On Aldous' Cover Time Conjecture

Draft George Andriopoulos, Vlad Margarint

September 24, 2020

Abstract

We prove Aldous' conjecture about the cover times of critical Galton-Watson trees. Let T_n be the uniform random tree chosen from the set of n labeled vertices denoted by $[n] := \{0, 1, ..., n-1\}$ with root at 0. Then, T_n has the conditional distribution of T given $|T| = n$, where T is the family tree of a branching process with offspring distributed as a $Poisson(1)$ random variable. Considering τ^n_{cov} to be the first time taken by the random walk to cover T_n, Aldous [1] showed that $n^{-3/2}\tau^n_{\text{cov}}$ is tight as $n \to \infty$ along some subsequence. This result shows that the limiting diffusion covers the continuum fractal in finite time. We purport to refine this to the distributional convergence of $n^{-3/2}\tau^n_{\text{cov}}$ to the first time taken by the limiting diffusion to cover the continuum fractal, where now τ^n_{cov} is the cover time variable of a critical Galton-Watson tree with finite variance and exponential moments, thus widening the class of examples Aldous is discussing in [1].

1 A Gromov-Hausdorff topology

Let (K, d_K) be a non-empty compact metric space. For a fixed $T > 0$, let X^K be a path in $D([0,T], K)$, the space of càdlàg functions, i.e. right-continuous functions with left-hand limits, from $[0,T]$ to K. We say that a function λ from $[0,T]$ onto itself is a time-change if it is strictly increasing and continuous. Let Λ denote the set of all time-changes. If $\lambda \in \Lambda$, then $\lambda(0) = 0$ and $\lambda(T) = T$. We equip $D([0,T], K)$ with the Skorohod metric d_{J_1} defined as follows:

$$d_{J_1}(x, y) := \inf_{\lambda \in \Lambda} \left\{ \sup_{t \in [0,T]} |\lambda(t) - t| + \sup_{t \in [0,T]} d_K(x(\lambda(t)), y(t)) \right\},$$

for $x, y \in D([0,T], K)$. The idea behind going from the uniform metric to the Skorohod metric d_{J_1} is to say that two paths are close if they are uniformly close in $[0,T]$, after allowing small perturbations of time. Moreover, $D([0,T], K)$ endowed with d_{J_1} becomes a separable metric space (see [4, Theorem 12.2]). Let $\mathcal{P}(K)$ denote the space of Borel probability measures on K. If

$\mu, \nu \in \mathcal{P}(K)$ we set

$$d_P(\mu,\nu) = \inf\{\varepsilon > 0 : \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon, \text{ for any } A \in \mathcal{M}(K)\},$$

where $\mathcal{M}(K)$ is the set of all closed subsets of K. This expression gives the standard Prokhorov metric between μ and ν. Moreover, it is known , see [8] Appendix A.2.5, that $(\mathcal{P}(K), d_P)$ is a Polish metric space, i.e. a complete and separable metric space, and the topology generated by d_P is exactly the topology of weak convergence, the convergence against bounded and continuous functionals.

Let π^K be a Borel probability measure on K and $L^K = (L_t^K(x))_{x\in K, t\in[0,T]}$ be a jointly continuous function of (t,x) taking nonnegative real values. Let $\mathbb{K}$ be the collection of quadruples (K, π^K, X^K, L^K). We say that two elements (K, π^K, X^K, L^K) and $(K', \pi^{K'}, X^{K'}, L^{K'})$ of $\mathbb{K}$ are equivalent if there exists an isometry $f : K \to K'$ such that

- $\pi^K \circ f^{-1} = \pi^{K'}$,
- $f \circ X^K = X^{K'}$, which is a shorthand of $f(X_t^K) = X_t^{K'}$, for every $t \in [0,T]$.
- $L_t^{K'} \circ f = L_t^K$, for every $t \in [0,T]$, which is a shorthand of $L_t^{K'}(f(x)) = L_t^K(x)$, for every $t \in [0,T]$, $x \in K$.

Not to overcomplicate our notation, we will often identify an equivalence class of $\mathbb{K}$ with a particular element of it. We now introduce a distance $d_{\mathbb{K}}$ on $\mathbb{K}$ by setting

$$\begin{aligned} & d_{\mathbb{K}}((K, \pi^K, X^K, L^K), (K', \pi^{K'}, X^{K'}, L^{K'})) \\ & := \inf_{Z,\phi,\phi',\mathcal{C}} \Bigg\{ d_P^Z(\pi^K \circ \phi^{-1}, \pi^{K'} \circ \phi'^{-1}) + d_{J_1}^Z(\phi(X_t^K), \phi'(X_t^{K'})) \\ & + \sup_{(x,x')\in\mathcal{C}} \left(d_Z(\phi(x), \phi'(x')) + \sup_{t\in[0,T]} |L_t^K(x) - L_t^{K'}(x')| \right) \Bigg\}, \end{aligned}$$

where the infimum is taken over all metric spaces (Z, d_Z), isometric embeddings $\phi : K \to Z$, $\phi' : K' \to Z$ and correspondences $\mathcal{C}$ between K and K'. A correspondence between K and K' is a subset of $K \times K'$, such that for every $x \in K$ there exists at least one x' in K' such that $(x,x') \in \mathcal{C}$ and conversely for every $x' \in K'$ there exists at least one $x \in K$ such that $(x,x') \in \mathcal{C}$. In the above expression d_P^Z is the standard Prokhorov distance between Borel probability measures on Z, and $d_{J_1}^Z$ is the Skorohod metric d_{J_1} between càdlàg paths on Z.

Proposition 1.1. $(\mathbb{K}, d_{\mathbb{K}})$ *is a separable metric space.*

Fix $T > 0$. Let $\tilde{\mathbb{K}}$ be the space of quadruples of the form (K, π^K, X^K, L^K), where K is a non-empty compact pointed metric space with distinguished vertex ρ, π^K is a Borel probability measure on K, $X^K = (X_t^K)_{t\in[0,K]}$ is a càdlàg path on K and $L^K = (L_t(x))_{x\in K, t\in[0,T]}$ is a jointly continuous positive real-valued function of (t, x). We say that two elements of $\tilde{\mathbb{K}}$, say (K, π^K, X^K, L^K) and $(K', \pi^{K'}, X^{K'}, L^{K'})$, are equivalent if and only there is a root-preserving isometry $f : K \to K'$, such that $f(\rho) = \rho'$, $\pi^K \circ f^{-1} = \pi^{K'}$, $f \circ X^K = X^{K'}$ and $L_t^{K'} \circ f = L_t^K$, for every $t \in [0, T]$. It is possible to define a metric on the equivalence classes of $\tilde{\mathbb{K}}$ by imposing in the definition of $d_{\mathbb{K}}$ that the infimum is taken over all correspondences that contain (ρ, ρ'). The incorporation of distinguished points to the extended Gromov-Hausdorff topology leaves the proof of Proposition 1.1 unchanged and it is possible to show that $(\tilde{\mathbb{K}}, d_{\tilde{\mathbb{K}}})$ is a separable metric space.

2 Graph theoretic framework

We continue by introducing the graph theoretic framework in which we work. Firstly, let $G = (V(G), E(G))$ be a finite connected graph with at least two vertices, where $V(G)$ denotes the vertex set of G and $E(G)$ denotes the edge set of G. We endow the edge set $E(G)$ with a symmetric weight function $\mu^G : V(G)^2 \to \mathbb{R}_+$ that satisfies $\mu^G_{xy} > 0$ if and only if $\{x, y\} \in E(G)$. Now, the weighted random walk associated with (G, μ^G) is the Markov chain $((X_t^G)_{t \ge 0}, \mathbf{P}_x^G, x \in V(G))$ with transition probabilities $(P_G(x, y))_{x,y \in V(G)}$ given by

$$P_G(x, y) := \frac{\mu^G_{xy}}{\mu^G_x},$$

where $\mu^G_x = \sum_{y \in V(G)} \mu^G_{xy}$. One can easily check that this Markov chain is reversible and has stationary distribution given by

$$\pi^G(A) := \frac{\sum_{x \in A} \mu^G_x}{\sum_{x \in V(G)} \mu^G_x},$$

for every $A \subseteq V(G)$. The process X^G has corresponding local times $(L_t^G(x))_{x \in V(G), t \ge 0}$ given by $L_0^G(x) = 0$, for every $x \in V(G)$, and, for $t \ge 1$

$$L_t^G(x) := \frac{1}{\mu^G_x} \sum_{i=0}^{t-1} \mathbf{1}_{\{X_i^G = x\}}, \tag{2.1}$$

for every $x \in V(G)$. The simple random walk on this graph is a Markov chain with transition probabilities $(P(x, y))_{x,y \in V(G)}$ given by

$$P(x, y) := 1/\deg(x),$$

where $\deg(x) := |\{y \in V(G) : \{x, y\} \in E(G)\}|$. The simple random walk is reversible and has stationary distribution given by

$$\pi(A) := \frac{\sum_{x \in A} \deg(x)}{2|E(G)|},$$

for every $A \subseteq V(G)$. It has corresponding local times as in (2.1) normalized by $\deg(x)$.

To endow G with a metric, we can choose d_G to be the shortest path distance, which collects the total weight accumulated in the shortest path between a pair of vertices in G. But this is not the most convenient choice in many cases. Another typical graph distance that arises from the view of G as an electrical network equipped with conductances $(\mu^G_{xy})_{\{x,y\}\in E(G)}$ is the so-called resistance metric. For $f, g : V(G) \to \mathbb{R}$ let

$$\mathcal{E}_G(f,g) := \frac{1}{2} \sum_{\substack{x,y\in V(G):\\ \{x,y\}\in E(G)}} (f(x)-f(y))(g(x)-g(y))\mu^G_{xy} \tag{2.2}$$

denote the Dirichlet form associated with the process X^G. Note that the sum in the expression above counts each edge twice. One can give the following interpretation of $\mathcal{E}_G(f,f)$ in terms of electrical networks. Given a voltage f on the network, the current flow I associated with f is defined as $I_{xy} := \mu^G_{xy}(f(x)-f(y))$, for every $\{x,y\} \in E(G)$. Then, the energy dissipation of a wire per unit time connecting x and y is $\mu^G_{xy}(f(x)-f(y))^2$. So, $\mathcal{E}_G(f,f)$ is the total energy dissipation of G. We define the resistance operator on disjoint sets $A, B \in V(G)$ through the formula

$$R_G(A,B)^{-1} := \inf\{\mathcal{E}_G(f,f) : f : V(G) \to \mathbb{R}, f|_A = 0, f|_B = 1\}. \tag{2.3}$$

Now, the distance on the vertices of G defined by $R_G(x,y) := R_G(\{x\},\{y\})$, for $x \neq y$, and $R_G(x,x) := 0$ is indeed a metric on the vertices of G. For a proof and a treatise on random walks on electrical networks see [12, Chapter 9].

Secondly, let (K, d_K) be a compact metric space and let π be a Borel probability measure of full support on (K, d_K). Take $((X_t)_{t\geq 0}, \mathbf{P}_x, x \in K)$ to be a π-symmetric Hunt process that admits jointly continuous local times $(L_t(y))_{y\in K, t\geq 0}$. A Hunt process is a strong Markov process that possesses useful properties such as the right-continuity and the existense of the left limits of sample paths (for definitions and other properties see [10, Appendix A.2]).

For a weighted random walk, X^G, associated with a graph G, and its weight function μ^G, we define its cover time variable as the first time in which all the states of the chain have been visited. Alternatively,

$$\tau^G_{\mathrm{cov}} := \inf\{t \geq 0 : \{X^G_0, \dots, X^G_t\} = V(G)\}. \tag{2.4}$$

the first time in which all the states of the chain have been visited. The mean of τ^G_{cov} from the worst-case initial state $t^G_{\mathrm{cov}} := \max_{x\in V(G)} \mathbf{E}_x \tau^G_{\mathrm{cov}}$ is known as the cover time of G. The order of growth of the cover time has been computed for various families of graphs.

Theorem 2.1 (**Aldous [1]**). *Let T be a critical Galton-Watson tree (with finite variance σ^2 for the offspring distribution). For any $\delta > 0$ there exists $A = A(\delta, \sigma^2) > 0$ such that*

$$P(A^{-1}n^{3/2} \leq t^T_{\mathrm{cov}} \leq An^{3/2} \,|\, |T| \in [n, 2n]) \geq 1 - \delta.$$

A connection has been made between the cover time and the maximum of the Gaussian Free Field (GFF) for any general graph.

Theorem 2.2 (**Ding, Lee, Peres [9]**). *For any finite connected graph $G = (V(G), E(G))$ with at least two vertices, using the notation $\asymp$ to denote equivalence up to universal constant factors,*

$$t_{\mathrm{cov}}^{G} \asymp |E(G)| \left(\mathbf{E} \max_{x \in V(G)} \eta_x \right)^2 ,$$

where $(\eta_x)_{x \in V(G)}$ is a centered Gaussian process with $\eta_{x_0} = 0$, for some $x_0 \in V(G)$, and

$$\left(\mathbf{E}(\eta_x - \eta_y)^2\right)_{x,y \in V(G)} = (R_G(\,x, y))_{x,y \in V(G)}.$$

Let us first introduce another version of the cover time variable. Let

$$\tilde{\tau}_{\mathrm{cov}}^{G} := \inf\{t \geq 0 : L_t^G(x) > 0,\ \forall x \in V(G)\}. \tag{2.5}$$

An observation to be made is that, at least in the discrete setting, the first time in which the local times of the random walk become strictly positive everywhere, gives the cover time variable of G. More specifically, $\tilde{\tau}_{\mathrm{cov}}^{G} = \tau_{\mathrm{cov}}^{G} + 1$, since $L_0^G(x) = 0$, $\forall x \in V(G)$. For the π-symmetric Hunt process X we define τ_{cov} and $\tilde{\tau}_{\mathrm{cov}}$ analogously to (2.4) and (2.5). One can check that $0 < \tau_{\mathrm{cov}} \leq \tilde{\tau}_{\mathrm{cov}} < \infty$, almost-surely, and therefore the expressions τ_{cov} and $\tilde{\tau}_{\mathrm{cov}}$ are non-trivial (see [3, Lemma 6.2] and cf. the proof of [3, Theorem 6.3]).

The following assumption encodes the information that, properly rescaled, the discrete state spaces, invariant measures, random walks, and local times, converge. This formulation will be described in terms of the extended Gromov-Hausdorff topology constructed above.

Assumption 1. *Let $(G^n)_{n \geq 1}$ be a sequence of finite connected graphs that have at least two vertices, such that*

$$((V(G^n), d_{G^n}, \rho^n), \pi^n, X^n, L^n) \longrightarrow ((K, d_K, \rho), \pi, X, L)$$

in the sense of the extended pointed Gromov-Hausdorff topology, for distinguished points $\rho^n \in V(G^n)$ and $\rho \in K$ at which X^n and X start respectively. In the above expression the definition of the discrete local times is extended to all positive times by linear interpolation.

The aim of the following lemmas is to establish a sufficient condition for Assumption 1 to hold, as well as to show that if Assumption 1 holds then we can isometrically embed the rescaled graphs, measures, random walks and local times into a common metric space such that they all converge to the relevant objects. To be more precise we formulate this last statement in the next lemma.

Lemma 2.3. *If Assumption 1 holds, then we can find isometric embeddings of $(V(G^n), d_{G^n})_{n \geq 1}$ and (K, d_K) into a common metric space (F, d_F) such that*

$$\lim_{n \to \infty} d_H^F(V(G^n), K) = 0, \quad \lim_{n \to \infty} d_P^F(\pi^n, \pi) = 0, \quad \lim_{n \to \infty} d_F(\rho^n, \rho) = 0, \tag{2.6}$$

and in such a way that if X^n is started from ρ^n, X started from ρ, then

$$(X_t^n)_{t\in[0,T]} \stackrel{(d)}{\to} (X_t)_{t\in[0,T]}. \tag{2.7}$$

Also, if the finite collections $(x_i^n)_{i=1}^k$ in $V(G^n)$, for $n \geq 1$, are such that $d_F(x_i^n, x_i) \to 0$, as $n \to \infty$, for some $(x_i)_{i=1}^k$ in K, then it holds that

$$(L_t^n(x_i^n))_{i=1,\dots,k,t\in[0,T]} \stackrel{(d)}{\to} (L_t(x_i))_{i=1,\dots,k,t\in[0,T]}. \tag{2.8}$$

For simplicity we have identified the measures and the random walks in $V(G^n)$ with their isometric embeddings in (F, d_F).

Lemma 2.4. *Suppose that* (2.6)-(2.8) *hold. Then so does Assumption 1.*

3 Local time equicontinuity

To check that Assumption 1 holds we need to verify that the convergence of the rescaled local times in (2.8), as suggested by Lemma 2.4. Due to work done in a more general framework in [7], we can weaken the local convergence statement of (2.8) and replace it by the equicontinuity condition of (3.2). In (2.3) we defined a resistance metric on a graph viewed as an electrical network. Next, we give the definition of a regular resistance form and its associated resistance metric for arbitrary non-empty sets, which is a combination of [7, Definition 2.1] and [7, Definition 2.2].

Definition 3.1. *Let K be a non-empty set. A pair $(\mathcal{E}, \mathcal{K})$ is called a regular resistance form on K if the following six conditions are satisfied.*

(i) *$\mathcal{K}$ is a linear subspace of the collection of functions $\{f : K \to \mathbb{R}\}$ containing constants, and $\mathcal{E}$ is a non-negative symmetric quadratic form on $\mathcal{K}$ such that $\mathcal{E}(f,f) = 0$ if and only if f is constant on K.*

(ii) *Let $\sim$ be an equivalence relation on $\mathcal{K}$ defined by saying $f \sim g$ if and only if the difference $f - g$ is constant on K. Then, $(\mathcal{K}/\sim, \mathcal{E})$ is a Hilbert space.*

(iii) *If $x \neq y$, there exists $f \in \mathcal{K}$ such that $f(x) \neq f(y)$.*

(iv) *For any $x, y \in K$,*

$$R(x,y) := \sup\left\{\frac{|f(x) - f(y)|^2}{\mathcal{E}(f,f)} : f \in \mathcal{K},\ \mathcal{E}(f,f) > 0\right\} < \infty. \tag{3.1}$$

(v) *For any $f \in \mathcal{K}$, $\bar{f} := (f \wedge 1) \vee 0 \in \mathcal{K}$, and $\mathcal{E}(\bar{f}, \bar{f}) \leq \mathcal{E}(f,f)$.*

(vi) *If $\mathcal{K} \cap C_0(K)$ is dense with respect to the supremum norm on $C_0(K)$, where $C_0(K)$ denotes the space of compactly supported, continuous (with respect to R) functions on K.*

It is the first five conditions that have to be satisfied in order for the pair $(\mathcal{E},\mathcal{K})$ to define a resistance form. If in addition the sixth condition is satisfied then $(\mathcal{E},\mathcal{K})$ defines a regular resistance form. Note that the fourth condition can be rewritten as

$$R(x,y)^{-1} = \inf\{\mathcal{E}(f,f) : f : K \to \mathbb{R}, f(x) = 0, f(y) = 1\},$$

and it can be proven that it is actually a metric on K (see [11, Proposition 3.3]). It also clearly resembles the effective resistance on $V(G)$ as defined in (2.3). More specifically, taking $\mathcal{K} := \{f : V(G) \to \mathbb{R}\}$ and $\mathcal{E}_G$ as defined in (2.2) one can prove that the pair $(\mathcal{E}_G,\mathcal{K})$ satisfies the six conditions of Definition 3.1, and therefore is a regular resistance form on $V(G)$ with associated resistance metric given by (2.3). For a detailed proof of this fact see [10, Example 1.2.5]. Finally, in this setting given a regular Dirichlet form, standard theory gives us the existence of an associated Hunt process $X = ((X_t)_{t\geq 0}, \mathbf{P}_x, x \in K)$ that is defined uniquely everywhere (see [10, Theorem 7.2.1] and [11, Theorem 9.9]).

Suppose that the discrete state spaces $(V(G^n))_{n\geq 1}$ are equipped with resistances $(R_{G^n})_{n\geq 1}$ as defined in (2.3) and that the limiting non-empty metric space K, that appears in Assumption 1, is equipped with a resistance metric R as in Definition 3.1, such that

- (K,R) is compact,
- π is a Borel probability measure of full support on (K,R),
- $X = ((X_t)_{t\geq 0}, \mathbf{P}_x, x \in K)$ admits jointly continuous (with respect to R) local times $L = (L_t(y))_{y\in K, t\geq 0}$.

In the following extra assumption we input the information encoded in the first three conclusions of Lemma 2.3.

Assumption 2. *Let $(G^n)_{n\geq 1}$ be a sequence of finite connected graphs that have at least two vertices, such that*

$$((V(G^n), R_{G^n}, \rho^n), \pi^n, X^n) \longrightarrow ((K,R,\rho), \pi, X)$$

in the sense of the extended pointed Gromov-Hausdorff topology, for distinguished points $\rho^n \in V(G^n)$ and $\rho \in K$ at which X^n and X start respectively. Furthermore, suppose that for every $\varepsilon > 0$ and $T > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbf{P}^n_{\rho^n} \left(\sup_{\substack{y,z\in V(G^n):\\ R_{G^n}(y,z)<\delta}} \sup_{t\in[0,T]} |L^n_t(y) - L^n_t(z)| \geq \varepsilon \right) = 0. \tag{3.2}$$

It is Assumption 2 we have to verify in the examples of random graphs we will consider later. As we prove below in the last lemma of this section, if Assumption 2 holds then the finite dimensional local times converge in distribution.

Lemma 3.2. *Suppose that Assumption 2 holds. Then, if the finite collections* $(x_i^n)_{i=1}^k$ *in* $V(G^n)$*, for* $n \geq 1$*, are such that* $d_F(x_i^n, x_i) \to 0$*, as* $n \to \infty$*, for some* $(x_i)_{i=1}^k$ *in* K*, then it holds that*

$$(L_t^n(x_i^n))_{i=1,\ldots,k,t\in[0,T]} \xrightarrow{(d)} (L_t(x_i))_{i=1,\ldots,k,t\in[0,T]}.$$

In other words, (2.8) *is satisfied.*

4 Cover time-scaling and distributional bounds

In this section, we show that under Assumption 1, and as a consequence of the local time convergence in Lemma 3.2, we are able to establish asymptotic bounds on the distribution of the cover times of the graphs in the sequence. The argument for the cover time-scaling was provided first in [6, Corollary 7.3] by restricting to the unweighted Sierpiński gasket graphs. Though, the argument is applicable to any other model as long as the relevant asssumptions are satisfied. Namely, it is an implication of local time equicontinuity for sequences of graphs for which the associated random walks admit a scaling limit.

Theorem 4.1. *Suppose that Assumption 1 holds and that* X *has (almost-surely) continuous paths. Then, for every* $t \in [0, T]$*,*

$$\limsup_{n\to\infty} \mathbf{P}_{\rho^n}^n \left(\beta_n^{-1}\tau_{\text{cov}}^n \leq t\right) \leq \mathbf{P}_\rho \left(\tau_{\text{cov}} \leq t\right), \tag{4.1}$$

$$\liminf_{n\to\infty} \mathbf{P}_{\rho^n}^n \left(\beta_n^{-1}\tau_{\text{cov}}^n \leq t\right) \geq \mathbf{P}_\rho \left(\tilde{\tau}_{\text{cov}} < t\right), \tag{4.2}$$

where $\mathbf{P}_\rho$ *is the law of* X *on* K*, started from* ρ*.*

Proof. Given that $(V(G^n), \alpha_n d_{G^n})_{n\geq 1}$ and (K, d_K) can be isometrically embedded into a common metric space (F, d_F) such that $\alpha_n X^n$ under $\mathbf{P}_{\rho^n}^n$ converges weakly to the law of X under $\mathbf{P}_\rho$ on $C([0,T], F)$, we can couple X^n started from ρ^n and X started from ρ into a common probability space such that $(\alpha_n X_{\beta_n t}^n)_{t\in[0,T]} \to (X_t)_{t\in[0,T]}$ in $C([0,T], F)$, almost-surely. Let $t \in [0,T]$. Suppose that $t < \tau_{\text{cov}}$. Then, there exists a $y \in K$ such that $y \notin \{X_s : 0 \leq s \leq t\}$. Since $X \in C([0,T], K)$, there exists $\varepsilon > 0$ such that

$$B_F(y, \varepsilon) \cap \{X_s : 0 \leq s \leq t\} = \emptyset.$$

We conclude that, for n large enough,

$$B_F\left(y, \varepsilon/2\right) \cap \{\alpha_n X_{\beta_n s}^n : 0 \leq s \leq t\} = \emptyset.$$

Thus, $\beta_n t \leq \tau_{\text{cov}}^n$ for large n and consequently $\tau_{\text{cov}} \leq \liminf_{n\to\infty} \beta_n^{-1}\tau_{\text{cov}}^n$, from which (4.1) follows. Assume now that $\tilde{\tau}_{\text{cov}} < t$. Since $(L_t(x))_{x\in K,t\in[0,T]}$ is increasing in t, it is the case that $L_t(x) > 0$, for every $x \in K$. From the continuity of $(L_t(x))_{x\in K,t\in[0,T]}$ in x, there exist an $\varepsilon > 0$ such that

$L_t(x) > \varepsilon$. As in the previous paragraph, using the Skorohod represantation theorem, we can assume that the conclusion of Lemma 3.2 holds in an almost-surely sense. From (2.6), for every $x^n \in V(G^n)$, there exists an $x \in K$ such that, for n large enough, $d_F(x^n, x) < 2\varepsilon$. Then, the local convergence at (2.8) implies that, for n large enough

$$\alpha_n L^n_{\beta_n t}(x^n) \geq L_t(x) - \varepsilon/2.$$

Therefore, for every $x^n \in V(G^n)$ and large enough n, it follows that $\alpha_n L^n_{\beta_n t}(x^n) > \varepsilon/2$. It holds that $\beta_n t \geq \tilde{\tau}^n_{\text{cov}}$ for large n and consequently

$$\limsup_{n\to\infty} \beta_n^{-1}\tau^n_{\text{cov}} = \limsup_{n\to\infty} \beta_n^{-1}(\tilde{\tau}^n_{\text{cov}} - 1) = \limsup_{n\to\infty} \beta_n^{-1}\tilde{\tau}^n_{\text{cov}} \leq \tilde{\tau}_{\text{cov}},$$

from which (4.2) follows. □

Corollary 4.2. *Let $\mathcal{T}_n$ be a critical Galton-Watson tree (with finite variance and exponential moments for the aperiodic offspring distribution) conditioned to have total progeny n. If τ^n_{cov} is the cover time variable of the simple random walk on $\mathcal{T}_n$, started from its root ρ^n, then*

$$n^{-3/2}\tau^n_{\text{cov}} \xrightarrow{(d)} [\tau_{\text{cov}}, \tilde{\tau}_{\text{cov}}],$$

(which a shorthand of (4.1) *and* (4.2)*), where $\tau_{\text{cov}}, \tilde{\tau}_{\text{cov}} \in (0,\infty)$ are defined according to* (2.4) *and* (2.5) *respectively.*

Proof. In [2], it was verified that Assumption 2 holds. More specifically, see [2, Section 4.1, (4.3)] and [2, Proposition 4.4]. As a consequence of Lemma 3.2, Assumption 1 holds as well, and the desired result follows from Theorem 4.1. □

5 Line-breaking construction of the CRT

5.1. $\mathbb{R}$-trees.

Definition 5.1. *A metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if for every $x, y \in \mathcal{T}$ the following hold.*

(i) There exists a unique isometry $f_{x,y} : [0, d(x,y)] \to \mathcal{T}$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x,y)) = y$.

(ii) If $q : [0,1] \to \mathcal{T}$ is continuous and injective such that $q(0) = x$ and $q(1) = y$, then $q([0,1]) = f_{x,y}([0, d(x,y)])$.

Definition 5.2. *A continuum tree is a triple $(\mathcal{T}, d, \mu)$, where $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree with leaves*

$$\mathcal{L}(\mathcal{T}) := \{x \in \mathcal{T} \setminus \{\rho\} : \mathcal{T} \setminus \{x\} \text{ is connected}\}$$

and μ is a probability measure on $\mathcal{T}$, which satisfies:

- *μ is non-atomic,*
- *μ is supported on $\mathcal{L}(\mathcal{T})$, i.e. $\mu(\mathcal{L}(\mathcal{T})) = 1$.*
- *For every $x \in \mathcal{T}$ of degree $k \geq 3$, let $\mathcal{T}_1$,...,$\mathcal{T}_k$ be the connected components of $\mathcal{T} \setminus \{x\}$. Then, $\mu(\mathcal{T}_i) > 0$, for all $1 \leq i \leq k$.*

Ever since Aldous, such trees have been viewed in several equivalent ways. In the case of the Brownian Continuum Random tree (CRT), we may view $\mathcal{T}$ as the $\mathbb{R}$-tree coded by a standard Brownian excursion. Another way to build $\mathcal{T}$ is inductively starting from a single line-segment $[\rho, \Sigma_0]$, grafting further line-segments $(J_{k-1}, \Sigma_k]$ to build trees $\mathcal{T}_k$ spanned by a growing number of finite points $\rho, \Sigma_0, ..., \Sigma_k,\ k \geq 1$, finally taking $\mathcal{T}$ to be the closure of $\bigcup_{k\geq 0} \mathcal{T}_k$. For the CRT, Aldous' autonomous construction describes the tree process $(\mathcal{T}_k, k \geq 0)$ by breaking $\mathbb{R}_+$ at the points $(R_k, k \geq 0)$ of a non-homogeneous Poisson point process with intensity tdt on $\mathbb{R}_+$. With $\mathcal{T}_0 = [0, R_0]$, each line-segment $(R_k, R_{k+1}]$ is grafted at a uniformly chosen point J_k (according to the normalized length measure of the structure $\mathcal{T}_k$ that has already been built). Note that the branch points $(J_k, k \geq 0)$ are distinct and the trees binary. Thinning of Poisson point processes shows that every line-segment receives a dense set of branch points.

Algorithm 5.3. Let $(R_k, k \geq 0)$ be the times of a non-homogeneous Poisson point process with intensity tdt on $\mathbb{R}_+$. For example,

$$\mathbb{P}(R_1 > t) = \mathbb{P}(\text{no point "fell" in } [0,t]) = e^{-\int_0^t s ds} = e^{-t^2/2}.$$

Equivalently, one could start with $(E_k, k \geq 1)$ independent random variables drawn from an $Exp(1/2)$, and set

$$R_k^2 = \sum_{i=1}^{k} E_i.$$

(Note that $P(R_1 > t) = P(E_1 > t^2) = e^{-t^2/2}$.) We grow discrete $\mathbb{R}$-trees $\mathcal{T}_k$ with a sequence of distinct branch points $(J_i)_{i=0}^{k-1}$ and $k+1$ leaves. The edge lengths between vertices are chosen according to the following procedure.

- Let $(\mathcal{T}_0, \rho)$ be isometric to $([0, R_0], 0)$, labeling the first leaf by 1.
- Given $(\mathcal{T}_j, (J_i)_{i=0}^{j-1})$, $0 \leq j \leq k$, and $L_k =: \text{Leb}(\mathcal{T}_k)$, select an edge $e \in E(\mathcal{T}_k)$ with probability proportional to its length.
- If an edge e is selected, sample J_k from the normalized length measure of e.
- Attach to $\mathcal{T}_k$ at J_k a new line-segment of length $R_{k+1} - R_k$ to form $\mathcal{T}_{k+1}$, labeling the new leaf by $k+1$.

At each step, Algorithm 5.3 induces a rooted (finite) discrete tree with edge-lengths. This object is determined by its combinatorial structure, which is a binary tree with a root that has degree 1, and a sequence of (random) lengths indexed by the set of edges of the combinatorial structure. Such a tree with edge-lengths is turned into an $\mathbb{R}$-tree by viewing its edges as line-segments.

6 Versions of the First Ray-Knight theorem

Let T_0 be the hitting time to ρ, which is the root of $\mathcal{T}$.

Lemma 6.1. *For every $x, y \in \mathcal{T}$,*

$$\mathbf{E}_x[L_{T_0}(y)] = 2d_{\mathcal{T}}(\rho, x \wedge y),$$

where $x \wedge y$ is the unique branch point between x, y and ρ.

For every $x \in \mathcal{T}$, $[[\rho, x]]$ is the path that connects the root to x, which we interpret as the ancestral line of vertex x. More specifically, we can define a partial order on the tree by setting $x \preceq x'$ (x is an ancestor of x') if and only if $x \in [[\rho, x']]$.

Theorem 6.2 (**Ray-Knight (I)**)**.** *Denote by $G := \{G_x : x \in \mathcal{T}\}$ a mean zero Gaussian process with covariance $d_{\mathcal{T}}(\rho, x \wedge y)$. Then, under $\mathbf{P}_{x_0} \times \mathbf{P}_{G,\bar{G}}$,*

$$\{L_{T_0}(y) : y \in [[\rho, x_0]]\} \overset{(d)}{=} \left\{G_y^2 + \bar{G}_y^2 : y \in [[\rho, x_0]]\right\},$$

where $(G, \bar{G})$ is a vector of independent copies of G.

Proof. Following the strategy in [14], we show that the two processes have the same finite-dimensional moment generating functions. Let $\rho \preceq y_1 \preceq \cdots \preceq y_n = x_0$, with $\preceq$ defined before. Let Σ be the matrix with elements $\Sigma_{i,j} = \mathbf{E}_{y_i}[L_{T_0}(y_j)]$, $i, j = 1, ..., n$. Let Λ be the matrix with elements $\Lambda_{i,j} = \lambda_i \delta_{i,j}$. Since T_0 is a terminal time, we apply Lemma 3.10.2 of [14]. For all $\lambda_1, ..., \lambda_n$ sufficiently small and $l = n$,

$$\mathbf{E}_{y_n} \exp\left(\sum_{i=1}^{n} \lambda_i L_{T_0}(y_i)\right) = \frac{\det(I - \widehat{\Sigma}\Lambda)}{\det(I - \Sigma\Lambda)},$$

where

$$\widehat{\Sigma}_{i,j} = \Sigma_{i,j} - \Sigma_{n,j}, \qquad i, j = 1, \ldots, n.$$

For $j \leqslant i \leqslant n$,

$$\begin{aligned}\widehat{\Sigma}_{ij} &= \mathbf{E}_{y_i}[L_{T_0}(y_j)] - \mathbf{E}_{y_n}[L_{T_0}(y_j)] \\ &= 2d_{\mathcal{T}}(\rho, y_i \wedge y_j) - 2d_{\mathcal{T}}(\rho, y_n \wedge y_j) = 2d_{\mathcal{T}}(\rho, y_j) - 2d_{\mathcal{T}}(\rho, y_j) = 0,\end{aligned}$$

i.e. $(I-\widehat{\Sigma}\Lambda)$ is an upper triangular matrix with ones on the diagonal. Therefore, $\det(I-\widehat{\Sigma}\Lambda)=1$. Given the covariance of our Guassian, we obtain that

$$\begin{aligned}\mathbf{E}_{y_n}\exp\left(\sum_{i=1}^n \lambda_i G_{y_i}^2\right) &= \frac{1}{(2\pi)^{n/2}(\det\Sigma)^{1/2}}\int e^{\sum_{i=1}^n \lambda_i y_i^2/2-\left(y,\Sigma^{-1}y\right)/2}dy \\ &= \frac{1}{(2\pi)^{n/2}(\det\Sigma)^{1/2}}\int e^{-\left(y,\left(\Sigma^{-1}-\Lambda\right)y\right)/2}dy \\ &= \left(\frac{\det\left(\Sigma^{-1}-\Lambda\right)^{-1}}{\det\Sigma}\right)^{1/2} \\ &= (\det(I-\Sigma\Lambda))^{-1/2}.\end{aligned}$$

The conclusion follows.

Remark 6.3. We notice that the root ρ of the CRT does not play any particular role: had we chosen to distinguish any other "fixed" point, the resulting object would have had the same law. In particular, for every $t\in[0,1]$, the random variables $(\mathcal{T},d_{\mathcal{T}},\rho)$ and $(\mathcal{T},d_{\mathcal{T}},p_e(t))$ have the same distribution.

Together with the exchangeability properties of the Brownian tree, one can go a step further and argue that the root is chosen uniformly at random in the tree. More specifically, the $\mathbb{R}$-tree encoded by a standard Brownian excursion comes with a canonical object, which is the probability measure given by the image of Lebesgue's measure on $[0,1]$ by the canonical projection $p_e:[0,1]\to\mathcal{T}$. This measure will be denoted by $\mu_{\mathcal{T}}$.

Let X denote a $\mu_{\mathcal{T}}$-distributed random variable. Then, $(\mathcal{T},d_{\mathcal{T}},\rho)$ and $(\mathcal{T},d_{\mathcal{T}},X)$ have the same distribution.

□

Theorem 6.4 (**Ray-Knight (II)**). *Denote by $G:=\{G_x:x\in\mathcal{T}\}$ a mean zero Gaussian process with covariance $d_{\mathcal{T}}(\rho,x\wedge y)$. Then, under $\mathbf{P}_b\times\mathbf{P}_{G,\bar{G}}$,*

$$\{L_{T_0}(y)+[(G_y-G_b)^2+(\bar{G}_y-\tilde{G}_b)^2]\mathbf{1}\{y\in\mathcal{T}[b]\}:y\in\mathcal{T}_2\}\overset{(d)}{=}\{G_y^2+\bar{G}_y^2:y\in\mathcal{T}_2\},$$

where $(G,\bar{G})$ is a vector of independent copies of G.

Proof. We follow again the approach in [14]. Let

$$y_0=\rho\preceq y_1\preceq\cdots\preceq y_l=b,\qquad y_{l+1},...,y_{n-1}\in\mathcal{T}[b],\qquad y_n=x_2.$$

We show that the moment generating functions of

$$\{L_{T_0}(y_i)+[(G_{y_i}-G_b)^2+(\bar{G}_{y_i}-\tilde{G}_b)^2]\mathbf{1}\{y_i\in\mathcal{T}[b]\}:i=1,...,n\}$$

and $\{G_{y_i}^2 + \bar{G}_{y_i}^2 : i = 1, ..., n\}$ are equal. Using Lemma 3.10.2 of [14], we have that

$$\mathbf{E}_{y_l} \exp\left(\sum_{i=1}^{n} \lambda_i L_{T_0}(y_i)\right) = \frac{\det(I - \widehat{\Sigma}\Lambda)}{\det(I - \Sigma\Lambda)},$$

where

$$\widehat{\Sigma}_{i,j} = \Sigma_{i,j} - \Sigma_{l,j}, \qquad i, j = 1, \ldots, n.$$

In order to complete the proof, we need to show

$$\mathbf{E}_{y_l} \exp\left(\sum_{i=l+1}^{n} \lambda_i (G_{y_i} - G_b)^2\right) = \frac{1}{\sqrt{\det(I - \widehat{\Sigma}\Lambda)}}. \tag{6.1}$$

For $1 \leqslant j \leqslant i \wedge l$,

$$\begin{aligned} \widehat{\Sigma}_{i,j} &= \mathbf{E}_{y_i}[L_{T_0}(y_j)] - \mathbf{E}_{y_l}[L_{T_0}(y_j)] \\ &= 2d_{\mathcal{T}}(\rho, y_i \wedge y_j) - 2d_{\mathcal{T}}(\rho, y_l \wedge y_j) = 2d_{\mathcal{T}}(\rho, y_j) - 2d_{\mathcal{T}}(\rho, y_j) = 0. \end{aligned}$$

For $l < i \leqslant n$ $(i \leqslant j)$,

$$\begin{aligned} \widehat{\Sigma}_{i,j} &= \mathbf{E}_{y_i}[L_{T_0}(y_j)] - \mathbf{E}_{y_l}[L_{T_0}(y_j)] \\ &= 2d_{\mathcal{T}}(\rho, y_i \wedge y_j) - 2d_{\mathcal{T}}(\rho, y_l \wedge y_j) = \begin{cases} 2d_{\mathcal{T}}(b, y_i), & y_i \preceq y_j, \\ 0, & y_i \in [[b, x_1]],\ y_j \in [[b, x_2]], \\ 0, & y_i \in [[b, x_2]],\ y_j \in [[b, x_1]]. \end{cases} \end{aligned}$$

For $l < j \leqslant n$ $(j \leqslant i)$,

$$\begin{aligned} \widehat{\Sigma}_{i,j} &= \mathbf{E}_{y_i}[L_{T_0}(y_j)] - \mathbf{E}_{y_l}[L_{T_0}(y_j)] \\ &= 2d_{\mathcal{T}}(\rho, y_i \wedge y_j) - 2d_{\mathcal{T}}(\rho, y_l \wedge y_j) = \begin{cases} 2d_{\mathcal{T}}(b, y_j), & y_j \preceq y_i, \\ 0, & y_i \in [[b, x_1]],\ y_j \in [[b, x_2]], \\ 0, & y_i \in [[b, x_2]],\ y_j \in [[b, x_1]], \end{cases} \end{aligned}$$

by symmetry. In any case,

$$\widehat{\Sigma}_{i,j} \in \{2d_{\mathcal{T}}(b, y_i), 2d_{\mathcal{T}}(b, y_j), 0\},$$

for $l < i \wedge j \leqslant n$.

We stress that $\widehat{\Sigma}$ has at least one non-zero entry, when $l < i \wedge j \leqslant n$. Since we are proving a version of the First Ray-Knight Theorem on $\mathcal{T}_2$, without loss of generality we are allowed to choose at least a pair of y_i's and y_j's that belongs to separate connected components emanating from b. Choosing all the y_i's and y_j's to belong either to $[[b, x_1]]$ or $[[b, x_2]]$ is covered by **Ray-Knight (I)**.

$$\det(I - \bar{\Sigma}\Lambda) = \det(I - \widehat{\Sigma}\Lambda) = 1,$$

where $\bar{\Sigma}$ is an $(n-l)\times(n-l)$ matrix with $\bar{\Sigma}_{i,j} = \widehat{\Sigma}_{i,j}$, $l < i \wedge j \leqslant n$. Equation (6.1) follows from this and the computation in the end of the proof of the previous version of Theorem 6.2.

□

7 Processes on $\mathcal{T}_k$

The Brownian motion on the CRT X has jointly continuous local times L. For $k \geq 1$, define the continuous additive functional

$$A_t^{(k)} := \int_{\mathcal{T}_k} L_t(x)\mu_{\mathcal{T}_k}(dx),$$

where $\mu_{\mathcal{T}_k}$ is the measure that satisfies

$$\mu_{\mathcal{T}_k}([[x,y]]) = \frac{d_{\mathcal{T}}(x,y)}{\mathrm{diam}\mathcal{T}_k}, \qquad \forall x, y \in \mathcal{T}_k.$$

We normalized $\mu_{\mathcal{T}_k}$ so that it is a probability measure. This was indeed possible given the finiteness of $\mathrm{diam}\mathcal{T}_k$, which in turn was a consequence of the compactness of $\mathcal{T}_k$.

Lemma 7.1 ([5])**.** *For $k \geq 1$, set*

$$\tau^{(k)}(t) := \inf\{s > 0 : A_s^{(k)} > t\}$$

to be the right-continuous inverse of the additive functional $A^{(k)}$. The time-changed process

$$B_t^{(k)} := X_{\tau^{(k)}(t)}$$

is the Brownian motion on $(\mathcal{T}_k, d_{\mathcal{T}}, \mu_{\mathcal{T}_k})$. Furthermore, $B^{(k)}$ admits jointly continuous local times $L^{(k)}$, such that

$$L_t^{(k)}(x) = L_{\tau^{(k)}(t)}(x), \qquad \forall t \geq 0, x \in \mathcal{T}_k.$$

Lemma 7.2 ([5])**.** *Fix $T > 0$. If the processes $(B_t^{(k)})_{k\geq 1}$ and X are coupled in such a way that Lemma 7.1 suggests, then a.s.,*

$$(B_t^{(k)})_{t\in[0,T]} \to (X_t)_{t\in[0,T]}, \qquad k \to \infty,$$

in $C([0,T],\mathcal{T})$.

8 The steps

Lemma 8.1. *Prove that*

$$\tau_{\mathrm{cov}}(\mathcal{T}_k) \to \tau_{\mathrm{cov}}(\mathcal{T}),$$

in probability or almost-surely.

Proof. For the proof of this Lemma I need to check [6].

□

Lemma 8.2 (Moduli of continuity for the local times in $\mathcal{T}$)**.** *Fix $T > 0$. Prove that*

$$|L_t^{(k)}(x) - L_t^{(k)}(t)| \leq C d_{\mathcal{T}}(x,y)^{1/2} \log d_{\mathcal{T}}(x,y)^{-1} \leq C d_{\mathcal{T}}(x,y)^{1/2-\varepsilon}, \qquad \forall t \in [0,T], x, y \in \mathcal{T}_k,$$

for some $\varepsilon \in (0, 1/2)$.

Proof. For the proof of this I need to check Chapters **7** and **9** in [13].

□

Conjecture 8.3. *Prove that*

$$L^{\mathcal{T}}_{\tau_{\mathrm{cov}}(\mathcal{T}_k)+\varepsilon}(x) \geq \varepsilon_k, \qquad \forall x \in \mathcal{T}_k.$$

Lemma 8.4. *Prove that*

$$L^{\mathcal{T}}_{\tau_{\mathrm{cov}}(\mathcal{T}_k)+\varepsilon}(x) > 0, \qquad \forall x \in \mathcal{T}.$$

Proof. Let us prove Lemma 8.4 assuming that Lemmas 8.1, 8.2 and conjecture 8.3 are true. We denote by $\pi_k : \mathcal{T} \to \mathcal{T}_k$ the projection of $\mathcal{T}$ onto $\mathcal{T}_k$. More precisely, for all $x \in \mathcal{T}$, $\pi_k(x)$ is the point of $\mathcal{T}_k$ which is the closest to x under $d_{\mathcal{T}}$. For every $x \in \mathcal{T}$, the moduli of continuity for the local times in $\mathcal{T}$ gives

$$\begin{aligned} L^{\mathcal{T}}_{\mathrm{cov}(\mathcal{T}_k)+\varepsilon}(x) &\geq L^{\mathcal{T}}_{\mathrm{cov}(\mathcal{T}_k)+\varepsilon}(\pi_k(x)) - C d_{\mathcal{T}}(x, \pi_k(x))^{1/2-\varepsilon} \\ &\geq \varepsilon_k - C d_{\mathcal{T}}(x, \pi_k(x))^{1/2-\varepsilon}. \end{aligned}$$

At this point, suppose that $d_H(\mathcal{T}_k, \mathcal{T}) \preceq \delta_k := \varepsilon_k^2$. From the previous equation, for every $x \in \mathcal{T}$, we deduce

$$L^{\mathcal{T}}_{\mathrm{cov}(\mathcal{T}_k)+\varepsilon}(x) \geq \varepsilon_k - C\delta_k^{1/2-\varepsilon} \geq \varepsilon_k - C\varepsilon_k^{1-2\varepsilon} > 0,$$

for k large enough.

□

Now, the preceding and Conjecture 8.1 along with the (almost-sure) joint continuity of the local times, will give us $L^{\mathcal{T}}_{\mathrm{cov}(\mathcal{T})}(x) > 0$, $\forall x \in \mathcal{T}$, which (by definition) yields $\tilde{\tau}_{\mathrm{cov}}(\mathcal{T}) \leq \tau_{\mathrm{cov}}(\mathcal{T})$. Since, the converse inequality holds, we deduce that $\tilde{\tau}_{\mathrm{cov}}(\mathcal{T}) = \tau_{\mathrm{cov}}(\mathcal{T})$-a.s.

Lemma 8.5.

$$L^{(k)}_{\tau_{\mathrm{cov}}(\mathcal{T}_k)+\varepsilon}(x) \geq \varepsilon_k > 0, \qquad x \in \mathcal{T}_k,$$

which gives

$$L^{\mathcal{T}}_{\tau_{\mathrm{cov}}(\mathcal{T}_k)+\varepsilon}(x) > 0, \forall x : d(x, \mathcal{T}_k) \leq \varepsilon_k^2.$$

So if

$$d(\mathcal{T}, \mathcal{T}_k) \leq \varepsilon_k^2.$$

asymptotically we are done.

9 Proof from [9] adapted accordingly in our setting

Lemma 9.1. *Consider a mean zero Gaussian process $\{G_x : x \in \mathcal{T}\}$ with covariance $\mathbf{E}_x[L_{T_0}(y)] = d_{\mathcal{T}}(x_0, b(x_0, x, y))$ and define the following quantity $\sigma := \sup_{x\in\mathcal{T}} \left(d_{\mathcal{T}}(x_0, x_0 \wedge x)\right)^{1/2}$, where $x_0 \wedge x$ is the least-common ancestor of x_0 and x. For $\eta > 0$,*

$$\mathbf{P}\left(\left|\sup_{x\in\mathcal{T}} G_x - \mathbf{E}\sup_{x\in\mathcal{T}} G_x\right| > \eta\right) \leqslant 2\exp\left(-\eta^2/2\sigma^2\right).$$

Note that by symmetry, sup can be replaced by inf in Lemma 9.1, which is one of the versions that we are going to use.

Let $M_{\rho,x_0} := \mathbf{E}\inf_{y\in[[\rho,x_0]]} G_y$. The strategy of proof is the following:

$$\mathbf{P}_\rho\left(\inf_{y\in[[\rho,x_0]]} L^{(1)}_{T_0}(y) \leqslant \delta\right) = \mathbf{P}_{G,\bar{G}}\left(\inf_{y\in[[\rho,x_0]]}(G_y^2 + \bar{G}_y^2) \leqslant \delta\right) \leqslant 2\mathbf{P}_{G,\bar{G}}\left(\inf_{y\in[[\rho,x_0]]} G_y^2 \leqslant \frac{\delta}{2}\right). \quad (9.1)$$

The last estimate follows from

$$\left\{\inf_{y\in[[\rho,x_0]]} G_y^2 > \frac{\delta}{2}\right\} \cap \left\{\inf_{y\in[[\rho,x_0]]} \bar{G}_y^2 > \frac{\delta}{2}\right\} \subseteq \left\{\inf_{y\in[[\rho,x_0]]}(G_y^2 + \bar{G}_y^2) > \delta\right\}.$$

Taking complements,

$$\left\{\inf_{y\in[[\rho,x_0]]}(G_y^2 + \bar{G}_y^2) \leqslant \delta\right\} \subseteq \left\{\inf_{y\in[[\rho,x_0]]} G_y^2 \leqslant \frac{\delta}{2}\right\} \cup \left\{\inf_{y\in[[\rho,x_0]]} \bar{G}_y^2 \leqslant \frac{\delta}{2}\right\}.$$

A union bound and the fact that G and $\bar{G}$ are equally distributed gives the desired inequality. Applying Lemma 9.1, for $\eta := M_{\rho,x_0} - \sqrt{\frac{\delta}{2}} > 0$ (remains to show that this choice of η is OK)

$$\mathbf{P}_{G,\bar{G}}\left(\inf_{y\in[[\rho,x_0]]} G_y \leqslant \sqrt{\frac{\delta}{2}}\right) \leqslant 2\exp\left(-\left(M_{\rho,x_0} - \sqrt{\delta/2}\right)^2 / 2d_{\mathcal{T}}(\rho, x_0)\right)$$

Therefore, we obtain that

$$\mathbf{P}_\rho\left(\inf_{y\in[[\rho,x_0]]} L^{(1)}_{T_0}(y) \leqslant \delta\right) \leqslant 4\exp\left(-\left(M_{\rho,x_0} - \sqrt{\delta/2}\right)^2 / 2d_{\mathcal{T}}(\rho, x_0)\right). \quad (9.2)$$

10 The version of Lemma 9.1 using the new version of Ray-Knight (II)

Let $G_{y,b} := G_y - G_b$ and $\bar{G}_{y,b} = \bar{G}_y - \bar{G}_b$ as in Theorem 6.4. Using it, we deduce

$$\begin{aligned}
\mathbf{P}_b\left(\inf_{y\in\mathcal{T}_2} L^{(2)}_{T_0}(y) \leqslant \delta\right) &\leqslant \mathbf{P}_b\left(\inf_{y\in[[\rho,b]]} L^{(2)}_{T_0}(y) \leqslant \delta\right) \\
&\quad + \sum_{i=1}^{2} \mathbf{P}_b\left(\inf_{y\in[[b,x_i]]} L^{(2)}_{T_0}(y) \leqslant \delta\right) \\
&\leqslant \mathbf{P}_{G,\bar{G}}\left(\inf_{y\in[[\rho,b]]} G_y^2 + \bar{G}_y^2 \leqslant \delta\right) \\
&\quad + \sum_{i=1}^{2} \mathbf{P}_{G,\bar{G}}\left(\inf_{y\in[[b,x_i]]} G_y^2 + \bar{G}_y^2 - \sup_{y\in[[b,x_i]]} G_{y,b}^2 + \bar{G}_{y,b}^2 \leqslant \delta\right).
\end{aligned}$$

We consider finding an upper bound for the right-hand side of the probability above. The first term of the preceding inequality can be bounded above using the same method as in Lemma 9.1. Then, our focus shifts to the remaining terms, which correspond to the behaviour of the infimum of the difference of two Gaussians, when the infimum is taken over disjoint arcs that belong to the component $\mathcal{T}[b]$ emanating from the unique branching b between the root ρ and the leaves labelled by x_1 and x_2. It sufficient to estimate one of the aforementioned probabilities. For each $i = 1, 2$, we have that

$$\begin{aligned}
&\mathbf{P}_{G,\bar{G}}\left(\inf_{y\in[[b,x_i]]} G_y^2 + \bar{G}_y^2 - \sup_{y\in[[b,x_i]]} G_{y,b}^2 + \bar{G}_{y,b}^2 \leqslant \delta\right) \\
\leqslant &\mathbf{P}_{G,\bar{G}}\left(\inf_{y\in[[b,x_i]]} G_y^2 + \bar{G}_y^2 \leqslant \frac{1+\delta}{2}\right) + \mathbf{P}_{G,\bar{G}}\left(\sup_{y\in[[b,x_i]]} G_{y,b}^2 + \bar{G}_{y,b}^2 \geqslant \frac{1-\delta}{2}\right) \\
\leqslant &2\mathbf{P}_{G,\bar{G}}\left(\inf_{y\in[[b,x_i]]} G_y^2 \leqslant \frac{1+\delta}{4}\right) \\
+&2\mathbf{P}_{G,\bar{G}}\left(\sup_{y\in[[b,x_i]]} G_{y,b}^2 \geqslant \frac{1-\delta}{4}\right)
\end{aligned}$$

11 Proof of the Conjecture 8.3 in the case $\mathcal{T}_2$

Proof.

$$\mathbb{P}\left(\inf_{x\in\mathcal{T}_2} L^{(2)}_{\tau^2_{\text{cov}}+\varepsilon}(x) < \varepsilon\right) \leqslant \mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau^2_{\text{cov}}+\varepsilon}(x) < \varepsilon\right) + \mathbb{P}\left(\inf_{x\in[[\rho,v_2]]} L^{(2)}_{\tau^2_{\text{cov}}+\varepsilon}(x) < \varepsilon\right)$$

Applying the law of total probability, we obtain that the right hand side is bounded by

$$
\begin{aligned}
&\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau^2_{\mathrm{cov}}+\varepsilon}(x) < \varepsilon | \tau^2_{\mathrm{cov}} = \tau_{v_1}\right)\mathbb{P}\left(\tau^2_{\mathrm{cov}} = \tau_{v_1}\right)\\
+&\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau^2_{\mathrm{cov}}+\varepsilon}(x) < \varepsilon | \tau^2_{\mathrm{cov}} \neq \tau_{v_1}\right)\mathbb{P}\left(\tau^2_{\mathrm{cov}} \neq \tau_{v_1}\right)\\
+&\mathbb{P}\left(\inf_{x\in[[\rho,v_2]]} L^{(2)}_{\tau^2_{\mathrm{cov}}+\varepsilon}(x) < \varepsilon | \tau^2_{\mathrm{cov}} = \tau_{v_2}\right)\mathbb{P}\left(\tau^2_{\mathrm{cov}} = \tau_{v_2}\right)\\
+&\mathbb{P}\left(\inf_{x\in[[\rho,v_2]]} L^{(2)}_{\tau^2_{\mathrm{cov}}+\varepsilon}(x) < \varepsilon | \tau^2_{\mathrm{cov}} \neq \tau_{v_2}\right)\mathbb{P}\left(\tau^2_{\mathrm{cov}} \neq \tau_{v_2}\right) \qquad (11.1)
\end{aligned}
$$

The goal of this section is to estimate (11.1). For this, we first bound (11.1) by

$$
\begin{aligned}
&\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}+\varepsilon}(x) < \varepsilon\right)\\
+&\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau^2_{\mathrm{cov}}+\varepsilon}(x) < \varepsilon | \tau^2_{\mathrm{cov}} \neq \tau_{v_1}\right)\mathbb{P}\left(\tau^2_{\mathrm{cov}} \neq \tau_{v_1}\right)\\
+&\mathbb{P}\left(\inf_{x\in[[\rho,v_2]]} L^{(2)}_{\tau_{v_2}+\varepsilon}(x) < \varepsilon\right)\\
+&\mathbb{P}\left(\inf_{x\in[[\rho,v_2]]} L^{(2)}_{\tau^2_{\mathrm{cov}}+\varepsilon}(x) < \varepsilon | \tau^2_{\mathrm{cov}} \neq \tau_{v_2}\right)\mathbb{P}\left(\tau^2_{\mathrm{cov}} \neq \tau_{v_2}\right). \qquad (11.2)
\end{aligned}
$$

For $i = 1, 2$, we introduce further the notation $\tau_b^{v_i} := \inf\{t \geqslant 0 : X_t^{v_i} = b\}$. We estimate the first term and the third is done similarly. Furthermore, let z be such that $X_\varepsilon^{v_1} = z$, i.e. $\tau_z^{v_1} = \varepsilon$. Then, for $i = 1$, we have the estimate

$$
\begin{aligned}
&\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}+\varepsilon}(x) < \varepsilon\right)\\
\leqslant&\int \mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b]]} L^{(2)}_{\tau_z}(x) < \varepsilon\right)\mu_{\mathcal{T}_2}(dz)\\
\leqslant&\int \mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b]]} L^{(2)}_{\tau_z}(x) < \varepsilon | \tau_b^{v_1} \leqslant \tau_z^{v_1}\right)\mathbb{P}\left(\tau_b^{v_1} \leqslant \tau_z^{v_1}\right)\mu_{\mathcal{T}_2}(dz)\\
+&\int \mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b]]} L^{(2)}_{\tau_z}(x) < \varepsilon | \tau_b^{v_1} > \tau_z^{v_1}\right)\mathbb{P}\left(\tau_b^{v_1} > \tau_z^{v_1}\right)\mu_{\mathcal{T}_2}(dz), \qquad (11.3)
\end{aligned}
$$

where in the first inequality we have used the additivity of the local times. We first estimate

$$
\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b]]} L^{(2)}_{\tau_z}(x) < \varepsilon | \tau_b^{v_1} > \tau_z^{v_1}\right)\mathbb{P}\left(\tau_b^{v_1} > \tau_z^{v_1}\right).
$$

Using Lemma 9.1 and Theorem 6.2 for both local times $L^{(2)}_{\tau_{v_1}}(x)$ and $L^{(2)}_{\tau_z}(x)$, we obtain that

$$
\begin{aligned}
&\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b]]} L^{(2)}_{\tau_z}(x) < \varepsilon | \tau_b^{v_1} > \tau_z^{v_1}\right)\mathbb{P}\left(\tau_b^{v_1} > \tau_z^{v_1}\right)\\
\leqslant&4\exp(-\varepsilon/8(M^2_{\rho,v_1} + \sigma^2_{\rho,v_1})) + 4\exp(-\varepsilon/8(\sup_{z\in\mathcal{T}_2} M^2_{v_1,z} + \sup_{z\in\mathcal{T}_2}\sigma^2_{v_1,z})), \qquad (11.4)
\end{aligned}
$$

where the indexing in M and σ denotes the fact that quantities M and σ in Lemma 9.1 are restricted to the corresponding subspaces indicated in a self-explanatory manner by the index.

Now, we estimate

$$\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b]]} L^{(2)}_{\tau_z}(x) < \varepsilon | \tau_b^{v_1} \leqslant \tau_z^{v_1}\right) \mathbb{P}\left(\tau_b^{v_1} \leqslant \tau_z^{v_1}\right).$$

As before, for $L^{(2)}_{\tau_{v_1}}(x)$ we use Theorem 6.2 along with Lemma 9.1. Since in this case, $\tau_b^{v_1} \leqslant \tau_z^{v_1}$, when considering the dynamics after the process reaches v_1, first, we use Theorem 6.2 for $L^{(2)}_{\tau_z}(x)$ on $[[v_1,b]]$, and Theorem 6.4 for the process restarted from the branching point b. In this manner, we obtain the estimate

$$\begin{aligned}
&\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(2)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b]]} L^{(2)}_{\tau_z}(x) < \varepsilon | \tau_b^{v_1} \leqslant \tau_z^{v_1}\right) \mathbb{P}\left(\tau_b^{v_1} \leqslant \tau_z^{v_1}\right) \\
\leqslant & 4\exp\left(-\varepsilon/8(M^2_{\rho,v_1} + \sigma^2_{\rho,v_1})\right) + 4\exp\left(-\varepsilon/8(M^2_{v_1,b} + \sigma^2_{v_1,b})\right) + 4\exp\left(-\varepsilon/8(M^2_{\rho,b} + \sigma^2_{\rho,b})\right). \quad (11.5)
\end{aligned}$$

For the second category of terms in (11.2), we use for pairs $(i,j) = (1,2)$ and $(i,j) = (2,1)$ the following estimate

$$\begin{aligned}
&\sum_{i\neq j} \mathbb{P}\left(\inf_{x\in[[\rho,v_i]]} L^{(2)}_{\tau_{v_i}+\tau_{v_i\to v_j}+\varepsilon}(x) < \varepsilon\right) \\
\leq & \sum_{i\neq j} \int \mathbb{P}\left(\inf_{x\in[[\rho,v_i]]} L^{(2)}_{\tau_{v_i}}(x) + \inf_{x\in[[]v_i,b]]} L^{(2)}_{\tau_{v_i\to b}}(x) + \inf_{x\in[[b,v_j]]} L^{(2)}_{\tau_{b\to v_j}}(x) + \inf L^{(2)}_{\tau_z^{v_1}}(x) < \varepsilon\right) \mu_{\mathcal{T}_2}(dz),
\end{aligned} \quad (11.6)$$

where we have used the additivity property of the local times. As before, we condition on the events $\tau_b^{v_1} \leqslant \tau_z^{v_1}$ and $\tau_b^{v_1} > \tau_z^{v_1}$. On each of the latter events, for $L^{(2)}_{\tau_{v_1}}(x)$ and $L^{(2)}_{\tau_{v_i\to b}}$ we use Theorem 6.2 along with Lemma 9.1 and for $L^{(2)}_{\tau_{b\to v_j}}$ we use Theorem 6.4 along with Lemma 9.1. Finally, we obtain the estimate

$$\begin{aligned}
&\sum_{i\neq j} \mathbb{P}\left(\inf_{x\in[[\rho,v_i]]} L^{(2)}_{\tau_{v_i}}(x) + \inf_{x\in[[]v_i,b]]} L^{(2)}_{\tau_{v_i\to b}} + \inf_{x\in[[b,v_j]]} L^{(2)}_{\tau_{b\to v_j}} + \inf L^{(2)}_{\tau_z^{v_1}}(x) < \varepsilon\right) \\
\leqslant & 8\exp\left(-\varepsilon/16(M^2_{\rho,b} + \sigma^2_{\rho,b})\right) + 8\exp(-\varepsilon/16(\sup_{z\in\mathcal{T}_2} M^2_{v_1,z} + \sup_{z\in\mathcal{T}_2} \sigma^2_{v_1,z})) \\
+ & 8\exp\left(-\varepsilon/16(M^2_{\rho,v_1} + \sigma^2_{\rho,v_1})\right) + 8\exp\left(-\varepsilon/16(M^2_{v_1,b} + \sigma^2_{v_1,b})\right) \\
+ & 8\exp\left(-\varepsilon/16(M^2_{\rho,v_1} + \sigma^2_{\rho,v_1})\right) + 8\exp\left(-\varepsilon/16(M^2_{v_1,b} + \sigma^2_{v_1,b})\right) \\
+ & 8\exp\left(-\varepsilon/16(M^2_{\rho,b} + \sigma^2_{\rho,b})\right) + 8\exp\left(-\varepsilon/16(M^2_{\rho,b} + \sigma^2_{\rho,b})\right).
\end{aligned}$$

□

12 Proof of the conjecture 8.3 in the case $\mathcal{T}_k$

▸

$$\begin{aligned}\mathbb{P}\left(\inf_{x\in\mathcal{T}_k} L^{(k)}_{\tau^k_{\text{cov}}+\varepsilon}(x)<\varepsilon_k\right) &\leq \sum_{l=1}^{k}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau^k_{\text{cov}}+\varepsilon}(x)<\varepsilon_k\right)\\ &\leq \sum_{l=1}^{k}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau^k_{\text{cov}}+\varepsilon}(x)<\varepsilon_k, \tau^k_{\text{cov}}=\tau_{v_l}\right)\\ &+\sum_{l=1}^{k}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau^k_{\text{cov}}+\varepsilon}(x)<\varepsilon_k, \tau^k_{\text{cov}}\neq\tau_{v_l}\right)\leq (f_\varepsilon(k)),\end{aligned}$$

$f(k)$-summable. Now, decomposing further,

$$\begin{aligned}&\sum_{l=1}^{k}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau^k_{\text{cov}}+\varepsilon}(x)<\varepsilon_k, \tau^k_{\text{cov}}=\tau_{v_l}\right)\\ &+\sum_{l=1}^{k}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau^k_{\text{cov}}+\varepsilon}(x)<\varepsilon_k, \tau^k_{\text{cov}}\neq\tau_{v_l}\right)\\ &\leq\sum_{l=1}^{k}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau_{v_l}+\varepsilon}(x)<\varepsilon_k\right)+\sum_{l\neq l'}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau^k_{\text{cov}}+\varepsilon}(x)<\varepsilon_k, \tau^k_{\text{cov}}=\tau_{v_{l'}}\right)\\ &\leq\sum_{l=1}^{k}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau_{v_l}+\varepsilon}(x)<\varepsilon_k\right)+\sum_{l\neq l'}\mathbb{P}\left(\inf_{x\in[\rho,v_l]} L^{(k)}_{\tau_{v_l}+\tau_{v_l\to v_{l'}}+\varepsilon}(x)<\varepsilon_k\right)\end{aligned}\tag{12.1}$$

Let $b^{(l)}$ be the branch point that is the closes in terms of tree distance to the vertex v_l. Also, let

$$\tau^{v_i}_{b^{(l)}}:=\inf\{t\geqslant 0: X^{v_i}_t=b^{(l)}\}.$$

Furthermore, let z^l be such that $X^{v_l}_\varepsilon=z^l$, i.e. $\tau^{v_l}_{z^l}=\varepsilon$ For each of the elements in the summation in the first term of (12.1) we condition in the following manner: We estimate the first term and the third is done similarly. Then, for $l=1$, we have the estimate

$$\begin{aligned}&\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(k)}_{\tau_{v_1}+\varepsilon}(x)<\varepsilon\right)\\ &\leqslant\int\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(k)}_{\tau_{v_1}}(x)+\inf_{x\in[[v_1,b^{(1)}]]} L^{(k)}_{\tau_{z^1}}(x)<\varepsilon\right)\mu_{\mathcal{T}_k}(dz)\\ &\leqslant\int\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(k)}_{\tau_{v_1}}(x)+\inf_{x\in[[v_1,b^{(1)}]]} L^{(k)}_{\tau_{z^1}}(x)<\varepsilon|\tau^{v_1}_{b^{(1)}}\leqslant\tau^{v_1}_{z^1}\right)\mathbb{P}\left(\tau^{v_1}_{b^{(1)}}\leqslant\tau^{v_1}_{z^1}\right)\mu_{\mathcal{T}_k}(dz)\\ &+\int\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(k)}_{\tau_{v_1}}(x)+\inf_{x\in[[v_1,b^{(1)}]]} L^{(k)}_{\tau_{z^1}}(x)<\varepsilon|\tau^{v_1}_{b^{(1)}}>\tau^{v_1}_{z^1}\right)\mathbb{P}\left(\tau^{v_1}_{b^{(1)}}>\tau^{v_1}_{z^1}\right)\mu_{\mathcal{T}_k}(dz),\end{aligned}\tag{12.2}$$

Using Lemma 9.1 and Theorem 6.2 for both local times $L^{(k)}_{\tau_{v_1}}(x)$ and $L^{(k)}_{\tau_{z^1}}(x)$, we obtain that the first term in (12.2) is estimated

$$\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(k)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b^{(1)}]]} L^{(k)}_{\tau_z}(x) < \varepsilon | \tau^{v_1}_{b^{(1)}} > \tau^{v_1}_{z^1}\right) \mathbb{P}\left(\tau^{v_1}_{b^{(1)}} > \tau^{v_1}_{z^1}\right)$$
$$\leqslant 4\exp(-\varepsilon_k/8(M^2_{\rho,v_1}+\sigma^2_{\rho,v_1})) + 4\exp(-\varepsilon_k/8(\sup_{z\in\mathcal{T}_k} M^2_{v_1,z^1} + \sup_{z\in\mathcal{T}_k}\sigma^2_{v_1,z^1})). \tag{12.3}$$

Note that this estimate is summed over l from 1 up to k.

Proposition 12.1 (Ray-Knight(k), to be proved)**.** *Denote by $G := \{G_x : x \in \mathcal{T}\}$ a mean zero Gaussian process with covariance $d_{\mathcal{T}}(\rho, x\wedge y)$. Then, under $\mathbf{P}_b \times \mathbf{P}_{G,\bar{G}}$,*

$$\left\{L_{T_0}(y) + (G^2_y + \bar{G}^2_y)\mathbf{1}\{y \in \mathcal{T}_k[b^{(l)}]\} : y \in \mathcal{T}_k\right\} \overset{(d)}{=} \{G^2_y + \bar{G}^2_y : y \in \mathcal{T}_k\},$$

where $(G,\bar{G})$ is a vector of independent copies of G and $\mathcal{T}_k[b^{(l)}]$ is the two sticks component emanating from the branch point $b^{(l)}$.

We now estimate the second term in (12.2). For $L^{(k)}_{\tau_{v_1}}(x)$ we use Theorem 6.2 along with Lemma 9.1. Since in this case, $\tau^{v_1}_{b^{(1)}} \leqslant \tau^{v_1}_{z^1}$, when considering the dynamics after the process reaches v_1, first, we use Theorem 6.2 for $L^{(k)}_{\tau_{z^1}}(x)$ on $[[v_1, b^{(1)}]]$, and $RK(k)$ for the process restarted from the branching point $b^{(1)}$. We obtain the estimate

$$\mathbb{P}\left(\inf_{x\in[[\rho,v_1]]} L^{(k)}_{\tau_{v_1}}(x) + \inf_{x\in[[v_1,b^{(1)}]]} L^{(k)}_{\tau_{z^1}}(x) < \varepsilon | \tau^{v_1}_{b^{(1)}} \leqslant \tau^{v_1}_{z^1}\right) \mathbb{P}\left(\tau^{v_1}_{b^{(1)}} \leqslant \tau^{v_1}_{z^1}\right)$$
$$\leqslant 4\exp\left(-\varepsilon/8(M^2_{\rho,v_1}+\sigma^2_{\rho,v_1})\right) + 4\exp\left(-\varepsilon/8(M^2_{v_1,b^{(1)}}+\sigma^2_{v_1,b^{(1)}})\right) + term(RK(k)). \tag{12.4}$$

Note that this estimate is summed over from $l = 1$ to k.

We know estimate the second term in (12.1).

We have that

$$\sum_{i\neq j} \mathbb{P}\left(\inf_{x\in[[\rho,v_i]]} L^{(k)}_{\tau_{v_i}+\tau_{v_i\to v_j}+\varepsilon}(x) < \varepsilon\right)$$
$$\leq \sum_{i\neq j}\int \mathbb{P}\left(\inf_{x\in[[\rho,v_i]]} L^{(k)}_{\tau_{v_i}}(x) + \inf_{x\in[[v_i,b^{(i)}]]} L^{(k)}_{\tau_{v_i\to b^{(i)}}}(x) + \cdots + \inf_{x\in[[b^{(j)},v_j]]} L^{(k)}_{\tau_{b^{(j)}\to v_j}}(x) + \inf L^{(k)}_{\tau^{v_j}_{z^j}}(x) < \varepsilon\right)\mu_{\mathcal{T}_k}(dz), \tag{12.5}$$

where we have used the additivity property of the local times.

We condition on the events $\tau^{v_j}_{b^{(j)}} \leqslant \tau^{v_j}_{z^j}$ and $\tau^{v_j}_{b^{(j)}} > \tau^{v_j}_{z^j}$. On each of the latter events, for $L^{(k)}_{\tau_{v_i}}(x)$ and $L^{(k)}_{\tau_{v_i\to b^{(i)}}} + \cdots$ we use Theorem 6.2 along with Lemma 9.1 and for $L^{(k)}_{\tau_{b^{(j)}\to v_j}}(x)$ we use Theorem 6.2 (with rerooting) along with Lemma 9.1. For the first event we use the same strategy with the difference that we use $RK(k)$ in the last term.

$$\sum_{i\neq j}\int \mathbb{P}\left(\inf_{x\in[[\rho,v_i]]} L^{(k)}_{\tau_{v_i}}(x)+\inf_{x\in[[v_i,b^{(i)}]]} L^{(k)}_{\tau_{v_i\to b^{(i)}}}(x)+\cdots+\inf_{x\in[[b^{(j)},v_j]]} L^{(k)}_{\tau_{b^{(j)}\to v_j}}(x)+\inf L^{(k)}_{\tau^{v_j}_{z^j}}(x)<\varepsilon\right)\mu_{\mathcal{T}_k}(dz)$$
$$\leqslant \sum_{i\neq j}\exp(-\varepsilon_k/16(\sup_{z\in\mathcal{T}_k} M^2_{v_i,z^i}+\sup_{z\in\mathcal{T}_k}\sigma^2_{v_i,z^i}))+RK(k)$$
$$+\sum_{i\neq j}\exp\left(-\varepsilon_k/16(M^2_{\rho,v_i}+\sigma^2_{\rho,v_i})\right)+\sum_{i\neq j}\left(\exp\left(-\varepsilon_k/16(M^2_{v_i,b^{(i)}}+\sigma^2_{v_i,b^{(i)}})\right)+\cdots+\exp\left(-\varepsilon_k/16(M^2_{b^{(j)},v_j}+\sigma^2_{b^{(j)},v_j})\right)\right)$$
$$+\sum_{i\neq j}\exp\left(-\varepsilon_k/16(M^2_{\rho,v_i}+\sigma^2_{\rho,v_i})\right)+\sum_{i\neq j}\left(\exp\left(-\varepsilon_k/16(M^2_{v_i,b^{(i)}}+\sigma^2_{v_i,b^{(i)}})\right)+\cdots+\exp\left(-\varepsilon_k/16(M^2_{b^{(j)},v_j}+\sigma^2_{b^{(j)},v_j})\right)\right)$$

.

References

[1] D. Aldous. Random walk covering of some special trees. *J. Math. Anal. Appl.*, 157(1):271–283, 1991.

[2] G. Andriopoulos. Convergence of blanket times for sequences of random walks on critical random graphs. *arXiv preprint arXiv:1810.07518*, 2018.

[3] M. T. Barlow and E. A. Perkins. Brownian motion on the Sierpiński gasket. *Probab. Theory Related Fields*, 79(4):543–623, 1988.

[4] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[5] D. A. Croydon. Convergence of simple random walks on random discrete trees to Brownian motion on the continuum random tree. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(6):987–1019, 2008.

[6] D. A. Croydon. Moduli of continuity of local times of random walks on graphs in terms of the resistance metric. *Trans. London Math. Soc.*, 2(1):57–79, 2015.

[7] D. A. Croydon, B. M. Hambly, and T. Kumagai. Time-changes of stochastic processes associated with resistance forms. *Electron. J. Probab.*, 22(82):1–41, 2017.

[8] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II*. Probability and its Applications (New York). Springer, New York, second edition, 2008. General theory and structure.

[9] J. Ding, J. R. Lee, and Y. Peres. Cover times, blanket times, and majorizing measures. *Ann. of Math. (2)*, 175(3):1409–1471, 2012.

[10] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011.

[11] J. Kigami. Resistance forms, quasisymmetric maps and heat kernel estimates. *Mem. Amer. Math. Soc.*, 216(1015):vi+132, 2012.

[12] D. A. Levin and Y. Peres. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2017. Second edition of [MR2466937], With contributions by Elizabeth L. Wilmer, With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson.

[13] M. B. Marcus and J. Rosen. Sample path properties of the local times of strongly symmetric Markov processes via Gaussian processes. *Ann. Probab.*, 20(4):1603–1684, 1992.

[14] M. B. Marcus and J. Rosen. *Markov processes, Gaussian processes, and local times*, volume 100. Cambridge University Press, 2006.