Up to excursion time  $t$ , there are  $\text{Poisson}(t)$  departures from  $\overleftarrow{\emptyset}$ , and by the equation above, each trip hits  $x$  independently with probability  $\frac{1}{\sigma_{|x|}^2}$ , thus there are  $\text{Poisson}\left(\frac{t}{\sigma_{|x|}^2}\right)$  arrivals on  $x$ . Upon arrival at  $x$ , the walk returns to  $\overleftarrow{\emptyset}$  in exponential time, with rate  $\mathbf{P}_x(\tau_{\overleftarrow{\emptyset}} < \tau_x^+ | \mathbf{T}_n) = \frac{1}{\sigma_{|x|}^2 \pi_n(x)}$ . Therefore, the total time spent at  $x$  has distribution  $\text{PG}\left(\frac{t}{\sigma_{|x|}^2}, \frac{1}{\sigma_{|x|}^2 \pi_n(x)}\right)$ . Recall that local times are normalized by  $\frac{1}{\pi_n}$ , and the result follows.

Conditioned at  $L_n^y(t)$ , the proof is similar. (Only to notice that the local times at both  $x, y$  are normalized.)  $\square$

**Lemma 2.4.** *Let  $a, b > 0$ , let  $X \sim \text{PG}(a, b)$ .*

(1) *The PG distribution has basic properties*

$$bX \sim \text{PG}(a, 1), \mathbb{E}(X) = \frac{a}{b}, \text{Var}(X) = \frac{2a}{b^2}, \mathbb{P}(X = 0) = e^{-a}.$$

(2) *If  $a > b$ , then*

$$\mathbb{P}(X \leq 1) \leq e^{2\sqrt{ab} - a - b}.$$

*Proof.* (1) is clear by definition. For (2), by Chernoff bounds, for any  $\theta > 0$ ,

$$\begin{aligned} \mathbb{P}(X \leq 1) &= \mathbb{P}(e^{-\theta X} \geq e^{-\theta}) \\ &\leq e^{\theta} \sum_{k=0}^{\infty} e^{-a} \frac{a^k}{k!} \left(1 + \frac{\theta}{b}\right)^{-k} = e^{\theta - \frac{\theta}{b+\theta}a}, \end{aligned}$$

and the result follows by choosing  $\theta = \sqrt{ab} - b$ .  $\square$

## 2.3 Ray-Knight Theorem

**Definition 2.5.** A Discrete Gaussian Free Field (DGFF) on a tree  $\mathbf{T}$  is a family of random variables  $(\eta_x)_{x \in \mathbf{T}}$  such that  $\eta_{\emptyset} = 0$ ,  $(\eta_x)_{x \neq \emptyset}$  are centered Gaussian variables with (both the effective resistance and the DGFF can be defined up to any scale, the factor  $\frac{1}{2}$  is taken in accordance to [8])

$$\mathbb{E}(\eta_x - \eta_y)^2 = \frac{1}{2} R_{\text{eff}}(x, y).$$

**Remark 2.6.** Since we have explicit relative resistances, if we attach an independent Gaussian variable  $N_y \sim \mathcal{N}\left(0, \frac{\lambda|y|}{2}\right)$  at each vertex  $y \in \mathbf{T} \setminus \{\emptyset\}$ , and let  $\eta_x = \sum_{y \preceq x} N_y$ , then  $(\eta_x)_{x \in \mathbf{T}}$  is a DGFF on  $\mathbf{T}$ .

**Theorem 2.7.** (Second Ray-Knight theorem, [14]) For any infinite tree  $\mathbf{T}$ , let  $(\eta_x)_{x \in \mathbf{T}}$  be a DGFF on  $\mathbf{T}$  independent of  $\mathbf{P}_w(\cdot | \mathbf{T}_n)$ . For any  $t > 0$ ,

$$\{L_n^x(t) + \eta_x^2 : x \in \mathbf{T}_n\} \stackrel{d}{=} \{(\eta_x + \sqrt{t})^2 : x \in \mathbf{T}_n\}.$$

In fact, Remark 2.6 indicates a direct proof of this theorem by induction.

**Lemma 2.8.** Let  $(\eta_x)_{x \in \mathbf{T}}$  be a DGFF on a tree  $\mathbf{T}$ , let  $n \geq 0$  and  $\mu \in \mathbb{R}$  such that

$$Z_n > 0, \log Z_n + \mu > 0,$$

we have

$$\mathbb{P}\left(\max_{|x|=n} \eta_x > \sigma_n \sqrt{\log Z_n + \mu}\right) \leq \frac{e^{-\mu}}{2\sqrt{\pi(\log Z_n + \mu)}}.$$

*Proof.* We first recall the Gaussian tail estimate for  $x > 0$  and  $X \sim \mathcal{N}(0, 1)$ ,

$$\mathbb{P}(X > x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(y+x)^2/2} dy \leq \frac{e^{-x^2/2}}{\sqrt{2\pi}} \int_0^\infty e^{-xy} dy = \frac{e^{-x^2/2}}{x\sqrt{2\pi}}.$$

Then by the union bound, when  $\log Z_n + \mu > 0$ ,

$$\mathbb{P}\left(\max_{|x|=n} \eta_x > \sigma_n \sqrt{\log Z_n + \mu}\right) \leq Z_n \frac{e^{-(\log Z_n + \mu)}}{2\sqrt{\pi(\log Z_n + \mu)}} = \frac{e^{-\mu}}{2\sqrt{\pi(\log Z_n + \mu)}}.$$

$\square$

### 3 Excursion time

As showed in the proof outline, the excursion time  $t_n^{cov}$  is compared to the quantity

$$t_n^\mu = \sigma_n^2(\log Z_n + \mu), \quad (3.6)$$

where we recall that

$$\sigma_n^2 = \frac{\lambda^{n+1} - 1}{\lambda - 1}.$$

To start with, we show that the first  $n - c \log n$  layers have negligible influences in terms of local times.

**Lemma 3.1.** *Let  $c > \frac{3}{\log \lambda}$ ,  $\mu \in \mathbb{R}$ , and  $\mathbf{T}$  be an infinite tree satisfying the conditions in Lemma 2.1. For  $n \rightarrow \infty$ , with probability  $1 - o(1)$  under  $\mathbf{P}_w(\cdot | \mathbf{T}_n)$ ,*

$$\max_{|x|=n-c \log n} \left| \frac{L_n^x(t_n^\mu) - t_n^\mu}{\sigma_n^2} \right| \lesssim \frac{1}{\sqrt{n}}.$$

*Proof.* Let  $(\eta_x)$  be a DGFF on  $\mathbf{T}$  independent of  $\mathbf{P}_w(\cdot | \mathbf{T}_n)$ . Denote the joint law of the DGFF and  $\mathbf{P}_w(\cdot | \mathbf{T}_n)$  by  $\mathbb{P}(\cdot | \mathbf{T}_n)$  in this proof.

It is guaranteed by Lemma 2.1 that  $\log Z_n + \mu > 0$  for  $n$  large enough. Then one can apply Lemma 2.8,

$$\mathbb{P} \left( \max_{|x|=n-c \log n} |\eta_x| \leq \sigma_{n-c \log n} \sqrt{\log Z_{n-c \log n} + \mu} \middle| \mathbf{T}_n \right) = 1 - o(1).$$

Moreover, by Theorem 2.7,

$$\mathbb{P} \left( \max_{|x|=n-c \log n} \left| \sqrt{L^x(t_n^\mu) + \eta_x^2} - \sqrt{t_n^\mu} \right| \leq \sigma_{n-c \log n} \sqrt{\log Z_{n-c \log n} + \mu} \middle| \mathbf{T}_n \right) = 1 - o(1).$$

Thus with probability  $1 - o(1)$ , for  $n$  large enough,

$$\begin{aligned} & \max_{|x|=n-c \log n} |L^x(t_n^\mu) - t_n^\mu| \\ &= \max_{|x|=n-c \log n} \left| \left( \sqrt{L^x(t_n^\mu) + \eta_x^2} - \sqrt{t_n^\mu} \right) \left( \sqrt{L^x(t_n^\mu) + \eta_x^2} + \sqrt{t_n^\mu} \right) - \eta_x^2 \right| \\ &\leq \max_{|x|=n-c \log n} \left| \sqrt{L^x(t_n^\mu) + \eta_x^2} - \sqrt{t_n^\mu} \right|^2 \\ &\quad + 2\sqrt{t_n^\mu} \max_{|x|=n-c \log n} \left| \sqrt{L^x(t_n^\mu) + \eta_x^2} - \sqrt{t_n^\mu} \right| + \max_{|x|=n-c \log n} |\eta_x| \\ &\lesssim \sigma_n^2 \lambda^{-c \log n / 2} (\log Z_n + \mu) \lesssim \sigma_n^2 n^{-1/2}, \end{aligned}$$

where the last line follows from Lemma 2.1 and the condition  $c > \frac{3}{\log \lambda}$ .  $\square$

Now we present the key observation that, non-visited vertices up to time  $t_n^\mu$  have distinct ancestors at generation  $n - c \log n$ .

**Lemma 3.2.** *Let  $c > \frac{3}{\log \lambda}$ ,  $\mu \in \mathbb{R}$ , and  $\mathbf{T}$  be an infinite tree satisfying the conditions in Lemma 2.1. There exists  $\epsilon > 0$  such that*

$$\max_{\substack{|y \wedge z| \geq n - c \log n + 1, \\ |y| = |z| = n, y \neq z}} \mathbf{P}_w \left( L_n^y(t_n^\mu) = L_n^z(t_n^\mu) = 0 \mid (L_n^x(t_n^\mu))_{|x|=n-c \log n}, \mathbf{T}_n \right) = o(Z_n^{-1-\epsilon}),$$

uniformly in any choices of local times at generation  $n - c \log n$  satisfying the conditions in Lemma 3.1.

*Proof.* For simplicity, the conditional probability  $\mathbf{P}_w(\cdot \mid (L_n^x(t_n^\mu))_{|x|=n-c \log n}, \mathbf{T}_n)$  is abbreviated as  $\mathbf{P}_w(\cdot \mid R_n)$ , and we write  $L^x$  for the local time  $L_n^x(t_n^\mu)$ .

Let  $\delta > 0$ . Fix two vertices  $y, z$  at generation  $n$  of  $\mathbf{T}_n$ , such that they have common ancestor  $w = y \wedge z$  after generation  $n - c \log n$ . Let  $|w| = n - s \geq n - c \log n + 1$ , and denote by  $x$  the ancestor of  $w$  with  $|x| = n - \log n$ . Then

$$\begin{aligned} & \mathbf{P}_w(L^y = L^z = 0 \mid R_n) \\ & \leq \mathbf{P}_w(L^y = L^z = 0, L^w \geq 2\delta t_n^\mu \mid R_n) + \mathbf{P}_w(L^w < 2\delta t_n^\mu \mid R_n) \\ & = \mathbf{P}_w(L^z = 0 \mid L^y = 0, L^w \geq 2\delta t_n^\mu, R_n) \mathbf{P}_w(L^y = 0, L^w \geq 2\delta t_n^\mu \mid R_n) + \mathbf{P}_w(L^w < 2\delta t_n^\mu \mid R_n) \\ & = \mathbf{P}_w(L^z = 0 \mid L^w \geq 2\delta t_n^\mu, R_n) \mathbf{P}_w(L^y = 0, L^w \geq 2\delta t_n^\mu \mid R_n) + \mathbf{P}_w(L^w < 2\delta t_n^\mu \mid R_n) \\ & \leq \mathbf{P}_w(L^z = 0 \mid L^w = 2\delta t_n^\mu, R_n) \mathbf{P}_w(L^y = 0 \mid R_n) + \mathbf{P}_w(L^w < 2\delta t_n^\mu \mid R_n). \end{aligned}$$

By Lemma 3.1, for  $n$  large enough we have

$$L^x > (1 - \delta)t_n^\mu,$$

moreover, by Lemma 2.3,

$$(L^w \mid L^x) \stackrel{d}{=} \text{PG} \left( \frac{L^x}{\sigma_{n-s}^2 - \sigma_{n-c \log n}^2}, \frac{1}{\sigma_{n-s}^2 - \sigma_{n-c \log n}^2} \right).$$

where denominator above is bounded by

$$\frac{\lambda - 1}{\lambda} \sigma_{n-s}^2 \leq \sigma_{n-s}^2 - \sigma_{n-s-1}^2 \leq \sigma_{n-s}^2 - \sigma_{n-c \log n}^2 \leq \sigma_{n-s}^2.$$

Therefore, for local times of the generation  $n - c \log n$  satisfying Lemma 3.1 and for  $n$  large enough,

$$\begin{aligned} & \mathbf{P}_w(L^w < 2\delta t_n^\mu \mid R_n) \\ & = \mathbf{P}_w \left( \text{PG} \left( \frac{L^x}{\sigma_{n-s}^2 - \sigma_{n-c \log n}^2}, \frac{1}{\sigma_{n-s}^2 - \sigma_{n-c \log n}^2} \right) < 2\delta t_n^\mu \mid R_n \right) \\ & \leq \mathbb{P} \left( \text{PG} \left( \frac{(1 - \delta)t_n^\mu}{\sigma_{n-s}^2}, \frac{\lambda/(\lambda - 1)}{\sigma_{n-s}^2} \right) \leq 2\delta t_n^\mu \right) \\ & \leq e^{-\lambda(\sqrt{1-\delta} - \sqrt{2\delta\lambda/(\lambda-1)})^2 (\log Z_n + \mu)}, \end{aligned}$$

where the notation  $\mathbb{P}(\text{PG}(a, b) \leq x)$  stands for the probability that a random variable distribution as  $\text{PG}(a, b)$  is smaller than  $x$ , and the last line is due to Lemma 2.4.

Similarly, study  $(L^z|L^w)$  and  $(L^y|L^x)$  by Lemma 2.3, for  $n$  large enough,

$$\begin{aligned} & \mathbf{P}_w(L^z = 0 | L^w = 2\delta t_n^\mu, R_n) \mathbf{P}_w(L^y = 0 | R_n) \\ &= e^{-\frac{2\delta t_n^\mu}{\sigma_n^2 - \sigma_{n-s}^2} - \frac{L^x}{\sigma_n^2 - \sigma_{n-c \log n}^2}} \leq e^{-\frac{2\delta t_n^\mu}{\sigma_n^2} - \frac{(1-\delta)t_n^\mu}{\sigma_n^2}} = e^{-(1+\delta)(\log Z_n + \mu)}. \end{aligned}$$

To show that the two probabilities above are bounded by  $o(Z_n^{-1-\epsilon})$ , it suffices to choose  $\delta, \epsilon > 0$  such that

$$\begin{aligned} \lambda \left( \sqrt{1-\delta} - \sqrt{2\delta\lambda/(\lambda-1)} \right)^2 &> 1 + \epsilon, \\ 1 + \delta &> 1 + \epsilon, \end{aligned} \tag{3.7}$$

which is always possible when  $\lambda > 1$ .  $\square$

The estimate above is enough to control the cover time of the  $n$ -th generation, however a Galton-Watson tree may have leaves in younger generations, in other words, the walk visits all vertices in the  $n$ -th generation does not guarantee that it covers  $\mathbf{T}_n$ . We treat leaves before the generation  $n-1$  separately.

**Lemma 3.3.** *Let  $\mu \in \mathbb{R}$ , and let  $\mathbf{T}$  be an infinite tree satisfying the conditions in Lemma 2.1, then*

$$\mathbf{P}_w(\exists x \in \mathbf{T}_n, |x| \leq n-1, L_n^x(t_n^\mu) = 0 | \mathbf{T}_n) = o(1).$$

*Proof.* By Lemma 2.3 and Lemma 2.4 (1), each  $x \in \mathbf{T}_n$  with  $|x| \leq n-1$  is not visited at excursion time  $t_n^\mu$  with probability

$$\mathbf{P}_w(L_n^x(t_n^\mu) = 0 | \mathbf{T}_n) = e^{-\sigma_{|x|}^{-2} t_n^\mu} \leq e^{-\lambda(\log Z_n + \mu)},$$

thus by the union bound,

$$\begin{aligned} \mathbf{P}_w(\exists x \in \mathbf{T}_n, |x| \leq n-1, L_n^x(t_n^\mu) = 0 | \mathbf{T}_n) &\leq e^{-\lambda(\log Z_n + \mu)} \sum_{|x| \leq n-1} 1 \\ &= e^{-\lambda(\log Z_n + \mu)} \sum_{i=0}^{n-1} Z_i, \end{aligned}$$

and the conclusion follows from (2.3).  $\square$

Returning to the  $n$ -th generation, by Lemma 3.2, the non-visited vertices are almost independent at excursion time  $t_n^\mu$ . Intuitively, the time to cover them is a binomial random variable of  $Z_{n-c\log n}$  trials, converging to a Poisson distribution. We conclude upon this intuition:

**Proposition 3.4.** *Let  $\mu \in \mathbb{R}$ , and let  $\mathbf{T}$  be an infinite tree satisfying the conditions in Lemma 2.1, recall  $t_n^\mu$  defined in (3.6), we have*

$$\#\{|x| = n, L_n^x(t_n^\mu) = 0\} \xrightarrow{d} \text{Poisson}(e^{-\mu}) \text{ under } \mathbf{P}_w(\cdot | \mathbf{T}_n), \quad (3.8)$$

$$\mathbf{P}_w \left( \frac{t_n^{\text{cov}}}{\sigma_n^2} - n \log m - \log W \leq \mu \middle| \mathbf{T}_n \right) \rightarrow e^{-e^{-\mu}}. \quad (3.9)$$

*Proof.* Since both the conclusions allow a  $o(1)$  error, we may assume the conditions in Lemma 3.1 at generation  $n - c \log n$  for an arbitrarily fixed  $c > \frac{3}{\log \lambda}$ . For simplicity, recall the abbreviation  $\mathbf{P}_w(\cdot | R_n)$  in the proof of Lemma 3.2, then it suffices replace  $\mathbf{P}_w(\cdot | \mathbf{T}_n)$  by  $\mathbf{P}_w(\cdot | R_n)$ . Moreover, denote by

$$F_n^\mu = \{|x| = n, L_n^x(t_n^\mu) = 0\}, E_n^\mu = \{|x| = n - c \log n + 1, \#(\mathbf{T}_n^x \cap F_n^\mu) = 1\},$$

the goal is to estimate  $F_n^\mu$ , which is achieved by comparison to  $E_n^\mu$ .

By Lemma 2.3, under  $\mathbf{P}_w(\cdot | R_n)$ , local times are distributed as

$$(L_n^y(t_n^\mu) | R_n) \sim \text{PG} \left( \frac{L_n^x(t_n^\mu)}{\sigma_n^2 - \sigma_{n-c\log n}^2}, \frac{1}{\sigma_n^2 - \sigma_{n-c\log n}^2} \right)$$

for each pair of vertices

$$x \prec y, |x| = n - c \log n, |y| = n.$$

Therefore, by Lemma 3.1,

$$\begin{aligned} \mathbf{E}_w(\#F_n^\mu | R_n) &= \sum_{|x|=n-c\log n} e^{-\frac{L_n^x(t_n^\mu)}{\sigma_n^2 - \sigma_{n-c\log n}^2} Z_n^x} \\ &= \sum_{|x|=n-c\log n} e^{-\log Z_n - \mu + o(1)} Z_n^x = e^{-\mu + o(1)}. \end{aligned}$$

Furthermore, by definition one has

$$0 \leq \#F_n^\mu - \#E_n^\mu \leq 2 \sum_{|x|=n-c\log n+1} \sum_{y, z \in \mathbf{T}_n^x} \mathbf{1}_{\{y, z \in F_n^\mu\}},$$

where  $\mathbf{E}_w(\mathbf{1}_{\{y,z \in F_n^\mu\}} | R_n) = o(Z_n^{-(1+\epsilon)})$  by Lemma 3.2.

Therefore, by (2.5),

$$\begin{aligned} \mathbf{E}_w(\#E_n^\mu | R_n) &= \mathbf{E}_w(\#F_n^\mu | R_n) + \mathbf{E}_w(\#E_n^\mu - \#F_n^\mu | R_n) \\ &= e^{-\mu+o(1)} + o\left(Z_n^{-(1+\epsilon)} \sum_{|x|=n-c \log n+1} (Z_n^x)^2\right) \\ &= e^{-\mu+o(1)} + o(1), \end{aligned}$$

where  $\epsilon$  is the parameter in Lemma 3.2.

Moreover, conditioned on local times of layer  $n - c \log n$ , the subtrees  $(\mathbf{T}^x)_{|x|=n-c \log n+1}$  are independent, thus for any  $\theta > 0$ , the Laplace transform of  $\#E_n^\mu$  is given by

$$\begin{aligned} &\mathbf{E}_w\left(e^{-\theta \#E_n^\mu} \mid R_n\right) \\ &= \prod_{|x|=n-c \log n+1} \mathbf{E}_w\left(e^{-\theta \mathbf{1}_{x \in E_n^\mu}} \mid R_n\right) \\ &= \prod_{|x|=n-c \log n+1} \left(1 - (1 - e^{-\theta}) \mathbf{P}_w(x \in E_n^\mu \mid R_n)\right) \\ &= \prod_{|x|=n-c \log n+1} e^{-(1-e^{-\theta}+o(1))\mathbf{P}_w(x \in E_n^\mu \mid R_n)} \\ &= e^{-(1-e^{-\theta}+o(1))\mathbf{E}_w(\#E_n^\mu \mid R_n)} \rightarrow e^{-(1-e^{-\theta})e^{-\mu}}, \end{aligned}$$

in other words

$$\#E_n^\mu \xrightarrow{d} \text{Poisson}(e^{-\mu}).$$

Finally, by Lemma 3.2 and the union bound again, we have

$$\mathbf{P}_w(\#E_n^\mu = \#F_n^\mu | R_n) = 1 - o(1),$$

thus  $\#F_n^\mu$  has the same distributional limit as  $\#E_n^\mu$ , completing the proof of (3.8).

As for (3.9),  $\{t_n^{\text{cov}} \leq t_n^\mu\}$  differs from  $\{\#F_n^\mu = 0\}$  by whether the leaves of  $\mathbf{T}_n$  in early generations are covered, which is controlled in Lemma 3.3. Therefore, by (3.8),

$$\mathbf{P}_w(t_n^{\text{cov}} \leq t_n^\mu | R_n) = \mathbf{P}_w(\#F_n^\mu = 0 | R_n) + o(1) \rightarrow e^{-e^{-\mu}},$$

then (3.9) follows from the asymptotic of  $t_n^\mu$ ,

$$\frac{t_n^\mu}{\sigma_n^2} - n \log m - \log W \rightarrow \mu.$$

□

**Remark 3.5.** (1) For the complete binary tree, we have  $m = 2$ . Take  $\lambda \rightarrow 1$ , then  $\mathbf{P}_w(\cdot|\mathbf{T}_n)$  converges to a simple random walk, our result gives (non-rigorously)

$$t_n^{\text{cov}} \approx n^2 \log 2 + O(n),$$

whereas the cover time on the binary tree of a simple random walk is (cf. eg. [10])

$$t_n^{\text{cov}} = n^2 \log 2 - n \log n + O(n).$$

Lack of the second order term  $n \log n$  is due to a difference in extremal landscapes: recall the notations in the proof of Proposition 3.4, in the case of a simple random walk [8], the set  $F_n^\mu$  is approximately identically distributed clusters indexed by  $E_n^\mu$ , whereas for the  $\lambda$ -biased walk, these clusters are single points instead.

(2) Following exactly the same structure of the proof, one can study the maximum of a DGFF ( $\eta_x$ ) (recall Definition 2.5),

$$\mathbb{P} \left( \max_{|x|=n} \eta_x \leq \sigma_n \sqrt{\log Z_n - \frac{1}{2} \log \log Z_n + \mu} \right) \rightarrow \exp \left( -\frac{e^{-\mu}}{2\sqrt{\pi}} \right).$$

(3) Comparing the cover time to the maximum of the corresponding DGFF, as suggested in [12], [13], one has

$$t_n^{\text{cov}} = \frac{\lambda^{n+1}}{\lambda - 1} (n \log m + O(1)),$$

$$\max_{|x|=n} \eta_x^2 = \frac{\lambda^{n+1}}{\lambda - 1} \left( n \log m - \frac{1}{2} \log n + O(1) \right).$$

This difference in second order is due to different tails of Gaussian and local time distributions. (Compare Lemma 2.4 (2) and Lemma 2.8.)

## 4 From excursion time to real time

**Lemma 4.1.** For any infinite tree  $\mathbf{T}$  satisfying the conditions in Lemma 2.1, let  $s_n = \sum_{i=0}^n \frac{Z_i}{\lambda^i}$ ,  $t > 0$ , then

$$\mathbf{E}_w(\tau_n(t)|\mathbf{T}_n) = \mathbf{E}_w \left( \sum_{x \in \mathbf{T}_n} \pi_n(x) L_n^x(t) \middle| \mathbf{T}_n \right) = 2ts_n, \quad (4.10)$$

$$\text{Var}_w(\tau_n(t)|\mathbf{T}_n) = o \left( \frac{t\lambda^n s_n^2}{n} \right). \quad (4.11)$$

*Proof.* The expected value (4.10) is clear using the fundamental estimates  $\mathbf{E}_w(L^x(t)|\mathbf{T}_n) = t$  and  $\sum_{|x|=k} \nu_x = Z_{k+1}$ . As for (4.11), for any  $x, y \in \mathbf{T}_n$ , conditioned at  $L_n^{x \wedge y}(t)$ , local times  $L_n^x(t)$  and  $L_n^y(t)$  are independent with the same expected value  $L_n^{x \wedge y}(t)$ , thus by Lemma 2.4 (1),

$$\text{Cov}_w(L_n^x(t), L_n^y(t)|\mathbf{T}_n) = \text{Var}_w(L_n^{x \wedge y}(t)|\mathbf{T}_n) = 2t\sigma_{|x \wedge y|}^2 \leq 2t \frac{\lambda}{\lambda - 1} \lambda^{|x \wedge y|},$$

then by (1.2),

$$\begin{aligned} & \text{Var}_w(\tau_n(t)|\mathbf{T}_n) \\ &= \text{Var}_w \left( \sum_{x \in \mathbf{T}_n} \pi(x) L_n^x(t) \middle| \mathbf{T}_n \right) \\ &= \sum_{x, y \in \mathbf{T}_n} \pi(x) \pi(y) \text{Cov}_w(L_n^x(t), L_n^y(t)|\mathbf{T}_n) \\ &\leq 2t \frac{\lambda}{\lambda - 1} \sum_{x, y \in \mathbf{T}_n} \lambda^{|x \wedge y| - |x| - |y|} \left( 1 + \frac{\nu_x}{\lambda} + \frac{\nu_y}{\lambda} + \frac{\nu_x \nu_y}{\lambda^2} \right), \end{aligned}$$

where  $\nu_x$  is the number of children for  $x \in \mathbf{T}_n$ .

Moreover,

$$\lambda^{|x \wedge y| - |x| - |y|} \frac{\nu_x}{\lambda} = \sum_{\substack{z \in \mathbf{T}_n \\ \bar{z} = x}} \lambda^{|z \wedge y| - |z| - |y|},$$

therefore the variance above is further bounded by

$$\text{Var}_w(\tau_n(t)|\mathbf{T}_n) \leq 8t \frac{\lambda}{\lambda - 1} \sum_{x, y \in \mathbf{T}_n} \lambda^{|x \wedge y| - |x| - |y|}.$$

Now it suffices to prove that

$$\sum_{x, y \in \mathbf{T}_n} \lambda^{|x \wedge y| - |x| - |y|} = o \left( \frac{\lambda^n s_n^2}{n} \right).$$

Indeed, fix any  $c > 0$ ,

$$\begin{aligned} & \sum_{x, y \in \mathbf{T}_n} \lambda^{|x \wedge y| - |x| - |y|} \\ &\leq \sum_{|x \wedge y| < n - c \log n} \lambda^{n - c \log n - |x| - |y|} + \sum_{|x \wedge y| \geq n - c \log n} \lambda^{n - |x| - |y|} \\ &\leq \lambda^{n - c \log n} \left( \sum_{x \in \mathbf{T}_n} \lambda^{-|x|} \right)^2 + \sum_{|x \wedge y| \geq n - c \log n} \lambda^{n - 2(n - c \log n)} \\ &\leq \frac{\lambda^n s_n^2}{n^2} + \lambda^{2c \log n - n} \sum_{|x| = n - c \log n} (Z_n^x)^2, \end{aligned}$$

where by (2.5) and definition of  $s_n$ ,

$$\begin{aligned} & \lambda^{2c \log n - n} \sum_{|x|=n-c \log n} (Z_n^x)^2 \\ & \leq \lambda^{2c \log n - n} Z_n^{1+\epsilon} \\ & \leq \lambda^{2c \log n - n} (s_n \lambda^n)^{1+\epsilon} = o\left(\frac{\lambda^n s_n^2}{n}\right). \end{aligned}$$

□

**Proposition 4.2.** Recall  $s_n$  from Lemma 4.1, and recall that  $\sigma_n^2 = \frac{\lambda^{n+1}-1}{\lambda-1}$ . Under (H), for any  $\mu \in \mathbb{R}$  and  $\mathbf{P}_{\text{GW}}(\cdot|\mathcal{S})$ -almost surely any tree  $\mathbf{T}$ ,

$$\mathbf{P}_w \left( \frac{T_n^{\text{cov}}}{2s_n \sigma_n^2} - n \log m - \log W \leq \mu \middle| \mathbf{T}_n \right) \rightarrow e^{-e^{-\mu}}.$$

*Proof.* By Lemma 2.1,

$$\log Z_n = n \log m + \log W + o(1),$$

therefore it suffices to show that (rigorously speaking, one should prove the following convergence for  $\mu \pm \epsilon$ , then take  $\epsilon \rightarrow 0$  to deduce the proposition),

$$\mathbf{P}_w \left( \frac{T_n^{\text{cov}}}{2s_n \sigma_n^2} - \log Z_n \leq \mu \middle| \mathbf{T}_n \right) \rightarrow e^{-e^{-\mu}}.$$

In other words (recall  $t_n^\mu$  defined in (3.6)), it suffices to prove that

$$\mathbf{P}_w \left( \frac{T_n^{\text{cov}}}{2s_n \sigma_n^2} \leq \frac{t_n^\mu}{\sigma_n^2} \middle| \mathbf{T}_n \right) \rightarrow e^{-e^{-\mu}}.$$

By (1.2), for any  $\alpha > 0$ ,

$$\begin{aligned} \mathbf{P}_w(T_n^{\text{cov}} \leq 2s_n t_n^\mu | \mathbf{T}_n) & \leq \mathbf{P}_w(\tau_n(t_n^{\text{cov}}) \leq 2s_n t_n^\mu | \mathbf{T}_n) \\ & \leq \mathbf{P}_w(t_n^{\text{cov}} \leq t_n^{\mu+\alpha} | \mathbf{T}_n) \\ & \quad + \mathbf{P}_w(t_n^{\text{cov}} > t_n^{\mu+\alpha}, |\tau_n(t_n^{\mu+\alpha}) - 2t_n^{\mu+\alpha} s_n| > 2s_n(t_n^{\mu+\alpha} - t_n^\mu) | \mathbf{T}_n). \end{aligned}$$

For the first term, by Proposition 3.4,

$$\mathbf{P}_w(t_n^{\text{cov}} \leq t_n^{\mu+\alpha} | \mathbf{T}_n) \leq (1 + o(1))e^{-e^{-\mu-\alpha}}.$$

For the second term, by Chebyshev's inequality and Lemma 4.1

$$\begin{aligned} & \mathbf{P}_w(t_n^{\text{cov}} > t_n^{\mu+\alpha}, |\tau_n(t_n^{\mu+\alpha}) - 2t_n^{\mu+\alpha} s_n| > 2s_n(t_n^{\mu+\alpha} - t_n^\mu) | \mathbf{T}_n) \\ & \leq \mathbf{P}_w(|\tau_n(t_n^{\mu+\alpha}) - 2t_n^{\mu+\alpha} s_n| > 2s_n(t_n^{\mu+\alpha} - t_n^\mu) | \mathbf{T}_n) \\ & = o\left(\frac{\frac{1}{n} \lambda^n t_n^{\mu+\alpha}}{(t_n^{\mu+\alpha} - t_n^\mu)^2}\right) = \alpha^{-2} o(1). \end{aligned}$$

Let  $\alpha \rightarrow 0+$ , we have

$$\limsup_{n \rightarrow \infty} \mathbf{P}_w(T_n^{cov} \leq 2s_n t_n^\mu | \mathbf{T}_n) \leq e^{-e^{-\mu}}.$$

Similarly, for any  $\alpha > 0$ , and any  $\beta(\alpha) > 0$  small enough,

$$\begin{aligned} & \mathbf{P}_w(T_n^{cov} \leq 2s_n t_n^\mu | \mathbf{T}_n) \\ & \geq \mathbf{P}_w(\tau_n(t_n^{cov} + \beta) \leq 2s_n t_n^\mu, t_n^{cov} \leq t_n^{\mu-\alpha} | \mathbf{T}_n) \\ & = \mathbf{P}_w(t_n^{cov} \leq t_n^{\mu-\alpha} | \mathbf{T}_n) - \mathbf{P}_w(\tau_n(t_n^{cov} + \beta) \geq 2s_n t_n^\mu, t_n^{cov} \leq t_n^{\mu-\alpha} | \mathbf{T}_n) \\ & \rightarrow e^{-e^{-\mu+\alpha}}. \end{aligned}$$

□

*Proof of Theorem 1.1.* By Proposition 4.2, it suffices to estimate  $s_n$ .

For  $1 < \lambda < m$ , one can use (2.4) to show that

$$s_n - \sum_{i=0}^n \frac{m^i}{\lambda^i} W = O\left(\frac{m^{n/2}}{\lambda^{n/2}} + \sum_{i=n/2}^n \frac{m^{i/2} \log n}{\lambda^i}\right) = o\left(\frac{m^n}{n\lambda^n}\right).$$

Therefore,  $s_n$  can be replaced by  $(\frac{m}{\lambda} - 1)^{-1} \frac{m^{n+1}}{\lambda^{n+1}} W$ , and the conclusion follows.

For  $\lambda > m$ , similarly, the difference between  $\sum_{i=0}^{\infty} \frac{Z_i}{\lambda^i}$  and  $s_n$  is negligible,

$$\sum_{i=0}^{\infty} \frac{Z_i}{\lambda^i} - s_n = \sum_{i=n+1}^{\infty} \frac{Z_i}{\lambda^i} = O\left(\frac{m^n}{\lambda^n}\right).$$

For  $\lambda = m$ ,  $s_n = \sum_{i=0}^n Z_i$  follows from its definition in Lemma 4.1. □

**Acknowledgement.** The author would like to thank Yueyun Hu for pointing out this topic, and providing valuable advice. I would also like to thank Yijun Wan and the reviewers for various suggestions on this paper.

## References

- [1] Yoshihiro Abe. Second-order term of cover time for planar simple random walk. *J. Theor. Probab.*, 2020.
- [2] David J. Aldous. Random walk covering of some special trees. *J. Math. Anal. Appl.*, 157(1):271–283, 1991.
- [3] David J. Aldous and James Fill. Reversible markov chains and random walks on graphs, 1995.

- [4] Pierre Andreatti and Pierre Debs. The number of generations entirely visited for recurrent random walks in a random environment. *J. Theor. Probab.*, 27(2):518–538, 2014.
- [5] David Belius and Nicola Kistler. The subleading order of two dimensional cover times. *Probab. Theory Rel.*, 167(1-2):461–552, 2017.
- [6] David Belius, Jay Rosen, and Ofer Zeitouni. Barrier estimates for a critical Galton-Watson process and the cover time of the binary tree. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(1):127–154, 2019.
- [7] David Belius, Jay Rosen, and Ofer Zeitouni. Tightness for the cover time of the two dimensional sphere. *Probab. Theory Relat. Fields*, 176(3-4):1357–1437, 2020.
- [8] Aser Cortines, Oren Luidor, and Santiago Saglietti. A scaling limit for the cover time of the binary tree. 2018. arXiv:1812.10101.
- [9] Amir Dembo, Yuval Peres, Jay Rosen, and Ofer Zeitouni. Cover times for Brownian motion and random walks in two dimensions. *Ann. of Math.*, 160(2):433–464, 2004.
- [10] Amir Dembo, Jay Rosen, and Ofer Zeitouni. Limit law for the cover time of a random walk on a binary tree. 2019. arXiv:1906.07276.
- [11] Jian Ding. On cover times for 2D lattices. *Electron. J. Probab.*, 17:no. 45, 18, 2012.
- [12] Jian Ding, James R. Lee, and Yuval Peres. Cover times, blanket times, and majorizing measures. *Ann. of Math. (2)*, 175(3):1409–1471, 2012.
- [13] Jian Ding and Ofer Zeitouni. A sharp estimate for cover times on binary trees. *Stoch. Proc. Appl.*, 122(5):2117–2133, 2012.
- [14] Nathalie Eisenbaum, Haya Kaspri, Michael B. Marcus, Jay Rosen, and Zhan Shi. A Ray-Knight theorem for symmetric Markov processes. *Ann. Probab.*, 28(4):1781–1796, 2000.
- [15] Uriel Feige. A tight lower bound on the cover time for random walks on graphs. *Random Struct. Algor.*, 6(4):433–438, 1995.
- [16] Uriel Feige. A tight upper bound on the cover time for random walks on graphs. *Random Struct. Algor.*, 6(1):51–54, 1995.

- [17] Christopher C. Heyde and Julian R. Leslie. Improved classical limit analogues for galton-watson processes with or without immigration. *B. Aust. Math. Soc.*, 5(2):145–155, 1971.
- [18] Harry Kesten and Bernt P Stigum. A limit theorem for multidimensional galton-watson processes. *Ann. Math. Statist.*, 37(5):1211–1223, 1966.
- [19] László Lovász. Random walks on graphs: a survey. In *Combinatorics, Paul Erdős is eighty*, pages 353–397. János Bolyai Math. Soc., Budapest, 1993.
- [20] Peter Matthews. Covering problems for Markov chains. *Ann. Probab.*, 16(3):1215–1228, 1988.
- [21] Alex Zhai. Exponential concentration of cover times. *Electron. J. Probab.*, 23:no. 32, 22, 2018.