Electronic Journal of Probability

Electron. J. Probab. **28** (2023), article no. 9, 1–64.
ISSN: 1083-6489 https://doi.org/10.1214/23-EJP901

Scaling limit of linearly edge-reinforced random walks on critical Galton-Watson trees*

George Andriopoulos† Eleanor Archer‡

Abstract

We prove an invariance principle for linearly edge reinforced random walks on γ-stable critical Galton-Watson trees, where $\gamma \in (1,2]$ and where the edge joining x to its parent has rescaled initial weight $d(O,x)^\alpha$ for some $\alpha \leq 1$. This corresponds to the recurrent regime of initial weights. We then establish fine asymptotics for the limit process. In the transient regime, we also give an upper bound on the random walk displacement in the discrete setting, showing that the edge reinforced random walk never has positive speed, even when the initial edge weights are strongly biased away from the root.

Keywords: random walk in random environment; Dirichlet distribution; reinforced random walks; Galton–Watson trees; diffusion in random environment; slow movement.
MSC2020 subject classifications: 60F17; 60K37; 60K50; 60J60.
Submitted to EJP on January 7, 2022, final version accepted on January 2, 2023.

1 Introduction

Linearly edge-reinforced random walks (LERRW) are a classical model of self-interacting processes introduced by Coppersmith and Diaconis in 1986 [29] (for a generalisation, we refer to the recent work of [17]). Given a rooted graph G endowed with initial edge weights, and a *reinforcement parameter* $\Delta > 0$, the model is defined as follows. The process $(X_n)_{n\geq 0}$ on G is started at the root and given $X_n = v$, the next edge traversed by X is chosen from the edges incident to v with probability proportional to their weights. After X has crossed the chosen edge, its weight is subsequently increased by Δ and the process repeats with the updated edge weights.

The non-Markovian nature of the LERRW makes it difficult to analyse. However, a remarkable result of Diaconis and Freedman [35] (for the recurrent case) and then

*The research of EA was supported by JSPS and ERC starting grant 676970 RANDGEOM.

†NYU-ECNU Institute of Mathematical Sciences, NYU Shanghai, Shanghai, 200062, China. E-mail: ga73@nyu.edu

‡Équipe Modal'X, Université Paris Nanterre, 200 avenue de la République, 92000 Nanterre, France. E-mail: eleanor.archer@parisnanterre.fr

Merkl and Rolles [59] (extension to the transient case) shows that the LERRW can be represented as a random walk in random environment (RWRE). This makes LERRW more tractable and for example leads to applications in Bayesian statistics [36, 15, 14, 16]. On trees this RWRE representation can be understood by noticing that the weights of the edges incident to each of the vertices evolve according to a Pólya urn model, independently for each vertex, see [62]. We will make this connection precise in Section 3.2.

The aim of this paper is to study LERRW on critical Galton-Watson trees by constructing its scaling limit for a range of initial weights and obtaining almost sure asymptotics for the limiting diffusion. Since critical Galton-Watson trees have a rich geometry, in particular having fractal properties, unbounded degrees, and non-uniform volume growth, we hope that these results are interesting in their own right as well as offering insight into the behaviour of LERRW on related critical random graphs such as uniform planar triangulations, high dimensional uniform spanning trees and critical percolation clusters.

Our first result, an invariance principle, extends an earlier work of Lupu, Sabot and Tarrès, who constructed the scaling limit of LERRW on $\mathbb{Z}$ as a diffusion on $\mathbb{R}$ [56]. In the Galton-Watson tree setting, in the special case where the initial weights are constant and identical for all edges and the critical Galton-Watson tree has finite variance, the scaling limit was previously constructed in [8].

We assume that the underlying Galton-Watson tree is critical with offspring distribution in the domain of attraction of a γ-stable law for some $\gamma \in (1,2]$. In the case $\gamma < 2$ this immediately entails that the offspring distribution ξ satisfies $\xi([k,\infty)) = k^{-\gamma}L(k)$ for some slowly-varying function L. If such a ξ is fixed and T_n is a ξ-Galton-Watson tree with n vertices endowed with uniform mass measure, it is well-known that there exists a sequence $(a_n)_{n=1}^{\infty}$ such that $n^{-1}a_nT_n$ converges in distribution to the *stable tree* of Le Gall and Le Jan [54] (see also [40, Chapter 1]). We denote the compact stable tree by $\mathcal{T}_\gamma^c$ and its root by O. In the case $\gamma = 2$ this is simply Aldous' Continuum Random Tree (see [5] and [7]). We give the formal constructions in Section 2.3.

To define our LERRW model, we fix a reinforcement parameter $\Delta > 0$, label the root of our Galton-Watson tree O_n and consider a class of initial weights $(\alpha_e^{(n)})_{e\in E(T_n)}$ parametrised by $\alpha \leq 1$ whereby, if $\overleftarrow{x}$ denotes the parent vertex of x,

$$\alpha^{(n)}_{\{\overleftarrow{x},x\}} = (na_n^{-1})^{1-\alpha}(d_n(O_n,x) + na_n^{-1})^{\alpha} \tag{1.1}$$

where d_n denotes the graph distance on T_n, cf. [69, Section 2.3]. Note that na_n^{-1} is the natural scale for branch lengths in $\mathcal{T}_\gamma^c$; the addition of of this term in is just to ensure that the weights and inverse weights are integrable at 0. Our first result is an invariance principle for this LERRW process.

Theorem 1.1. *Let $X^{(n)}$ be a discrete-time LERRW on T_n, started at O_n, with initial weights as in* (1.1) *and reinforcement parameter Δ. Denote its law by $P_{O_n}^{(\alpha^{(n)})}$. For every $\alpha \leq 1$, there exists a stochastic process X defined on $\mathcal{T}_\gamma^c$ and started at O, with law P_O, such that the following convergence holds jointly with respect to the Gromov-Hausdorff topology, and the topology of weak convergence of measures on the càdlàg path space on T_n (itself endowed with the Skorohod-J_1 topology):*

$$\left(a_n n^{-1}T_n, P_{O_n}^{(\alpha^{(n)})}\left(\left(n^{-1}a_nX^{(n)}_{\lfloor 2n^2a_n^{-1}t\rfloor}\right)_{t\geq 0} \in \cdot\right)\right) \xrightarrow{(d)} \left(\mathcal{T}_\gamma^c, P_O\left((X_t)_{t\geq 0}\in\cdot\right)\right).$$

The law of X will be defined explicitly in Section 3.3. To establish scaling limits for tree-indexed random walks, one generally needs tighter control on the input laws compared to the $\mathbb{Z}^d$ case, cf. [47, Theorem 2] and [57, Theorem 1] which require finiteness of higher moments in order to prove convergence of tree-indexed random walks to tree-indexed Brownian motion (known as snake processes). This is because all the extra branches present in the tree present more opportunities for large displacements of the random walk, leading to larger fluctuations without the higher moment assumption. However, we will see in the LERRW setting that the random environment becomes increasingly concentrated as $n \to \infty$ without making any higher moment assumptions.

The analogous result previously proved by Lupu, Sabot and Tarrès for the LERRW in dimension one [56] used a high-level representation of a LERRW as a RWRE that was established in [34]. They encode the corresponding random environment in a random walk process and show that this converges to a time-changed Brownian motion on $\mathbb{R}$ under the appropriate rescaling. The law of the limiting LERRW diffusion is analogously encoded by this Brownian motion. We will use the same strategy to prove Theorem 1.1. In contrast to [56] and [8], we directly use the observation of Pemantle that the random environment can be represented by independent Dirichlet random variables at each vertex [62], rather than the (more complicated) mixing measure obtained in [34], which makes our proof more elementary.

In the second half of this paper, we obtain asymptotics for the limiting diffusion appearing in Theorem 1.1. To understand the long-time asymptotics of LERRW and its scaling limit it is more natural to study these processes on Galton-Watson trees conditioned to survive and their continuum counterparts, rather than on compact trees. In the discrete setting Kesten [48] showed that this conditioning can be achieved by conditioning a Galton-Watson tree to have a single branch that survives to infinity, known as the backbone, and attaching finite Galton-Watson trees to the backbone. We denote this tree by T_∞. This construction also has a natural analogue in the continuum known as sin-trees. These were constructed by Aldous in the case $\gamma = 2$ [6, Section 2.5] and Duquesne in the case $\gamma < 2$ [39]. We denote the infinite stable tree by $\mathcal{T}_\gamma^\infty$. We give the formal constructions in Sections 2.4 and 5.

It seems intuitively clear that increasing Δ should make the LERRW "more recurrent" whereas increasing α should make it "more transient". Since the infinite tree T_∞ contains a unique path to infinity, it follows directly from the corresponding result of Takeshima [70] (see also [3]) for LERRW on the infinite half-line that the LERRW is almost surely recurrent if $\sum_{n\geq 1} n^{-\alpha} = \infty$, and almost surely transient otherwise. However, although the value of Δ cannot change the transience and recurrence properties of X, we will see in our next theorems that both α and Δ affect the time it takes for X to escape a ball. In the case $\alpha < 1$, which we refer to as "strongly recurrent", the reinforcement gives an exponential bias towards the root, so that X only moves away from the root at logarithmic speed. In what follows, we denote by $\mathbf{P}$ the probability measure for Kesten's tree T_∞ and for the infinite stable tree $\mathcal{T}_\gamma^\infty$.

Theorem 1.2. *[Strongly recurrent regime $\alpha < 1$]. Let $(X_t)_{t\geq 0}$ be the limiting diffusion of Theorem 1.1 on the infinite stable tree $\mathcal{T}_\gamma^\infty$ when $\alpha < 1$. If $\Delta > 0$, we have that $\mathbf{P} \times P_O$-almost surely,*

$$\limsup_{t\to\infty} \frac{d(O, X_t)}{(\log t)^{\frac{1}{1-\alpha}}} = \left(\frac{1-\alpha}{\Delta}\right)^{\frac{1}{1-\alpha}},$$

where by d we denote the tree metric on $\mathcal{T}_\gamma^\infty$.

Remark 1.3. Takei [69] also proved similar results for the discrete-time LERRW on $\mathbb{Z}$. Lupu, Sabot and Tarrès also proved a result analogous to Theorem 1.2 in the case $\alpha = 0$ for the limiting diffusion on $\mathbb{R}$ [56, Proposition 4.9]; this corresponds to taking the initial occupation profile function L_0 equal to 1 everywhere, except possibly a compact interval. Our long-time asymptotics results cover a larger class of initial occupation profiles, and we additionally have to take care of fluctuations in the underlying tree $\mathcal{T}_\gamma^\infty$.

In the critical case $\alpha = 1$, the slow-down effect from the reinforcement is almost balanced by the increasing initial edge weights, so we lose the exponential attraction towards the root and no longer see the slow movement. Instead the attraction is of polynomial order and this is reflected in the results of Theorem 1.4. Because the reinforcement effect and the volumes of balls in the tree both grow on a polynomial scale, we see that there are two regimes depending on which one has the dominant effect.

Theorem 1.4. *[Critical regime $\alpha = 1$]. Let $(X_t)_{t\geq 0}$ be the limiting diffusion of Theorem 1.1 on the infinite stable tree $\mathcal{T}_\gamma^\infty$ when $\alpha = 1$.*

(i) If $\Delta \leq \frac{2\gamma-1}{\gamma-1}$, then for any $\beta > \frac{1}{2}$, we have that $\mathbf{P} \times P_O$-almost surely,

$$0 < \limsup_{t\to\infty} \frac{d(O,X_t)}{t^{\frac{\gamma-1}{2\gamma-1}} e^{-(\log t)^\beta}} \qquad \textit{and} \qquad \limsup_{t\to\infty} \frac{d(O,X_t)}{t^{\frac{\gamma-1}{2\gamma-1}} e^{(\log t)^\beta}} < \infty.$$

(ii) If $\Delta > \frac{2\gamma-1}{\gamma-1}$, then for any $\beta > \frac{1}{2}$, we have that $\mathbf{P} \times P_O$-almost surely,

$$0 < \limsup_{t\to\infty} \frac{d(O,X_t)}{t^{1/\Delta} e^{-(\log t)^\beta}} \qquad \textit{and} \qquad \limsup_{t\to\infty} \frac{d(O,X_t)}{t^{1/\Delta} e^{(\log t)^\beta}} < \infty.$$

In the transient regime when $\alpha > 1$, the resistance techniques used to prove Theorem 1.1 no longer provide the machinery to identify the LERRW scaling limit since the rescaled resistances degenerate in this regime. However, we can work in the discrete setting and obtain an upper bound on the displacement of the random walk.

Theorem 1.5. *[Transient regime $\alpha > 1$]. Let $(Y_n)_{n\geq 0}$ be a discrete-time LERRW on Kesten's tree T_∞ with offspring distribution in the domain of attraction of a γ-stable law, with initial weights $\alpha_{\{\overleftarrow{x},x\}} = d_{T_\infty}(O_\infty, x)^\alpha$. Then, almost surely, for any $A > \frac{1}{2}$,*

$$d_{T_\infty}(O_\infty, Y_n) \leq n^{\frac{2\gamma-1}{2\gamma}} e^{(\log n)^A},$$

for all sufficiently large n. Consequently, Y does not have positive speed.

It is somewhat more delicate to obtain lower bounds on the displacement since this requires uniform control of the environment. However, we believe that it should also be possible to get a comparable lower bound using discrete analogues of the arguments used to obtain this control in the recurrent case.

Despite the bias of our initial weights away from the root, the result of Theorem 1.5 contrasts strongly to those obtained on b-regular trees with constant initial weights by Collevecchio [28], who proved that the speed is positive when b is large enough or under an appropriate moment condition for the LERRW. This was later extended to all $b \geq 2$ by Aidekon [2]. In our case, Theorem 1.5 holds because the LERRW is slowed down by the time spent in dead ends of the tree; this effect is particularly pronounced because of the fractal properties of the tree (i.e. many dead ends en route to and from other dead ends). By contrast, in regular trees there are no dead ends, and exponential volume growth, which gives a much stronger bias away from the root.

In the same spirit, due to the uniqueness of the path to infinity in T_∞, the LERRW on T_∞ can be viewed as a LERRW on $\mathbb{Z}^+$ slowed down by extra excursions into subtrees hanging off this path. This can be quantified: on $\mathbb{Z}^+$, with the same initial weights, it is possible to show that $\sup_{m\leq n} |Y_m| \geq n^{\frac{1}{2}-o(1)}$ (see Remark 8.5). Since $\gamma > 1$, the LERRW is really slower on the trees we consider (note however that it is not true that $T_\infty \to \mathbb{Z}^+$ in any sense as $\gamma \downarrow 1$, in fact due to the heavy tails, the Galton-Watson trees undergo a condensation phenomenon, e.g. see [51]).

Here we briefly record some notation that we will use throughout the paper.

T_n	critical Galton-Watson tree satisfying (2.1)
d_n	graph distance on T_n
μ_n	uniform probability measure on the vertices of T_n
R_n	distorted graph distance on T_n (3.8)
ν_n	distorted measure on the vertices of T_n (3.8)
O_n	root of T_n
T_∞	critical Galton-Watson tree conditioned to survive, satisfying (2.1)
d_{T_∞}	graph distance on T_∞
R	distorted graph distance on T_∞
ν	distorted measure on the vertices of T_∞
O_∞	root of T_∞
$\mathcal{T}_\gamma^c$	compact stable tree
$\mathcal{T}_\gamma^\infty$	infinite stable tree
μ	canonical measure on $\mathcal{T}_\gamma^c$ or $\mathcal{T}_\gamma^\infty$ (2.5)
ν_ϕ	distorted measure on $\mathcal{T}_\gamma^c$ or $\mathcal{T}_\gamma^\infty$(3.13)
d	canonical metric on $\mathcal{T}_\gamma^c$ or $\mathcal{T}_\gamma^\infty$ (2.4)
R_ϕ	distorted metric on $\mathcal{T}_\gamma^c$ or $\mathcal{T}_\gamma^\infty$ (3.12)
O	root of $\mathcal{T}_\gamma^c$ or $\mathcal{T}_\gamma^\infty$
a_n	scaling factor for $T_n : n^{-1}a_nT_n \overset{(d)}{\to} \mathcal{T}_\gamma^c$
$\mathbf{P}$	probability measure for the mm-spaces $T_n, T_\infty, \mathcal{T}_\gamma^c, \mathcal{T}_\gamma^\infty$
$\alpha_e^{(n)}$	initial weight for LERRW for $e \in E(T_n)$
Δ	reinforcement parameter for LERRW on T_n
$\mathbf{b}_v^{(n)}$	parameter for RWDE at $v \in T_n$
$\mathbb{P}^{(\mathbf{b})}$	law of Dirichlet random environment on T_n with (vector) parameter $\mathbf{b}$
$X^{(n)}$	discrete-time LERRW on T_n
$P_{O_n}^{(\alpha)}$	quenched law of $X^{(n)}$ with initial weights α
$Z^{(n)}$	continuous-time constant speed RWDE on T_n
$\tilde{P}_{O_n,\omega}$	quenched law of $Z^{(n)}$
$\tilde{P}_{O_n}^{(\mathbf{b})}$	annealed law of $Z^{(n)}$
X	scaling limit of $X^{(n)}$ as in Theorem 1.1
$\tilde{P}_{O,\phi}$	quenched law of X (over the environment)
P_O	annealed law of X (over the environment)
Y	discrete-time LERRW on T_∞
$d^{(\alpha)}(s,t)$	$\begin{cases} d(0,s)^{1-\alpha} + d(0,t)^{1-\alpha} - 2d(0,s\wedge t)^{1-\alpha}, & \text{if } \alpha < 1, \\ \log(d(0,s)+1) + \log(d(0,t)+1) - 2\log(d(0,s\wedge t)+1), & \text{if } \alpha = 1, \end{cases}$
$\phi^{(\alpha)}$	Gaussian process defined such that $\mathbb{E}\big[\lvert\phi^{(\alpha)}(s) - \phi^{(\alpha)}(t)\rvert^2\big] = d^{(\alpha)}(s,t)$
$\mathbb{P}$	law of $\phi^{(\alpha)}$

1.1 Organisation

The paper is organised as follows. In Section 2 we give background on critical Galton-Watson trees and the topologies used in this paper. In Section 3 we make the connection between LERRW and RWRE precise, and construct a candidate for the scaling limit of LERRW. In Section 4 we prove that the resistance metrics and measures characterising the associated RWRE converge to those characterising the aforementioned limit candidate, which proves Theorem 1.1. The second half of the paper is devoted to the proofs of Theorems 1.2 - 1.5: in Section 5 we establish some properties of $\mathcal{T}_\gamma^\infty$ and its associated Gaussian potential, then in Sections 6 - 8 we respectively prove Theorems 1.2 - 1.5.

1.2 Acknowledgements

We would like to thank David Croydon for introducing us to this topic, and Andrea Collevecchio for useful comments on the manuscript.

2 Preliminaries

2.1 Critical Galton-Watson trees

To prove the convergence we will view our Galton-Watson trees as plane trees using the following Ulam-Harris labelling notation for discrete trees. To define these, we follow [60] and first introduce the set

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n.$$

By convention, $\mathbb{N}^0 = \{\emptyset\}$. If $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_m) \in \mathcal{U}$, we also let $uv = (u_1, \ldots, u_n, v_1, \ldots, v_m)$ denote the concatenation of u and v.

Definition 2.1. *A plane tree T is a finite subset of $\mathcal{U}$ such that*

(i) $\emptyset \in T$.

(ii) If $v \in T$ and $v = uj$ for some $j \in \mathbb{N}$, then $u \in T$.

(iii) For every $u \in T$, there exists a number $k_u(T) \geq 0$ such that $uj \in T$ if and only if $1 \leq j \leq k_u(T)$.

We let $\mathbb{T}$ denote the set of all plane trees. If $T \in \mathbb{T}$ and $u \in T$ we also define $\tau_u(T) = \{v \in \mathcal{U} : uv \in T\}$ to be the subtree emanating from u. In this paper we are interested in random plane trees, more specifically *Galton-Watson* trees. To define these, first let ξ be a probability measure on $\mathbb{Z}^{\geq 0}$. We will refer to ξ as the *offspring distribution*.

Definition 2.2. *A Galton-Watson tree with offspring distribution ξ is a plane tree T with law $\mathbb{P}_\xi$ satisfying the following properties:*

(i) $\mathbb{P}_\xi(k_\emptyset = j) = \xi(j)$ *for all* $j \in \mathbb{Z}^{\geq 0}$.

(ii) For every $j \geq 1$ with $\xi(j) > 0$, the shifted trees $\tau_1(T), \ldots, \tau_j(T)$ are independent under the conditional probability $\mathbb{P}_\xi(\cdot \mid k_\emptyset = j)$, with law $\mathbb{P}_\xi$.

In other words, a Galton-Watson tree with offspring distribution ξ is a branching process with a single root $\emptyset$, where the trees emanating from each vertex are independently distributed according to $\mathbb{P}_\xi$. It is shown in [60, Section 3] that for any probability measure ξ on $\mathbb{Z}^{\geq 0}$, there is a unique probability measure $\mathbb{P}_\xi$ on $\mathbb{T}$ satisfying the above two properties.

We will restrict to Galton-Watson trees with a critical aperiodic offspring distribution in the domain of attraction of a γ-stable law for some $\gamma \in (1, 2]$, by which we mean that $\sum_{k=0}^{\infty} k\xi(k) = 1$ and there exists an increasing sequence $a_n \uparrow \infty$ such that, if $(\xi^{(i)})_{i=1}^{\infty}$ are i.i.d. copies of ξ, then

$$\frac{\sum_{i=1}^{n} \xi^{(i)} - n}{a_n} \xrightarrow{(d)} Z_\gamma, \tag{2.1}$$

as $n \to \infty$, where Z_γ is a γ-stable random variable, i.e. can be normalised so that $\mathbb{E}\left[e^{-\lambda Z_\gamma}\right] = e^{-\lambda^\gamma}$ for all $\lambda > 0$. In the finite variance case we always have $\gamma = 2$. It is shown in [24, Section 8.3.2] that necessarily $a_n = n^{\frac{1}{\gamma}} L(n)$ for some slowly-varying function L. Equivalently, $\xi([n, \infty)) = n^{-\gamma} L(n)$ (but not necessarily with the same L). Throughout the paper we will make the assumption that $\gamma \in (1, 2]$, and let $(a_n)_{n=1}^{\infty}$ be the sequence appearing in (2.1).

A Galton-Watson tree T conditioned to have n vertices can be coded by a random walk $(W_m^T)_{0\le m\le n}$ called the Lukasiewicz path; this is defined by setting $W_0^T=0$, then for $1\le m\le n-1$ listing the vertices $u_0,u_1,\dots,u_{n-1}$ in lexicographical order and setting $W_{m+1}^T=W_m^T+k_{u_m}(T)-1$. It is not hard to see that $W_m^T\ge 0$ for all $0\le m\le n-1$, and $W_n^T=-1$.

A discrete tree T conditioned to have n vertices can be also characterised by its contour function. To introduce the contour function, we imagine that the tree is embedded in the plane in such a way that each edge has length one. Consider a particle that is placed at the root of the tree at time $i=0$ and then traverses the tree, moving continuously along the edges at unit speed (respecting the left-right order of the vertices), until all vertices have been explored and the particle has returned back to the root. Then, we denote by $C^T(i)$ the distance to the root of the position of the particle at time i. More specifically, letting x_i denote the i-th visited vertex by the particle, set

$$C^T(i)=d_T(x_0,x_i),\qquad 0\le i<2n.$$

By convention, we also set $x_{2n}=x_0$, and $C^T(2n)=0$. Naturally, we extend C^T by linear interpolation between integer times. Note that the particle visits every vertex apart from the root a number of times given by its degree.

2.2 Gromov-Hausdorff topology and correspondences

In this paper we will be interested in pointed Gromov-Hausdorff-Prokhorov (GHP) convergence between pointed compact metric measure (mm) spaces. Accordingly, let $\mathbb{K}^c$ denote the set of quadruples (X,d,μ,O) such that (X,d) is a compact metric space, μ is a finite Borel measure of full support, and O is a distinguished point, which will play the role of the root. Given two pointed mm-spaces (X,d,μ,O) and (X',d',μ',O'), the pointed GHP distance between them is defined as

$$\inf\left\{d_H^F(\phi(X),\phi'(X'))\vee d_P^F(\phi_*\mu,\phi'_*\mu')\vee\delta(\phi(O),\phi'(O'))\right\},\tag{2.2}$$

where the infimum is taken over all isometric embeddings $\phi:X\to F$, $\phi':X'\to F$ into some common metric space (F,δ), where d_H^F denotes the Hausdorff distance between two subsets of F and where d_P denotes the Prokhorov distance between finite Borel measures on F. For a definition of the latter, see [23, Chapter 1]. If the second term is omitted from the right-hand side of (2.2), this is known simply as the pointed Gromov-Hausdorff (GH) distance, and is denoted by d_{GH}. We say that two spaces (X,d,μ,O) and (X',d',μ',O') in $\mathbb{K}^c$ are equivalent if there exists a measure and a root-preserving isometry between them. It is well-known, see [1, Theorem 2.3], that the pointed GHP distance defines a metric on the space of equivalence classes of $\mathbb{K}^c$, and that this is a Polish metric space.

To prove Theorem 1.1, we will use the helpful notion of *correspondences*. A correspondence $\mathcal{R}$ between two metric spaces (M,R) and (M',R') is a subset of $M\times M'$ such that for every $x\in M$, there exists $y\in M'$ with $(x,y)\in\mathcal{R}$, and similarly for every $y\in M'$, there exists $x\in M$ with $(x,y)\in\mathcal{R}$.

It is straightforward to show (e.g. see [26, Theorem 7.3.25]) that

$$d_{GH}((M,R,O),(M',R',O'))=\frac{1}{2}\inf_{\mathcal{R}}\sup_{(x,x'),(y,y')\in\mathcal{R}}|R(x,y)-R'(x',y')|,$$

where the infimum is taken over all correspondences $\mathcal{R}$ between (M,R) and (M',R') that contain the pair of distinguished points (O,O'). The quantity on the right-hand side above is known as the distortion of $\mathcal{R}$, and is denoted by $\mathsf{dis}(\mathcal{R})$.

In this paper, we will prove GHP convergence by first proving GH convergence using correspondences, and then prove Prokhorov convergence between the measures on the canonical embedding induced by the correspondence.

2.3 Stable trees

If T_n denotes a discrete Galton-Watson tree conditioned to have n vertices with critical aperiodic offspring distribution ξ in the domain of attraction of a γ-stable law, endowed with the graph metric d_n, it is well-known that it admits a compact metric space scaling limit, known as the γ-*stable tree*. We denote this by $\mathcal{T}_\gamma^c$. More precisely, it is shown in [37, Theorem 3.1] that

$$n^{-1}a_nT_n \to \mathcal{T}_\gamma^c \tag{2.3}$$

in the GH topology as $n \to \infty$, where a_n is defined as in (2.1). The limiting space $\mathcal{T}_\gamma^c$ can be formally defined from a γ-stable Lévy excursion which plays the role of a continuum Lukasiewicz path (e.g. see [41] for details). Given a spectrally positive γ-stable Lévy excursion $X^{(\gamma)}$, we define the height function $H^{(\gamma)}$ to be the continuous modification of the process

$$H^{(\gamma)}(t) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^t \mathbb{1}\{X_s^{(\gamma)} < \inf_{r\in[s,t]} X_r^{(\gamma)} + \varepsilon\}ds.$$

The limit above exists in probability, see [40, Lemma 1.1.3].

For $s,t \in [0,1]$, let $m_{H^{(\gamma)}}(s,t) = \inf_{r\in[s\wedge t, s\vee t]} H^{(\gamma)}(r)$ and $d : [0,1]\times[0,1] \to \mathbb{R}_+$ be defined by

$$d(s,t) = H^{(\gamma)}(s) + H^{(\gamma)}(t) - 2m_{H^{(\gamma)}}(s,t). \tag{2.4}$$

It is obvious that d is symmetric and satisfies the triangle inequality. One can introduce the equivalence relation $s \sim t$ if and only if $H^{(\gamma)}(s) = H^{(\gamma)}(t) = m_{H^{(\gamma)}}(s,t)$, which is equivalent to $d(s,t) = 0$. The quotient space $([0,1]/\sim, d)$ is the γ-stable tree $\mathcal{T}_\gamma^c$, which can be proven to be almost surely a compact metric space [41, Theorem 2.1]. In addition, this construction provides a natural way to define a canonical (non-atomic) probability measure associated with it, μ, which has full support. Denote by $p_{H^{(\gamma)}} : [0,1] \to \mathcal{T}_\gamma^c$ the canonical projection. For every $A \in \mathcal{B}(\mathcal{T}_\gamma^c)$, we let

$$\mu(A) = \ell(\{t \in [0,1] : p_{H^{(\gamma)}}(t) \in A\}) \tag{2.5}$$

denote the image measure on $\mathcal{T}_\gamma^c$ of Lebesgue measure ℓ on $\mathbb{R}$ by the canonical projection $p_{H^{(\gamma)}}$.

We briefly outline how to prove (2.3) since we will use a similar strategy in Section 4. Using Skorohod's Representation theorem, we can assume that we are working on a probability space under which the contour function of T_n, which is normalised by setting:

$$C^{(n)}(t) = \left(n^{-1}a_nC^n(2nt); 0 \le t \le 1\right)$$

where C^n is the contour function of T_n, converges almost surely (with respect to the uniform norm) to the height function constructed by a spectrally positive γ-stable Lévy excursion. The convergence in distribution

$$C^{(n)} \xrightarrow{(d)} H^{(\gamma)} \tag{2.6}$$

was originally shown by Duquesne in [38, Theorem 3.1]. We construct the related correspondence

$$\mathcal{R}_n = \left\{(x_{\lfloor 2nt\rfloor}, p_{H^{(\gamma)}}(t)); 0 \le t \le 1\right\}, \tag{2.7}$$

where x_i is the i-th visited vertex in the exploration of the outline of T_n and $p_{H^{(\gamma)}}(t)$ is the equivalence class of t with respect to the relation $\sim$. It is straightforward to show that the distortion of this correspondence between $n^{-1}a_nT_n$ and $\mathcal{T}_\gamma^c$ is upper bounded by $2||C^{(n)} - H^{(\gamma)}||$, where $||C^{(n)} - H^{(\gamma)}||$ stands for the uniform norm of $C^{(n)} - H^{(\gamma)}$. Since (2.6) holds, the convergence of the metric spaces follows. In fact, if μ_n denotes the uniform probability measure on its vertices, it is the case that

$$(T_n, n^{-1}a_nd_n, \mu_n) \to (\mathcal{T}_\gamma^c, d, \mu) \tag{2.8}$$

in distribution with respect to the GHP distance between compact mm-spaces, see [53, Theorem 4.2], which is a corollary of the result originally proved in [37]. Although the uniform probability measure μ_n on the vertices of T_n was not considered in [37], it is not difficult to extend the result in (2.3) to include it since the work regarding the convergence in (2.8) has already been done.

2.4 Infinite critical trees

A critical Galton-Watson tree is almost surely finite, so to study the long-time asymptotics of LERRW on critical trees and more specifically to prove Theorems 1.2 - 1.5, we will instead consider the model on a Galton-Watson tree *conditioned to survive.* In the discrete setting such a model is naturally described by *Kesten's tree*, denoted T_∞ and defined as follows.

Definition 2.3. *[49]. Let ξ be a critical offspring distribution, and define its size-biased version ξ^* by*

$$\xi^*(n) = n\xi(n).$$

The ***Kesten's tree*** *T_∞ associated to the probability distribution ξ is a two-type Galton-Watson tree distributed as follows:*

- *Individuals are either normal or special.*
- *The root of T_∞ is special.*
- *A normal individual produces only normal individuals according to ξ.*
- *A special individual produces individuals according to the size-biased distribution ξ^*. Of these, one of them is chosen uniformly at random to be special, and the rest are normal.*

Almost surely, the special vertices form a unique one-ended infinite backbone of T_∞. Aldous in [4] coined the term *sin-trees* for such trees, since they have a single infinite spine. Although a critical Galton-Watson tree is almost surely finite, Kesten [49, Lemma 1.14] showed that T_∞ arises as the local limit of a critical Galton-Watson tree with offspring distribution ξ as its height goes to infinity.

Kesten's construction has been imitated in the continuum by Duquesne in [39], who constructs continuum sin-trees and shows that these arise as the appropriate local limit of compact continuum trees conditioned on being large. Duquesne's construction involves defining two height functions from two independent Lévy processes in the same way as done for the compact case. In the stable case, we denote the infinite stable sin-tree by $\mathcal{T}_\gamma^\infty$. It is also possible to show that $\mathcal{T}_\gamma^\infty$ is the scaling limit of a discrete tree constructed from a critical aperiodic γ-stable offspring distribution as constructed in Definition 2.3 (e.g. by following a similar strategy to [9]). As in the discrete case, $\mathcal{T}_\gamma^\infty$ has a single infinite path to infinity, to which compact stable trees are grafted.

3 Connection between LERRW and a RWRE on trees

The fine mesh limit of the LERRW in dimension one, which was called linearly reinforced motion (LRM), was introduced in [56]. In fact, the authors construct the continuous space limit of the vertex-reinforced jump process (VRJP) out of a convergent

Bass-Burdzy flow [20]. To obtain LRM as the continuous space limit of the LERRW, they then use the close relation between the VRJP and the LERRW that was established in [64], namely that the LERRW can be represented in terms of a VRJP with independent gamma conductances. The principles used in [56] are broadly the same as those used in this paper.

In the setting of [56] an appropriate potential related to the VRJP converges. This yields a characterisation of the limiting LRM as a diffusion in random potential that contains a Wiener term and a drift (the motion drifts towards the places it has already visited many times), or equivalently describes the LRM as a scale-transformed and time-changed diffusion in a random environment, cf. [67, 25, 66, 27, 61].

In our setting, we work directly on the trees and view them as electrical networks equipped with a so-called resistance metric and a measure that we will specify in Section 3.1. The latter provide the natural scale functions and speed measures of a RWRE associated with the LERRW in a representation of Pemantle, see Section 3.2. Then we are able to use the main result of [31] to yield convergence of the RWRE as a consequence of the convergence of the resistance metric and the speed measure. Since the resistance metric and the speed measure are expressed in terms of the potential of the RWRE, their convergence can be deduced from the convergence of the aforementioned potential, see Sections 4.1 and 4.2.

The limiting resistance metric and speed measure on the limiting stable tree are distortions of its canonical metric and uniform mass measure. These distortions are expressed in terms of an exponential that includes a tree-indexed Gaussian term and a drift. The Gaussian term corresponds to a tree-indexed process that is essentially a Brownian snake on the stable tree. This yields a characterisation of the limiting diffusion X as a diffusion in random potential on the stable tree, see Section 3.3. We note here that the notion of a "scale change" on general separable real trees was formalised in [11].

3.1 Random walk in a random environment

In order to study a LERRW on T_n we will use a representation of a LERRW as a RWRE. We briefly introduce the formalism of a RWRE here.

In a rooted metric tree (T, d_T, O), we define for $u, v \in T$ the path intervals

$$[u,v] = \{z \in T : d_T(u,z) + d_T(z,v) = d_T(u,v)\},$$

and

$$(u,v) = [u,v] \setminus \{u,v\}, \quad [u,v) = [u,v] \setminus \{v\}, \quad (u,v] = [u,v] \setminus \{u\}.$$

We say that u and v are connected by an edge, denoted $u \sim v$, precisely when $u \neq v$ and $[u,v] = \{u,v\}$. We also define a partial order on T by setting $u \preceq v$ (u is an ancestor of v) if and only if $u \in [O,v]$. We write $u \prec v$ if $u \preceq v$ and $u \neq v$. Finally, the unique $z \in T$ for which $[O,u] \cap [O,v] = [O,z]$ is written $u \wedge v$. We call $u \wedge v$ the most recent common ancestor of u and v.

For a fixed tree T with root O, a RWRE on T can be constructed by assigning each edge $e \in E(T)$ a random conductance. At each time, if a random walk is currently at vertex v it moves to one of the neighbours of v with a probability proportional to the conductance of the edge joined to v. Rather than defining the full set of conductances $(c(e))_{e \in E(T)}$, we can equivalently assign a single conductance to a single special "root edge" e_{root} and then sample a sequence of random variables $(\mathbf{W}_v)_{v \in T}$, so that

$$\mathbf{W}_v = (W_v^u)_{u \sim v}$$

gives the ratios of the conductances of edges emanating from v. We will follow this approach as it allows us to formalise the connection to Pólya urns outlined in the introduction.

We now assume that an environment $\omega = \big(c(e_{\text{root}}), \mathbf{W} = (\mathbf{W}_v)_{v \in T}\big)$ is fixed. For each non-root vertex v, let $\overleftarrow{v}$ denote the parent of v. Let

$$\rho_v(\omega) = \frac{W_{\overleftarrow{v}}^{\overleftarrow{\overleftarrow{v}}}}{W_{\overleftarrow{v}}^{v}}, \tag{3.1}$$

which in our case represents the transition probability from $\overleftarrow{v}$ to the ancestor $\overleftarrow{\overleftarrow{v}}$ of $\overleftarrow{v}$, divided by the transition probability from $\overleftarrow{v}$ to the descendant v of $\overleftarrow{v}$.

We add a new vertex $\overleftarrow{O}$, which we call the base and attach it to the root with a new edge. This edge will play the role of e_{root} and we call the new tree the planted tree. To keep the notation simple, even if the statements are expressed in terms of this new planted tree, say T^*, we still phrase them in terms of T. It follows from (3.1) that the conductance of an edge $\{\overleftarrow{v}, v\}$ is given by

$$c(\{\overleftarrow{v}, v\}) = c(\{\overleftarrow{O}, O\}) \cdot e^{-\sum_{O \prec u \preceq v} \log \rho_u(\omega)}. \tag{3.2}$$

This motivates the definition of the potential on T as

$$V_\omega(x) = \begin{cases} \displaystyle\sum_{O \prec v \preceq x} \log \rho_v(\omega), & \text{for } x \neq O, \\ 0, & \text{for } x = O. \end{cases}$$

By defining the RWRE from conductances in this way, we have in fact defined it via an electrical network and can therefore take advantage of electrical network theory (e.g. see [55, Chapter 9] for an introduction). The conductance of an edge $\{\overleftarrow{v}, v\}$ is given by (3.2), and consequently its electrical resistance is given by

$$R(\{\overleftarrow{v}, v\}) = c(\{\overleftarrow{v}, v\})^{-1} = c(\{\overleftarrow{O}, O\})^{-1} \cdot e^{V_\omega(v)}.$$

Moreover, since our graph is a tree, the electrical resistance between any two vertices is given by

$$R(u, v) = \sum_{e \in E_{u,v}} R(e) = c(\{\overleftarrow{O}, O\})^{-1} \cdot \sum_{x \in [u,v] \setminus \{u \wedge v\}} e^{V_\omega(x)}, \tag{3.3}$$

where in the first sum, $E_{u,v}$ is the set of edges contained in $[u, v]$. Clearly, R defines a metric on T (this also follows from a more general result of Tetali [71]). We will take the measure

$$\mu(x) = \begin{cases} c(\{\overleftarrow{O}, O\}) \cdot \left(e^{-V_\omega(x)} \mathbb{1}_{\{x \neq O\}} + \displaystyle\sum_{y \sim x :\, \overleftarrow{y} = x} e^{-V_\omega(y)} \right), & \text{for } x \neq \overleftarrow{O}, \\ 0, & \text{for } x = \overleftarrow{O}. \end{cases} \tag{3.4}$$

This is the stationary reversible measure for a stochastic process associated with (T, R, O). This stochastic process has generator which acts on $L^2(T, \mu)$ as follows:

$$(\Delta f)(x) = \frac{1}{\mu(x)} \sum_{y \in T : y \sim x} c(\{x, y\})(f(y) - f(x))$$

(see [30, Section 2] for more details on the correspondence between graphs equipped with edge conductances and a measure, and stochastic processes). This stochastic process is a continuous-time random walk Z having $\mathsf{exp}(1)$ holding time at each non-root vertex, and at each jump time, the random walk traverses an edge incident to its previous state, chosen with probability proportional to the conductance of each of the available edges. At the base, Z jumps from $\overleftarrow{O}$ to O with rate $c(\{\overleftarrow{O}, O\})/\mu(\overleftarrow{O}) = \infty$, i.e. the expected holding time at $\overleftarrow{O}$ is 0. The random walk transitions from the base to the root with probability 1. Moreover, the overall rate at which Z jumps from O to its collection of children is 1.

Remark 3.1. The electrical resistance between x and y can be alternatively defined by

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

where

$$\mathcal{E}(f,g) = \frac{1}{2} \sum_{\substack{x,y \in T: \\ x \sim y}} c(\{x,y\})(f(x) - f(y))(g(x) - g(y))$$

is a quadratic form on T. In fact, $\mathcal{E}$ is a regular Dirichlet form on $L^2(T, \mu)$, and corresponds to electrical energy. For such functions f and g, we note that

$$\mathcal{E}(f,g) = -\sum_{x \in T} (\Delta f)(x) g(x) \mu(x).$$

We will not directly use the machinery of Dirichlet forms in this paper.

The RWRE that we use in this paper will be defined by a collection of Dirichlet random variables taking the role of $(\mathbf{W}_v)_{v \in T}$.

Definition 3.2 (Dirichlet distribution)**.** *For a finite set I and positive real numbers $(b_i)_{i \in I} \in (0, +\infty)^I$, the Dirichlet distribution $\mathcal{D}((b_i)_{i \in I})$ on*

$$\Sigma_I = \left\{ (u_i)_{i \in I} \in (0,1)^{|I|} : \sum_{i \in I} u_i = 1 \right\}$$

has density

$$\frac{\Gamma(\sum_{i \in I} b_i)}{\prod_{i \in I} \Gamma(b_i)} \prod_{i \in I} u_i^{b_i - 1} \mathbb{1}_{\Sigma_I}((u_i)_{i \in I}) \prod_{j \in I \setminus \{j_0\}} du_j,$$

for any choice of $j_0 \in I$.

The Dirichlet distribution can alternatively be defined as the law of a normalised Gamma vector. One can indeed check that the following lemma holds.

Lemma 3.3. *[63, Section 0.3.2]. Let $(W_i)_{i \in I}$ be independent random variables such that, for $i \in I$,*

$$W_i \sim \Gamma(b_i, 1), \text{ i.e. with density } \frac{1}{\Gamma(b_i)} w^{b_i - 1} e^{-w} \mathbb{1}_{(0,+\infty)}(w) dw. \tag{3.5}$$

Then,

$$(U_i)_{i \in I} = \frac{1}{\sum_{i \in I} W_i} \cdot (W_i)_{i \in I} \sim \mathcal{D}((b_i)_{i \in I}). \tag{3.6}$$

Furthermore, $(U_i)_{i \in I}$ is independent of $\sum_{i \in I} W_i$.

3.2 Representation of LERRW as a random walk in random conductances

In this section we outline the connection between LERRW, Pólya urns and Dirichlet random variables on trees. This connection was first used in the context of (non-directed) edge reinforced random walk on trees by Pemantle [62] where, due to the absence of cycles, the process evolves independently at each vertex.

Given a tree T, let $\alpha = (\alpha_e)_{e\in E(T)} \in (0,+\infty)^{E(T)}$ be a sequence of positive initial weights on the edges, and let O denote the root of T. We consider edges to be undirected, so that $e = \{x, y\} = \{y, x\}$ denotes the edge joining two vertices x and y.

Definition 3.4. *The discrete-time linearly edge reinforced random walk (LERRW) on T with initial weights α and starting at O is the process on T with law $P_O^{(\alpha)}$ defined by: $P_O^{(\alpha)}$-a.s., $X_0 = O$ and, for all $n \geq 0$, for all edges $e \in E(T)$,*

$$P_O^{(\alpha)}(X_{n+1} = v | X_0, ..., X_n) = \frac{N_{\{w,v\}}(n)}{\sum_{y:y\sim w} N_{\{w,y\}}(n)} \mathbb{1}_{\{X_n=w\}},$$

*where $N_e(n) = \alpha_e + \#\{0 \leq k \leq n-1 : (X_k, X_{k+1}) = e\}\Delta$. We call $\Delta > 0$ the **reinforcement parameter**.*

In other words, at time n the walk jumps along a neighbouring edge e chosen with probability proportional to its current weight $N_e(n)$. This weight is initially equal to α_e and then increases by Δ each time the edge e is traversed.

Consider a vertex $v \in T$ and suppose v has $\#v$ offspring in T. Let $e_0(v)$ denote the edge joining v to its parent, and $e_1(v), ..., e_{\#v}(v)$ denote the edges joining v to each of its offspring. When the LERRW arrives at v for the first time, it must have arrived via the parent, so the weights of these edges will be respectively given by the components of the vector

$$(\alpha_{e_0(v)} + \Delta, \alpha_{e_1(v)}, ..., \alpha_{e_{\#v}(v)}).$$

Moreover, since T is a tree, if the LERRW exits v by edge $e_i(v)$, it must also return to v via $e_i(v)$, so that at the next visit to v the weight of $e_i(v)$ will have increased by 2Δ and all other weights will have remained the same. Moreover, this holds independently of what happens to the LERRW between its successive traversals of $e_i(v)$. The same logic applies on subsequent visits to v, with the weights updating each time.

It follows that the decisions of the LERRW process are ruled by independent Pólya urns, one per vertex, where outgoing edges play the role of colours, and $\alpha_{e_i(v)}$ determines the initial number of balls of colour i. Since the asymptotic proportion of colours in such an urn model follows the Dirichlet distribution, conditional on which the draws are independent and identically distributed as Bernoulli random variables with success probability given by the asymptotic proportion of the drawn colour, the reinforced walk may equivalently be obtained by assigning a Dirichlet random variable $\omega_{(x,\cdot)}$ to each vertex x, which then defines the transition probabilities of the walk every time it is at x. This is the definition of the random walk in Dirichlet random environment (RWDE) that we define below. To get this equivalent representation of LERRW we have to average over all the Dirichlet random variables.

Given $v \in T$, let the positive initial weights from v to its neighbouring vertices $\overleftarrow{v}, v_1, ..., v_{\#v}$ be denoted by the vector

$$\alpha_v = (\alpha_{e_0(v)}, \alpha_{e_1(v)}, ..., \alpha_{e_{\#v}(v)}),$$

with $\alpha_{e_0(v)}$ being the positive weight to the parent of v. We define the set of environments on T as

$$\Omega_T = \left\{ (\omega_v = (\omega_v^e)_{v\in e})_{v\in T} \in \prod_{v\in T} \Sigma_{\deg v} \right\}, \qquad \text{where } \Sigma_j = \left\{ (u_i)_{i\leq j} \in (0,1)^j : \sum_{i\leq j} u_i = 1 \right\},$$

for $j \geq 1$. (So for $v \in T$, $\Sigma_{\deg v}$ is the $(\deg v)$-dimensional simplex). We shall denote by ω a random environment sampled from Ω_T.

Definition 3.5. *For $\omega = (\omega_v)_{v\in T} \in \Omega_T$, the quenched random walk in random environment ω starting at O is the Markov chain on T starting at O and with transition* *probabilities* $(\omega_v)_{v\in T}$*. We denote its law by* $\tilde{P}_{O,\omega}$*. Thus,* $\tilde{P}_{O,\omega}$*-a.s.,* $X_0 = O$ *and for all* $n \geq 0$*, for all edges* $e \in E(T)$*,*

$$\tilde{P}_{O,\omega}((X_n, X_{n+1}) = e | X_0, ..., X_n) = \omega_v^e \mathbb{1}_{\{X_n = v\}}.$$

We define the Dirichlet distribution on T with parameters $(\mathbf{b}_v = (b_v^e)_{v\in e})_{v\in T}$ as the product distribution:

$$\mathbb{P}^{(\mathbf{b})} = \prod_{v\in T} \mathcal{D}\left(\mathbf{b}_v\right).$$

where each $\mathcal{D}\left(\mathbf{b}_v\right)$ is as in Definition 3.2.

Now let us consider the joint law $\tilde{P}_O^{(\mathbf{b})}$ of (ω, X) on $\Omega_T \times V(T)^{\mathbb{N}}$ such that $\omega \sim \mathbb{P}^{(\mathbf{b})}$ and the conditional distribution of X given ω is $\tilde{P}_{O,\omega}$. Its law is thus

$$\tilde{P}_O^{(\mathbf{b})}\left(X \in \cdot\right) = \int \tilde{P}_{O,\omega}(X \in \cdot)\mathbb{P}^{(\mathbf{b})}(d\omega).$$

Recall from Definition 3.4 that $P_O^{(\alpha)}$ denotes the law of a LERRW started at O with initial edge weights $\alpha = (\alpha_e)_{e\in E(T)}$. The next result can be found in [62, Section 3] and formalises the representation via Pólya urns.

Lemma 3.6. *[62, Lemma 2]. Fix a tree* T*. If* $\alpha = (\alpha_e)_{e\in E(T)}$ *is a collection of positive initial weights for a LERRW, then for any* $k \geq 0$ *and any* $O = x_0, x_1, ..., x_k \in T$*, we have that*

$$P_O^{(\alpha)}(X_1 = x_1, ..., X_k = x_k) = \tilde{P}_O^{(\mathbf{b})}(X_1 = x_1, ..., X_k = x_k),$$

where $\mathbf{b} = (\mathbf{b}_v)_{v\in T}$ *is a collection of vectors of positive weights such that if* $v \in T$ *with* $\#v$ *offspring,*

$$\mathbf{b}_v = \left((\alpha_{e_0(v)} + \Delta)/(2\Delta), \alpha_{e_1(v)}/(2\Delta), \cdots, \alpha_{e_{\#v}(v)}/(2\Delta)\right).$$

Remark 3.7. On general (non-tree) graphs this representation of non-directed LERRW is not available since a random walk may return to a vertex v via a different edge from that which it used to leave v. Therefore, the urn models at each vertex do not update independently of each other. However, on general graphs, non-directed LERRWs can still be seen as random walks in an explicit correlated random environment. For instance, see [64]. Additionally, it is possible to represent a directed LERRW on general graphs as a RWDE, see for example [65, Lemma 2].

In this paper we want to consider a LERRW on the *random* tree T_n. We consider the quenched law of LERRW on T_n, by which we mean that we first sample T_n and then run a LERRW on T_n started at $\overleftarrow{O_n}$ according to Definition 3.4. In order to obtain a non-trivial scaling limit, we will consider a LERRW on T_n with initial weights given by (1.1) and reinforcement parameter Δ. In light of Lemma 3.6, we will construct the scaling limit of LERRW on T_n by instead constructing the scaling limit of the corresponding RWDE on T_n, which can be achieved by representing it as an electrical network endowed with a resistance metric and a measure as outlined in Section 3.1. This is equivalent to sampling a random environment $\omega = (c(e_{\text{root}}), \mathbf{W}^{(n)})$ and defining a resistance metric R_n and a measure ν_n from ω as in (3.3) and (3.4). The edge $\{\overleftarrow{O_n}, O_n\}$ will play the role of e_{root}.

To sample the random environment, we therefore define the set of parameters $\mathbf{b}^{(n)} = (\mathbf{b}_v^{(n)})_{v\in T_n}$ such that, if $v \in T_n$ has $\#v$ offspring,

$$\mathbf{b}_v^{(n)} = \left((\alpha_{e_0(v)}^{(n)} + \Delta)/(2\Delta), \alpha_{e_1(v)}^{(n)}/(2\Delta), \ldots, \alpha_{e_{\#v}(v)}^{(n)}/(2\Delta)\right). \tag{3.7}$$

We will sample the Dirichlet distribution with parameter $\mathbf{b}_v^{(n)}$ using Lemma 3.3. Since we are only interested in the ratios of the Dirichlet weights, we can work directly with the non-normalised Gamma weights, so we assume that our random environment $\omega_n = (c(\{\overleftarrow{O_n}, O_n\}), \mathbf{W}^{(n)})$ is obtained by sampling $(\mathbf{W}_v^{(n)})_{v \in T_n}$ according to (3.5), and without loss of generality take $c(\{\overleftarrow{O_n}, O_n\}) = 1$ for simplicity.

In the setup of Section 3.1, we therefore have that

$$\rho_v(\omega_n) = \frac{W_{\overleftarrow{v}}^{\overleftarrow{\overleftarrow{v}}}}{W_{\overleftarrow{v}}^{v}}, \qquad V_{\omega_n}(x) = \sum_{O_n \prec v \preceq x} \log \rho_v(\omega_n) \mathbb{1}\{x \neq O_n, \overleftarrow{O_n}\},$$
$$R^{\omega_n}(u,v) = \sum_{x \in [u,v] \setminus \{u \wedge v\}} e^{V_{\omega_n}(x)},$$

$$\nu^{\omega_n}(x) = \begin{cases} e^{-V_{\omega_n}(x)} \mathbb{1}\{x \neq O_n\} + \sum_{y \sim x : \overleftarrow{y} = x} e^{-V_{\omega_n}(y)} & \text{if } x \neq \overleftarrow{O_n}, \\ 0 & \text{if } x = \overleftarrow{O_n}. \end{cases} \tag{3.8}$$

For simplicity, we henceforth suppress the dependence on ω_n and write $\rho_v, V^{(n)}, R_n$ and ν_n instead of $\rho_v(\omega_n)$, V_{ω_n}, R^{ω_n} and ν^{ω_n} respectively. Moreover, we let $Z^{(n,\mathrm{pl})}$ denote the stochastic process associated with the triple (T_n^*, R_n, ν_n), which was called the planted tree, see Section 3.1. Since $Z^{(n,\mathrm{pl})}$ is defined on the planted tree, we define $Z^{(n)}$ to be its trace on the unplanted tree T_n; that is, we set

$$Z^{(n)}(t) = \begin{cases} Z^{(n,\mathrm{pl})}(t), & \text{if } Z^{(n,\mathrm{pl})}(t) \neq \overleftarrow{O_n}, \\ O_n, & \text{if } Z^{(n,\mathrm{pl})}(t) = \overleftarrow{O_n}. \end{cases}$$

It follows from our definitions that $Z^{(n)}(t)$ is a continuous-time random walk on T_n, with exp(1) holding times at each vertex, and that at each jump time, the transition probabilities from a vertex are given by the Dirichlet weights of (3.6).

Finally, in light of Lemma 3.6, in order to connect this with LERRW we need to consider the law of $Z^{(n)}$ annealed over the Dirichlet weights. We denote this process $Z^{(n)}$ and its law $\tilde{P}_{O_n}^{(\mathbf{b}^{(n)})}$, so that

$$\tilde{P}_{O_n}^{(\mathbf{b}^{(n)})}\left(Z^{(n)} \in \cdot\right) = \int \tilde{P}_{O_n,\omega_n}(Z^{(n)} \in \cdot) \mathbb{P}^{(\mathbf{b}^{(n)})}(d\omega_n). \tag{3.9}$$

It follows from Lemma 3.6 that under $\tilde{P}_{O_n}^{(\mathbf{b}^{(n)})}$, $Z^{(n)}$ has the law of a *quenched* (with respect to the randomness of T_n) continuous-time LERRW on T_n, with initial weights (1.1) and reinforcement parameter Δ.

3.3 Candidate for the scaling limit

To construct the scaling limit of (T_n, R_n, ν_n) we need to imitate the definitions of (3.8) in the continuum. To this end, we take $\mathcal{T}_\gamma^c$ as in Section 2.3, denote its root by O and for all $\sigma, \sigma' \in \mathcal{T}_\gamma^c$ we set (cf. (2.4)):

$$d^{(\alpha)}(\sigma,\sigma') = \begin{cases} (d(O,\sigma)+1)^{1-\alpha} + (d(O,\sigma')+1)^{1-\alpha} - 2(d(O,\sigma\wedge\sigma')+1)^{1-\alpha}, & \text{if } \alpha < 1, \\ \log(d(O,\sigma)+1) + \log(d(O,\sigma')+1) - 2\log(d(O,\sigma\wedge\sigma')+1), & \text{if } \alpha = 1, \end{cases}$$

where $\sigma \wedge \sigma'$ denotes the most recent common ancestor of σ and σ'. This is a metric on $\mathcal{T}_\gamma^c$.

Definition 3.8 (Snake process). *Let $(\phi^{(\alpha)}(\sigma))_{\sigma\in\mathcal{T}_\gamma^c}$ be the $\mathbb{R}$-valued Gaussian process with law $\mathbb{P}$ whose distribution, given $\mathcal{T}_\gamma^c$ is characterised by*

$$\mathbb{E}\phi^{(\alpha)}(\sigma)=0, \qquad \mathrm{Cov}(\phi^{(\alpha)}(\sigma),\phi^{(\alpha)}(\sigma'))=d^{(\alpha)}(O,\sigma\wedge\sigma'). \tag{3.10}$$

Since

$$\mathbb{E}|\phi^{(\alpha)}(\sigma)-\phi^{(\alpha)}(\sigma')|^2=d^{(\alpha)}(\sigma,\sigma'),$$

which is a metric on $\mathcal{T}_\gamma^c$, it can be seen by the same arguments as in [41, Section 6] that such a process $\phi^{(\alpha)}$ exists, and that it has a continuous modification, with bounded sample paths, $\mathbf{P}\times\mathbb{P}$-almost surely. In the continuum, $\phi^{(\alpha)}$ will play the same role as the potential V_{ω_n} in (3.8).

To capture that our diffusions are processes on "natural scale" it is desirable to introduce the notion of a length measure on $\mathcal{T}_\gamma^c$ which extends Lebesgue measure on $\mathbb{R}$. For real trees it was first presented in [44] and later extended to any separable 0-hyperbolic metric space in [13]. Denote the skeleton of $(\mathcal{T}_\gamma^c,d)$ by

$$(\mathcal{T}_\gamma^c)^0=\bigcup_{x\in\mathcal{T}_\gamma^c}(O,x).$$

Recalling that $\mathcal{T}_\gamma^c$ is a separable pointed metric space, observe that if $D\subset\mathcal{T}_\gamma^c$ is a dense countable subset, then the previous definition is still the same when the union is taken over points in D. In particular, $(\mathcal{T}_\gamma^c)^0\in\mathcal{B}(\mathcal{T}_\gamma^c)$ and $\mathcal{B}(\mathcal{T}_\gamma^c)|_{(\mathcal{T}_\gamma^c)^0}=\sigma(\{(x,y);x,y\in D\})$. Therefore, by Carathéodory's Extension Theorem, there exists a unique σ-finite measure λ^γ on $\mathcal{T}_\gamma^c$, called the length measure, such that $\lambda^\gamma(\mathcal{T}_\gamma^c\setminus(\mathcal{T}_\gamma^c)^0)=0$ and

$$\lambda^\gamma((O,x])=d(O,x), \qquad x\in\mathcal{T}_\gamma^c. \tag{3.11}$$

Alternatively, the length measure is the trace onto the skeleton of the γ-stable tree of the one-dimensional Hausdorff measure on it.

Now fix the reinforcement parameter $\Delta>0$ and let $\mathcal{W}$ be the space of continuous functions $\mathcal{T}_\gamma^c\to\mathbb{R}$ vanishing at the root, and let Ω be the space of processes $\mathbb{R}_+\to\mathcal{T}_\gamma^c$. Given a realisation $\phi^{(\alpha)}\in\mathcal{W}$, we firstly define

$$R_\phi(x,y)=\begin{cases}\displaystyle\int_{[x,y]}(d(O,z)+1)^{-\alpha}e^{\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(z)+\frac{\Delta}{1-\alpha}[(d(O,z)+1)^{1-\alpha}-1]}\lambda^\gamma(\mathrm{d}z), & \text{if } \alpha<1,\\ \displaystyle\int_{[x,y]}e^{\sqrt{4\Delta}\phi^{(1)}(z)+(\Delta-1)\log(d(O,z)+1)}\lambda^\gamma(\mathrm{d}z), & \text{if } \alpha=1,\end{cases} \tag{3.12}$$

for all $x,y\in\mathcal{T}_\gamma^c$, and secondly define ν_ϕ to be the measure which is absolutely continuous with respect to the measure μ defined in (2.5) with density given by

$$(\mathrm{d}\nu_\phi/\mathrm{d}\mu)(x)=\begin{cases}(d(O,x)+1)^{\alpha}e^{-\left[\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(x)+\frac{\Delta}{1-\alpha}[(d(O,x)+1)^{1-\alpha}-1]\right]}, & \text{if } \alpha<1,\\ e^{-\left[\sqrt{4\Delta}\phi^{(1)}(x)+(\Delta-1)\log(d(O,x)+1)\right]}, & \text{if } \alpha=1.\end{cases} \tag{3.13}$$

We let $\mathcal{T}_{\gamma,\alpha}$ denote the random mm-space $(\mathcal{T}_\gamma^c,R_\phi,\nu_\phi)$, let $X=(X_t)_{t\geq0}$ denote the process canonically associated with it via the resistance form of [31, Definition 2.1], and let $\tilde{P}_{O,\phi}$ denote the quenched law of X when started from O. Finally, given $\mathcal{T}_\gamma^c$, we denote by P_O the corresponding annealed probability measure of the process on $\mathcal{W}\times\Omega$ defined by

$$P_O\left((X_t)_{t\geq0}\in\cdot\right)=\int\tilde{P}_{O,\phi}\left((X_t)_{t\geq0}\in\cdot\right)\mathbb{P}(d\phi^{(\alpha)}). \tag{3.14}$$

Remark 3.9. We can also view $((X_t)_{t\geq 0}, \tilde{P}_{O,\phi})$ as an ν_ϕ-Brownian motion on $(\mathcal{T}_\gamma^c, R_\phi)$ as characterised in [11, Proposition 1.9]. In this setting, $(X_t)_{t\geq 0}$ is a diffusion process with regular Dirichlet form

$$\mathcal{E}_\phi(u,v) = \begin{cases} \frac{1}{2}\int \nabla_{R_\phi} u(z)\nabla_{R_\phi} v(z)(d(O,z)+1)^{-\alpha} e^{\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(z)+\frac{\Delta}{1-\alpha}[(d(O,x)+1)^{1-\alpha}-1]}\lambda^\gamma(\mathrm{d}z), \\ \frac{1}{2}\int \nabla_{R_\phi} u(z)\nabla_{R_\phi} v(z) e^{\sqrt{4\Delta}\phi^{(1)}(z)+(\Delta-1)\log(d(O,z)+1)}\lambda^\gamma(\mathrm{d}z), \end{cases}$$

for every $u,v \in \mathcal{D}(\mathcal{E}_\phi)$, where $\mathcal{D}(\mathcal{E}_\phi) = \{u \in \mathcal{A} : \nabla u \in L^2(\lambda)\} \cap L^2(\nu_\phi) \cap \mathcal{C}_\infty(\mathcal{T}_\gamma^c)$. As usual, by $\mathcal{C}_\infty(\mathcal{T}_\gamma^c)$ we refer to the space of continuous functions which vanish at infinity. A function u belongs to $\mathcal{A}$ if and only if u is locally absolutely continuous. We need to stress that the gradients of u and v correspond to gradients of functions from $(\mathcal{T}_\gamma^c, R_\phi)$, i.e. $\nabla_{R_\phi} f$, of $f \in \mathcal{D}(\mathcal{E}_\phi)$, is the function, which is unique up to

$$\lambda^{(\mathcal{T}_\gamma^c, R_\phi)}(\mathrm{d}z) = \begin{cases} (d(O,x)+1)^{-\alpha} e^{\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(z)+\frac{\Delta}{1-\alpha}[(d(O,x)+1)^{1-\alpha}-1]}\lambda^\gamma(\mathrm{d}z), & \text{if } \alpha<1, \\ e^{\sqrt{4\Delta}\phi^{(1)}(z)+(\Delta-1)\log(d(O,z)+1)}\lambda^\gamma(\mathrm{d}z), & \text{if } \alpha=1, \end{cases}$$

-zero sets that satisfies

$$\int_{[x_1,x_2]} \nabla_{R_\phi} f(z)\lambda^{(\mathcal{T}_\gamma^c, R_\phi)}(\mathrm{d}z) = f(x_2)-f(x_1), \qquad x_1,x_2 \in \mathcal{T}_\gamma^c, \qquad f \in \mathcal{D}(\mathcal{E}_\phi).$$

We will not directly use the theory of Dirichlet forms in this paper (though it is hidden behind the result we apply from [31]).

Remark 3.10 (Connection to 1d results)**.** The analogue of [56, Proposition 1.9] (which is written with a scaling factor of 2^n rather than na_n^{-1}) equally applies to the LERRW on a single branch of $n^{-1}a_n T_n$ with initial weights

$$w_0^{(n)}(x, x-n^{-1}a_n) = \frac{1}{2}na_n^{-1}L_0(x)L_0(x-n^{-1}a_n).$$

Comparing with (1.1), we see that we can formulate our process in the setting of [56, Proposition 1.9] by taking $L_0(x)^2 \approx 2\Delta^{-1}(|x|+1)^\alpha$. In this case, $S_0(x)$ as defined by [56, Equation (1.5)] satisfies $S_0(x) \approx \frac{\Delta[(|x|+1)^{1-\alpha}-1]}{2(1-\alpha)}$ if $\alpha<1$, and $S_0(x) \approx \frac{\Delta\log(|x|+1)}{2}$ if $\alpha=1$.

According to [56, Proposition 1.9], such a rescaled LERRW on $\mathbb{Z}$ converges to a diffusion in the random potential

$$2\sqrt{2}W(S_0(x)) + 2|S_0(x)| - 2\log(L_0(x)) - \log\Delta,$$

cf. [56, Equation (1.8)], where W is a standard Brownian motion on $\mathbb{R}$ (the final term of $\log\Delta$ above appears since we have rescaled everything by Δ to fit into the framework of [56]). In particular, we can identify $W((|x|+1)^{1-\alpha}-1)$ with $\phi^{(\alpha)}(|x|)$ and $W(\log(|x|+1))$ with $\phi^{(1)}(|x|)$, with a slight abuse of notation. Substituting the above values of $L_0(x)$ and $S_0(x)$ we obtain a limiting potential of the form

$$\begin{aligned} &2\sqrt{2}W\left(\frac{\Delta[(|x|+1)^{1-\alpha}-1]}{2(1-\alpha)}\right) + \frac{\Delta[(|x|+1)^{1-\alpha}-1]}{1-\alpha} - \log(2(|x|+1)^\alpha) \\ &= \sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(|x|) + \frac{\Delta[(|x|+1)^{1-\alpha}-1]}{1-\alpha} - \alpha\log(|x|+1) - \log 2 \end{aligned}$$

when $\alpha<1$, and

$$\begin{aligned} &2\sqrt{2}W\left(\frac{\Delta\log(|x|+1)}{2}\right) + \Delta\log(|x|+1) - \log(2(|x|+1)) \\ &= \sqrt{4\Delta}\phi^{(1)}(|x|) + (\Delta-1)\log(|x|+1) - \log 2 \end{aligned}$$

when $\alpha=1$, which is consistent with (3.12) up to the constant $\log 2$ (but note that the effect of adding a constant cancels out when inserted into both (3.12) and (3.13)).

4 Scaling limit of (T_n, R_n, ν_n) and the LERRW $X^{(n)}$

Our goal in this section is to prove the following proposition. The space (T_n, R_n, ν_n) is as defined just below (3.8).

Proposition 4.1. *Under the joint law $\mathbf{P} \times \mathbb{P}$ and with initial weights as in* (1.1)*, the following convergence holds with respect to the GHP topology as $n \to \infty$:*

$$(T_n, (na_n^{-1})^{-1} R_n, (2n)^{-1}\nu_n, O_n) \overset{(d)}{\to} (\mathcal{T}_\gamma^c, R_\phi, \nu_\phi, O).$$

Throughout the section, we will work pointwise on the probability space $(\mathbf{\Omega}, \mathcal{F}, \mathbf{P})$ (on which we defined (T_n, d_n, μ_n)), and where the convergence of (2.3) holds almost surely. This means that most of the statements that follow should be written conditionally on T_n or on $\mathcal{T}_\gamma^c$. To make the arguments clearer to follow, we have not written this explicitly in the statements or the proofs, and instead ask the reader to keep this in mind throughout. There is one specific case (Proposition 4.8) where we need to restrict to a certain set $A_{n,\varepsilon} \subset \mathbf{\Omega}$, and in this case we make it explicit.

We also let $V^{(n,\alpha)}(2nt) = V_{\omega_n}(x_{\lfloor 2nt \rfloor})$, where V_{ω_n} is as defined in (3.8). At times we will abuse notation and write $V^{(n,\alpha)}(x)$ in place of $V^{(n,\alpha)}(2nt)$, where t is such that $x = x_{\lfloor 2nt \rfloor}$.

We will prove Proposition 4.1 in two main steps. We first apply Skorohod Representation and assume that (2.6) holds almost surely on $(\mathbf{\Omega}, \mathcal{F}, \mathbf{P})$. On this space, the set of all pairs $(x_{\lfloor 2nt \rfloor}, p_{H_{(\gamma)}}(t))$ defines a correspondence between $n^{-1}a_n T_n$ and $\mathcal{T}_\gamma^c$ with distortion going to 0; see (2.7) and Section 2.2 for details.

We first prove the following claim.

Claim 4.2. *Take $\alpha \le 1$. We have for almost every $\omega \in \mathbf{\Omega}$ that*

$$\left(V^{(n,\alpha)}(2nt)\right)_{t\in[0,1]} \overset{(d)}{\to} \begin{cases} \left(\frac{\Delta[(d(O,t)+1)^{1-\alpha}-1]}{1-\alpha} - \alpha \log\left(d(O,t)+1\right) + \sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(t)\right)_{t\in[0,1]}, \\ \left((\Delta-1)\log(d(O,t)+1) + \sqrt{4\Delta}\phi^{(1)}(t)\right)_{t\in[0,1]}, \end{cases} \tag{4.1}$$

with respect to the topology of uniform convergence on $C([0,1])$.

Then, we apply Skorohod Representation a second time to work on a probability space where the convergence of (4.1) also holds almost surely. On this space, we show that the rescaled resistances and measures defined in (3.8) converge to the limit candidates suggested in (3.12) and (3.13).

4.1 The limiting potential

In this subsection we establish (4.1) above.

4.1.1 Proof of Claim 4.2

Throughout this section, recall that $(\mathbf{\Omega}, \mathcal{F}, \mathbf{P})$ is a probability space where the convergence of (2.6) holds almost surely. This also implies that the convergence of (2.3) holds almost surely (see [53, Theorem 4.2]). For $n \ge 1$ and $x \in T_n$ we define the shorthand

$$\Delta_n = \begin{cases} \Delta(na_n^{-1})^{-(1-\alpha)} & \text{if } \alpha < 1, \\ \Delta & \text{if } \alpha = 1, \end{cases} \qquad |x| = d_n(O_n, x).$$

Note though that the reinforcement parameter for the process defined in Section 3.2 is still Δ, not Δ_n. If $x \in T_n$, then the Dirichlet weights for the model have parameters given by (3.7), so that by (1.1) the weight at x is given by

$$\mathbf{b}_x^{(n)} = \left(\frac{(|x|+na_n^{-1})^\alpha + \Delta_n}{2\Delta_n}, \frac{(|x|+1+na_n^{-1})^\alpha}{2\Delta_n}, \frac{(|x|+1+na_n^{-1})^\alpha}{2\Delta_n}, \dots, \frac{(|x|+1+na_n^{-1})^\alpha}{2\Delta_n} \right).$$

For each $x \in T_n \setminus \{O_n\}$, as the ratio of two independent Gamma random variables, we observe that

$$\rho_x := \frac{W_{\overleftarrow{x}}^{\overleftarrow{\overleftarrow{x}}}}{W_{\overleftarrow{x}}^{x}} \sim \frac{\Gamma(b_{\overleftarrow{x}}^{(\overleftarrow{x},\overleftarrow{\overleftarrow{x}})},1)}{\Gamma(b_{\overleftarrow{x}}^{(\overleftarrow{x},x)},1)} \sim \beta' \left(\frac{\left((|x|-1)+na_n^{-1}\right)^\alpha + \Delta_n}{2\Delta_n}, \frac{(|x|+na_n^{-1})^\alpha}{2\Delta_n} \right),$$

and $(\rho_x)_{x \in T_n \setminus \{O_n\}}$ is a sequence of independent random variables. Here $\beta'(a,b)$ denotes the beta prime distribution with positive parameters a and b, and has probability density function

$$x^{a-1}(1+x)^{-a-b}/B(a,b), \qquad x > 0,$$

where B is the beta function.

We start by giving some preliminary claims about the expectation and variance of the quantities $\log \rho_x$ summed along a branch. The proofs are elementary but are included in the appendix.

Claim 4.3. *(i) For all $x \in T_n$ we have that $|\mathbb{E}[\log \rho_x]| \le \frac{2\alpha}{(|x|+na_n^{-1})} + \frac{4\Delta_n}{(|x|+na_n^{-1})^\alpha}$.*

(ii) For almost every $\omega \in \mathbf{\Omega}$, it holds uniformly over $t \in [0,1]$ as $n \to \infty$ that:

$$\sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \mathbb{E}[\log \rho_x] \to \frac{\Delta[(d(O,t)+1)^{1-\alpha}-1]}{1-\alpha} - \alpha \log\left(d(O,t)+1\right) \qquad \text{if } \alpha < 1,$$

$$\sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \mathbb{E}[\log \rho_x] \to (\Delta - 1)\log(d(O,t)+1), \qquad \text{if } \alpha = 1.$$

Proof. See Lemma A.1 in the Appendix. □

Claim 4.4. *For almost every $\omega \in \mathbf{\Omega}$, it holds uniformly over $t \in [0,1]$ as $n \to \infty$ that:*

$$\sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \mathrm{Var}(\log \rho_x) \to \frac{4\Delta[(d(O,t)+1)^{1-\alpha}-1]}{1-\alpha} \qquad \text{if } \alpha < 1,$$

$$\sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \mathrm{Var}(\log \rho_x) \to 4\Delta \log(d(O,t)+1) \qquad \text{if } \alpha = 1.$$

Proof. See Lemma A.2 in the Appendix. □

Claim 4.5. *For all $\alpha \le 1$ and almost every $\omega \in \mathbf{\Omega}$, it holds for all $x \in T_n$ and all $1 \le k \le (na_n^{-1})^{1/2}$ that*

$$\mathbb{E}\left[e^{k \log \rho_x}\right] \le \exp\left\{ (1+3\Delta)\left(\frac{2k\alpha}{|x|+na_n^{-1}} + \frac{3k^2\Delta_n}{(|x|+na_n^{-1})^\alpha} \right) \right\}.$$

$$\mathbb{E}\left[e^{k \log(\rho_x^{-1})}\right] \le \exp\left\{ (1+3\Delta)\left(\frac{2k\alpha}{|x|+na_n^{-1}} + \frac{3k^2\Delta_n}{(|x|+na_n^{-1})^\alpha} \right) \right\}.$$

Proof. See Lemma A.3 in the Appendix. □

We will prove the convergence of Claim 4.2 in two steps. Firstly we consider the recentred processes defined by

$$M_i^{(n,\alpha)} = \sum_{O_n \prec x \preceq x_i} (\log \rho_x - \mathbb{E}[\log \rho_x]) \tag{4.2}$$

for $0 \leq i \leq 2n$, and extended to non-integer time indices by interpolation. We then prove that $\left(M_{2nt}^{(n,\alpha)}\right)_{t\in[0,1]}$ converges to an appropriate snake process as follows. Firstly, by applying a martingale CLT along finitely many branches of T_n, we deduce that the finite dimensional marginals of $M^{(n,\alpha)}$ converge to those of the snake process. For this part of the proof, the martingales in question are indexed by the branch lengths, and not by the time interval $[0,1]$. Then, to extend this finite-dimensional convergence to full convergence in $C([0,1])$, we verify Kolmogorov's tightness condition, this time considering the whole process indexed by the time interval $[0,1]$.

We also set

$$A(t) = \begin{cases} \dfrac{4\Delta[(d(O,t)+1)^{1-\alpha} - 1]}{1-\alpha} & \text{if } \alpha < 1, \\ 4\Delta \log(d(O,t)+1) & \text{if } \alpha = 1. \end{cases}$$

Proposition 4.6 (Finite dimensional convergence)**.** *For each $n \geq 1$, let $(M_m^{(n,\alpha)})_{m\geq 0}$ be as in* (4.2)*, and for each $1 \leq m \leq 2n$ let $Z_m^{(n,\alpha)} = M_m^{(n,\alpha)} - M_{m-1}^{(n,\alpha)}$.*

1. *Almost surely on $\mathbf{\Omega}$, it holds for each $t \in (0,1)$ that*

$$\sum_{m:x_m \preceq x_{\lfloor 2nt \rfloor}} \mathbb{E}\Big[(Z_m^{(n,\alpha)})^2\Big] \to A(t),$$

   as $n \to \infty$.

2. *Almost surely on $\mathbf{\Omega}$, it holds $\forall \varepsilon > 0, \forall t \in (0,1)$,*

$$\sum_{m:x_m \preceq x_{\lfloor 2nt \rfloor}} \mathbb{E}\Big[(Z_m^{(n,\alpha)})^2 \mathbb{1}\{|Z_m^{(n,\alpha)}| > \varepsilon\}\Big] \to 0,$$

   as $n \to \infty$.

3. *Almost surely on $\mathbf{\Omega}$, for any $k \geq 1$ and any sequence $0 < t_1 < \ldots < t_k < 1$, it holds that*

$$(M_{2nt_i}^{(n,\alpha)})_{1\leq i\leq k} \to \begin{cases} \left(\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(t_i)\right)_{1\leq i\leq k} & \text{if } \alpha < 1, \\ \left(\sqrt{4\Delta}\phi^{(1)}(t_i)\right)_{1\leq i\leq k} & \text{if } \alpha = 1, \end{cases} \tag{4.3}$$

   jointly in distribution as $n \to \infty$.

Proof. 1. This is Theorem 4.4.

2. First note that, for any $t \in (0,1)$,

$$\begin{aligned} &\sum_{m:x_m \preceq x_{\lfloor 2nt \rfloor}} \mathbb{E}\Big[(Z_m^{(n,\alpha)})^2 \mathbb{1}\{|Z_m^{(n,\alpha)}| > \varepsilon\}\Big] \\ &= \sum_{m:x_m \preceq x_{\lfloor 2nt \rfloor}} \int_{\varepsilon^2}^{\infty} \mathbb{P}\Big((Z_x^{(n,\alpha)})^2 > y\Big)\, dy \\ &= \sum_{m:x_m \preceq x_{\lfloor 2nt \rfloor}} \int_{\varepsilon^2}^{\infty} \mathbb{P}(|\log \rho_{x_m} - \mathbb{E}[\log \rho_{x_m}]\,| > \sqrt{y})\, dy. \end{aligned} \tag{4.4}$$

Also note that it follows from Claim 4.3(i) and Claim 4.5 that almost surely on $\mathbf{\Omega}$, we have for all $1 \leq k \leq (na_n^{-1})^{1/2}$ and all $x \in T_n$,

$$\mathbb{E}\left[e^{k(\log \rho_x - \mathbb{E}[\log \rho_x])}\right] \vee \mathbb{E}\left[e^{k(\log \rho_x^{-1} - \mathbb{E}[\log \rho_x^{-1}])}\right] \leq e^{3(1+3\Delta)\left(\frac{2k\alpha}{|x|+na_n^{-1}} + \frac{3k^2\Delta_n}{(|x|+na_n^{-1})^\alpha}\right)}. \tag{4.5}$$

Taking $k = (na_n^{-1})^{1/2}$, the latter expression in (4.4) is upper bounded by

$$\begin{aligned} &n \sup_{x:|x|\geq 1} \int_{\varepsilon^2}^{\infty} 1 \wedge \left\{\mathbb{E}\left[e^{(na_n^{-1})^{\frac{1}{2}}(\log \rho_x - \mathbb{E}[\log \rho_x])}\right] e^{-(na_n^{-1})^{\frac{1}{2}}\sqrt{y}}\right\} dy \\ &\quad + n \sup_{x:|x|\geq 1} \int_{\varepsilon^2}^{\infty} 1 \wedge \left\{\mathbb{E}\left[e^{(na_n^{-1})^{\frac{1}{2}}\left(\log \rho_x^{-1} - \mathbb{E}[\log \rho_x^{-1}]\right)}\right] e^{-(na_n^{-1})^{\frac{1}{2}}\sqrt{y}}\right\} dy \\ &\leq 2n \sup_{x:|x|\geq 1} \int_{\varepsilon^2}^{\infty} 1 \wedge \left\{\exp\left\{3(1+3\Delta)\left(\frac{2\alpha}{(na_n^{-1})^{1/2}} + 3\Delta\right) - (na_n^{-1})^{\frac{1}{2}}\sqrt{y}\right\}\right\} dy. \end{aligned}$$

For any $\varepsilon > 0$, there exists $N_\varepsilon < \infty$ such that the minimum in the final integrand is equal to the second expression for all $y \in (\varepsilon^2, \infty)$. By dominated convergence, this integral therefore tends to 0 as $n \to \infty$, as required. (In fact, for fixed $\varepsilon > 0$, this upper bound is also uniform over $t \in (0, 1)$).

3. Without loss of generality we can assume that $t_1, \ldots, t_k \in \frac{1}{2n}\mathbb{Z}$; the general case follows since we interpolate and the contribution of a single jump goes to zero in probability as $n \to \infty$ (for example by Markov's inequality). Note that, for fixed $t \in (0, 1)$, parts 1 and 2 exactly verify the conditions given in [46, Corollary 3.1 and Theorem 3.2] to imply that
$$M_{2nt}^{(n,\alpha)} \to N(0, A(t)), \tag{4.6}$$
as $n \to \infty$. Therefore, given such a sequence $0 < t_1 < \ldots < t_k < 1$, for each $k \geq 1$ we can define an *augmented sequence* obtained by adding the time indices corresponding to all the most recent common ancestors of pairs of vertices in the original sequence. Given this augmented sequence, we can then sum the contributions along the relevant branch segments between vertices of the form x_{t_i} and x_{t_j}, where t_i and t_j are such that there is no $\ell \leq k$ with $t_i \preceq t_\ell \preceq t_j$. Since the process evolves independently along distinct branch segments and the sum of independent Gaussians is Gaussian, the finite-dimensional result then follows from (4.6). □

In order to strengthen the above convergence, we will verify Kolmogorov's tightness condition. We first give a preliminary lemma. For this, for $\alpha < 1$ we first define $d_n^{(\alpha)}(s,t)$ to be equal to

$$\begin{aligned} &(na_n^{-1})^{-(1-\alpha)}(d_n(O_n, x_{\lfloor 2nt \rfloor}) + na_n^{-1})^{1-\alpha} \\ &+(na_n^{-1})^{-(1-\alpha)}(d_n(O_n, x_{\lfloor 2ns \rfloor}) + na_n^{-1})^{1-\alpha} - 2(na_n^{-1})^{-(1-\alpha)}(d_n(O_n, x_{\lfloor 2ns \rfloor \wedge \lfloor 2nt \rfloor}) + na_n^{-1})^{1-\alpha}. \end{aligned} \tag{4.7}$$

Note that $d_n^{(\alpha)}(s,t) = d^{(\alpha)}(s,t) + o(1)$, where the $o(1)$ is uniform over all $s, t \in [0,1]$ by (2.8). Also let $D_n = (na_n^{-1})^{-1}\mathsf{Diam}(T_n)$. We first give a useful lemma.

Lemma 4.7. *Almost surely on $\mathbf{\Omega}$, we have for all $n \geq 1$ and all $s, t \in [0,1]$ that*

$$d_n^{(\alpha)}(s,t) \leq \begin{cases} (1-\alpha)(D_n+1)^{-\alpha} d_n^{(0)}(s,t) & \text{if } \alpha \leq 0, \\ 2 d_n^{(0)}(s,t)^{1-\alpha} & \text{if } \alpha \in (0,1). \end{cases}$$

Proof. By breaking at the most recent common ancestor, it is enough to prove this when $s \preceq t$. If $\alpha \leq 0$, we just write

$$\begin{aligned} d_n^{(\alpha)}(s,t) &\leq (1-\alpha)(na_n^{-1})^{-1} d_n(x_{\lfloor 2ns \rfloor}, x_{\lfloor 2nt \rfloor})[(na_n^{-1})^{-1} d_n(O_n, x_{\lfloor 2nt \rfloor}) + 1]^{-\alpha} \\ &\leq (1-\alpha)(D_n+1)^{-\alpha} d_n^{(0)}(s,t). \end{aligned}$$

If instead $\alpha \in (0,1)$, we treat two cases:

1. If $d_n(O_n, x_{\lfloor 2ns \rfloor}) \geq \frac{1}{2} d_n(O_n, x_{\lfloor 2nt \rfloor})$, then we obtain that

$$\begin{aligned} d_n^{(\alpha)}(s,t) &\leq (1-\alpha)(na_n^{-1})^{-1} d_n(x_{\lfloor 2ns \rfloor}, x_{\lfloor 2nt \rfloor})[(na_n^{-1})^{-1} d_n(O_n, x_{\lfloor 2ns \rfloor})]^{-\alpha} \\ &\leq (1-\alpha) d_n^{(0)}(s,t)^{1-\alpha}. \end{aligned}$$

2. If $d_n(O_n, x_{\lfloor 2ns \rfloor}) < \frac{1}{2} d_n(O_n, x_{\lfloor 2nt \rfloor})$, then we get

$$d_n^{(\alpha)}(s,t) \leq (na_n^{-1})^{-(1-\alpha)} d_n(O_n, x_{\lfloor 2nt \rfloor})^{1-\alpha} \leq 2^{1-\alpha} d_n^{(0)}(s,t)^{1-\alpha}. \qquad \square$$

We are now ready to verify the tightness condition.

Proposition 4.8 (Kolmogorov's condition)**.** *For every $\varepsilon > 0, n \geq 1$ there exist $p > 0, q > 1$, $C_{\varepsilon,p,q} < \infty$ and an event $A_{n,\varepsilon} \subset \mathbf{\Omega}$ with $\mathbf{P}(A_{n,\varepsilon}) \geq 1-\varepsilon$, such that on the event $A_{n,\varepsilon}$, we have for all $s,t \in [0,1]$ that*

$$\mathbb{E}\left[\frac{|M_{2nt}^{(n,\alpha)} - M_{2ns}^{(n,\alpha)}|^p}{|s-t|^q} \,\middle|\, A_{n,\varepsilon} \right] < C_{\varepsilon,p,q}.$$

In particular, Kolmogorov's tightness condition is satisfied on the event $A_{n,\varepsilon}$ and for almost every $\omega \in \mathbf{\Omega}$, the convergence of (4.3) *holds in distribution on the space $C([0,1])$ equipped with the uniform topology.*

Proof. First note that it follows from [57, Lemma 1.4] and [50, Theorem 2] that for any $\gamma' < \frac{\gamma-1}{\gamma}$ and $\varepsilon > 0$, we can choose $D_\varepsilon, C_{\varepsilon,\gamma'} < \infty$ such that with probability $1-\varepsilon$, we have that

$$\begin{aligned} &d_n^{(0)}(s,t) = (na_n^{-1})^{-1} d_n(x_{2nt}, x_{2ns}) \leq C_{\varepsilon,\gamma'} |t-s|^{\gamma'} \text{ for all } s,t \in [0,1], \\ &D_n := \sup_{t\in[0,1]} (na_n^{-1})^{-1} d_n(O_n, x_{2nt}) \leq D_\varepsilon - 1. \end{aligned} \tag{4.8}$$

(Note that [57, Lemma 1.4] actually states a result using the lexicographical ordering rather than the contour ordering as we use, but since the difference in the labels of two fixed vertices can only decrease by at most a factor of 2 in the contour ordering, this immediately implies the same result with the contour ordering and therefore really implies the first line of (4.8)).

Case $\alpha < 1$**.** Assume for now that $2ns$ and $2nt$ are integers and that x_{2ns} is an ancestor of x_{2nt}. This will then extend to the general case by breaking paths at the most recent common ancestor and since $M^{(n,\alpha)}$ is defined by interpolation. Note that by (4.5), we almost surely have for all $1 \leq k \leq (na_n^{-1})^{1/2}$ and all $x \in T_n$ that

$$\mathbb{E}\Big[e^{k(\log \rho_x - \mathbb{E}[\log \rho_x])}\Big] \leq \exp\left\{ 3\,(1+3\Delta)\left(\frac{2k\alpha}{|x| + na_n^{-1}} + \frac{3k^2 \Delta_n}{(|x| + na_n^{-1})^\alpha} \right) \right\}.$$

Therefore, for all $1 \leq k \leq (na_n^{-1})^{1/2}$ we deduce that there exists $C_\Delta < \infty$ such that (using also that, given T_n, the sequence $(\rho_x)_{x\in T_n}$ is independent)

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{\sum_{x_{ns}\prec x\preceq x_{nt}}(\log\rho_x - \mathbb{E}[\log\rho_x])\right\}\right]\\
&\quad\leq \exp\left\{\sum_{x_{ns}\prec x\preceq x_{nt}} 3\,(1+3\Delta)\left(\frac{2k\alpha}{|x|+na_n^{-1}}+\frac{3k^2\Delta_n}{(|x|+na_n^{-1})^{\alpha}}\right)\right\}\\
&\quad\leq \exp\left\{\frac{C_\Delta k^2(na_n^{-1})^{-(1-\alpha)}}{1-\alpha}\left((d_n(O_n,x_{nt})+na_n^{-1})^{1-\alpha}-(d_n(O_n,x_{ns})+na_n^{-1})^{1-\alpha}\right)\right\}\\
&\qquad\times\exp\left\{\frac{C_\Delta k\alpha}{1-\alpha}\left(\log((d_n(O_n,x_{nt})+na_n^{-1})^{1-\alpha})-\log((d_n(O_n,x_{ns})+na_n^{-1})^{1-\alpha})\right)\right\}\\
&\quad\leq \exp\left\{\frac{C_\Delta k(k+\alpha)}{1-\alpha}d_n^{(\alpha)}(s,t)\right\},
\end{aligned} \tag{4.9}$$

where $d_n^{(\alpha)}(s,t)$ is given by (4.7). Applying Lemma 4.7, we deduce that this is upper bounded by

$$\begin{cases}\exp\left\{C_\Delta k(k+\alpha)D_n^{-\alpha}d_n^{(0)}(s,t)\right\} & \text{if } \alpha\leq 0,\\ \exp\left\{\frac{2C_\Delta k(k+\alpha)}{1-\alpha}d_n^{(0)}(s,t)^{1-\alpha}\right\} & \text{if } \alpha\in(0,1).\end{cases} \tag{4.10}$$

Assume that $d_n^{(0)}(s,t)\leq 1$. Set

$$k = d_n^{(0)}(s,t)^{-\frac{(1-\alpha)\wedge 1}{3}} \leq (na_n^{-1})^{\frac{1}{3}}.$$

In the case $\alpha\leq 0$, we deduce from a Chernoff bound that there exists $C_{\varepsilon,\Delta,\alpha}<\infty$ such that on the event $A_{n,\varepsilon}$, we have for all $y>0$ that

$$\begin{aligned}
\mathbb{P}\left(\sum_{x_{ns}\prec x\preceq x_{nt}}(\log\rho_x-\mathbb{E}[\log\rho_x])>y\right) &\leq e^{C_\Delta k^3 D_n^{-\alpha}d_n^{(0)}(s,t)-yk}\leq e^{C_\Delta D_n^{-\alpha}-yk}\\
&\leq C_{\varepsilon,\Delta,\alpha}e^{-yd_n^{(0)}(s,t)^{-\frac{(1-\alpha)\wedge 1}{3}}},
\end{aligned} \tag{4.11}$$

where (4.10) was used to provide the first bound in the inequality above. Moreover, by Claim 4.5 the same result holds on replacing ρ_x by ρ_x^{-1}, meaning that we can instead consider the absolute value of the sum in the bound above.

We obtain the same upper bound in the case $\alpha>0$ (modifying the constant $C_{\varepsilon,\Delta,\alpha}$ a bit if necessary). We deduce that, on the event $A_{n,\varepsilon}$, for all $p>1$ there exists $C_{p,\varepsilon,\Delta,\alpha}<\infty$ such that for all $s,t\in[0,1]$ with $d_n^{(0)}(s,t)\leq 1$,

$$\begin{aligned}
\mathbb{E}\left[\left|\sum_{x_{ns}\prec x\preceq x_{nt}}(\log\rho_x-\mathbb{E}[\log\rho_x])\right|^p\right] &\leq \int_0^\infty \mathbb{P}\left(\sum_{x_{ns}\prec x\preceq x_{nt}}|\log\rho_x-\mathbb{E}[\log\rho_x]|>y^{1/p}\right)dy\\
&\leq \int_0^\infty 2C_{\varepsilon,\Delta,\alpha}e^{-y^{1/p}d_n^{(0)}(s,t)^{-\frac{(1-\alpha)\wedge 1}{3}}}dy\\
&\leq \int_0^\infty 2C_{\varepsilon,\Delta,\alpha}d_n^{(0)}(s,t)^{\frac{(1-\alpha)\wedge 1}{3}}pu^{p-1}e^{-u}du\\
&\leq 2C_{p,\varepsilon,\Delta,\alpha}d_n^{(0)}(s,t)^{\frac{(1-\alpha)\wedge 1}{3}}.
\end{aligned} \tag{4.12}$$

By replacing $C_{p,\varepsilon,\Delta,\alpha}$ with $D_\varepsilon C_{p,\varepsilon,\Delta,\alpha}$ this extends to the case $d_n^{(\alpha)}(s,t) > 1$ on the event $A_{n,\varepsilon}$. Finally, this extends to the general case where $2ns$ and $2nt$ are not integers and x_{2ns} is no longer an ancestor of x_{2nt} by breaking paths at the most recent common ancestor and since $M^{(n,\alpha)}$ is defined by interpolation.

We now suppress the dependence on Δ and α since these are assumed to be fixed. We deduce that, on the event $A_{n,\varepsilon}$, we have for all $\gamma' < \frac{\gamma-1}{\gamma}$ that there exists $C_{p,\varepsilon,\gamma'}$ (also applying (4.8)) such that for all $s,t \in [0,1]$,

$$\mathbb{E}\Big[|M_{2nt}^{(n,\alpha)} - M_{2ns}^{(n,\alpha)}|^p\Big] \leq 2C_{p,\varepsilon} d_n^{(0)}(s,t)^{\frac{(1-\alpha)\wedge 1}{3}} \leq C_{p,\varepsilon,\gamma'}|s-t|^{\frac{[(1-\alpha)\wedge 1]\gamma'}{3}}.$$

This proves the result by taking p large enough.

Case $\alpha = 1$**.** Again we can assume, without loss of generality, that $2ns$ and $2nt$ are integers and that x_{2ns} is an ancestor of x_{2nt}. This time, repeating the arguments that led to (4.9) gives that there exists $C_\Delta < \infty$ such that

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{\sum_{x_{ns}\prec x\preceq x_{nt}} (\log\rho_x - \mathbb{E}[\log\rho_x])\right\}\right] \\
&\quad\leq \exp\left\{\sum_{x_{ns}\prec x\preceq x_{nt}} 3\,(1+3\Delta)\,\frac{k(2+3k\Delta)}{|x|+na_n^{-1}}\right\} \\
&\quad\leq \exp\left\{C_\Delta k(k+1)\left(\log((d_n(O_n,x_{nt})+na_n^{-1})) - \log((d_n(O_n,x_{ns})+na_n^{-1}))\right)\right\} \\
&\quad\leq \exp\left\{C_\Delta k(k+1) d_n^{(0)}(s,t)\right\}.
\end{aligned}$$

Again we assume that $d_n^{(0)}(s,t) \leq 1$ and set

$$k = d_n^{(0)}(s,t)^{-\frac{1}{3}} \leq (na_n^{-1})^{\frac{1}{3}}.$$

As in (4.11), this implies that there exists $C_{\varepsilon,\Delta} < \infty$ such that for all $y > 0$,

$$\mathbb{P}\left(\sum_{x_{ns}\prec x\preceq x_{nt}} (\log\rho_x - \mathbb{E}[\log\rho_x]) > y\right) \leq C_{\varepsilon,\Delta} e^{-y d_n^{(0)}(s,t)^{-\frac{1}{3}}}.$$

Again by Claim 4.5 we can replace ρ_x by ρ_x^{-1} in this argument, so we can continue as in (4.12) to deduce that on the event $A_{n,\varepsilon}$, we have that for any $p > 1$ there exists $C_{p,\varepsilon,\Delta} < \infty$ such that for all $s,t \in [0,1]$ with $d_n^{(0)}(s,t) \leq 1$,

$$\mathbb{E}\left[\left|\sum_{x_{ns}\prec x\preceq x_{nt}} (\log\rho_x - \mathbb{E}[\log\rho_x])\right|^p\right] \leq 2C_{p,\varepsilon,\Delta} d_n^{(0)}(s,t)^{\frac{1}{3}}.$$

The proof then proceeds as in the case $\alpha < 1$. ☐

Proof of Claim 4.2. We write the proof in the case $\alpha < 1$. First set

$$W^{(n,\alpha)}(2nt) = V^{(n,\alpha)}(2nt) - \frac{\Delta[(d(O,t)+1)^{1-\alpha} - 1]}{1-\alpha} + \alpha\log(d(O,t)+1).$$

Then for every $t \in [0,1]$,

$$W^{(n,\alpha)}(2nt) - \mathbb{E}\Big[W^{(n,\alpha)}(2nt)\Big] = V^{(n,\alpha)}(2nt) - \mathbb{E}\Big[V^{(n,\alpha)}(2nt)\Big] = M_{2nt}^{(n,\alpha)},$$

so it follows from (4.3) and Proposition 4.8 that

$$\left(W^{(n,\alpha)}(2nt) - \mathbb{E}\left[W^{(n,\alpha)}(2nt)\right]\right)_{t\in[0,1]} \to \left(\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(t)\right)_{t\in[0,1]}, \tag{4.13}$$

uniformly on $C([0,1])$. Also, it follows from Claim 4.3 that

$$\mathbb{E}\left[W^{(n,\alpha)}(2nt)\right] \to 0, \tag{4.14}$$

uniformly on $C([0,1])$. Claim 4.2 therefore follows on combining (4.13) and 4.14.

The proof in the case $\alpha = 1$ follows similarly. □

4.2 Convergence of the metric measure spaces

In Section 4.1, we showed that, almost surely on $\mathbf{\Omega}$, the relevant processes converge in distribution with respect to the Skorohod-J_1 topology and therefore with respect to the uniform topology since the limit process is continuous. Therefore, applying Skorohod's Representation theorem again (the space of continuous functions on $[0,1]$ is separable), we will work on a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ where these functionals additionally converge almost surely with respect to the uniform topology.

Proof of Proposition 4.1. ***Part 1: convergence of metrics.*** In the first part of the proof we will show that the distortion of the natural correspondence, defined by

$$\mathsf{dis}(\mathcal{R}_n) = \sup\{|n^{-1}a_n R_n(x_{\lfloor 2ns\rfloor}, x_{\lfloor 2nt\rfloor}) - R_\phi(s,t)| : (x_{\lfloor 2ns\rfloor}, s), (x_{\lfloor 2nt\rfloor}, t) \in \mathcal{R}_n\}, \tag{4.15}$$

converges to 0 as $n \to \infty$, almost surely on Ω'. It is enough to prove this just in the case $t = 0$ since

$$\begin{aligned} |n^{-1}a_n R_n(x_{\lfloor 2ns\rfloor}, x_{\lfloor 2nt\rfloor}) - R_\phi(s,t)| \leq{}& 2|n^{-1}a_n R_n(x_0, x_{\lfloor 2nr\rfloor}) - R_\phi(0,r)| \\ &+ |n^{-1}a_n R_n(x_0, x_{\lfloor 2ns\rfloor}) - R_\phi(0,s)| \\ &+ |n^{-1}a_n R_n(x_0, x_{\lfloor 2nt\rfloor}) - R_\phi(0,t)| \\ &+ 2n^{-1}a_n R_n(x_{r_n}, x_{\lfloor 2nr\rfloor}), \end{aligned}$$

where $(x_{\lfloor 2nr\rfloor}, p_{H^{(\gamma)}}(r)) \in \mathcal{R}_n$, where $r \in [s,t]$ is any time between s and t at which the minimum of $H^{(\gamma)}$ is achieved, and where r_n is the index of the most recent common ancestor of $x_{\lfloor 2ns\rfloor}$ and $x_{\lfloor 2nt\rfloor}$. We will therefore first establish (4.15) when $t = 0$ and then show that $n^{-1}a_n R_n(x_{r_n}, x_{\lfloor 2nr\rfloor}) \to 0$, uniformly over $s, t \in [0,1]$.

For the former, first note that by the definition in (3.8):

$$n^{-1}a_n R_n(x_0, x_{\lfloor 2ns\rfloor}) = n^{-1}a_n \sum_{x_0 \prec x \preceq x_{\lfloor 2ns\rfloor}} e^{V^{(n,\alpha)}(x)} = 2a_n \int_{A_s^{(n)}} e^{V^{(n,\alpha)}(2nr)} d\ell(r), \tag{4.16}$$

where $A_s^{(n)} = \{r < s : \inf_{u\in(\lfloor r\rfloor, s]} C^{(n)}(u) > C^{(n)}(r) \text{ and } x_{\lfloor 2nr\rfloor} \neq x_0\} \cup \{s\}$ (to avoid double counting repeat contour visits) and ℓ is Lebesgue measure on $\mathbb{R}$.

By the definition in (3.12):

$$R_\phi(0,s) = \begin{cases} \displaystyle\int_{(O,x_s]} (d(O,z)+1)^{-\alpha} e^{\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(z) + \frac{\Delta}{1-\alpha}[(d(O,z)+1)^{1-\alpha}-1]} \lambda^\gamma(\mathrm{d}z) & \text{if } \alpha < 1, \\ \displaystyle\int_{(O,x_s]} e^{\sqrt{4\Delta}\phi^{(1)}(z) + (\Delta-1)\log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) & \text{if } \alpha = 1. \end{cases} \tag{4.17}$$

For $r \in (0,1)$, let us set

$$h^{(n,\alpha)}(r) = e^{V^{(n,\alpha)}(2nr)}$$
$$h^{(\alpha)}(r) = \begin{cases} (d(O,r)+1)^{-\alpha} e^{\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(r) + \frac{\Delta}{1-\alpha}[(d(O,r)+1)^{1-\alpha}-1]}, & \text{if } \alpha < 1, \\ e^{\sqrt{4\Delta}\phi^{(1)}(r) + (\Delta-1)\log(d(O,r)+1)}, & \text{if } \alpha = 1. \end{cases}$$

Combining (4.16) and (4.17), we deduce that for any pair $(x_{\lfloor 2ns \rfloor}, p_{H^{(\gamma)}}(s)) \in \mathcal{R}_n$,

$$\begin{aligned} \left| n^{-1} a_n R_n(x_0, x_{\lfloor 2ns \rfloor}) - R_\phi(0,s) \right| &\leq \left| 2a_n \int_{A_s^{(n)}} h^{(n,\alpha)}(r) d\ell(r) - 2a_n \int_{A_s^{(n)}} h^{(\alpha)}(r) d\ell(r) \right| \\ &\quad + \left| 2a_n \int_{A_s^{(n)}} h^{(\alpha)}(r) d\ell(r) - \int_{(O,x_s]} h^{(\alpha)}(z) \lambda^\gamma(\mathrm{d}z) \right| \\ &\leq 2a_n ||h^{(n,\alpha)} - h^{(\alpha)}|| \ell(A_s^{(n)}) \\ &\quad + ||h^{(\alpha)}|| \sup_{s\in[0,1]} \left| 2a_n \int_{A_s^{(n)}} d\ell(r) - \int_{(O,x_s]} \lambda^\gamma(\mathrm{d}z) \right|. \end{aligned} \tag{4.18}$$

There are two steps to show that almost surely on Ω', the right-hand side converges to 0 uniformly over $s \in [0,1]$. First note that $\ell(A_s^{(n)}) = \frac{1}{2n} d_n(O_n, x_{\lfloor 2ns \rfloor}) = \frac{1}{2a_n} C^{(n)}(s)$. Therefore, almost surely on Ω',

$$2a_n ||h^{(n,\alpha)} - h^{(\alpha)}|| \ell(A_s^{(n)}) = ||h^{(n,\alpha)} - h^{(\alpha)}|| C^{(n)}(s) \xrightarrow{n\to\infty} 0, \tag{4.19}$$

uniformly over $s \in [0,1]$, by (2.6) and (4.1) (recall that we are working on a probability space where the convergences of (2.6) and (4.1) hold almost surely). Secondly, using (3.11) it is not hard to see that

$$\sup_{s\in[0,1]} \left| 2a_n \int_{A_s^{(n)}} d\ell(r) - \int_{(O,x_s]} \lambda^\gamma(\mathrm{d}z) \right| = ||C^{(n)} - H^{(\gamma)}||,$$

which also goes to 0 almost surely as $n \to \infty$, by (2.6) (which holds almost surely on Ω').

This therefore shows that each individual term on the right hand side of (4.18) converges to 0 uniformly in $s \in [0,1]$, completing the first part of the proof about the convergence to 0 of the distortion of the correspondence $\mathcal{R}_n$ as defined in (4.15) when $t = 0$.

The second part of the proof is to show that

$$\sup_{s,t\in[0,1]} n^{-1} a_n R_n(x_{r_n^{s,t}}, x_{\lfloor 2nr^{s,t} \rfloor}) \to 0, \tag{4.20}$$

where $r^{s,t} = s \wedge t$ and $r_n^{s,t}$ is the index of the most recent common ancestor of $x_{\lfloor 2ns \rfloor}$ and $x_{\lfloor 2nt \rfloor}$. However, note that (writing $r = r^{s,t}$ and $r_n = r_n^{s,t}$)

$$\begin{aligned} n^{-1} a_n R_n(x_{r_n}, x_{\lfloor 2nr \rfloor}) &= n^{-1} a_n \sum_{x \in [x_{r_n}, x_{\lfloor 2nr \rfloor}] \setminus \{x_{r_n} \wedge x_{\lfloor 2nr \rfloor}\}} e^{V^{(n,\alpha)}(x)} \\ &\leq (n^{-1} a_n d_n(x_{r_n}, x_{\lfloor 2nr \rfloor}) \sup_{x\in T_n} e^{V^{(n,\alpha)}(x)}. \end{aligned}$$

By (2.6) (which holds almost surely on Ω'), $n^{-1} a_n d_n(x_{r_n}, x_{\lfloor 2nr \rfloor}) \to 0$ uniformly over $s,t \in [0,1]$, almost surely on Ω'. Similarly, by (4.1), for almost every $\omega \in \Omega'$ we have that $\sup_{x\in T_n} e^{V^{(n,\alpha)}(x)}$ is bounded by a constant (which may depend on ω, but that can nevertheless be upper bounded on a set of probability $1-\varepsilon$ for any $\varepsilon > 0$). Therefore, we deduce that (4.20) holds almost surely on Ω', as required.

Part 2: convergence of measures.

Recall from (3.8) that for a non-root vertex $x \in T_n$ with $\#x$ offspring,

$$(2n)^{-1}\nu_n(x) = \frac{1}{2n}\sum_{i=0}^{\#x} e^{-V^{(n,\alpha)}(x_i)} = \frac{1}{2n}\sum_{i=0}^{\#x} e^{-V^{(n,\alpha)}(2nt_{x_i})},$$

where t_{x_i} is the minimal t such that $(x_i, p_{H^{(\gamma)}}(t)) \in \mathcal{R}_n$, where x_0 denotes the parent of x and $(x_i)_{i=1}^{\#x}$ denotes its children. Therefore, for any set A_n of vertices in T_n, we have that

$$(2n)^{-1}\nu_n(A_n) = \frac{1}{2n}\sum_{x\in A_n}\sum_{i=0}^{\#x} e^{-V^{(n,\alpha)}(2nt_{x_i})}.$$

We introduce an intermediate measure by setting:

$$\tilde{\nu}_n(x) = \sum_{\substack{0\leq i\leq 2n:\\ x_i = x}} e^{-V^{(n,\alpha)}(x_i)} = \begin{cases}(\deg x)e^{-V^{(n,\alpha)}(x)}, & \text{if } x \neq O_n,\\ (\deg x + 1)e^{-V^{(n,\alpha)}(x)}, & \text{if } x = O_n.\end{cases}$$

Therefore, for a subset $A_n \subset T_n$,

$$(2n)^{-1}\tilde{\nu}_n(A_n) = \frac{1}{2n}\sum_{x\in A_n}\sum_{\substack{0\leq i\leq 2n:\\ x_i = x}} e^{-V^{(n,1)}(2nu_x)},$$

where u_x is the minimal u such that $(x, p_{H^{(\gamma)}}(u)) \in \mathcal{R}_n$. We will now show that it suffices to consider $\tilde{\nu}_n$ in place of ν_n to obtain the Prokhorov limit. For $s \in (0,1)$, let us set

$$g^{(n,\alpha)}(s) = e^{-V^{(n,\alpha)}(2ns)}.$$

Note that, if u_x and t_{x_i} are defined as above, then

$$\Delta_{g^{(n,\alpha)}} := \sup_{x\in T_n, x_i\sim x}\left\{g^{(n,\alpha)}(u_x) - g^{(n,\alpha)}(t_{x_i})\right\} \to 0, \tag{4.21}$$

by the convergence of (4.1) and since $[0,1]$ is compact. Therefore, if $A_n \subset T_n$, then

$$\begin{aligned}\left|(2n)^{-1}\nu_n(A_n) - (2n)^{-1}\tilde{\nu}(A_n)\right| &= \left|\frac{1}{2n}\sum_{x\in A_n}\sum_{i=0}^{\#x} g^{(n,\alpha)}(t_{x_i}) - \frac{1}{2n}\sum_{x\in A_n}\sum_{i=0}^{\#x} g^{(n,\alpha)}(u_x)\right|\\ &\leq \frac{1}{2n}\sum_{x\in A_n}\sum_{i=0}^{\#x}\Delta_{g^{(n,\alpha)}}\\ &= \frac{\Delta_{g^{(n,\alpha)}}}{2n}\left(\sum_{x\in A_n}\mathsf{deg}(x) + \mathbb{1}\{O_n \in A_n\}\right)\\ &\leq \frac{\Delta_{g^{(n,\alpha)}}}{2n}\cdot 2n \to 0\end{aligned}$$

as $n\to\infty$. Therefore, letting

$$r_n = \mathsf{dis}(\mathcal{R}_n)$$

as in (4.15), it is sufficient to show that under the canonical Gromov-Hausdorff embedding $F_n = T_n \sqcup \mathcal{T}_\gamma^c$,

$$d_P^{F_n}((2n)^{-1}\tilde{\nu}_n, \nu_\phi) \leq r_n,$$

and hence converges to 0 as $n \to \infty$ (by Part 1 of this proof). We proceed as follows. First, take a subset $B \subset \mathcal{T}_\gamma^c$, and for $s \in (0,1)$ set

$$g^{(\alpha)}(s) = \begin{cases} (d(O,s)+1)^\alpha e^{-\left[\sqrt{\frac{4\Delta}{1-\alpha}}\phi^{(\alpha)}(s)+\frac{\Delta}{1-\alpha}[(d(O,s)+1)^{1-\alpha}-1]\right]}, & \text{if } \alpha < 1, \\ e^{-[\sqrt{4\Delta}\phi^{(1)}(s)+(\Delta-1)\log(d(O,s)+1)]}, & \text{if } \alpha = 1. \end{cases}$$

Note that $||g^{(n,\alpha)} - g^{(\alpha)}|| \to 0$ as $n \to \infty$ by the convergence of (4.1) and (2.6). Let $B' = p_{H^{(\gamma)}}^{-1}(B)$, $I_{n,i} = [\frac{i}{2n}, \frac{i+1}{2n})$ for $0 \le i < 2n$ and

$$B_n = \cup_{0 \le i < 2n}\{x \in T_n : \exists s \in B' \text{ such that } s \in I_{n,i} \text{ and } x = x_i\},$$

where x_i corresponds to the i-th node of T_n in contour exploration order. Clearly $B' \subset \cup_{i:x_i \in B_n} I_{n,i}$, and so

$$\begin{aligned}
\nu_\phi(B) &= \int_{B'} g^{(\alpha)}(s)ds \\
&\le \sum_{i:x_i\in B_n}\int_{I_{n,i}} g^{(\alpha)}(s)ds \\
&\le \sum_{i:x_i\in B_n}\int_{I_{n,i}} |g^{(n,\alpha)}(s)-g^{(\alpha)}(s)|ds + \sum_{i:x_i\in B_n}\int_{I_{n,i}} g^{(n,\alpha)}(s)ds \\
&\le \sup_{s\in(0,1)}|g^{(n,\alpha)}(s)-g^{(\alpha)}(s)| + \frac{1}{2n}\sum_{x\in B_n}\sum_{0\le i\le 2n} g^{(n,\alpha)}\left(\frac{i}{2n}\right)\mathbb{1}\{x_i = x\} \\
&= o(1) + (2n)^{-1}\tilde{\nu}(B_n).
\end{aligned}$$

Now note that if $x \in B_n$, then there exists $i \le 2n$ and $s \in B'$ with $s \in I_{n,i}$ so that $(x_{\lfloor 2ns \rfloor}, p_{H^{(\gamma)}(s)}) \in \mathcal{R}_n$. This further entails that $B_n \subset B^{r_n}$, and so

$$\nu_\phi(B) \le o(1) + (2n)^{-1}\tilde{\nu}(B^{r_n}). \tag{4.22}$$

We now prove the reverse statement. Let A_n be a set of vertices in T_n, and let

$$A_n' = \bigcup_{i:x_i\in A_n} I_{n,i}.$$

Let $A_n'' = p_{H^{(\gamma)}}(A_n')$. For any $x \in A_n''$, there exists $t \in A_n'$ with $x = p_{H^{(\gamma)}}(t)$ and $t \in I_{n,i}$ for some i with $x_i \in A_n$. Moreover, this entails that $i = \lfloor 2nt \rfloor$, and hence $(x_{\lfloor 2nt \rfloor}, p_{H^{(\gamma)}}(t))$ is in $\mathcal{R}_n$. It follows that $x \in A_n^{r_n}$, i.e. $A_n'' \subset A_n^{r_n}$; hence it follows from (4.21) that

$$\begin{aligned}
(2n)^{-1}\tilde{\nu}(A_n) = \frac{1}{2n}\sum_{x\in A_n}\sum_{\substack{0\le i\le 2n:\\ x_i = x}} g^{(n,\alpha)}(u_x) &= \sum_{x\in A_n}\sum_{\substack{0\le i\le 2n:\\ x_i = x}}\int_{I_{n,i}} g^{(n,\alpha)}\left(\frac{1}{2n}\lfloor 2ns \rfloor\right)ds \\
&= \sup_{s\in(0,1)}|g^{(n,\alpha)}(s)-g^{(\alpha)}(s)| + \int_{A_n'} g^{(\alpha)}(s)ds \\
&= o(1) + \nu_\phi(A_n'') \le o(1) + \nu_\phi(A_n^{r_n}).
\end{aligned}$$

Together with (4.22), this entails that

$$d_P^{F_n}((2n)^{-1}\tilde{\nu}_n, \nu_\phi) \le r_n.$$

The desired result follows. □

In the following corollary, applying [31, Theorem 7.2], yields the quenched convergence of the LERRW to a mixture of diffusions, with the limit law of this process given by the annealed law P_O as defined in (3.14).

Corollary 4.9. *Let $X^{(n)}$ have the law of a discrete-time LERRW on T_n, with initial weights* (1.1) *and reinforcement parameter Δ. Let X be as defined by* (3.14)*. Then,*

$$P_{O_n}^{(\alpha^{(n)})}\left(\left(n^{-1}a_nX^{(n)}_{\lfloor 2n^2a_n^{-1}t\rfloor}\right)_{t\geq 0}\in\cdot\right)\xrightarrow{(d)}P_O\left((X_t)_{t\geq 0}\in\cdot\right).$$

Proof. Recall that we are working on a probability space where the convergence of (2.6) holds almost surely. It follows that almost surely on this probability space, there exists a canonical embedding into the common metric space $n^{-1}a_nT_n\sqcup\mathcal{T}_\gamma^c$ in which the GHP distance between the random elements in Proposition 4.1 goes to zero as $n\to\infty$. We work pointwise on this probability space so that we only need to consider the randomness of the Dirichlet weights.

We first consider a continuous-time LERRW with $\mathsf{exp}(1)$ holding times at each vertex. By our choice of metric and measure, it follows that the stochastic process associated with the discrete space (T_n,R_n,ν_n) has the quenched law of a RWDE also with $\mathsf{exp}(1)$ holding times at each vertex. Therefore, by Lemma 3.6 the continuous-time LERRW corresponds to the annealed law of this RWDE exactly as considered in (3.9). It therefore follows directly from [31, Theorem 7.2] (the non-explosion condition (39) appearing there is satisfied since all the spaces are compact) that this continuous-time LERRW converges in distribution to X as defined by (3.14). The extension to the discrete-time LERRW in place of the continuous-time one then follows by a straightforward application of the law of large numbers on the time index. □

Clearly Theorem 1.1 follows directly from Corollary 4.9.

5 Properties of $\mathcal{T}_\gamma^\infty$ and its Gaussian potential

In order to understand the long-time behaviour of a LERRW it is more natural to consider it on an infinite critical Galton-Watson tree T_∞ as introduced in Section 2.4. Analogously to Section 3.2, one can define a RWRE on T_∞ determined by $(\rho_v)_{v\in T_\infty},V,R$ and ν on T_∞, as defined in (3.8). Once again, the law of the LERRW can be represented as a RWRE as in Theorem 3.6.

With such offspring distribution as in (2.1), Kesten's tree T_∞ is well-known to converge to a non-compact version of a stable tree coded by two independent Lévy excursions and an appropriate immigration measure under rescaling [39]; these are known as stable sin-trees and we will denote them by $\mathcal{T}_\gamma^\infty$ for $\gamma\in(1,2)$ (in fact due to uniform re-rooting invariance one can also construct the same objects with only the Lévy processes, but we do not explore this here). Analogously to $\mathcal{T}_\gamma^c$, one can define a canonical metric and a measure on $\mathcal{T}_\gamma^\infty$ using the canonical projection from the coding Lévy processes, and additionally define a snake process $\phi^{(\alpha)}$ satisfying all the properties of Theorem 3.8. Given these, we can define a resistance metric R_ϕ and a measure ν_ϕ on $\mathcal{T}_\gamma^\infty$ exactly as in (3.12) and (3.13). (To keep the notation light, we only write "∞" explicitly on the spaces T_∞ and $\mathcal{T}_\gamma^\infty$ and not on all the metrics and measures).

For non-compact mm-spaces, GHP convergence extends to Gromov-Hausdorff-vague (GHv) convergence which is equivalent to GHP convergence of metric balls $B(O_n,r)\to B(O,r)$ for almost every $r>0$ [12, Definition 5.8]. The results of Theorem 4.1 extend straightforwardly to these balls and therefore we deduce the following proposition. The proofs to extend to the infinite-dimensional setting are not illuminating, but can be carried out using the same strategy as in the proof of [9, Theorem 1.2].

Proposition 5.1. *Let T_∞ be Kesten's tree as defined in Theorem 2.3, with offspring distribution satisfying* (2.1)*. As $n\to\infty$, with initial weights as in* (1.1) *and reinforcement parameter Δ,*

$$(T_\infty,n^{-1}a_nR_n,(2n)^{-1}\nu_n,O_\infty)\xrightarrow{(d)}(\mathcal{T}_\gamma^\infty,R_\phi,\nu_\phi,O)$$

in the GHv topology.

Exactly as in Corollary 4.9, it has the following immediate corollary, by [31, Theorem 7.2]. Note that condition (40) of [31, Theorem 7.2] is satisfied as a consequence of the unique spine to infinity.

Corollary 5.2. *Let $X^{(n)}$ be the discrete-time LERRW on T_∞, with initial weights as in* (1.1) *and reinforcement parameter Δ. Let X be the diffusion on $\mathcal{T}_\gamma^\infty$ defined analogously to* (3.14)*. Then*

$$\left(n^{-1}a_n X^{(n)}_{\lfloor 2n^2 a_n^{-1} t\rfloor}\right)_{t\geq 0} \xrightarrow{(d)} (X_t)_{t\geq 0}.$$

Theorem 5.2 shows that the quenched scaling limit of LERRW, which we denote $(X_t)_{t\geq 0}$, can be represented as a mixture of diffusions in different random environments parametrised by different realisations of the functional $\phi^{(\alpha)}$. In particular, for a fixed realisation of $\mathcal{T}_\gamma^\infty$ we do not have a pointwise correspondence between realisations of the law of $\phi^{(\alpha)}$ and realisations of $(X_t)_{t\geq 0}$, but instead we must average over $\phi^{(\alpha)}$ to get the equality in distribution. Nevertheless, we can transfer almost sure results for $\phi^{(\alpha)}$ directly to $(X_t)_{t\geq 0}$ to prove Theorems 1.2 and 1.4. In this section we establish some almost sure properties of $\mathcal{T}_\gamma^\infty$ and $\phi^{(\alpha)}$. For some of the following propositions, the stated results are well-known but we could not find them explicitly written in the literature (for example Proposition 5.5). Therefore, we have provided an outline of the proofs but omitted some details (which would be long to justify and somewhat tangential to the main purpose of this paper).

Throughout this section, the notation d and μ refer to the measure and metric on the non-compact tree $\mathcal{T}_\gamma^\infty$.

Proposition 5.3. *$\mathbf{P}$-almost surely, for any $\varepsilon > 0$,*

$$\limsup_{r\to\infty}\left(\frac{\mu(B(O,r))}{r^{\frac{\gamma}{\gamma-1}}(\log r)^{\frac{1+\varepsilon}{\gamma-1}}}\right) = 0, \qquad \liminf_{r\downarrow 0}\left(\frac{\mu(B(\rho,r))}{r^{\frac{\gamma}{\gamma-1}}(\log\log r^{-1})^{\frac{1}{\gamma-1}}}\right) = \gamma - 1.$$

Proof. Set $r_m = 2^m$. By monotone convergence (applied twice), scaling invariance of $\mathcal{T}_\gamma^\infty$, then monotone convergence again,

$$\begin{aligned}
&\mathbf{P}\left(\limsup_{m\to\infty}\left(\frac{\mu(B(O,r_m))}{(r_m)^{\frac{\gamma}{\gamma-1}}(\log r_m)^{\frac{1+\varepsilon}{\gamma-1}}}\right)\leq 1\right)\\
&= \mathbf{P}\left(\lim_{k\to\infty}\lim_{K\to\infty}\sup_{k\leq m\leq K}\left(\frac{\mu(B(O,r_m))}{(r_m)^{\frac{\gamma}{\gamma-1}}(\log r_m)^{\frac{1+\varepsilon}{\gamma-1}}}\right)\leq 1\right)\\
&= \lim_{k\to\infty}\lim_{K\to\infty}\mathbf{P}\left(\sup_{k\leq m\leq K}\left(\frac{\mu(B(O,r_m))}{(r_m)^{\frac{\gamma}{\gamma-1}}(\log r_m)^{\frac{1+\varepsilon}{\gamma-1}}}\right)\leq 1\right)\\
&= \lim_{k\to\infty}\lim_{K\to\infty}\mathbf{P}\left(\sup_{k\leq m\leq K}\left(\frac{\mu(B(O,r_m^{-1}))}{(r_m^{-1})^{\frac{\gamma}{\gamma-1}}(\log r_m)^{\frac{1+\varepsilon}{\gamma-1}}}\right)\leq 1\right)\\
&= \mathbf{P}\left(\limsup_{m\to\infty}\left(\frac{\mu(B(O,r_m^{-1}))}{(r_m^{-1})^{\frac{\gamma}{\gamma-1}}(\log r_m)^{\frac{1+\varepsilon}{\gamma-1}}}\right)\leq 1\right).
\end{aligned}$$

By [42, Theorem 1.4], the final probability on the RHS is 1 when considered on the compact stable tree $\mathcal{T}_\gamma^c$ instead of $\mathcal{T}_\gamma^\infty$. Combining again with scaling invariance, it therefore follows from [39, Theorem 1.3] (which shows that $\mathcal{T}_\gamma^\infty$ is a local limit of compact stable trees) that this probability is also 1 on $\mathcal{T}_\gamma^\infty$. Finally, since $\varepsilon > 0$ was arbitrary we can replace 1 with and $\delta > 0$ and then with 0, and extend more generally to $r \to 0$ since $\mu(B(O,r)) \leq \mu(B(O,2^{\lceil \log_2 r\rceil}))$.

The second statement follows by the same proof using [43, Proposition 1.1] for $\mathcal{T}_\gamma^c$. □

It follows from [39, Proposition 1.1] that we can define a local time measure $(L^{(r)})_{r>0}$ on $\mathcal{T}_\gamma^\infty$ such that for any compactly supported continuous function $g : (0, \infty) \to \mathbb{R}$,

$$\int_{\mathcal{T}_\gamma^\infty} g(d(O,x))\mu(dx) = \int_0^\infty L^{(r)} g(r) dr. \tag{5.1}$$

The following result is a direct consequence of [42, Theorem 1.5] and [58, Theorem 8.8]. (In fact the result of [42, Theorem 1.5] is for the limit $r \downarrow 0$ on the compact $\mathcal{T}_\gamma^c$, but transfers to the limit $r \to \infty$ on $\mathcal{T}_\gamma^\infty$ exactly as in the proof of Proposition 5.3).

Proposition 5.4. *For any* $\varepsilon > 0$*,*

$$\limsup_{r\to\infty} \left(\frac{L^{(r)}}{r^{\frac{1}{\gamma-1}} (\log r)^{\frac{1+\varepsilon}{\gamma-1}}} \right) = 0.$$

Similarly, we have the following.

Proposition 5.5. $\mathbf{P}$*-almost surely, for any* $\varepsilon > 0$*,*

$$\limsup_{r\to\infty} \left(\frac{\mu(B(O,r+1) \setminus B(O,r))}{r^{\frac{1}{\gamma-1}} (\log r)^{\frac{1+\varepsilon}{\gamma-1}}} \right) = 0.$$

Proof. It follows from (5.1) (and continuous approximation) that

$$\mu(B(O,r+1) \setminus B(O,r)) = \int_r^{r+1} L^{(s)} ds,$$

for all r almost surely. The result therefore follows directly from Proposition 5.4 on integrating $L^{(s)}$ over $s \in [r, r+1]$. □

The construction of $\mathcal{T}_\gamma^\infty$ in [39] codes $\mathcal{T}_\gamma^\infty$ by two Lévy processes plus two stable subordinators that represent immigration. The result of [39, Proposition 1.3] shows that $\mathcal{T}_\gamma^\infty$ is also the local limit of compact stable trees conditioned on their height going to infinity. Since compact stable trees satisfy the property of uniform re-rooting invariance, it is also possible to consider them to be coded by two halves of a Lévy bridge, rather than a Lévy excursion (by applying the Vervaat transform to the Lévy excursion that codes them). By taking a limit using this bridge representation, it follows that $\mathcal{T}_\gamma^\infty$ can in fact be constructed solely by two Lévy processes, by imagining that the time 0 corresponds to the "tip" rather than the "base" of $\mathcal{T}_\gamma^\infty$. The argument proceeds exactly the same as in the proof in [9, Section 5.1]; we do not repeat it here, but just give the construction.

Construction of $\mathcal{T}_\gamma^\infty$

1. Let X and X' be independent γ-stable, spectrally positive Lévy processes.
2. Define a function $X^\infty : \mathbb{R} \to \mathbb{R}$ by

$$X_t^\infty = \begin{cases} X_t & \text{if } t \geq 0, \\ -X'_{-t^-} & \text{if } t < 0. \end{cases}$$

3. Given $s < t$, let $I_{s,t} = \inf_{r \in [s,t)} X_r^\infty$. Say that $s \preceq t$ if $X_{s-}^\infty \leq I_{s,t}$ and that $s \prec t$ if $s \preceq t$ and $s \neq t$. Also set $s \wedge t = \sup\{r \in \mathbb{R} : r \preceq t \text{ and } r \preceq s\}$.

4. If $s \preceq t$ set

$$d(s,t) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \int_s^t \mathbb{1}\{X_r \leq I_{s,t} + \varepsilon\} dr.$$

For all s, t this limit exists in probability, see [40, Lemma 1.1.3]. Then, for general $s, t \in \mathbb{R}$, set

$$d(s,t) = d(s \wedge t, s) + d(s \wedge t, t).$$

Finally, define an equivalence relation $\sim$ on $\mathbb{R}$ by setting $s \sim t$ if and only if $d(s,t) = 0$. We define

$$\mathcal{T}_\gamma^\infty = (\mathbb{R}/\sim, d)$$

and denote by π the canonical projection $\mathbb{R} \to \mathcal{T}_\gamma^\infty$, and set the root to be equal to $\pi(0)$.

We will use this construction and in particular excursion theory for the Lévy processes to prove some results about $\mathcal{T}_\gamma^\infty$. (This is more convenient than using the construction of [39] since we don't have to deal with the immigration). In what follows we will let $N(\cdot)$ denote the Itô excursion measure for X. This can be thought of as the "law" of an excursion of X, though it is not normalisable (see [33, Section 3.1.1] for a concise explanation).

For a subset $A \subset \mathcal{T}_\gamma^\infty$ and a pseudometric $\tilde{d}$ on A, let $D(A, \varepsilon, \tilde{d})$ denote the ε-packing number of $(A, \tilde{d})$; in other words, $D(A, \varepsilon, \tilde{d})$ is the maximal size of a collection of points $(x_i)_{i \leq D(A,\varepsilon,\tilde{d})}$ contained in A such that $\tilde{d}(x_i, x_j) \geq \varepsilon$ for all $i \neq j$.

Proposition 5.6. *For any $p > \frac{2\gamma}{\gamma-1}$, there exists a constant $C_p < \infty$ such that for any $r, t > 0$,*

$$\mathbf{E}\Big[D(B(O,r),t,d)^{\frac{1}{p}}\Big] \leq C_p (rt^{-1})^{\frac{\gamma}{\gamma-1}(1+\frac{1}{p})} \vee 1.$$

Remark 5.7. This is not an optimal result, but is sufficient for our purposes.

Proof of Proposition 5.6. Without loss of generality we can assume that $r > t$; otherwise $D(B(O,r),t,d) = 1$. By scaling invariance, it is sufficient to show the result when $r = 1$.

We claim the following: for any $x \in \mathbb{R}$, $\delta > 0$, there exist $C < \infty$, $c > 0$ such that

$$\begin{aligned} \mathbf{P}(B(O,1) \not\subset \pi([-\lambda,\lambda])) &\leq C\lambda^{-\frac{(1-\delta)(\gamma-1)}{\gamma}}, \\ \mathbf{P}\Big(\mathsf{Diam}(\pi([x - t^{\frac{\gamma}{\gamma-1}}\lambda^{-1}, x])) > t\Big) &\leq e^{-c\lambda^{\frac{\gamma-1}{\gamma^2}}}. \end{aligned} \tag{5.2}$$

The result then follows since if $B(O,1) \subset \pi([-\lambda,\lambda])$, we can divide the interval $[-\lambda,\lambda]$ up into $\lceil 2t^{\frac{-\gamma}{\gamma-1}}\lambda^2 \rceil$ covering intervals of length $t^{\frac{\gamma}{\gamma-1}}\lambda^{-1}$, and with high probability the diameter of each of these intervals is at most t. In particular, the probability that this does not happen is upper bounded by

$$C\lambda^{-\frac{(1-\delta)(\gamma-1)}{\gamma}} + 4t^{\frac{-\gamma}{\gamma-1}}\lambda^2 e^{-c\lambda^{\frac{\gamma-1}{\gamma^2}}}.$$

Any t-packing can have at most one point in each of these intervals (in fact, we have bounded the covering number), so we deduce that

$$\mathbf{P}\Big(D(B(O,1),t,d) \geq 4t^{-\frac{\gamma}{\gamma-1}}\lambda^2\Big) \leq C\lambda^{-\frac{(1-\delta)(\gamma-1)}{\gamma}} + 4t^{\frac{-\gamma}{\gamma-1}}\lambda^2 e^{-c\lambda^{\frac{\gamma-1}{\gamma^2}}}. \tag{5.3}$$

The result then follows by writing

$$\mathbf{E}\Big[D(B(O,r),t,d)^{\frac{1}{p}}\Big] = \int_0^\infty \mathbf{P}(D(B(O,r),t,d) \geq x^p)\,dx$$

and performing the appropriate change of variables to apply the tail bound of (5.3).

By scaling invariance, for the general claim we just need to replace t with $r^{-1}t$, so we see that the claim holds for any $p > \frac{2\gamma}{\gamma-1}$.

We therefore just need to prove (5.2). We will use excursion theory for the Lévy process X' coding the left side of $\mathcal{T}_\gamma^\infty$ (i.e. on the negative real line). We therefore let $\overline{X'}_t = \sup_{s \leq t} X'_s$ denote the running supremum process of X', and let $(L(t))_{t\geq 0}$ denote the local time of $\overline{X'} - X'$ at zero, normalised so that $\mathbf{E}\big[\exp\{-\lambda \overline{X'}_{L^{-1}(t)}\}\big] = e^{-t\lambda^{\gamma-1}}$ (this is well-defined, e.g. [22, Section VIII]). Moreover, we have by [22, Section VIII, Lemma 1] that L^{-1} is a stable subordinator of index $1-\frac{1}{\gamma}$, and [33, Proposition 3.1(ii)] that the measure $\sum_{\overline{X'}_s > \overline{X'}_{s^-}} \delta(L(s), \Delta_s(X'))$ is a Poisson point measure with intensity $dl \cdot C_\gamma x^{-\gamma} dx$.

First bound. To prove the first statement, first note that $X^\infty_{-t} = -X'_{t^-}$ for $t \geq 0$. It follows that new suprema of X' correspond to backwards minima of X^∞ from $t=0$. Therefore, $L = \inf\{s \geq 0 : L(s) \geq 1\}$ corresponds to the ancestor of 0 on the infinite backbone with $d(O, \pi(L)) = 1$, and setting $R = \inf\{t \geq 0 : X_t \leq -X'_L\}$ it follows that

$$B(O,1) \subset \pi([-L,R]).$$

Let S denote a stable subordinator with Lévy measure $C_\gamma x^{-\gamma} dx$. We then have that

1. $\mathbf{P}(L > \lambda) = \mathbf{P}(L(\lambda) \leq 1) = \mathbf{P}\big(L^{-1}(1) \geq \lambda\big) \leq c\lambda^{\frac{-(\gamma-1)}{\gamma}}$.

2. By [22, Section VIII, Propositions 3 and 4],

$$\begin{aligned}
\mathbf{P}(R > \lambda) &\leq \mathbf{P}\Big(X'_L \geq \lambda^{\frac{1-\delta}{\gamma}}\Big) + \mathbf{P}\Big(\inf_{s\leq\lambda} X_s \geq -\lambda^{\frac{1-\delta}{\gamma}}\Big) \\
&\leq \mathbf{P}\Big(S_1 \geq \lambda^{\frac{1-\delta}{\gamma}}\Big) + \mathbf{P}\Big(\inf_{s\leq 1} X_s \geq -\lambda^{\frac{-\delta}{\gamma}}\Big) \\
&\leq C\lambda^{-\frac{(1-\delta)(\gamma-1)}{\gamma}} + Ce^{-c\lambda^\delta}.
\end{aligned}$$

This establishes the first bound in (5.2).

Second bound. We will prove the diameter bound for $x=0$; the proof is the same for arbitrary x. Set $x_t = \inf\{s \geq 0 : d(0,-s) > t\}$. $\mathbf{P}$-almost surely, there is a unique path Γ_t from O to $\pi(x_t)$ in $\mathcal{T}_\gamma^\infty$ of length exactly t. Moreover, the interval $[-x_t, 0]$ codes all the subtrees grafted to one side of this path. Each of these complete subtrees are coded by the Itô excursion measure N, and moreover, x_t is equal to the sum of the lengths of all of these Itô excursions. In fact, the subtrees grafted to this side of the path Γ_t form a Poisson point process on this path. Let $M_{t,\lambda}$ denote the number of subtrees of lifetime at least $t^{\frac{\gamma}{\gamma-1}}\lambda^{-1}$ grafted to Γ_t at a point within distance $\frac{t}{2}$ of O. The subtrees grafted to Γ_t are concentrated in groups at certain hubs along Γ_t, and it follows from [21, Corollary 1] that $M_{t,\lambda}$ stochastically dominates a Poisson random variable with parameter $S_{\frac{t}{2}} N(\sigma > t^{\frac{\gamma}{\gamma-1}}\lambda^{-1}, H < \frac{t}{2})$, where S is a $(\gamma-1)$-stable subordinator by [45, Proposition 5.6] (here σ denotes the lifetime of a Lévy excursion, and H the height of a tree it codes). Moreover, since

$$\mathbf{P}\Big(S_{t/2} \leq ct^{\frac{1}{\gamma-1}}\lambda^{-p}\Big) \leq e\mathbf{E}\Big[e^{-c^{-1}\lambda^p S_1}\Big] \leq e^{1-c'\lambda^{p(\gamma-1)}}$$

and

$$N(\sigma > t^{\frac{\gamma}{\gamma-1}}\lambda^{-1}, H < t) \geq ct^{-\frac{1}{\gamma-1}}\lambda^{\frac{1}{\gamma}} - c't^{-\frac{1}{\gamma-1}},$$

this parameter is lower bounded by $c\lambda^{\frac{1}{\gamma}-p}$ with probability at least $1 - Ce^{-c\lambda^{p(\gamma-1)}}$. Therefore,

$$\mathbf{P}(M_{t,\lambda} = 0) \leq Ce^{-c\lambda^{p(\gamma-1)}} + e^{-c\lambda^{\frac{1}{\gamma}-p}}.$$

On the event $M_{t,\lambda} > 0$, it follows that $x_t \geq t^{\frac{\gamma}{\gamma-1}}\lambda^{-1}$, so taking $p = \gamma^{-2}$, we deduce the second result of (5.2), which completes the proof. □

We can extend the definitions of $d^{(\alpha)}$ and $\phi^{(\alpha)}$ from Section 3.3 to the infinite tree $\mathcal{T}_\gamma^c$. Just as we did there, for $\sigma, \sigma' \in \mathcal{T}_\gamma^\infty$, we set

$$d^{(\alpha)}(\sigma,\sigma') = \begin{cases} (d(O,\sigma)+1)^{1-\alpha} + (d(O,\sigma')+1)^{1-\alpha} - 2(d(O,\sigma\wedge\sigma')+1)^{1-\alpha}, & \text{if } \alpha < 1, \\ \log(d(O,\sigma)+1) + \log(d(O,\sigma')+1) - 2\log(d(O,\sigma\wedge\sigma')+1) & \text{if } \alpha = 1. \end{cases}$$

We also define $\phi^{(\alpha)}$ exactly as in Theorem 3.8, but on $\mathcal{T}_\gamma^\infty$ instead of $\mathcal{T}_\gamma^c$. This means that $\phi^{(\alpha)}$ is a centred Gaussian process and that for all $s, t \in \mathbb{R}$,

$$\mathbb{E}[|\phi^{(\alpha)}(s) - \phi^{(\alpha)}(t)|^2] = d^{(\alpha)}(s,t).$$

These are well-defined on $\mathcal{T}_\gamma^\infty$ by the same considerations as in the compact case. We now transfer the result of Proposition 5.6 to the metric $d^{(\alpha)}$.

Corollary 5.8. *For any $\alpha < 1$ and any $\eta > 0$, there exists $P > 0$ such that for all $p \geq P$ there exists $\tilde{C}_p < \infty$ such that*

$$\mathbf{E}\left[D\left(B(O,r),\delta,\sqrt{d^{(\alpha)}}\right)^{\frac{1}{p}}\right] \leq \tilde{C}_p((r+1)^{1-\alpha}\delta^{-2})^{\eta} \vee 1.$$

Proof. **Case 1:** $0 < \alpha < 1$**.** Note that, since $1-\alpha \in (0,1)$ in this case, we can assume that $r > \left(\frac{\delta^2}{2}\right)^{\frac{1}{1-\alpha}}$, otherwise $D\left(B(O,r),\delta,\sqrt{d^{(\alpha)}}\right) = 1$.

Firstly, suppose that $p > \frac{2\gamma q}{\gamma-1}$ for some $q > 1$. Then, by Jensen's inequality and Proposition 5.6, there exists $\varepsilon > 0$ such that

$$\mathbf{E}\left[D(B(O,r),t,d)^{\frac{1}{p}}\right] \leq \mathbf{E}\left[D(B(O,r),t,d)^{\frac{1}{\frac{2\gamma}{\gamma-1}+\varepsilon}}\right]^{\frac{1}{q}} \leq \left(C_p(rt^{-1})^{\frac{\gamma}{\gamma-1}+\frac{1}{2}}\right)^{\frac{1}{q}}. \tag{5.4}$$

We now set $B = B(O,r)$, measured with respect to the metric d, and observe the following: if $x, y \in B$, then $\left(\frac{d^{(\alpha)}(x,y)}{2}\right)^{\frac{1}{1-\alpha}} \leq d(x,y)$. Therefore, for any $\delta > 0$, a δ-packing of B with respect to $d^{(\alpha)}$ is a $\left(\frac{\delta}{2}\right)^{\frac{1}{1-\alpha}}$-packing of B with respect to d, so that

$$D\left(B,\delta,\sqrt{d^{(\alpha)}}\right) \leq D\left(B,\left(\frac{\delta^2}{2}\right)^{\frac{1}{1-\alpha}},d\right).$$

Taking p and q as above and combining with (5.4), we therefore deduce that

$$\mathbf{E}\left[D\left(B,\delta,\sqrt{d^{(\alpha)}}\right)^{\frac{1}{p}}\right] \leq \mathbf{E}\left[D\left(B,\left(\frac{\delta^2}{2}\right)^{\frac{1}{1-\alpha}},d\right)^{\frac{1}{p}}\right] \leq \tilde{C}_{p,\gamma}(r\delta^{\frac{-2}{1-\alpha}})^{\frac{1}{q}\left(\frac{\gamma}{\gamma-1}+\frac{1}{2}\right)}.$$

In particular, we can choose p and therefore q large enough so that this final exponent is less than η, which proves the result.

Case 2: $\alpha < 0$**.** Again set $B = B(O, r)$, measured with respect to the metric d, and now observe the following: if $x, y \in B$, then $d^{(\alpha)}(x, y) \le (1-\alpha)(r+1)^{|\alpha|} d(x, y)$. Therefore, for any $\delta > 0$, a δ-packing of B with respect to $d^{(\alpha)}$ is a $\frac{\delta}{(1-\alpha)(r+1)^{|\alpha|}}$-packing of B with respect to d, so that

$$D\left(B, \delta, \sqrt{d^{(\alpha)}}\right) \le D\left(B, \frac{\delta^2}{(\alpha-1)(r+1)^{|\alpha|}}, d\right).$$

Taking p and q as above and combining with (5.4), we therefore deduce that

$$\begin{aligned}\mathbf{E}\left[D\left(B, \delta, \sqrt{d^{(\alpha)}}\right)^{\frac{1}{p}}\right] &\le \mathbf{E}\left[D\left(B, \frac{\delta^2}{(1-\alpha)(r+1)^{|\alpha|}}, d\right)^{\frac{1}{p}}\right] \\ &\le \tilde{C}_p((1-\alpha)(r+1)^{1-\alpha}\delta^{-2})^{\frac{1}{q}\left(\frac{\gamma}{\gamma-1}+\frac{1}{2}\right)}.\end{aligned}$$

In particular, we can choose p and q large enough so that this final exponent is less than η, which proves the result. □

In the critical case the packing number instead grows logarithmically in r, as we see in the following proposition.

Proposition 5.9. *Take $\alpha = 1$. For any $p > \frac{1}{\gamma-1}$ and any $\varepsilon > 0$, there exists an event A_ε with $\mathbf{P}(A_\varepsilon^c) < \varepsilon$ and a constant $c_{\gamma,p,\varepsilon} < \infty$ such that for any $r > 0$,*

$$\mathbf{E}\left[D(B(O, r), \delta, \sqrt{d^{(1)}})^{\frac{1}{p}} \mathbb{1}\{A_\varepsilon\}\right] \le c_{\gamma,p,\varepsilon}\delta^{-\frac{2}{p}}(\lceil(\log(r+1))\rceil)^{\frac{\gamma+\varepsilon}{p(\gamma-1)}}.$$

Proof. Given $\delta > 0$, we construct a δ-covering of $B(O, r)$ with respect to the metric $d^{(1)}$, the size of which gives an upper bound for the δ-packing number. In particular, for $i \ge 0$ we let $S_\delta^{(i)}$ denote the set of vertices at distance $e^{\frac{i\delta}{2}} - 1$ from the root that have descendants at distance $e^{\frac{(i+1)\delta}{2}} - 1$ from the root. We claim that $\cup_{i=0}^{\lceil 2\delta^{-1}(\log(r+1))\rceil} S_\delta^{(i)}$ is a δ-covering of $B(O, r)$ with respect to the metric $d^{(1)}$. Indeed, if $x \in B(O, r)$ with $d(O, x) \ge e^\delta - 1$ then set $i_x = \lfloor 2\delta^{-1}\log(d(O, x) + 1)\rfloor$, so that i_x is the maximal i such that $e^{\frac{i\delta}{2}} - 1 \le d(O, x)$. Note that $i_x \ge 2$ since we are assuming that $d(O, x) \ge e^\delta - 1$. Then let A_x be the ancestor of x at distance $e^{\frac{\delta}{2}i_x} - 1$ from the root, and let A'_x be the ancestor of x at distance $e^{\frac{\delta}{2}(i_x-1)} - 1$ from the root. Then, since $d(O, x) \le e^{\frac{\delta}{2}(i_x+1)} - 1$, we have that

$$d^{(1)}(A'_x, x) \le d^{(1)}(A_x, A'_x) + d^{(1)}(A_x, x) \le \delta,$$

which implies that $x \in S_\delta^{(i_x-1)}$. If instead $d(O, x) \le e^\delta - 1$ then $x \in S_\delta^{(0)}$.

We therefore turn to bounding $|S_\delta^{(i)}|$ for each i. Recall that the size of generation $e^{\frac{i\delta}{2}} - 1$ can be formally measured by the local time measure of (5.1). Moreover, conditionally on the total local time measure at level $e^{\frac{i\delta}{2}} - 1$, denoted $L^{(e^{\frac{i\delta}{2}}-1)}$, being equal to $l^{(i)}$, it follows from [33, Proposition 3.1(ii)] that the number of subtrees emanating from level $e^{\frac{i\delta}{2}} - 1$ that reach level $e^{\frac{(i+1)\delta}{2}} - 1$ is one more than a Poisson random variable with parameter $2l^{(i)}N(H > e^{\frac{i\delta}{2}}(e^{\frac{\delta}{2}} - 1))$, where N is the Itô excursion measure, so that $N(H > t) = c_\gamma t^{-\frac{1}{\gamma-1}}$ (see [45, Proposition 5.6]).

By Theorem 5.4, it follows that we can choose $K_\varepsilon < \infty$ such that

$$A_\varepsilon := \{\forall r > 0 : L^{(r)} \le K_\varepsilon r^{\frac{1}{\gamma-1}}(\log r)^{\frac{1+\varepsilon}{\gamma-1}}\}$$

satisfies $\mathbf{P}(A_\varepsilon^c) < \varepsilon$. On this event, $l^{(i)} \le K_\varepsilon e^{\frac{i\delta}{2(\gamma-1)}}(i\delta)^{\frac{1+\varepsilon}{\gamma-1}}$ for all $i \ge 0$, so that

$$|S^{(i)}| \overset{s.d.}{\preceq} \mathsf{Poi}\left(c_\gamma K_\varepsilon e^{\frac{i\delta}{2(\gamma-1)}}(i\delta)^{\frac{1+\varepsilon}{\gamma-1}}\left(e^{\frac{i\delta}{2}}(e^{\frac{\delta}{2}} - 1)\right)^{-\frac{1}{\gamma-1}}\right) = \mathsf{Poi}\left(c_\gamma K_\varepsilon (i\delta)^{\frac{1+\varepsilon}{\gamma-1}}(e^{\frac{\delta}{2}} - 1)^{-\frac{1}{\gamma-1}}\right),$$

where the symbol $\overset{s.d.}{\preceq}$ denotes stochastic domination between random variables. Therefore, if $p > 1$ and $\delta > 0$ is sufficiently small,

$$\begin{aligned}\mathbf{E}\left[\left(\sum_{i=0}^{\lceil 2\delta^{-1}\log(r+1)\rceil}|S_\delta^{(i)}|\right)^{\frac{1}{p}}\right] &\leq \mathbf{E}\left[\sum_{i=0}^{\lceil 2\delta^{-1}(\log(r+1))\rceil}|S_\delta^{(i)}|\right]^{\frac{1}{p}} \\ &\leq \left(c_\gamma K_\varepsilon (e^{\frac{\delta}{2}}-1)^{-\frac{1}{\gamma-1}}\sum_{i=0}^{\lceil 2\delta^{-1}\log(r+1)\rceil}(i\delta)^{\frac{1+\varepsilon}{\gamma-1}}\right)^{\frac{1}{p}} \\ &\leq c_{\gamma,p,\varepsilon}\delta^{-\frac{1}{p}}(\lceil\log(r+1)\rceil)^{\frac{\gamma+\varepsilon}{p(\gamma-1)}}.\end{aligned}$$

□

Proposition 5.10. *(i)* $\mathbf{P}\times\mathbb{P}$*-almost surely, for any* $\alpha < 1, \beta > \frac{1}{2}$*,*

$$\lim_{r\uparrow\infty}\sup_{x\in B(O,r)^c}\frac{|\phi^{(\alpha)}(x)|}{d(O,x)^{(1-\alpha)\beta}} = 0.$$

(ii) $\mathbf{P}\times\mathbb{P}$*-almost surely, for any* $\beta > \frac{1}{2}$*,*

$$\lim_{r\uparrow\infty}\sup_{x\in B(O,r)^c}\frac{|\phi^{(1)}(x)|}{(\log(d(O,x)+1))^{\beta}} = 0$$

Proof. **Case** $\alpha < 1$**.** The result is a consequence of [72, Theorem 2.2.4 and Corollary 2.2.5]. First, fix some $p > \frac{2\gamma}{\gamma-1}$, and note that by Gaussianity, we have for any $s,t > 0$ that $\mathbf{P}$-almost surely,

$$\mathbb{E}\left[\frac{|\phi^{(\alpha)}(s)-\phi^{(\alpha)}(t)|^p}{d^{(\alpha)}(s,t)^{\frac{p}{2}}}\right] \leq C_p^p,$$

where C_p is a deterministic constant. Rearranging, we deduce that

$$\mathbb{E}\Big[|\phi^{(\alpha)}(s)-\phi^{(\alpha)}(t)|^p\Big]^{\frac{1}{p}} \leq C_p\sqrt{d^{(\alpha)}(s,t)},$$

which verifies the first condition of [72, Theorem 2.2.4] in the case $\psi(x) = x^p$ (in which case the ψ-Orlicz norm in Theorem 2.2.4 there coincides with the usual L^p-norm). Now choose some $\delta \in (0, p-\frac{2\gamma}{\gamma-1})$. Working with the pseudometric $\sqrt{d^{(\alpha)}(\cdot,\cdot)}$, in which case $\mathsf{Diam}(B(O,r-1)) \leq 2\sqrt{r^{1-\alpha}}$ for all $r \geq 1$, we therefore deduce from [72, Corollary 2.2.5] and Fubini's theorem that for any $p > 1$,

$$\begin{aligned}\mathbf{E}\left[\mathbb{E}\left[\sup_{s,t\in B(O,r-1)}|\phi^{(\alpha)}(s)-\phi^{(\alpha)}(t)|\right]\right] &\leq \mathbf{E}\left[\mathbb{E}\left[\sup_{s,t\in B(O,r-1)}|\phi^{(\alpha)}(s)-\phi^{(\alpha)}(t)|^p\right]^{\frac{1}{p}}\right] \\ &\leq \mathbf{E}\left[K_p\int_0^{2\sqrt{r^{1-\alpha}}}D\left(B(O,r-1),\varepsilon,\sqrt{d^{(\alpha)}}\right)^{\frac{1}{p}}d\varepsilon\right] \\ &= K_p\int_0^{2\sqrt{r^{1-\alpha}}}\mathbf{E}\left[D\left(B(O,r-1),\varepsilon,\sqrt{d^{(\alpha)}}\right)^{\frac{1}{p}}\right]d\varepsilon.\end{aligned}$$

By Corollary 5.8, we can also assume that p is large enough that we can upper bound this by

$$K_p\int_0^{2\sqrt{r^{1-\alpha}}}\tilde{C}_p(r^{1-\alpha}\varepsilon^{-2})^{\frac{1}{3}}d\varepsilon = K_p'(r^{1-\alpha})^{\frac{1}{3}}\left[\varepsilon^{\frac{1}{3}}\right]_0^{2\sqrt{r^{1-\alpha}}} = \tilde{K}_p\sqrt{r^{1-\alpha}}.$$

Therefore, if $\beta > \frac{1}{2}$, we can choose some $\eta < \beta - \frac{1}{2}$ and apply Markov's inequality to get that

$$\begin{aligned}
&\mathbf{P} \times \mathbb{P}\left(\sup_{s,t \in B(O,r-1)} |\phi^{(\alpha)}(s) - \phi^{(\alpha)}(t)| \geq r^{(1-\alpha)(\beta-\eta)} \right) \\
&\leq \mathbf{E}\left[\mathbb{E}\left[\sup_{s,t \in B(O,r)} |\phi^{(\alpha)}(s) - \phi^{(\alpha)}(t)| \right]\right] r^{-(1-\alpha)(\beta-\eta)} \\
&\leq \tilde{K}_p \sqrt{r^{1-\alpha}} r^{-(1-\alpha)(\beta-\eta)} \\
&= \tilde{K}_p r^{-(1-\alpha)(\beta-\eta-\frac{1}{2})}.
\end{aligned}$$

Therefore, applying Borel-Cantelli along the subsequence $r_n = 2^n$, we deduce that

$$\sup_{s,t \in B(O,r_n-1)} |\phi^{(\alpha)}(s) - \phi^{(\alpha)}(t)| < r_n^{(1-\alpha)(\beta-\eta)}$$

for all sufficiently large n. Then, if $x \in B(O, r_{n+1}-1) \setminus B(O, r_n - 1)$, and $d(O,x) = r$, we have that

$$\begin{aligned}
\frac{\phi^{(\alpha)}(x)}{d(O,x)^{(1-\alpha)\beta}} \leq r^{-(1-\alpha)\beta} \sup_{s,t \in B(O,r_{n+1}-1)} |\phi^{(\alpha)}(s) - \phi^{(\alpha)}(t)| &< r^{-(1-\alpha)\beta} r_{n+1}^{(1-\alpha)(\beta-\eta)} \\
&\leq 2^{(1-\alpha)(\beta-\eta)} r^{-(1-\alpha)\eta} \to 0,
\end{aligned}$$

as $r \to \infty$, which proves the result.

Case $\alpha = 1$**.** As above, we can apply [72, Theorem 2.2.4] with the function $\psi(x) = x^p$. Again let $\beta > \frac{1}{2}, \varepsilon > 0$ and choose some $\eta < \frac{2}{3}\left(\beta - \frac{1}{2}\right)$. This time we instead apply Proposition 5.9, work on the event A_ε and choose p large enough that $\frac{1+\varepsilon}{p(\gamma-1)} < \eta/2$ to deduce that

$$\begin{aligned}
\mathbf{E}\left[\mathbb{E}\left[\sup_{s,t \in B(O,r)} |\phi^{(1)}(s) - \phi^{(1)}(t)| \right] \mathbb{1}\{A_\varepsilon\} \right] &\leq \int_0^{\sqrt{\log(r+1)}} c_{\gamma,p,\varepsilon} \delta^{-\frac{2}{p}} \lceil (\log(r+1)) \rceil^{\frac{\gamma+\varepsilon}{p(\gamma-1)}} d\delta \\
&= c_{\gamma,p,\varepsilon} \sqrt{\log(r+1)} \lceil (\log(r+1)) \rceil^{\frac{\eta}{2}}.
\end{aligned}$$

Therefore, if $\beta > \frac{1}{2}$, we can choose some $\eta < \frac{2}{3}\left(\beta - \frac{1}{2}\right)$ and apply Markov's inequality to get that

$$\begin{aligned}
&\mathbf{P} \times \mathbb{P}\left(\sup_{s,t \in B(O,r)} |\phi^{(1)}(s) - \phi^{(1)}(t)| \geq (\log(r+1))^{\beta-\eta}, \text{ and } A_\varepsilon \right) \\
&\leq \mathbf{E}\left[\mathbb{E}\left[\sup_{s,t \in B(O,r)} |\phi^{(1)}(s) - \phi^{(1)}(t)| \right] \mathbb{1}\{A_\varepsilon\} \right] (\log(r+1))^{-(\beta-\eta)} \\
&\leq c_{\gamma,p,\varepsilon} (\log(r+1))^{\frac{1}{2}+\frac{\eta}{2}-(\beta-\eta)} \\
&= c_{\gamma,p,\varepsilon} (\log(r+1))^{-(\beta-\frac{3\eta}{2}-\frac{1}{2})}.
\end{aligned}$$

Therefore, applying Borel-Cantelli along the subsequence $r_n = 2^{2^n} - 1$, we deduce that on the event A_ε, i.e. with probability at least $1-\varepsilon$, $\sup_{s,t \in B(O,r_n)} |\phi^{(1)}(s) - \phi^{(1)}(t)|$ is upper bounded by $(\log(r_n+1))^{\beta-\eta}$ for all sufficiently large n. Then, if $x \in B(O, r_{n+1}) \setminus B(O, r_n)$

and $d(O,x) = r \in [r_n, r_{n+1}]$, we have that

$$\begin{aligned}\frac{\phi^{(1)}(x)}{(\log(d(O,x)+1))^\beta} &\le (\log(r+1))^{-\beta} \sup_{s,t\in B(O,r_{n+1})} |\phi^{(1)}(s) - \phi^{(1)}(t)| \\ &< (\log(r+1))^{-\beta}(\log(r_{n+1}+1))^{\beta-\eta} \\ &\le 2(\log(r+1))^{-\beta}(\log(r_n+1))^{\beta-\eta} \\ &\le 2(\log(r+1))^{-\eta} \\ &\to 0,\end{aligned}$$

as $r \to \infty$. Since $\varepsilon > 0$ was arbitrary, this proves the result. □

6 Reinforced strongly recurrent regime: $\alpha < 1, \Delta > 0$

In this section we will prove Theorem 1.2. Recall from (3.14) that, given a realisation of $\mathcal{T}_\gamma^\infty$, the limiting diffusion X has the annealed law of a diffusion in a random potential $\phi^{(\alpha)}$, i.e.

$$P_O\left((X_t)_{t\ge 0} \in \cdot\right) = \int \tilde{P}_{O,\phi}\left((X_t)_{t\ge 0} \in \cdot\right) \mathbb{P}(d\phi^{(\alpha)}).$$

We will in fact prove the analogue of Theorem 1.2 for the *quenched* law of this diffusion, i.e. $\mathbf{P} \times \mathbb{P} \times \tilde{P}_{O,\phi}$- almost surely. This clearly implies the same result for the annealed process, and therefore for X.

6.1 Volume and resistance growth in $\mathcal{T}_{\gamma,\alpha}$

We start with some asymptotics for balls with respect to the distorted metric R_ϕ. Here B_ϕ denotes the open ball measured with respect to R_ϕ, as defined by (3.12):

$$R_\phi(x,y) = \int_{[x,y]} e^{B_\Delta \phi^{(\alpha)}(z) + A_\Delta[(d(O,z)+1)^{1-\alpha}-1] - \alpha \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z),$$
$$\nu_\phi(A) = \int_A e^{-\left[B_\Delta \phi^{(\alpha)}(x) + A_\Delta[(d(O,z)+1)^{1-\alpha}-1] - \alpha \log(d(O,x)+1)\right]} \mu(\mathrm{d}x),$$

where $B_\Delta = \sqrt{\frac{4\Delta}{1-\alpha}}$ and $A_\Delta = \frac{\Delta}{1-\alpha}$.

Proposition 6.1. $\mathbf{P} \times \mathbb{P}$*-almost surely, for any* $\varepsilon > 0$ *we have that*

$$B_\phi(O, e^{(1-\varepsilon)A_\Delta r^{1-\alpha}}) \subset B(O,r) \subset B_\phi(O, e^{(1+\varepsilon)A_\Delta r^{1-\alpha}})$$

for all sufficiently large r.

Proof. Take some $\beta \in (\frac{1}{2}, 1)$ and some $\varepsilon > 0$. Also set $b = 1-\alpha$. By Proposition 5.10(i), there almost surely exists $R < \infty$ such that

$$\sup_{x \in B(O,R)^c} \frac{|\phi^{(\alpha)}(x)|}{d(O,x)^{b\beta}} \le 1.$$

Without loss of generality, we also assume that

- $A_\Delta[(R+1)^b - 1] - B_\Delta R^{b\beta} - \alpha \log(R+1) \ge (1 - \frac{\varepsilon}{2}) A_\Delta R^b$,
- $\left(\frac{R-1}{R}\right)^b \ge \frac{1-\varepsilon}{1-\frac{\varepsilon}{2}}$,
- $A_\Delta[(R+1)^b - 1] + B_\Delta R^{b\beta} - \alpha \log(R+1) \le (1+\varepsilon) A_\Delta R^b$,
- $A_\Delta b R^{b-1} \ge \mathbb{1}\{b > 1\}$,
- $\frac{b^{-1}R^{1-b}}{A_\Delta} \le e^{\varepsilon A_\Delta R^b}$, which we will just use in the case $b \le 1$.

Now take some $r > R+1$. We start by showing that $B(O,r)^c \subset B_\phi(O, e^{(1-\varepsilon)A_\Delta r^b})^c$ for all sufficiently large r. Indeed, assume that $d(O,x) > r$, and let y be the point on the path from O to x such that $d(O,y) = R$ and $d(x,y) = d(O,x) - R$. (Since $\mathcal{T}_\gamma^\infty$ is a length space, y almost surely exists and is unique.) Moreover, $d(O,z) > R$ for all $z \in [y,x]$ by our choice of R, which implies that $A_\Delta d(O,z)^b - B_\Delta d(O,z)^{b\beta} - \alpha \log d(O,z) \geq (1-\frac{\varepsilon}{2})A_\Delta d(O,z)^b$ for all such z. We therefore have that

$$\begin{aligned}
R_\phi(O,x) &\geq \int_{[y,x]} e^{A_\Delta[(d(O,z)+1)^b - 1] - B_\Delta d(O,z)^{b\beta} - \alpha \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \\
&\geq \int_R^r e^{(1-\frac{\varepsilon}{2})A_\Delta s^b} ds \\
&\geq e^{(1-\frac{\varepsilon}{2})A_\Delta (r-1)^b} \\
&\geq e^{(1-\varepsilon)A_\Delta r^b}.
\end{aligned}$$

For the second inclusion, we keep β, R and y as above, and now assume that x is such that $d(O,x) < r$. Also set

$$\tilde{r}_c = \inf\left\{ r \geq R : \sup_{w \in \partial B(O,R)} \int_{[O,w]} e^{B_\Delta \phi^{(\alpha)}(z) + A_\Delta[(d(O,z)+1)^b - 1] - \alpha \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \right.$$
$$\left. \leq \frac{\varepsilon}{1+\varepsilon} e^{(1+\varepsilon)A_\Delta r^b} \right\}.$$

(This is always possible since $\phi^{(\alpha)}$ is continuous with bounded sample paths, so is bounded on the compact set $B(O,R)$, and so

$$\sup_{w \in \partial B(O,R)} \int_{[O,w]} e^{B_\Delta \phi^{(\alpha)}(z) + A_\Delta d(O,z)^b - \alpha \log d(O,z)} \lambda^\gamma(\mathrm{d}z)$$

is almost surely finite).

We then have for all $r > \tilde{r}_c \vee R$ and all $x \in B(O,r)$ that

$$\begin{aligned}
R_\phi(O,x) &\leq \int_{[O,y]} e^{B_\Delta \phi^{(\alpha)}(z) + A_\Delta[(d(O,z)+1)^b - 1] - \alpha \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \\
&\quad + \int_{[y,x]} e^{B_\Delta \phi^{(\alpha)}(z) + A_\Delta[(d(O,z)+1)^b - 1] - \alpha \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \\
&\leq \frac{\varepsilon}{1+\varepsilon} e^{(1+\varepsilon)A_\Delta r^b} + \int_R^r e^{(1+\varepsilon)A_\Delta s^b} ds.
\end{aligned}$$

In the case $b > 1$ we bound this above by (using that $A_\Delta b R^{b-1} \geq 1$):

$$\frac{\varepsilon}{1+\varepsilon} e^{(1+\varepsilon)A_\Delta r^b} + \frac{1}{1+\varepsilon} \int_R^r (1+\varepsilon) A_\Delta b s^{b-1} e^{(1+\varepsilon)A_\Delta s^b} ds \leq e^{(1+\varepsilon)A_\Delta r^b}.$$

If $b \leq 1$, we instead bound this above by (using that $\frac{b^{-1}R^{1-b}}{A_\Delta} \leq e^{\varepsilon A_\Delta R^b}$):

$$\frac{\varepsilon}{1+\varepsilon} e^{(1+\varepsilon)A_\Delta r^b} + \frac{b^{-1} r^{1-b}}{A_\Delta(1+\varepsilon)} \int_R^r (1+\varepsilon) A_\Delta b s^{b-1} e^{(1+\varepsilon)A_\Delta s^b} ds \leq e^{(1+2\varepsilon)A_\Delta r^b},$$

which completes the proof since ε was arbitrary. □

Proposition 6.2. *For any $\alpha < 1$, we have that $\mathbf{P} \times \mathbb{P}$-almost surely, for any $\varepsilon > 0$,*

$$r^{1-\varepsilon} \leq R_\phi(O, B_\phi(O,r)^c) \leq r$$

for all sufficiently large r.

Proof. Fix $\varepsilon > 0$, and let $M_\phi(r)$ be the smallest cardinality of a set of points in

$$B_\phi(O,r) \setminus B_\phi(O,r^{1-\varepsilon})$$

such that any path passing from O to $B_\phi(O,r)^c$ must pass through one of the points in the set. Similarly to [19, Lemma 4.5], we note that since any such set is a cutset, it follows from the parallel law for resistance that

$$R_\phi(O, B_\phi(O,r)^c) \geq \frac{r^{1-\varepsilon}}{M_\phi(r)}. \tag{6.1}$$

Now choose $\hat{\varepsilon}$ small enough that $(1+\hat{\varepsilon})(1-\varepsilon) < 1-\hat{\varepsilon}$. By Proposition 6.1, it follows that any cutset separating

$$B\left(O, \left(\frac{(1-\varepsilon)\log r}{A_\Delta(1-\hat{\varepsilon})}\right)^{\frac{1}{1-\alpha}}\right) \text{ from } B\left(O, \left(\frac{\log r}{A_\Delta(1+\hat{\varepsilon})}\right)^{\frac{1}{1-\alpha}}\right)^c$$

is also a cutset separating $B_\phi(O,r^{1-\varepsilon})$ from $B_\phi(O,r)^c$, so that $M_\phi(r) \leq M(r)$, where $M(r)$ is the smallest cardinality of a set of points separating

$$B\left(O, \left(\frac{(1-\varepsilon)\log r}{A_\Delta(1-\hat{\varepsilon})}\right)^{\frac{1}{1-\alpha}}\right) \text{ from } B\left(O, \left(\frac{\log r}{A_\Delta(1+\hat{\varepsilon})}\right)^{\frac{1}{1-\alpha}}\right)^c.$$

By scaling invariance of the stable tree, $M(r) \overset{(d)}{=} M(e^{A_\Delta(1+\hat{\varepsilon})})$ for all $r>0$, so we bound the latter quantity, and set

$$\delta = 1 - \left(\frac{(1+\hat{\varepsilon})(1-\varepsilon)}{(1-\hat{\varepsilon})}\right)^{\frac{1}{1-\alpha}},$$

so that we are in fact counting subtrees from level $1-\delta$ to level 1.

The "size" of generation $1-\delta$ can be formally measured by the local time measure of Proposition 5.4. Moreover, conditional on the total local time measure at level $1-\delta$ being equal to L, it follows from the construction on page 31 that the number of subtrees emanating from level $1-\delta$ that reach level 1 is one more than a Poisson random variable with parameter $2LN(H>\delta)$, where N is the Itô excursion measure, so that $N(H>\delta) = c_\gamma \delta^{-\frac{1}{\gamma-1}}$ for a deterministic $c_\gamma \in (0,\infty)$ [45, Proposition 5.6]. Moreover, $M(e^{A_\Delta(1+\hat{\varepsilon})})$ is equal to the number of such subtrees. Again letting L denote the total local time measure at level $1-\delta$, we have from (a minor adaptation of) [42, Proposition 5.2] that there exists $c>0$ such that $\mathbf{P}(L>\lambda) \leq c\lambda^{-(\gamma-1)}$, uniformly over $\delta \in [0,1]$. From a Chernoff bound, we therefore deduce that

$$\mathbf{P}\Big(M(e^{A_\Delta(1+\hat{\varepsilon})}) > \lambda\Big) \leq \mathbf{P}\left(L > \frac{\delta^{\frac{1}{\gamma-1}}}{2c_\gamma e}\lambda\right) + \mathbf{P}\left(\text{Poisson}\left(\frac{\lambda}{e}\right) \geq \lambda\right) \leq c_\delta \lambda^{-(\gamma-1)} + e^{-\lambda}. \tag{6.2}$$

We therefore deduce from Borel-Cantelli and (6.1) that if we take $r_n = 2^n$ and $\lambda_n = (\log r_n)^{\frac{1+\varepsilon}{\gamma-1}}$, then $\mathbf{P} \times \mathbb{P}$-almost surely, we have that

$$R_\phi(O, B_\phi(O,r_n)^c) \geq \frac{r_n^{1-\varepsilon}}{(\log r_n)^{\frac{1+\varepsilon}{\gamma-1}}} \geq r_n^{1-2\varepsilon}$$

for all sufficiently large n. By monotonicity, this implies that $R_\phi(O, B_\phi(O,r)^c) \geq r^{1-3\varepsilon}$ for all sufficiently large r. Since ε was arbitrary and the upper bound $R_\phi(O, B_\phi(O,r)^c) \leq r$ holds trivially, this proves the result. □

Proposition 6.3. $\mathbf{P} \times \mathbb{P}$*-almost surely,* $\nu_\phi(\mathcal{T}_\gamma^\infty) < \infty$.

Proof. Fix $\varepsilon > 0$ and $\beta \in (\frac{1}{2}, 1)$. By Propositions 5.3, 5.10(i) and since $\phi^{(\alpha)}$ is continuous, we can $\mathbf{P} \times \mathbb{P}$-almost surely choose $R \geq 1$ and $V, \varepsilon, p < \infty$ so that

- $\sup_{x \in B(O,R)} \left\{ e^{-\left[B_\Delta \phi^{(\alpha)}(x) + A_\Delta[(d(O,x)+1)^{1-\alpha} - 1] - \alpha \log(d(O,x)+1)\right]} \right\} \leq p$,
- $\sup_{x \in B(O,R)^c} \frac{B_\Delta |\phi^{(\alpha)}(x)| + \alpha \log(d(O,x)+1)}{d(O,x)^{(1-\alpha)\beta}} < \varepsilon$,
- $\sup_{r \geq 1} \left\{ \frac{\mu(B(O,r+1))}{r^{\frac{\gamma}{\gamma-1}} (\log r)^{\frac{1+\varepsilon}{\gamma-1}}} \right\} \leq V$.

We then have that

$$\begin{aligned}
&\nu_\phi(\mathcal{T}_\gamma^\infty) \\
&\leq p\mu\big(B(O,R)\big) \\
&+ \sum_{r=R}^{\infty} \sup_{x \in B(O,r+1) \setminus B(O,r)} \left\{ e^{-\left[B_\Delta \phi^{(\alpha)}(x) + A_\Delta[(d(O,x)+1)^{1-\alpha} - 1] - \alpha \log(d(O,x)+1)\right]} \right\} \mu\big(B(O,r+1)\big) \\
&\leq pVR^{\frac{\gamma}{\gamma-1}} (\log R)^{\frac{1+\varepsilon}{\gamma-1}} + V \sum_{r=R}^{\infty} e^{\varepsilon(r+1)^{(1-\alpha)\beta} - A_\Delta(r^{1-\alpha} - 1)} r^{\frac{\gamma}{\gamma-1}} (\log r)^{\frac{1+\varepsilon}{\gamma-1}} =: C(p,V,R,\varepsilon),
\end{aligned}$$

where $C(p,V,R,\varepsilon)$ is a finite constant. □

6.2 Proof of Theorem 1.2

We now have all the tools to prove Theorem 1.2.

Proof of Theorem 1.2. Again set $b = 1 - \alpha$. We work pointwise on the probability space Ω' on which $\mathcal{T}_\gamma^\infty$ and the environment $\phi^{(\alpha)}$ are defined. Recall from (3.14) that given $\mathcal{T}_\gamma^c$ and $\phi^{(\alpha)}$, $\tilde{P}_{O,\phi}$ denotes the (quenched) law of the corresponding RWRE. We will show that the statement of Theorem 1.2 holds $\tilde{P}_{O,\phi}$ almost surely on Ω'. This then transfers to the annealed law via (3.14). Since the quenched law of the LERRW limit is equal to the annealed law of the RWRE, this proves the result.

Throughout we assume that $\mathcal{T}_\gamma^c$ and $\phi^{(\alpha)}$ are fixed and therefore just write $\tilde{P}$ in place of $\tilde{P}_{O,\phi}$.

We first show that, $\tilde{P}$-almost surely,

$$\limsup_{t \to \infty} \frac{d(O, X_t)}{(\log t)^{1/b}} \leq A_\Delta^{-1/b}.$$

Letting $T_{B(O,r)}$ denote the exit time of X from $B(O,r)$, we will in fact show that, almost surely,

$$T_{B(O,r(1+\varepsilon))} \geq e^{(1-\varepsilon)A_\Delta r^b},$$

for all sufficiently large r. For notational convenience, set $t = e^{A_\Delta r^b}$ and $\tau_r = T_{B(O,r(1+\varepsilon))}$. Also let T_r be a geometric random variable, with parameter p_r given by

$$p_r = \sup_{x \in \partial B(O,1)} \tilde{P}_x(\tau_r < \tau_O),$$

where τ_O is the hitting time of O (p_r is well-defined by [11, Proposition 1.9]). We can then write

$$\tilde{P}(\tau_r \leq t) \leq \tilde{P}\left(\sum_{i=1}^{T_r - 1} \xi_i \leq t \right),$$

where $(\xi_i)_{i=1}^{T_r-1}$ are the times to travel from $\partial B(O,1)$ to O and back to $\partial B(O,1)$, conditional on not hitting $\partial B(O,r(1+\varepsilon))$ first, from a "best case" starting point on $\partial B(O,1)$. Letting $L(t)=\log t$, we then have that

$$\tilde{P}(\tau_r \le t) \le \tilde{P}(T_r \le tL(t)) + \tilde{P}\left(\sum_{i=1}^{tL(t)-1} \xi_i \le t\right). \tag{6.3}$$

To bound the first term, note that since $\partial B(O,1)$ is a cutset separating $\partial B(O,r(1+\varepsilon))$ from O, we can apply Propositions 6.1 and 6.2 to deduce that $\mathbf{P}\times\mathbb{P}$-almost surely for any $\varepsilon>0$, we have for all sufficiently large r that

$$\begin{aligned} p_r \le \frac{1}{1+R_\phi(\partial B(O,1),\partial B(O,(1+\varepsilon)r))} &\le \frac{1}{R_\phi(O,B(O,(1+\varepsilon)r)^c)} \\ &\le \frac{1}{R_\phi(O,B_\phi(O,e^{A_\Delta r^b(1+\frac{1}{2}\varepsilon)})^c)} \\ &\le e^{-A_\Delta r^b(1+\frac{1}{4}\varepsilon)}. \end{aligned}$$

Since $t=e^{A_\Delta r^b}$, we therefore have for all sufficiently large r that $p_r tL(t) \le A_\Delta r^b e^{-\frac{1}{4}A_\Delta \varepsilon r^b}$, and hence converges to zero as $r\to\infty$. It follows that for all sufficiently large r,

$$\tilde{P}(T_r \le tL(t)) = 1-(1-p_r)^{tL(t)} \le 1-e^{-2p_r tL(t)} \le (1+\delta)p_r tL(t) \le 2A_\Delta r^b e^{-\frac{\varepsilon}{4}A_\Delta r^b}. \tag{6.4}$$

Now take any $s>0$. To bound the second term of (6.3), we follow the approach of [52, Section 4] and use the Markov property to write

$$\tilde{E}_O\left[T_{B(O,1)}\right] \le s + \tilde{E}_{X_s}\left[T_{B(O,1)}\mathbb{1}\{T_{B(O,1)>s}\}\right] \le s + \sup_{x\in B(O,1)} \tilde{E}_x\left[T_{B(O,1)}\right]\tilde{P}\left(T_{B(O,1)}>s\right). \tag{6.5}$$

Letting $g_B(\cdot,\cdot)$ denote the Green's function for the diffusion killed on exiting the ball $B(O,1)$ (e.g. as in [52, Equation (4.7)]), we can then write (also using [52, Equations (4.5) and (4.6)])

$$\begin{aligned} \sup_{x\in B(O,1)} \tilde{E}_x\left[T_{B(O,1)}\right] &= \sup_{x\in B(O,1)} \int_{B(O,1)} g_B(x,y)\nu_\phi(dy) \\ &\le \sup_{x\in B(O,1)} \int_{B(O,1)} R_\phi(x,B(O,1)^c)\nu_\phi(dy) \\ &\le \sup_{x\in B(O,1)} R_\phi(x,B(O,1)^c)\cdot \nu_\phi(B(O,1)). \end{aligned}$$

Note that $\nu_\phi(B(O,1))$ is $\mathbf{P}\times\mathbb{P}$-almost surely finite and non-zero by Proposition 6.3, and the same holds for $\sup_{x\in B(O,1)} R_\phi(x,B(O,1)^c)$ by Proposition 6.1. It therefore follows from (6.5) that there $\mathbf{P}\times\mathbb{P}$-almost surely exist constants $c_1<c_2\in(0,\infty)$ such that for all $s\ge 0$,

$$c_1 \le s + c_2\tilde{P}\left(T_{B(O,1)}>s\right).$$

Rearranging, we see that

$$\tilde{P}\left(T_{B(O,1)}\le s\right) \le \frac{1}{c_2}s + \frac{c_2-c_1}{c_2}.$$

To return to (6.3), note that the sequence $(\xi_i)_{i=1}^{tL(t)}$ in the sum there stochastically dominates a sequence of independent copies of $T_{B(O,1)}$ (since the above bounds hold $\mathbf{P}\times\mathbb{P}$-almost surely, we can work pointwise on the probability space so that the tree and the environment are fixed, so we are only considering randomness under $P(\cdot)$). Therefore, letting $\left(\tilde{\xi}_i\right)_{i=1}^{tL(t)}$ denote independent copies of $T_{B(O,1)}$ (conditional on our particular realisations of $\mathcal{T}_\gamma^\infty$ and $\phi^{(\alpha)}$), and recalling again that $t = e^{A_\Delta r^b}$, we have from [18, Lemma 3.14] that $\mathbf{P} \times \mathbb{P}$-almost surely there exist $c_3, c_4 \in (0, \infty)$ such that

$$\tilde{P}\left(\sum_{i=1}^{tL(t)-1} \xi_i \leq t\right) \leq \tilde{P}\left(\sum_{i=1}^{tL(t)-1} \tilde{\xi}_i \leq t\right) \leq \exp\left\{-tL(t)\left(c_3 - c_4 L(t)^{-\frac{1}{2}}\right)\right\} \leq e^{-c_3 e^{A_\Delta r^b}} \tag{6.6}$$

for all sufficiently large r. In particular, substituting the bounds in (6.4) and (6.6) back into (6.3) we have that, $\mathbf{P} \times \mathbb{P}$-almost surely,

$$\tilde{P}\left(\tau_r \leq e^{A_\Delta r^b}\right) \leq 2e^{-\frac{\varepsilon}{4} A_\Delta r^b} + e^{-c_3 e^{A_\Delta r^b}}$$

for all sufficiently large r. Therefore, applying Borel-Cantelli along the sequence of integer r, we deduce that $\mathbf{P} \times \mathbb{P} \times \tilde{P}$-almost surely,

$$T_{B(O,r(1+\varepsilon))} \geq e^{A_\Delta \lfloor r^b \rfloor} \geq e^{(1-\varepsilon) A_\Delta r^b}$$

for all sufficiently large r, or in other words,

$$d(O, X_t) \leq \frac{1+\varepsilon}{A_\Delta^{1/b}(1-\varepsilon)^{1/b}} (\log t)^{\frac{1}{b}}$$

for all sufficiently large t. Since ε was arbitrary, this proves the upper bound. To prove the lower bound, i.e. that

$$\limsup_{t \to \infty} \frac{d(O, X_t)}{(\log t)^{\frac{1}{b}}} \geq A_\Delta^{-1/b},$$

we take $\varepsilon > 0$ and set $f(t) = A_\Delta^{-1/b}(1-\varepsilon)(\log t)^{\frac{1}{b}}$. Then, again using [52, Equations (4.5), (4.6) and (4.7)] and Markov's inequality, we have that

$$\begin{aligned}\tilde{P}_O(T_{B(O,f(t))} \geq t) \leq \tilde{E}_O[T_{B(O,f(t))}]t^{-1} &\leq R_\phi(O, B(O, f(t))^c)\nu_\phi(B(O, f(t)))t^{-1} \\ &\leq R_\phi(O, B(O, f(t))^c)\nu_\phi(\mathcal{T}_\gamma^\infty)t^{-1}.\end{aligned}$$

By Proposition 6.1, $\nu_\phi(\mathcal{T}_\gamma^\infty)$ is almost surely finite. Now choose some $\delta = \delta(\varepsilon) > 0$ small enough that $(1+\delta)(1-\varepsilon)^b < 1$. By Propositions 6.1 and 6.2, we therefore have that, almost surely for all sufficiently large t,

$$\tilde{P}_O(T_{B(O,f(t))} \geq t) \leq \nu_\phi(\mathcal{T}_\gamma^\infty)e^{(1+\delta)A_\Delta f(t)^b}t^{-1} = \nu_\phi(\mathcal{T}_\gamma^\infty)e^{((1+\delta)(1-\varepsilon)^b - 1)\log t},$$

which vanishes as $t \to \infty$. Applying this along with Fatou's Lemma, we therefore deduce that

$$\tilde{P}_O\left(\frac{d(O, X_t)}{(\log t)^{1/b}} > A_\Delta^{-1/b}(1-\varepsilon) \text{ i.o.}\right) = \tilde{P}_O\big(T_{B(O,f(t))} < t \text{ i.o.}\big) \tag{6.7}$$

$$\geq \limsup_{t \to \infty} \tilde{P}_O\big(T_{B(O,f(t))} < t\big) = 1. \tag{6.8}$$

Since ε was arbitrary, this completes the proof. □

7 Reinforced critical regime: $\alpha = 1, \Delta > 0$

In this section we prove Theorem 1.4. The strategy is the same as in Section 6, but due to the logarithmic factors we obtain different estimates.

In contrast to the previous section, this time

$$R_\phi(x,y) = \int_{[x,y]} e^{B_\Delta \phi^{(1)}(z) + A_\Delta \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z),$$
$$\nu_\phi(A) = \int_A e^{-\left[B_\Delta \phi^{(1)}(z) + A_\Delta \log(d(O,z)+1)\right]} \mu(\mathrm{d}x),$$

where $B_\Delta = \sqrt{4\Delta}$ and $A_\Delta = \Delta - 1$.

7.1 Volume and resistance growth in $\mathcal{T}_{\gamma,0}$

Again it is natural to start with some volume and resistance estimates.

Proposition 7.1. $\mathbf{P} \times \mathbb{P}$*-almost surely, for any* $\varepsilon > 0$ *we have that*

$$B_\phi(O, r^{A_\Delta+1} e^{-(\log(r+1))^{\frac{1}{2}+\varepsilon}}) \subset B(O,r) \subset B_\phi(O, r^{A_\Delta+1} e^{(\log(r+1))^{\frac{1}{2}+\varepsilon}})$$

for all sufficiently large r.

Proof. Take some $\beta \in (\frac{1}{2},1), \beta' \in (\frac{1}{2}, \beta)$ and some $\varepsilon > 0$. By Proposition 5.10(ii), there $\mathbf{P} \times \mathbb{P}$-almost surely exists $R < \infty$ such that $(\log R)^\beta < \frac{\varepsilon}{2} \log R$ and

$$\sup_{x \in B(O,R)^c} \frac{B_\Delta |\phi^{(1)}(x)|}{(\log(d(O,x)+1)^{\beta'}} \leq 1.$$

Take some $r > 2R+1$. We start by proving the first inclusion by showing the reverse inclusion for the complements. Indeed, assume that $d(O,x) > r$, and let y be the point on the path from O to x such that $d(O,y) = R$, and $d(x,y) = d(O,x) - R$. Note that this almost surely exists and is unique. Moreover, $d(O,z) > R$ for all $z \in [y,x]$ by our choice of R, which implies that $\log(d(O,z)+1) - (\log(d(O,z)+1))^\beta \geq (1 - \frac{\varepsilon}{2}) \log(d(O,z)+1)$ for all such z.

Similarly to Theorem 6.1, provided that r is sufficiently large (depending only on R and Δ) we have that

$$\begin{aligned} R_\phi(O,x) &\geq \int_{[y,x]} e^{B_\Delta \phi^{(1)}(z) + A_\Delta \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \\ &\geq e^{-(\log(r+1))^{\beta'}} \int_{[y,x]} e^{A_\Delta \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \geq e^{-(\log(r+1))^\beta} r^{A_\Delta+1}. \end{aligned}$$

For the second inclusion, we keep R, y, β and β' as above, and now assume that $d(O,x) < r$. Also set

$$\begin{aligned} \tilde{r}_c = \inf \Bigg\{ r \geq R : &\sup_{w \in \partial B(O,R)} \int_{[O,w]} e^{B_\Delta \phi^{(1)}(z) + A_\Delta \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \\ &\leq \left(1 - \frac{1}{A_\Delta + 1}\right) r^{A_\Delta+1} e^{(\log(r+1))^{\frac{1}{2}}} \Bigg\}. \end{aligned}$$

(This is always possible since $\phi^{(1)}$ is continuous with bounded sample paths, so for any fixed $R > 0$, $\phi^{(1)}$ is bounded on the compact set $B(O,R)$, and so the supremum is almost surely finite.)

We then have for all sufficiently large (depending on R and Δ) r and all $x \in B(O,r)$ that

$$\begin{aligned} R_\phi(O,x) &\leq \sup_{w\in\partial B(O,R)} \int_{[O,w]} e^{B_\Delta \phi^{(1)}(z)+A_\Delta \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \\ &\quad + \int_{[y,x]} e^{B_\Delta \phi^{(1)}(z)+A_\Delta \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \\ &\leq \left(1-\frac{1}{A_\Delta+1}\right) r^{A_\Delta+1} e^{(\log(r+1))^{\frac{1}{2}}} + e^{(\log(r+1))^{\beta'}} \int_{[y,x]} e^{A_\Delta \log(d(O,z)+1)} \lambda^\gamma(\mathrm{d}z) \\ &\leq r^{A_\Delta+1} e^{(\log(r+1))^{\beta}}, \end{aligned}$$

which completes the proof. □

Proposition 7.2. *For $\alpha = 1$, we have for any $\varepsilon > 0$ that $\mathbf{P}\times\mathbb{P}$-almost surely,*

$$re^{-(\log r)^{\frac{1}{2}+\varepsilon}} \leq R_\phi(O, B_\phi(O,r)^c) \leq r$$

for all sufficiently large r.

Proof. The proof is similar to that of Theorem 6.2. Fix $\varepsilon > 0$, and let $M_\phi(r)$ be the smallest cardinality of a set of points in

$$B_\phi(O,r) \setminus B_\phi(O, re^{-3(\log r)^{\frac{1}{2}+\varepsilon}}),$$

such that any path passing from O to $B_\phi(O,r)^c$ must pass through one of the points in the set. Since any such set is a cutset, it follows from the parallel law for resistance that

$$R_\phi(O, B_\phi(O,r)^c) \geq \frac{re^{-3(\log r)^{\frac{1}{2}+\varepsilon}}}{M_\phi(r)}.$$

It follows from Proposition 7.1 that $M_\phi(r) \leq M(r)$ almost surely for all sufficiently large r, where $M(r)$ is the smallest cardinality of a set of points separating

$$B\left(O, r^{\frac{1}{A_\Delta+1}} e^{-2(\log(r+1))^{\frac{1}{2}+\varepsilon}}\right) \text{ from } B\left(O, r^{\frac{1}{A_\Delta+1}} e^{-(\log(r+1))^{\frac{1}{2}+\varepsilon}}\right)^c,$$

almost surely for all sufficiently large r. Let $(L^{(s)})_{s\geq 0}$ be as in (5.1). The same arguments that led to (6.2) entail that

$$\begin{aligned} &\mathbf{P}(M(r) > 1) \\ &\leq \mathbf{P}\left(L^{(r^{\frac{1}{A_\Delta+1}} e^{-2(\log(r+1))^{\frac{1}{2}+\varepsilon}})} > \left(r^{\frac{1}{A_\Delta+1}} e^{-2(\log(r+1))^{\frac{1}{2}+\varepsilon}}\right)^{\frac{1}{\gamma-1}} e^{\frac{k}{(\gamma-1)}(\log(r+1))^{\frac{1}{2}+\varepsilon}}\right) \\ &+ \mathbf{P}\left(\text{Poisson}\left(c_\gamma e^{\frac{-(\log(r+1))^{\frac{1}{2}+\varepsilon}}{2(\gamma-1)}}\right) \geq 1\right) \\ &\leq c'_\gamma e^{-(\log r)^{\frac{1}{2}+\varepsilon}/2(\gamma-1)}. \end{aligned}$$

We therefore deduce from Borel-Cantelli that if we take $r_n = 2^n$, then $\mathbf{P}\times\mathbb{P}$-almost surely, we have that

$$R_\phi(O, B_\phi(O,r_n)^c) \geq r_n e^{-3(\log r_n)^{\frac{1}{2}+\varepsilon}}$$

for all sufficiently large n. This then implies that $R_\phi(O, B_\phi(O,r)^c) \geq re^{-(\log r)^{\frac{1}{2}+2\varepsilon}}$ for all sufficiently large r by monotonicity. Since ε was arbitrary and the upper bound $R_\phi(O, B_\phi(O,r)^c) \leq r$ holds trivially, this is enough to prove the result. □

Proposition 7.3. *Take any $\beta > \frac{1}{2}$. Then $\mathbf{P} \times \mathbb{P}$-almost surely,*

- *If $A_\Delta \leq \frac{\gamma}{\gamma-1}$, then $e^{-(\log(r+1))^\beta} r^{\frac{\gamma}{\gamma-1}-A_\Delta} \leq \nu_\phi(B(O,r)) \leq r^{\frac{\gamma}{\gamma-1}-A_\Delta} e^{(\log(r+1))^\beta}$ for all sufficiently large r.*
- *If $A_\Delta > \frac{\gamma}{\gamma-1}$ then $\nu_\phi(\mathcal{T}_\gamma^\infty) < \infty$.*

Proof. Again the proof is similar to that of Theorem 6.3. We treat the case $A_\Delta < \frac{\gamma}{\gamma-1}$ first, starting with the upper bound. Take some $\beta \in (\frac{1}{2}, 1)$. By Proposition 5.10$(ii)$ and since $\phi^{(1)}$ is continuous, we can $\mathbf{P} \times \mathbb{P}$-almost surely choose $R \geq 1$ and $\varepsilon, p < \infty$ so that

- $\sup_{x \in B(O,R)} \left\{ e^{-B_\Delta \phi^{(1)}(x) - A_\Delta \log(d(O,x)+1)} \right\} \leq p$,
- $\sup_{x \in B(O,R)^c} \frac{B_\Delta |\phi^{(1)}(x)|}{(\log(d(O,x)+1)^\beta} < \varepsilon$.

We then have from Proposition 5.5 that for any $\delta > 0$, almost surely for all sufficiently large r,

$$\begin{aligned}
\nu_\phi(B(O,r)) &\leq \sup_{x \in B(O,R)} \left\{ e^{-B_\Delta \phi^{(1)}(x) - A_\Delta \log(d(O,x)+1)} \right\} \mu\big(B(O,R)\big) \\
&\qquad + \sum_{n=R}^{r} \sup_{x \in B(O,n) \setminus B(O,n-1)} \left\{ e^{-B_\Delta \phi^{(1)}(x) - A_\Delta \log(d(O,x)+1)} \right\} \mu\big(B(O,n)\big) \\
&\leq p\mu\big(B(O,R)\big) + \sum_{n=R}^{r} n^{-A_\Delta} e^{\varepsilon (\log n)^\beta} n^{\frac{1}{\gamma-1}} (\log n)^{\frac{1+\varepsilon}{\gamma-1}} \\
&\leq C_{p,R,\varepsilon} + r^{\frac{\gamma}{\gamma-1} - A_\Delta} e^{(\log r)^{\beta+\delta}}.
\end{aligned} \tag{7.1}$$

Since $\beta, \delta > 0$ were arbitrary this implies the upper bound. If $A_\Delta = \frac{\gamma}{\gamma-1}$ we obtain a factor of $\log r$ from the sum, but this is dominated by the $e^{(\log r)^{\beta+\delta}}$ term, so the result still holds.

The lower bound is simpler since

$$\begin{aligned}
\nu_\phi(B(O,r)) &\geq \inf_{x \in B(O,r)} \left\{ e^{-B_\Delta \phi^{(1)}(x) - A_\Delta \log(d(O,x)+1)} \right\} \mu(B(O,r)) \\
&\geq (r+1)^{-A_\Delta} e^{-\varepsilon(\log(r+1))^\beta} \frac{r^{\frac{\gamma}{\gamma-1}}}{\log r}
\end{aligned}$$

almost surely for all sufficiently large r by Theorem 5.3. This proves the result since β was arbitrary. Clearly the result holds trivially if $A = \frac{\gamma}{\gamma-1}$.

If $A_\Delta > \frac{\gamma}{\gamma-1}$ the calculation in (7.1) shows that $\nu_\phi(\mathcal{T}_\gamma^\infty) < \infty$ almost surely. □

7.2 Proof of Theorem 1.4

We now have all the tools to prove Theorem 1.4.

Proof of Theorem 1.4. As in the proof of Theorem 1.2, we assume that $\mathcal{T}_\gamma^c$ and $\phi^{(1)}$ are fixed and therefore just write $\tilde{P}$ or $\tilde{P}_O$ in place of $\tilde{P}_{O,\phi}$.

Take $\beta' \in (\frac{1}{2}, 1)$ and assume first that $A_\Delta \leq \frac{\gamma}{\gamma-1}$. Letting $g_B(\cdot,\cdot)$ denote the Green's function for the diffusion killed on exiting the ball $B(O,r)$ we firstly have from Propositions 7.1 and 7.3 that, $\mathbf{P} \times \mathbb{P}$-almost surely for all sufficiently large r,

$$\begin{aligned}
\tilde{E}_O\left[T_{B(O,r)}\right] = \int_{B(O,r)} g_B(O,y) \nu_\phi(dy) &\leq \int_{B(O,r)} R_\phi(O, B(O,r)^c) \nu_\phi(dy) \\
&\leq r^{\frac{2\gamma-1}{\gamma-1}} e^{2(\log(r+1))^{\beta'}}.
\end{aligned}$$

Therefore, for any $\beta > \frac{1}{2}$ we can choose $\beta' \in (\frac{1}{2}, \beta)$ so that Markov's inequality gives

$$\tilde{P}_O\Big(T_{B(O,r)} \geq r^{\frac{2\gamma-1}{\gamma-1}} e^{(\log(r+1))^\beta}\Big) \leq e^{-(\log(r+1))^\beta/2}$$

for all sufficiently large r. In particular, Fatou's Lemma gives that

$$\tilde{P}_O\Big(d(O, X_t) \geq t^{\frac{\gamma-1}{2\gamma-1}} e^{-(\log(t+1))^\beta} \text{ i.o.}\Big) \geq \limsup_{r\to\infty} \tilde{P}_O\Big(T_{B(O,r)} \leq r^{\frac{2\gamma-1}{\gamma-1}} e^{(\log(r+1))^\beta}\Big) = 1.$$

If instead $A_\Delta > \frac{\gamma}{\gamma-1}$, letting $V = \nu_\phi(\mathcal{T}_\gamma^\infty)$ the same calculation gives

$$\tilde{E}_O\big[T_{B(O,r)}\big] \leq V r^{A_\Delta+1} e^{(\log(r+1))^{\beta'}},$$

so we similarly deduce from Markov's inequality and Fatou's Lemma that

$$\tilde{P}_O\Big(d(O, X_t) \geq t^{\frac{1}{A_\Delta+1}} e^{-(\log(t+1))^\beta} \text{ i.o.}\Big) = 1.$$

For the upper bound, we use [31, Lemma 4.2(b)] which states that for any $t > 0$ and any $\delta \in (0, R_\phi(O, B_\phi(O,r)^c))$,

$$\tilde{P}_O\big(T_{B_\phi(O,r)} \leq t\big) \leq 4\left[\frac{\delta}{R_\phi(O, B_\phi(O,r)^c)} + \frac{t}{\nu_\phi\left(B_\phi(O,\delta)\right)\left(R_\phi(O, B_\phi(O,r)^c) - \delta\right)}\right]. \quad (7.2)$$

Set $\delta = re^{-(\log r)^\beta}$ for some $\beta > \frac{1}{2}$, and choose $\beta' \in (\frac{1}{2}, \beta)$. Note that

- $B_\phi(O,r) \subset B(O, r^{\frac{1}{A_\Delta+1}} e^{(\log(r+1))^{\beta'}})$ eventually almost surely by Proposition 7.1,
- $R_\phi(O, B_\phi(O,r)^c) \geq re^{-(\log r)^{\beta'}}$ eventually almost surely by Proposition 7.2,
- If $A_\Delta \leq \frac{\gamma}{\gamma-1}$, then almost surely, $\nu_\phi(B_\phi(O,\delta)) \geq \delta^{\frac{1}{A_\Delta+1}\left(\frac{\gamma}{\gamma-1} - A_\Delta\right)} e^{-(\log(\delta+1))^\beta}$ for all sufficiently large δ by Propositions 7.1 and 7.3. If $A_\Delta \geq \frac{\gamma}{\gamma-1}$, then there exists (a random) $V > 0$ such that $\nu_\phi(B_\phi(O,\delta)) \geq V$ for all $\delta \geq 1$.

In the case $A_\Delta \leq \frac{\gamma}{\gamma-1}$, we take some $\eta > 0$ and set $t = r^{\frac{2\gamma-1}{(\gamma-1)(A_\Delta+1)}} e^{-(\log r)^{\beta+2\eta}}$. We then deduce from (7.2) that, on the almost sure events above and provided r is sufficiently large,

$$\begin{aligned}
&\tilde{P}_O\left(T_{B\left(O, r^{\frac{1}{A_\Delta+1}} e^{(\log(r+1))^\beta}\right)} \leq t\right) \\
&\leq \tilde{P}_O\big(T_{B_\phi(O,r)} \leq t\big) \\
&\leq 4\left[\frac{re^{-(\log r)^\beta}}{re^{-(\log r)^{\beta'}}} + \frac{r^{\frac{2\gamma-1}{(\gamma-1)(A_\Delta+1)}} e^{-(\log r)^{\beta+2\eta}}}{r^{\frac{\gamma}{(\gamma-1)(A_\Delta+1)} - \frac{A_\Delta}{A_\Delta+1}} e^{-(\log r)^{\beta+\eta}} re^{-(\log r)^{\beta'}}}\right] \leq 8e^{-(\log r)^{\beta'}}.
\end{aligned}$$

In particular, if $r_n = 2^n$ we have from Borel-Cantelli that

$$\tilde{P}_O\left(T_{B\left(O, r_n^{\frac{1}{A_\Delta+1}} e^{(\log(r_n+1))^\beta}\right)} \leq r_n^{\frac{2\gamma-1}{(\gamma-1)(A_\Delta+1)}} e^{-(\log r_n)^{\beta+2\eta}} \text{ i.o.}\right) = 0.$$

If $r \in [r_n, r_{n+1}]$, this implies that

$$T_{B\left(O, r^{\frac{1}{A_\Delta+1}} e^{(\log(r+1))^\beta}\right)} \geq T_{B\left(O, r_n^{\frac{1}{A_\Delta+1}} e^{(\log(r_n+1))^\beta}\right)} \geq r^{\frac{2\gamma-1}{(\gamma-1)(A_\Delta+1)}} e^{-(\log r)^{\beta+3\eta}}$$

for all sufficiently large r. By inverting this relation we deduce the result. If instead $A_\Delta \geq \frac{\gamma}{\gamma-1}$ we instead take $t = re^{-(\log r)^{\beta+2\eta}}$ and (7.2) instead gives that

$$\begin{aligned}\tilde{P}_O\left(T_{B\left(O, r^{\frac{1}{A_\Delta+1}} e^{(\log(r+1))^\beta}\right)} \leq re^{-(\log r)^{\beta+2\eta}}\right) &\leq \tilde{P}_O\left(T_{B_\phi(O,r)} \leq re^{-(\log r)^{\beta+2\eta}}\right) \\ &\leq 4\left[\frac{re^{-(\log r)^\beta}}{re^{-(\log r)^{\beta'}}} + \frac{re^{-(\log r)^{\beta+2\eta}}}{Ve^{-(\log r)^{\beta+\eta}} re^{-(\log r)^{\beta'}}}\right] \\ &\leq C_V e^{-(\log r)^{\beta'}},\end{aligned}$$

where $C_V > 0$ is a (random) constant that depends on V. The final result can be deduced from Borel-Cantelli exactly as above. □

8 Reinforced transient regime: $\alpha > 1, \Delta > 0$

In the transient regime resistance does not provide the right framework to characterise the scaling limit of LERRW, since the resistance between pairs of points in the correspondence collapses to zero under any rescaling. However, we can still use resistance when working directly with the unrescaled process to understand its exit times from a ball of radius r.

Accordingly, we let $(Y_t)_{t\geq 0}$ denote a constant speed continuous-time LERRW on T_∞ as given by Definition 2.3 with offspring distribution satisfying (2.1) (the final result transfers directly to a discrete-time LERRW by the strong law of large numbers applied to the time index). For convenience, we assume that the function a_n appearing in (2.1) is of the form $a_n = cn^{1/\gamma}$; however one can also incorporate a slowly-varying correction and "pull it through" in all the proofs that follow. More precisely, $(Y_t)_{t\geq 0}$ has the annealed law of a continuous-time random walk in the Dirichlet random environment of (3.7) on the infinite tree T_∞ with $\mathsf{exp}(1)$ holding time at each vertex and initial weights given by

$$\alpha_{\{\overleftarrow{x}, x\}} = d_{T_\infty}(O_\infty, x)^\alpha \tag{8.1}$$

(so we just take $\alpha_{\{\overleftarrow{x},x\}}$ instead of $\alpha^{(n)}_{\{\overleftarrow{x},x\}}$ in (3.7)). We will work primarily with the RWRE in this section and denote the law of the environment by $\mathbb{P}$. The discrete setup of Section 4.2 still holds, so that for $y \in T_\infty$, cf. (3.8), we can write

$$V(y) = \sum_{O_\infty \prec v \preceq y} \log \rho_v, \qquad R(\overleftarrow{y}, y) = e^{V(y)}, \qquad \nu(y) = e^{-V(y)} \mathbb{1}\{y \neq O_\infty\} + \sum_{z:\, \overleftarrow{z} = y} e^{-V(z)}. \tag{8.2}$$

Let $O_\infty = b_0, b_1, \ldots$ denote the backbone vertices of T_∞ (these are the special vertices of Definition 2.3), ordered by their distance from the root. Given $i \geq 1$, also let T^i denote the (infinite) subtree of T_∞ rooted at b_i, and for any integer $r > 0$ let $B_{T^i}(b_i, r)$ denote the ball of radius r around b_i *in this subtree*, defined with respect to the graph metric. We use the notation B^R to denote a ball defined with respect to the metric R defined above.

We start with a brief lemma on the structure of $T_\infty \setminus T^r$.

Proposition 8.1. $\mathbf{P}$*-almost surely, for any* $\varepsilon > 0$*,*

$$\limsup_{r\to\infty} \frac{\mathit{Diam}(T_\infty \setminus T^r)}{r(\log r)^{\gamma+\varepsilon}} < \infty.$$

Proof. Let T denote an *unconditioned* Galton-Watson tree with the same offspring distribution as T_∞. By conditioning on the heights of the subtrees grafted to the infinite backbone, it follows from Definition 2.3, [10, Lemma A.2], [68, Theorem 2] and a union bound that there exists $c < \infty$ such that for any $p > 0, r \geq 1, \lambda \geq 1$,

$$\mathbf{P}(\mathsf{Diam}(T_\infty \setminus T^r) \geq 2r\lambda) \leq \mathbf{P}\left(\sum_{v \prec b_r} \deg v \geq r^{\frac{1}{\gamma-1}}\lambda^p\right) + r^{\frac{1}{\gamma-1}}\lambda^p \mathbf{P}(\mathsf{Height}(T) \geq r\lambda)$$
$$\leq c\lambda^{-p(\gamma-1)} + cr^{\frac{1}{\gamma-1}}\lambda^p (r\lambda)^{-\frac{1}{\gamma-1}}.$$

Taking $p = \frac{1}{\gamma(\gamma-1)}$ gives an upper bound of order $\lambda^{-\frac{1}{\gamma}}$. Now set $\hat{\varepsilon} = \frac{\varepsilon}{2}$, apply Borel-Cantelli with $\lambda_r = (\log r)^{\gamma+\hat{\varepsilon}}$ along the subsequence $r_n = 2^n$, then use monotonicity and divide through by $(\log r)^{\hat{\varepsilon}}$ for the result. □

The difference with the recurrent regime is that typical terms of the form $\log \rho_v$ will now be negative. In particular, we show in the Appendix in Lemmas A.4 and A.5 that, as $d_{T_\infty}(O_\infty, y) \to \infty$,

$$\mathbb{E}\left[\sum_{O_\infty \prec v \preceq y} \log \rho_v\right] = -\alpha \log(d_{T_\infty}(O_\infty, y)) + O(1), \quad \mathsf{Var}\left(\sum_{O_\infty \prec v \preceq y} \log \rho_v\right) = O(1). \quad (8.3)$$

We start with the following bound on the behaviour of the potential on T_∞.

Lemma 8.2. *For any $\delta > 0$, $\mathbf{P}$-almost surely,*

(i)

$$\mathbb{P}\left(\left|\sum_{O_\infty \prec v \preceq b_n} \log \rho_v + \alpha \log(d_{T_\infty}(O_\infty, b_n))\right| \geq (\log d_{T_\infty}(O_\infty, b_n))^{\frac{1}{2}+\delta} \textit{ i.o.}\right) = 0.$$

(ii) As $m \to \infty$,

$$\inf_{y: b_m \preceq y} \mathbb{P}\left(\left|\sum_{O_\infty \prec u \preceq v} \log \rho_u + \alpha \log(d_{T_\infty}(O_\infty, v))\right|\right.$$
$$\left. < (\log d_{T_\infty}(O_\infty, v))^{\frac{1}{2}+\delta} \textit{ for all } b_m \preceq v \prec y\right) \to 1.$$

Proof. (i) For $y \in T_\infty$, let $W_y = \sum_{O_\infty \prec v \preceq y} \log \rho_v + \alpha \log(d_{T_\infty}(O_\infty, y))$. It follows from Chebyshev's inequality and the bounds of (8.3) that there exists a constant c such that for any $y \in T_\infty$ and any $A > 0$,

$$\mathbb{P}\left(|W_y| \geq (\log d_{T_\infty}(O_\infty, y))^A\right) \leq c(\log d_{T_\infty}(O_\infty, y))^{-2A}. \quad (8.4)$$

Applying Borel-Cantelli, we therefore deduce that for any $\delta > 0$,

$$\mathbb{P}\left(|W_{b_{2^m}}| \geq (\log d_{T_\infty}(O_\infty, b_{2^m}))^{\frac{1+2\delta}{2}} \text{ i.o.}\right) = 0. \quad (8.5)$$

Moreover, if $r \in [2^m, 2^{m+1}]$ and $\delta < 1$, then

$$|(\log d_{T_\infty}(O_\infty, b_{2^{m+1}}))^{\frac{1+\delta}{2}} - (\log d_{T_\infty}(O_\infty, b_r))^{\frac{1+\delta}{2}}| \leq \log 2,$$

so that

$$\mathbb{P}\left(|W_{b_{2^{m+1}}}| \geq (\log(d_{T_\infty}(O_\infty, b_{2^{m+1}})))^{\frac{1+\delta}{2}} \;\middle|\; |W_{b_r}| \geq (\log(d_{T_\infty}(O_\infty, b_r)))^{\frac{1+2\delta}{2}}\right)$$
$$\geq \mathbb{P}\left(\sum_{b_r \prec v \preceq b_{2^{m+1}}} \log \rho_v \geq -\frac{1}{2}(\log(d_{T_\infty}(O_\infty, b_r)))^{\frac{1+2\delta}{2}}\right) \to 1,$$

as $m \to \infty$ by (8.3) and Chebyshev's inequality. It therefore follows that

$$\frac{1}{2}\mathbb{P}\Big(|W_{b_r}| \geq (\log(d_{T_\infty}(O_\infty, b_r)))^{\frac{1+2\delta}{2}} \text{ i.o.}\Big)$$
$$\leq \mathbb{P}\Big(\big|W_{b_{2^{m+1}}}\big| \geq (\log(d_{T_\infty}(O_\infty, b_{2^{m+1}})))^{\frac{1+\delta}{2}} \text{ i.o.}\Big) = 0.$$

The last equality holds from (8.5). This proves (i) since $\delta > 0$ was arbitrary.

(ii) We first prove (ii) when y is on the backbone. First take $m \geq 1$ and note that

$$\mathbb{P}\left(\left|\sum_{O_\infty \prec v \preceq b_k} \log \rho_v + \alpha \log(d_{T_\infty}(O_\infty, b_k))\right| < (\log d_{T_\infty}(O_\infty, b_k))^{\frac{1}{2}+\delta} \text{ for all } k \geq m\right) \to 1,$$

as $m \to \infty$ by part (i). Now note that if $y \in T_\infty$ with $b_k \preceq y$ for some $k \geq m$ and y is not on the backbone, then it follows from the Dirichlet construction of Section 3.2 that $R(b_k, y)$ and $(\sum_{u \prec v} \log \rho_u)_{v \in [b_m, y]}$ are distributed as $R(b_k, b_{d_{T_\infty}(O_\infty, y)})$ and $(\sum_{u \prec v} \log \rho_u)_{v \in [b_m, b_{d_{T_\infty}(O_\infty, y)}]}$ respectively. Part (ii) follows. □

Recall that our aim in this section is to prove Theorem 1.5. Our strategy to do this will be to divide the ball $B(O_\infty, r)$ up into smaller sets which must all be traversed by the RWRE of (3.7) before it can exit $B(O_\infty, r)$. We then apply a well-known chaining technique to bound the sum of all the exit times from these smaller sets. We then transfer this back to the LERRW using (3.9).

Before doing this, we give some definitions of good and typical vertices and establish some of their properties.

Definition 8.3. *Fix some $\delta > 0$ and then fix some $A' > A > \frac{1}{2} + \delta$. To ease notation in the rest of this section, we view these parameters as fixed and do not record them as indices.*

1. *Given $r, \lambda > 1$, we define a sequence $(r_i)_{i=0}^\infty$ by $r_i = r_i(r, \lambda) = \frac{r}{2}(1 + i\lambda^{-1})$ and set $\tilde{r} = \tilde{r}(r, \lambda) = \frac{r}{2} e^{-(\log r)^{A'}} \lambda^{-1}$.*
2. *We say that a vertex $y \in T_\infty$ is **m-good** if $b_m \preceq y$ and*

$$\left|\sum_{O_\infty \prec u \preceq v} \log \rho_u + \alpha \log(d_{T_\infty}(O_\infty, v))\right| < (\log d_{T_\infty}(O_\infty, v))^{\frac{1}{2}+\delta} \qquad \text{for all } b_m \preceq v \prec y,$$

*and that y is **m-bad** otherwise.*

3. *For $c, C, m \in [0, \infty), r > 2m$ and $\lambda > 1$, we say that an index $0 \leq i < \lambda$ is (m, r, c, C, λ)-**typical** if*

$$\frac{c}{2} r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}} \leq |\{y \in B_{T^{r_i}}(b_{r_i}, \tilde{r}) : y \text{ m-good}\}| < C r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}}.$$

We will use the fact that for typical indices i, we can control (in probability) the time for the RWRE to exit a ball of the form $B_{T^{r_i}}(b_{r_i}, \tilde{r})$. Moreover, it is highly likely that most indices are typical. We make this precise in the following lemma.

Lemma 8.4. *Take $A' > A$ as in Definition 8.3, and $A'' > A'$.*

(i) *$\mathbf{P} \times \mathbb{P}$-almost surely, there exist deterministic $C_\alpha > 0$, $\tilde{C}_\alpha < \infty$ such that*

$$C_\alpha \lambda^{-1} r^{1-\alpha} e^{-(\log r)^A} \leq R(b_{r_i}, b_{r_{i+1}}) \leq \tilde{C}_\alpha \lambda^{-1} r^{1-\alpha} e^{(\log r)^A},$$

for all $\lambda > 1$, all $1 \leq i < \lambda$ and all sufficiently large r.

(ii) Moreover, let B^R denote the ball with respect to the metric R defined in (8.2). Then for any $\varepsilon > 0$, it holds for all sufficiently large r, any $\lambda > 1$ and any $0 \leq i < \lambda$ that with $\mathbf{P} \times \mathbb{P}$-probability at least $1-\varepsilon$,

$$\nu\left(B_{T^{r_i}}(b_{r_i}, \tilde{r}) \cap B^R(b_{r_i}, \tilde{r}e^{(\log r)^A})\right) \geq r^\alpha e^{-(\log r)^A} r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A''}}.$$

(iii) Let $N = N(m,r,c,C,\lambda)$ denote the number of (m,r,c,C,λ)-typical i in $\{0,\ldots,\lambda-1\}$. Then for any $\varepsilon > 0$, there exist deterministic $c > 0, C < \infty, m < \infty$ and an event A_ε such that, $\mathbf{P}$-almost surely, $\mathbb{P}(A_\varepsilon^c) < \varepsilon$ and $\mathbb{P}\big(N \leq \frac{\lambda}{2}$ and $A_\varepsilon\big) \leq e^{-\frac{\lambda}{6}}$.

Proof. (i) To prove point (i), note from Lemma 8.2(i) that we almost surely have for all sufficiently large r and all $\lambda > 1$ that

$$\begin{aligned} R(b_{r_i}, b_{r_{i+1}}) = \sum_{b_{r_i} \prec y \preceq b_{r_{i+1}}} e^{V(y)} &\geq \sum_{b_{r_i} \prec y \preceq b_{r_{i+1}}} e^{-\alpha \log(d_{T_\infty}(O_\infty, y)) - (\log d_{T_\infty}(O_\infty, y))^A} \\ &\geq \frac{{r_i}^{1-\alpha} - r_{i+1}^{1-\alpha}}{\alpha - 1} e^{-(\log r_{i+1})^A} \\ &\geq C_\alpha r^{1-\alpha} \lambda^{-1} e^{-(\log r)^A}. \end{aligned} \tag{8.6}$$

The calculation for the upper bound follows similarly.

(ii) For the second point, take some $i \geq 1$ and some $y \in B_{T^{r_i}}(b_{r_i}, \tilde{r})$ and choose some $0 < \varepsilon \ll \frac{1}{96}$. Note that, if y is m-good, then similarly to (8.6), it holds that

$$\begin{aligned} R(b_{r_i}, y) &\leq \frac{r_i^{1-\alpha} - d_{T_\infty}(O_\infty, y)^{1-\alpha}}{\alpha - 1} e^{(\log d_{T_\infty}(O_\infty, y))^A} \\ &\leq C'_\alpha r^{1-\alpha} \lambda^{-1} e^{-(\log r)^{A'}} e^{(\log d_{T_\infty}(O_\infty, y))^A} \\ &\leq r^{1-\alpha} \lambda^{-1} e^{-(\log r)^{\frac{1}{2}(A+A')}}, \\ \nu(y) &\geq e^{-V(y)} \geq d_{T_\infty}(O_\infty, y)^\alpha e^{-(\log d_{T_\infty}(O_\infty, y))^{\frac{1}{2}+\delta}} \geq r^\alpha e^{-(\log r)^A}. \end{aligned} \tag{8.7}$$

Therefore, by Lemma 8.2(ii), which says that y is m-good whp provided $d_{T_\infty}(O_\infty, y)$ and m are sufficiently large, we $\mathbf{P}$-almost surely have that for all sufficiently large m, if we fix some $r > 2m, i \geq 1$ and $y \in B_{T^{r_i}}(b_{r_i}, \tilde{r})$, then

$$\begin{aligned} &\mathbb{P}\Bigg(\Bigg| \sum_{O_\infty \prec u \preceq v} \log \rho_u + \alpha \log(d_{T_\infty}(O_\infty, v))\Bigg| \\ &\qquad \geq (\log d_{T_\infty}(O_\infty, v))^{\frac{1}{2}+\delta} \text{ for some } b_m \preceq v \prec y\Bigg) &\leq \varepsilon, \\ &\mathbb{P}\Big(R(b_{r_i}, y) \geq r^{1-\alpha}\lambda^{-1} e^{-(\log r)^{\frac{1}{2}(A+A')}}\Big) &\leq \varepsilon \\ &\mathbb{P}\Big(\nu(y) \leq r^\alpha e^{-(\log r)^A}\Big) &\leq \varepsilon. \end{aligned}$$

Now fix m large enough satisfying the above, and take any $r > 2m$. By Markov's inequality and the inequalities above, it $\mathbf{P}$-almost surely holds for all sufficiently large r and any $i \geq 1$ that

$$\mathbb{P}\left(|\{y \in B_{T^{r_i}}(b_{r_i}, \tilde{r}) : y \text{ is m-bad }\}| \geq \frac{1}{2}|B_{T^{r_i}}(b_{r_i}, \tilde{r})|\right) \leq 6\varepsilon. \tag{8.8}$$

Moreover, since $B_{T^{r_i}}(b_{r_i}, \tilde{r}) \overset{(d)}{=} B(O_\infty, \tilde{r})$ and $A'' > A'$, we have from [32, Lemmas 2.5 and 2.7] that

$$\mathbb{P}\Big(|B_{T^{r_i}}(b_{r_i}, \tilde{r})| \notin \Big[r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A''}}, r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A''}}\Big]\Big) \leq \varepsilon.$$

Therefore the result follows by a union bound over the relevant events above.

(iii) As above, we have from [32, Lemmas 2.5 and 2.7] that there exist deterministic $c > 0$, $C < \infty$ such that, for any $r > 1, \lambda > 1, i \geq 1$,

$$\begin{aligned}
&\mathbb{P}\Big(|B_{T^{r_i}}(b_{r_i}, \tilde{r})| \notin \Big[c r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}}, C r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}}\Big]\Big) \\
&\leq \mathbb{P}\Big(|B_{T^{r_i}}(b_{r_i}, \tilde{r})| \notin \Big[2^{-\frac{\gamma}{\gamma-1}} c \tilde{r}^{\frac{\gamma}{\gamma-1}}, 2^{\frac{\gamma}{\gamma-1}} C \tilde{r}^{\frac{\gamma}{\gamma-1}}\Big]\Big) \leq \frac{1}{8} - 6\varepsilon.
\end{aligned} \tag{8.9}$$

We deduce from a union bound over the events in (8.8) and (8.9) that, provided m is sufficiently large,

$$\mathbb{P}(i \text{ is not } (m, r, c, C, \lambda)\text{-typical }) \leq \frac{1}{8}.$$

These events are not independent for distinct i. However, by Lemma 8.2 we can choose m so that the probability appearing in Lemma 8.2(ii) is at least $1 - \varepsilon$, and define the event A_ε to be the corresponding high probability event. Then on the event A_ε we can control $V(b_{r_i})$ and $R(O_\infty, b_{r_i})$ for all i (for sufficiently large r), conditionally on which the tail bounds above hold independently for each i, so if $N = |\{i : i \text{ is } (m, r, c, C, \lambda)\text{-typical }\}|$, then N stochastically dominates a $\mathsf{Bin}(2\lambda, \frac{7}{8})$ random variable. In particular this establishes the claim. □

In the proof of Theorem 1.5 we assume that $\varepsilon > 0$ has been fixed and c, C are as in Lemma 8.4(iii).

Proof of Theorem 1.5. We proceed in three steps:

1. Firstly we couple T_∞ and its Dirichlet weights with a modified model $\hat{T}_\infty = \hat{T}_\infty(r, \lambda)$ such that (using a hat to denote analogous quantities in this model), letting $\hat{\tau}_i$ denote the time to hit $b_{r_{i+1}}$ starting from b_{r_i}, each $\hat{\tau}_i$ is stochastically dominated by its analogous quantity in T_∞. We then show that, with probability at least $1 - \varepsilon$, we can obtain comparable upper and lower bounds for $\mathbb{E}[\hat{\tau}_i]$ in the modified model for all typical indices i. We use these bounds to show that there $\mathbf{P}$-almost surely exists a constant $\kappa_\alpha \in (0, \infty)$ and an event A_ε such that $\mathbb{P}(A_\varepsilon^c) < \varepsilon$ (the same event as in Lemma 8.4(iii)) and such that, provided $r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}} \geq r\lambda^{-1}$, we have $\mathbf{P}$-almost surely have for all $0 \leq i < \lambda$ that
$$\tilde{P}_{b_{r_i}, \omega}(\hat{\tau}_i \leq s) \mathbb{1}\{A_\varepsilon\} \leq 1 - \kappa_\alpha e^{-\frac{3\gamma-2}{\gamma-1}(\log r)^{A'}} + \kappa_\alpha \frac{s}{r^{\frac{2\gamma-1}{\gamma-1}} \lambda^{\frac{-(2\gamma-1)}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}}} \tag{8.10}$$
for each (m, r, c, C, λ)-typical i.
2. We then apply a chaining result of [18] to deduce that, $\mathbf{P}$-almost surely,
$$\tilde{P}_\omega\left(\sum_{i=1}^N \hat{\tau}_i < t, N \geq \frac{\lambda}{2}\right) \mathbb{1}\{A_\varepsilon\} \leq \exp\left\{-e^{(1-\frac{\delta}{4})(\log r)^{A'+\delta}}\right\}.$$
3. Using the stochastic domination, we then transfer this result to the analogous sum $\sum_{i=1}^N \tau_i$ on the original tree. Using this tail bound, we bound the hitting time of b_r with a Borel-Cantelli argument.

Step 1. Fix $\varepsilon > 0$ and $\delta > 0$, and choose m' large enough that the probability in Lemma 8.2 is at least $1-\varepsilon$. Let A_ε be the corresponding high probability event. For any $r > 2m'$, we can proceed as follows.

Set $m = \lfloor \frac{r}{2} \rfloor$. The bound of Lemma 8.2$(ii)$ still holds since this can only increase the value of m, and for the rest of the proof we work on the event A_ε, which in particular implies that

$$\left| \sum_{O_\infty \prec u \preceq v} \log \rho_u + \alpha \log(d_{T_\infty}(O_\infty, v)) \right| < (\log d_{T_\infty}(O_\infty, v))^{\frac{1}{2}+\delta} \text{ for all } b_m \preceq v \prec b_r.$$

Given $r > 1, \lambda > 1$ and T_∞, we define $\hat{T}_\infty = \hat{T}_\infty(r, \lambda)$ from T_∞ as follows:

- First remove all m-bad vertices and their incident edges from T_∞.
- Then also remove all vertices in $\bigcup_{i=0}^{\lambda-1} (T^{r_i} \setminus T^{r_{i+1}}) \setminus \left(\bigcup_{i=0}^{\lambda-1} B_{T^{r_i}}(b_{r_i}, \tilde{r}) \cup \bigcup_{j=r_i}^{r_{i+1}} b_j \right)$.

(For the vertices and edges that are retained in $\hat{T}_\infty$, they retain the measures and resistances that they previously had in T_∞). The resulting structure is $\hat{T}_\infty$. Note that the definition of m-good and Lemma 8.2 ensures that $\hat{T}_\infty$ is really a connected tree for all sufficiently large r.

We use hat notation to denote all analogous quantities in $\hat{T}_\infty$. For each $i < \lambda$, we consider the subtree $\hat{T}^{r_i}$ (rooted at b_{r_i}) and let $\hat{\tau}_i$ denote the exit time from $\hat{T}^{r_i} \setminus \hat{T}^{r_{i+1}}$ *in the subtree* $\hat{T}^{r_i}$ for a random walk started at b_{r_i} (this means that the random walk can only exit $\hat{T}^{r_i} \setminus \hat{T}^{r_{i+1}}$ through $b_{r_{i+1}}$; since the random walk eventually has to pass through b_{r_i} to exit $\hat{T}_\infty \setminus \hat{T}^{r_i}$ and further excursions back towards the root and in subtrees can only slow the random walk down, it is clearly sufficient to restrict to this set, and clearly $\hat{\tau}_i$ is stochastically dominated by τ_i for all i).

If r is sufficiently large and $y \in B_{\hat{T}^{r_i}}(b_{r_i}, \tilde{r})$ is m-good, then by (8.7) and Lemma 8.4(i), we have that

$$\frac{R(b_{r_i}, y)}{R(b_{r_i}, b_{r_{i+1}})} \leq \frac{r^{1-\alpha}\lambda^{-1}e^{-(\log r)^{\frac{1}{2}(A+A')}}}{C_\alpha r^{1-\alpha}\lambda^{-1}e^{-(\log r)^A}} \leq \frac{1}{2}, \tag{8.11}$$

since $A' > A$. Now let $\hat{A} = \{y \in B_{T^{r_i}}(b_{r_i}, \tilde{r}) : y \ m-\text{good}\}$, let $\hat{B} = \hat{T}^{r_i} \setminus \hat{T}^{r_{i+1}}$ and let $g_{\hat{A}}(\cdot,\cdot)$, $g_{\hat{B}}(\cdot,\cdot)$ respectively denote the Green's function for the diffusion killed on exiting $\hat{A}$ and $\hat{B}$ (e.g. as in [52, Equation (4.7)]). It then follows as in [52, Equation (4.8)] (also using (8.11)) that

$$\frac{g_{\hat{B}}(b_{r_i}, y)}{g_{\hat{B}}(b_{r_i}, b_{r_i})} \geq 1 - \sqrt{2}.$$

Recall that $\hat{\tau}_i$ is the exit time of $\hat{B}$. Consequently, since $g_{\hat{B}}(b_{r_i}, b_{r_i}) = R(b_{r_i}, b_{r_{i+1}})$ by [52, Equation (4.5)], if i is (m, r, c, C, λ)-typical, we have from Definition 8.3, Lemma 8.4(i) and (8.7) that

$$\begin{aligned}
\tilde{E}_{b_{r_i},\omega}[\hat{\tau}_i]\mathbb{1}\{A_\varepsilon\} &\geq \sum_{y \in \hat{A}} g_{\hat{B}}(b_{r_i}, y)\nu(y) \\
&\geq (1-\sqrt{2}) \sum_{y \in \hat{A}} R(b_{r_i}, b_{r_{i+1}})\nu(y) \\
&\geq (1-\sqrt{2})\frac{C}{2} r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}} \cdot C_\alpha r^{1-\alpha}\lambda^{-1}e^{-(\log r)^A} \cdot r^\alpha e^{-(\log r)^A} \\
&\geq r^{\frac{2\gamma-1}{\gamma-1}} \lambda^{\frac{-(2\gamma-1)}{\gamma-1}} e^{-\frac{3\gamma-2}{\gamma-1}(\log r)^{A'}}.
\end{aligned} \tag{8.12}$$

We would like to obtain a similar upper bound. First note that for any $z \in \cup_{i=0}^{\lambda-1} B_{\hat{T}^{r_i}}(b_{r_i}, \tilde{r})$, it also follows from (8.11) and Lemma 8.4(i) that

$$R(z, b_{r_{i+1}}) \leq \frac{3}{2} R(b_{r_i}, b_{r_{i+1}}) \leq \frac{3}{2} \tilde{C}_\alpha r^{1-\alpha} \lambda^{-1} e^{(\log r)^A}.$$

Similarly if $z \in \cup_{j=r_i}^{r_{i+1}} b_j$, we have that $R(z, b_{r_{i+1}}) \leq R(b_{r_i}, b_{r_{i+1}})$ and therefore the same upper bound holds. Therefore, since $g_{\hat{B}}(z,y) \leq g_{\hat{B}}(y,y)$ for all such z and $g_{\hat{B}}(y,y) = R(y, b_{r_{i+1}})$ by [52, Equation (4.6)] and by [52, Equation (4.5)] respectively, if i is (m,r,c,C,λ)-typical we can use Definition 8.3 and (8.7) to deduce that there exists a constant $K_\alpha < \infty$ such that, whenever $r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}} \geq r\lambda^{-1}$,

$$\begin{aligned}
\tilde{E}_{z,\omega}[\hat{\tau}_i] \, \mathbb{1}\{A_\varepsilon\} &= \sum_{y \in \hat{B}} g_{\hat{B}}(z,y)\nu(y) \\
&\leq \sum_{y \in \hat{B}} g_{\hat{B}}(y,y)\nu(y) \\
&\leq (|\hat{A}| + (r_{i+1} - r_i)) \sup_{y \in \hat{B}} \{R(y, b_{r_{i+1}}) \cdot \nu(y) : y \in \hat{A}\} \\
&\leq C(r^{\frac{\gamma}{\gamma-1}} \lambda^{\frac{-\gamma}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}} + r\lambda^{-1}) \cdot \frac{3}{2} \tilde{C}_\alpha r^{1-\alpha} \lambda^{-1} e^{(\log r)^A} \cdot r^\alpha e^{(\log r)^A} \\
&\leq K_\alpha r^{\frac{2\gamma-1}{\gamma-1}} \lambda^{\frac{-(2\gamma-1)}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}},
\end{aligned} \tag{8.13}$$

By the arguments of [52, Lemma 4.2] and applying (8.12) and (8.13), we therefore have on the event A_ε that for any $s > 0$,

$$\begin{aligned}
r^{\frac{2\gamma-1}{\gamma-1}} \lambda^{\frac{-(2\gamma-1)}{\gamma-1}} e^{-\frac{3\gamma-2}{\gamma-1}(\log r)^{A'}} &\leq \tilde{E}_{b_{r_i},\omega}[\hat{\tau}_i] \leq s + \tilde{P}_{b_{r_i},\omega}(\hat{\tau}_i > s) \sup_{z \in B_{T^{r_i}}(b_{r_i}, \tilde{r})} \tilde{E}_{z,\omega}[\hat{\tau}_i] \\
&\leq s + \tilde{P}_{b_{r_i},\omega}(\hat{\tau}_i > s) \, K_\alpha r^{\frac{2\gamma-1}{\gamma-1}} \lambda^{\frac{-(2\gamma-1)}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}}.
\end{aligned}$$

Rearranging we deduce (8.10). **Step 2.** Since (8.10) holds for all (m,r,c,C,λ)-typical indices i, and recalling that $N = N(m,r,c,C,\lambda)$ denotes the number of (m,r,c,C,λ)-typical indices, we therefore have from [18, Lemma 3.14] that on the event A_ε, it holds for all $t > 0$ that

$$\begin{aligned}
\log\left(\tilde{P}_\omega\left(\sum_{i=1}^N \hat{\tau}_i \leq t\right)\right) &\leq 2\left(\frac{Nt\kappa_\alpha}{r^{\frac{2\gamma-1}{\gamma-1}} \lambda^{\frac{-(2\gamma-1)}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}} (1 - \kappa_\alpha e^{-\frac{3\gamma-2}{\gamma-1}(\log r)^{A'}})}\right)^{\frac{1}{2}} \\
&\quad - N \log\left((1 - \kappa_\alpha e^{-\frac{3\gamma-2}{\gamma-1}(\log r)^{A'}})^{-1}\right).
\end{aligned}$$

Taking $\lambda = e^{(\log r)^{A'+\delta}}$ (which satisfies the conditions required in step 1 provided that r is sufficiently large) and $t = r^{\frac{2\gamma-1}{\gamma-1}} \lambda^{\frac{-(2\gamma-1)}{\gamma-1}} e^{-\frac{\gamma}{\gamma-1}(\log r)^{A'}} (1 - e^{-\frac{3\gamma-2}{\gamma-1}(\log r)^{A'}}) \lambda^{1-\delta}$ gives that, for all sufficiently large r,

$$\begin{aligned}
&\tilde{P}_\omega\left(\sum_{i=1}^N \hat{\tau}_i < t, N \geq \frac{\lambda}{2}\right) \mathbb{1}\{A_\varepsilon\} \\
&\quad \leq \exp\left\{4\kappa_\alpha \lambda^{1-\frac{\delta}{2}} - \frac{\kappa_\alpha}{2} \lambda e^{-\frac{3\gamma-2}{\gamma-1}(\log r)^{A'}}\right\} \qquad \leq \exp\left\{-e^{(1-\frac{\delta}{4})(\log r)^{A'+\delta}}\right\}.
\end{aligned}$$

Consequently, combining with Lemma 8.4(iii) and a union bound we deduce that, on the event A_ε,

$$\tilde{P}_\omega\left(\sum_{i=1}^{N}\hat{\tau}_i < t\right) \leq \tilde{P}_\omega\left(N < \frac{\lambda}{2}\right) + \tilde{P}_\omega\left(\sum_{i=1}^{N}\hat{\tau}_i < t, N \geq \frac{\lambda}{2}\right)$$
$$\leq \exp\left\{\frac{-e^{(\log r)^{A'+\delta}}}{12}\right\} + \exp\left\{-e^{(1-\frac{\delta}{4})(\log r)^{A'+\delta}}\right\}.$$

Step 3. Since adding back the m-bad vertices and then adding the times to traverse balls corresponding to non-typical i can only increase the total exit time, the same tail bound clearly holds for the exit time of $T_\infty \setminus T^{r_i}$. By Proposition 8.1, we can also choose R large enough that the event

$$B_\varepsilon := \left\{\sup_{r\geq R}\frac{\mathsf{Diam}(T_\infty \setminus T^r)}{r(\log r)^{\gamma+\varepsilon}} < 1\right\}$$

satisfies $\mathbf{P}(B_\varepsilon) \geq 1-\varepsilon$.

We deduce from steps 1 and 2 that, on the events A_ε and B_ε,

$$\tilde{P}_\omega\left(T_{B(O_\infty, r(\log r)^{\gamma+\varepsilon})} \leq r^{\frac{2\gamma}{2\gamma-1}}e^{-(\log r)^{A'+2\delta}}\right) \leq \tilde{P}_\omega\left(\sum_{i=1}^{N}\hat{\tau}_i \leq r^{\frac{2\gamma}{2\gamma-1}}e^{-(\log r)^{A'+2\delta}}\right)$$
$$\leq 2e^{-e^{(1-\frac{\delta}{4})(\log r)^{A'+\delta}}}.$$

Consequently, it follows from Borel-Cantelli along the subsequence $r_n = 2^n$ and monotonicity of T_r that almost surely,

$$d_{T_\infty}\left(O_\infty, Y_{r^{\frac{2\gamma}{2\gamma-1}}e^{-(\log r)^{A'+2\delta}}}\right)\mathbb{1}\{A_\varepsilon, B_\varepsilon\} \leq r(\log r)^{\gamma+\varepsilon},$$

for all sufficiently large r. Since A_ε and B_ε both have $\mathbf{P}\times\mathbb{P}$-probability at least $1-\varepsilon$, this implies that with overall probability at least $1-2\varepsilon$, we have that

$$d_{T_\infty}(O_\infty, Y_t) \leq t^{\frac{2\gamma-1}{2\gamma}}e^{(\log t)^{A'+3\delta}}$$

for all sufficiently large t. Since $A' > \frac{1}{2}$, $\delta > 0$ and $\varepsilon > 0$ were arbitrary, this implies the result for the annealed law $\tilde{P}$, and therefore for the LERRW. □

Remark 8.5 (Application to LERRW on $\mathbb{Z}_+$)**.** We can also apply these results to LERRW on the infinite half line with these initial weights. Note that Lemma 8.2(i) also applies in this setting, to give that, for any $\delta > 0$,

$$\mathbb{P}\left(\left|\sum_{0<i\leq r}\log\rho_i + \alpha\log r\right| \geq (\log r)^{\frac{1}{2}+\delta} \text{ i.o.}\right) = 0.$$

In particular, this implies that $\mathbf{P}\times\mathbb{P}$-almost surely, there exist constants $c>0, C<\infty$ such that for all sufficiently large r,

$$r^{\alpha+1}e^{-(\log r)^{\frac{1}{2}+\delta}} \leq \nu([0,r]) \leq r^{\alpha+1}e^{(\log r)^{\frac{1}{2}+\delta}}$$
$$c \leq R(0,[0,r]^c) \leq C.$$

Moreover, again using [52, Equation (4.5)], we deduce that, if T_r is the exit time from $[0,r]$, then on these events it holds that

$$cr^{\alpha+1}e^{-(\log r)^{\frac{1}{2}+\delta}} \leq \sum_{y\in[0,\frac{r}{2}]}R(y,[0,r]^c)\nu(y) \leq \tilde{E}_{0,\omega}[T_r] \leq \sum_{y\in[0,r]}R(0,[0,r]^c)\nu(y)$$
$$\leq Cr^{\alpha+1}e^{(\log r)^{\frac{1}{2}+\delta}}.$$

By Markov's inequality, Borel-Cantelli along the sequence $r_n = 2^n$ and monotonicity we therefore get that $T_r \leq r^2 e^{(\log r)^{\frac{1}{\alpha+1}+\delta}}$ eventually almost surely, or in other words that $\sup_{s \leq t} d(O_\infty, Y_s) \geq t^{\frac{1}{2}} e^{-(\log t)^{\frac{1}{\alpha+1}+\delta}}$ eventually almost surely.

A Appendix

In the appendix we prove Claims 4.3, 4.4 and 4.5, regarding the expectation, variance and moment generating function of the random variables $(\log \rho_x)_{x \in T_n}$.

In the appendix we always work conditionally on T_n. In particular, we will work pointwise on the probability space $(\mathbf{\Omega}, \mathcal{F}, \mathbf{P})$ (on which we defined (T_n, d_n, μ_n)), and on which the convergence of (2.3) holds almost surely. This means that most of the statements that follow should really be written conditionally on T_n. To make the arguments clearer to follow, we have not written this explicitly in the statements or the proofs, and instead ask the reader to keep this in mind throughout.

Moreover, a statement of the form $a_n = a + o(1)$ almost surely on Ω means that $a_n \to a$ almost surely and therefore that the $o(1)$ term is not necessarily bounded uniformly on Ω. However, it will always be true that for any $\varepsilon > 0$, the $o(1)$ term can be bounded uniformly on a set of probability at least $1 - \varepsilon$. The same holds for $O(\cdot)$ terms.

Recall that we simplified notation by writing

$$\Delta_n = \begin{cases} \Delta (n a_n^{-1})^{-(1-\alpha)}, & \text{if } \alpha < 1, \\ \Delta, & \text{if } \alpha = 1, \end{cases} \qquad |x| = d_n(O_n, x).$$

A.1 Properties of the digamma function

The mean and variance of terms of the form $\log \rho_x$ can be expressed in terms of the digamma function, defined for $z > 0$ by $\psi(z) = \Gamma'(z)/\Gamma(z)$, where Γ is the gamma function.

Note that, for each $n \geq 1$ and each $x \in T_n$, we have that $\rho_x = \frac{1-p_x}{p_x}$, where p_x has the beta distribution with some positive parameters (a_x, b_x). This entails that $\rho_x \sim \beta'(a_x, b_x)$ and that $\frac{1}{\rho_x} \sim \beta'(b_x, a_x)$, where β' is the beta prime distribution.

For a random variable ρ with a beta prime distribution with positive parameters (a, b), the following formulae were respectively established in [70, Equations (4.1) and (4.5)]:

$$\mathbb{E}[\log \rho] = \psi(b) - \psi(a), \qquad \text{Var}\,(\log \rho) = \psi'(a) + \psi'(b). \tag{A.1}$$

In our case, we have for each $n \geq 1$ and each $x \in T_n$, that

$$\rho_x \sim \beta' \left(\frac{\left((|x|-1) + n a_n^{-1}\right)^\alpha + \Delta_n}{2\Delta_n}, \frac{(|x| + n a_n^{-1})^\alpha}{2\Delta_n} \right).$$

and $(\rho_x)_{x \in T_n \setminus \{O_n\}}$ is a sequence of independent random variables. Next, we record three properties of the digamma function that we will use throughout this section.

- Proved in [69, Equation (6.6)]:

$$\psi\left(z + \frac{1}{2}\right) - \psi(z) \sim \frac{1}{2z} \text{ as } z \to \infty. \tag{A.2}$$

- Proved in [70, Equation (4.11)]: for $s, t > 0$,

$$-\frac{1}{t} \leq [\psi(t) - \psi(s)] - \log\left(\frac{t}{s}\right) \leq \frac{1}{s}. \tag{A.3}$$

- Proved in [70, Equation (4.12)] and [69, Equation (6.5)]:

$$\psi'(z) \sim \frac{1}{z} \text{ as } z \to \infty. \tag{A.4}$$

A.2 Claim 4.3: expectation of the potential when $\alpha \leq 1$

Lemma A.1. *(i) For all $x \in T_n$, we have that*

$$|\mathbb{E}[\log \rho_x]\,| \leq \frac{2\alpha}{\left(|x| + na_n^{-1}\right)} + \frac{4\Delta_n}{\left(|x| + na_n^{-1}\right)^{\alpha}}.$$

(ii) For almost every $\omega \in \mathbf{\Omega}$, it holds uniformly over $t \in [0,1]$ as $n \to \infty$ that:

$$\sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \mathbb{E}[\log \rho_x] \to \frac{\Delta[(d(O,t)+1)^{1-\alpha} - 1]}{1-\alpha} - \alpha \log\left(d(O,t)+1\right) \qquad \text{if } \alpha < 1,$$

$$\sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \mathbb{E}[\log \rho_x] \to (\Delta - 1)\log(d(O,t)+1), \qquad \text{if } \alpha = 1.$$

Proof. (i) By (A.1) and (A.3), we have that

$$\begin{aligned}|\mathbb{E}[\log \rho_x]\,| &= \left|\psi\left(\frac{(|x| + na_n^{-1})^{\alpha}}{2\Delta_n}\right) - \psi\left(\frac{\left((|x|-1) + na_n^{-1}\right)^{\alpha} + \Delta_n}{2\Delta_n}\right)\right| \\ &\leq \log\left(\left|\frac{(|x| + na_n^{-1})^{\alpha}}{\left((|x|-1) + na_n^{-1}\right)^{\alpha} + \Delta_n}\right|\right) + \frac{2\Delta_n}{\left((|x|-1) + na_n^{-1}\right)^{\alpha} + \Delta_n} \\ &\leq \frac{2\left(|x| + na_n^{-1}\right)^{\alpha} - \left((|x|-1) + na_n^{-1}\right)^{\alpha} + 4\Delta_n}{\left(|x| + na_n^{-1}\right)^{\alpha} + \Delta_n} \\ &\leq \frac{2\alpha}{\left(|x| + na_n^{-1}\right)} + \frac{4\Delta_n}{\left(|x| + na_n^{-1}\right)^{\alpha}}.\end{aligned}$$

(ii) First note that it follows from (A.1) that for all $t \in [0,1]$,

$$\sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \mathbb{E}[\log(\rho_x)] \tag{A.5}$$

$$\begin{aligned}&= \sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \left\{\psi\left(\frac{\left((|x|-1) + na_n^{-1}\right)^{\alpha} + \Delta_n}{2\Delta_n}\right) - \psi\left(\frac{(|x| + na_n^{-1})^{\alpha}}{2\Delta_n}\right)\right\} \\ &= \psi\left(\frac{\left(na_n^{-1}\right)^{\alpha} + \Delta_n}{2\Delta_n}\right) - \psi\left(\frac{(|x_{\lfloor 2nt \rfloor}| + na_n^{-1})^{\alpha}}{2\Delta_n}\right) \\ &\quad + \sum_{i=1}^{|x_{\lfloor 2nt \rfloor}|-1} \left\{\psi\left(\frac{\left(i + na_n^{-1}\right)^{\alpha}}{2\Delta_n} + \frac{1}{2}\right) - \psi\left(\frac{(i + na_n^{-1})^{\alpha}}{2\Delta_n}\right)\right\}. \end{aligned}\tag{A.6}$$

First, we assume that $\alpha \leq 1$. For the first term in (A.5), we use (A.3) and note that for almost every $\omega \in \Omega$,

$$\begin{aligned}\log\left(\frac{\left(na_n^{-1}\right)^{\alpha} + \Delta_n}{2\Delta_n}\right) - \log\left(\frac{(|x_{\lfloor 2nt \rfloor}| + na_n^{-1})^{\alpha}}{2\Delta_n}\right) &= \log\left(\frac{(na_n^{-1})^{\alpha} + \Delta_n}{(|x_{\lfloor 2nt \rfloor}| + na_n^{-1})^{\alpha}}\right) \\ &= -\alpha \log(d(O,t)+1) + o(1),\end{aligned}$$

as $n \to \infty$ by (2.8). Thus, by (A.3), for almost every $\omega \in \mathbf{\Omega}$ the first term in (A.5) satisfies

$$\left|\psi\left(\frac{\left(na_n^{-1}\right)^{\alpha} + \Delta_n}{2\Delta_n}\right) - \psi\left(\frac{(|x_{\lfloor 2nt \rfloor}| + na_n^{-1})^{\alpha}}{2\Delta_n}\right) + \alpha \log(d(O,t)+1)\right| \leq \frac{2\Delta_n}{(na_n^{-1})^{\alpha}} + o(1)$$

as $n \to \infty$, and therefore

$$\psi\left(\frac{\left(na_n^{-1}\right)^{\alpha}+\Delta_n}{2\Delta_n}\right)-\psi\left(\frac{(|x_{\lfloor 2nt\rfloor}|+na_n^{-1})^{\alpha}}{2\Delta_n}\right)=-\alpha\log(d(O,t)+1)+o(1),$$

as $n \to \infty$.

For the second term, since $\inf_{i\geq 0}\frac{(n^{-1}a_n+i)^{\alpha}}{2\Delta_n}\to\infty$ as $n\to\infty$, it follows from (A.2) that for all $i \geq 0$,

$$\psi\left(\frac{\left(i+na_n^{-1}\right)^{\alpha}}{2\Delta_n}+\frac{1}{2}\right)-\psi\left(\frac{(i+na_n^{-1})^{\alpha}}{2\Delta_n}\right)=\frac{\Delta_n}{(i+na_n^{-1})^{\alpha}}(1+o(1)),$$

where the $o(1)$ term is uniform over all $i \geq 0$ (but may depend on ω).

This implies that for almost every $\omega \in \mathbf{\Omega}$ it holds for all $t \in [0,1]$ that

$$\sum_{i=na_n^{-1}+1}^{|x_{\lfloor 2nt\rfloor}|-1+na_n^{-1}}\left\{\psi\left(\frac{i^{\alpha}}{2\Delta_n}+\frac{1}{2}\right)-\psi\left(\frac{i^{\alpha}}{2\Delta_n}\right)\right\}$$

$$\sim\sum_{i=na_n^{-1}+1}^{|x_{\lfloor 2nt\rfloor}|-1+na_n^{-1}}\frac{\Delta_n}{i^{\alpha}}$$

$$\sim\begin{cases}\dfrac{\Delta_n}{1-\alpha}\left[(|x_{\lfloor 2nt\rfloor}|-1+na_n^{-1})^{1-\alpha}-(na_n^{-1})^{1-\alpha}\right] & \text{if } \alpha<1,\\ \Delta\left[\log(|x_{\lfloor 2nt\rfloor}|-1+na_n^{-1})-\log(na_n^{-1})\right] & \text{if } \alpha=1,\end{cases}$$

$$\sim\begin{cases}\dfrac{\Delta}{1-\alpha}\left[\left[\dfrac{(|x_{\lfloor 2nt\rfloor}|-1)+na_n^{-1}}{na_n^{-1}}\right]^{1-\alpha}-1\right] & \text{if } \alpha<1,\\ \Delta\log(n^{-1}a_n|x_{\lfloor 2nt\rfloor}|-n^{-1}a_n+1) & \text{if } \alpha=1,\end{cases}$$

$$\sim\begin{cases}\dfrac{\Delta\left[(d(0,t)+1)^{1-\alpha}-1\right]}{1-\alpha} & \text{if } \alpha<1,\\ \Delta\log(d(0,t)+1) & \text{if } \alpha=1.\end{cases}$$

Since $\sup_{t\in[0,1]} d(0,t)$ is bounded for almost every $\omega \in \mathbf{\Omega}$, this implies the result. ☐

A.3 Claim 4.4: variance of the potential when $\alpha \leq 1$

Proposition A.2. *For almost every $\omega \in \mathbf{\Omega}$, it holds uniformly over $t \in [0,1]$ as $n \to \infty$ that:*

$$\sum_{O_n\prec x\preceq x_{\lfloor 2nt\rfloor}}\mathrm{Var}[\log(\rho_x)]\to\frac{4\Delta[(d(O,t)+1)^{1-\alpha}-1]}{1-\alpha}\qquad \text{if } \alpha<1,$$

$$\sum_{O_n\prec x\preceq x_{\lfloor 2nt\rfloor}}\mathrm{Var}[\log(\rho_x)]\to 4\Delta\log(d(O,t)+1),\qquad \text{if } \alpha=1.$$

Proof. It follows from (A.1) that for all $t \in [0,1]$,

$$\sum_{O_n\prec x\preceq x_{\lfloor 2nt\rfloor}}\mathrm{Var}(\log\rho_x)=\sum_{i=1}^{|x_{\lfloor 2nt\rfloor}|}\left\{\psi'\left(\frac{\left((i-1)+na_n^{-1}\right)^{\alpha}}{2\Delta_n}+\frac{1}{2}\right)+\psi'\left(\frac{(i+na_n^{-1})^{\alpha}}{2\Delta_n}\right)\right\}.$$

First we assume that $\alpha \leq 1$. Since both arguments in the derivatives of the digamma function blow up uniformly over $i \geq 0$ as $n \to \infty$, it follows from (A.4) that for all $i \geq 1$,

$$\psi'\left(\frac{(i+na_n^{-1})^\alpha}{2\Delta_n}\right) + \psi'\left(\frac{\left((i-1)+na_n^{-1}\right)^\alpha}{2\Delta_n} + \frac{1}{2}\right) = \frac{4\Delta_n}{(i+na_n^{-1})^\alpha}(1+o(1)),$$

as $n \to \infty$, where the $o(1)$ term is uniform over all $i \geq 1$. This implies that

$$\sum_{O_n \prec x \preceq x_{\lfloor 2nt \rfloor}} \mathrm{Var}(\log \rho_x) \sim \sum_{i=1+na_n^{-1}}^{|x_{\lfloor 2nt \rfloor}|+na_n^{-1}} \frac{4\Delta_n}{i^\alpha}$$

$$\sim \begin{cases} \frac{4\Delta_n}{1-\alpha}\left[(|x_{\lfloor 2nt \rfloor}| + na_n^{-1})^{1-\alpha} - (1+na_n^{-1})^{1-\alpha}\right] & \text{if } \alpha < 1, \\ 4\Delta\left[\log(|x_{\lfloor 2nt \rfloor}| + na_n^{-1}) - \log(1+na_n^{-1})\right] & \text{if } \alpha = 1, \end{cases}$$

$$\sim \begin{cases} \frac{4\Delta}{1-\alpha}\left[\left[\frac{|x_{\lfloor 2nt \rfloor}| + na_n^{-1}}{na_n^{-1}}\right]^{1-\alpha} - \left[\frac{1+na_n^{-1}}{na_n^{-1}}\right]^{1-\alpha}\right] & \text{if } \alpha < 1, \\ 4\Delta\left[\log(n^{-1}a_n|x_{\lfloor 2nt \rfloor}| + 1) - \log(n^{-1}a_n + 1)\right] & \text{if } \alpha = 1, \end{cases}$$

$$\sim \begin{cases} \frac{4\Delta\left[(d(0,t)+1)^{1-\alpha} - 1\right]}{1-\alpha} & \text{if } \alpha < 1, \\ 4\Delta\log(d(0,t)+1) & \text{if } \alpha = 1. \end{cases}$$

Again, since $\sup_{t\in[0,1]} d(0,t)$ is bounded for almost every $\omega \in \mathbf{\Omega}$, this implies the result. □

A.4 Claim 4.5: exponential moments when $\alpha \leq 1$

Lemma A.3. *For almost every $\omega \in \mathbf{\Omega}$, it holds for all $x \in T_n$ and all $1 \leq k \leq (na_n^{-1})^{1/2}$ that*

$$\mathbb{E}\left[e^{k\log\rho_x}\right] \leq \exp\left\{(1+3\Delta)\left(\frac{2k\alpha}{|x|+na_n^{-1}} + \frac{3k^2\Delta_n}{(|x|+na_n^{-1})^\alpha}\right)\right\},$$

$$\mathbb{E}\left[e^{k\log(\rho_x^{-1})}\right] \leq \exp\left\{(1+3\Delta)\left(\frac{2k\alpha}{|x|+na_n^{-1}} + \frac{3k^2\Delta_n}{(|x|+na_n^{-1})^\alpha}\right)\right\}.$$

Proof. First recall that, since

$$\rho_x \sim \beta'\left(\frac{(|x|-1+na_n^{-1})^\alpha + \Delta_n}{2\Delta_n}, \frac{(|x|+na_n^{-1})^\alpha}{2\Delta_n}\right),$$

we also have that

$$\frac{1}{\rho_x} \sim \beta'\left(\frac{(|x|+na_n^{-1})^\alpha}{2\Delta_n}, \frac{(|x|-1+na_n^{-1})^\alpha + \Delta_n}{2\Delta_n}\right).$$

Consequently (using the formula for the k^{th} moment of the beta prime distribution), it holds for any $k \geq 1$ that

$$\mathbb{E}\left[\rho_x^k\right] = \prod_{j=0}^{k-1} \frac{(|x|-1+na_n^{-1})^\alpha + (2j+1)\Delta_n}{(|x|+na_n^{-1})^\alpha - (2j+2)\Delta_n}, \tag{A.7}$$

$$\mathbb{E}\left[\rho_x^{-k}\right] = \prod_{j=0}^{k-1} \frac{(|x| + na_n^{-1})^\alpha + 2j\Delta_n}{(|x| - 1 + na_n^{-1})^\alpha - (2j+1)\Delta_n}. \tag{A.8}$$

Now note from (A.7) that

$$\begin{aligned}
&\mathbb{E}\left[e^{k \log \rho_x}\right] \\
&= \mathbb{E}\left[\rho_x^k\right] \\
&= \prod_{j=0}^{k-1} \frac{(|x| - 1 + na_n^{-1})^\alpha + (2j+1)\Delta_n}{(|x| + na_n^{-1})^\alpha - (2j+2)\Delta_n} \\
&\leq \prod_{j=0}^{k-1} \left(1 + \left[\frac{|(|x| - 1 + na_n^{-1})^\alpha - (|x| + na_n^{-1})^\alpha|}{(|x| + na_n^{-1})^\alpha} + \frac{(4j+3)\Delta_n}{(|x| + na_n^{-1})^\alpha}\right]\left[1 + \frac{2(2j+2)\Delta_n}{(|x| + na_n^{-1})^\alpha}\right]\right\} \\
&\leq \exp\left\{(1+3\Delta)\left(\frac{2k\alpha}{|x| + na_n^{-1}} + \frac{3k^2\Delta_n}{(|x| + na_n^{-1})^\alpha}\right)\right\}.
\end{aligned}$$

Similarly, note from (A.8) that

$$\begin{aligned}
&\mathbb{E}\left[e^{k \log(\rho_x^{-1})}\right] \\
&= \mathbb{E}\left[(\rho_x)^{-k}\right] \\
&= \prod_{j=0}^{k-1} \frac{(|x| + na_n^{-1})^\alpha + 2j\Delta_n}{(|x| - 1 + na_n^{-1})^\alpha - (2j+1)\Delta_n} \\
&\leq \prod_{j=0}^{k-1} \left(1 + \left[\frac{|(|x| + na_n^{-1})^\alpha - (|x| - 1 + na_n^{-1})^\alpha|}{(|x| - 1 + na_n^{-1})^\alpha} + \frac{(4j+1)\Delta_n}{(|x| - 1 + na_n^{-1})^\alpha}\right]\left[1 + \frac{2(2j+1)\Delta_n}{(|x| + na_n^{-1})^\alpha}\right]\right) \\
&\leq \exp\left\{(1+3\Delta)\left(\frac{2k\alpha}{|x| + na_n^{-1}} + \frac{3k^2\Delta_n}{(|x| + na_n^{-1})^\alpha}\right)\right\}.
\end{aligned}$$

□

A.5 Expectation of the potential when $\alpha > 1$

In the transient regime $\alpha > 1$ considered in Section 8, we do not rescale the initial weights so for each $x \in T_\infty$ we have

$$\rho_x \sim \beta'\left(\frac{((|x| - 1))^\alpha + \Delta}{2\Delta}, \frac{|x|^\alpha}{2\Delta}\right).$$

Lemma A.4. *Assume $\alpha > 1$. Then for almost every $\omega \in \mathbf{\Omega}$, the following holds. As $|y| \to \infty$, we have that*

$$\sum_{O_\infty \prec x \preceq y} \mathbb{E}[\log \rho_x] = -\alpha \log |y| + O(1).$$

Proof. First note that it follows from (A.1) that for any $y \in T_\infty$,

$$\begin{aligned}
\sum_{O_\infty \prec x \preceq y} \mathbb{E}[\log(\rho_x)] &= \sum_{O_\infty \prec x \preceq y} \left\{\psi\left(\frac{(|x| - 1)^\alpha + \Delta}{2\Delta}\right) - \psi\left(\frac{|x|^\alpha}{2\Delta}\right)\right\} \\
&= \psi\left(\frac{1}{2}\right) - \psi\left(\frac{|y|^\alpha}{2\Delta}\right) \\
&\quad + \sum_{i=1}^{|y|-1}\left\{\psi\left(\frac{i^\alpha}{2\Delta} + \frac{1}{2}\right) - \psi\left(\frac{i^\alpha}{2\Delta}\right)\right\}.
\end{aligned} \tag{A.9}$$

For the first term in (A.9), we use (A.3) which implies that, as $|y| \to \infty$,

$$\psi\left(\frac{1}{2}\right) - \psi\left(\frac{|y|^\alpha}{2\Delta}\right) = \log\left(\frac{\Delta}{|y|^\alpha}\right) + O(1) = -\alpha \log|y| + O(1).$$

For the second term, since $\frac{i^\alpha}{2\Delta} \to \infty$ as $i \to \infty$, it follows from (A.2) that as $i \to \infty$,

$$\psi\left(\frac{i^\alpha}{2\Delta} + \frac{1}{2}\right) - \psi\left(\frac{i^\alpha}{2\Delta}\right) = \frac{\Delta}{i^\alpha}(1 + o(1)).$$

This implies that

$$\sum_{i=1}^{|y|-1} \left\{\psi\left(\frac{i^\alpha}{2\Delta} + \frac{1}{2}\right) - \psi\left(\frac{i^\alpha}{2\Delta}\right)\right\} \sim \sum_{i=1}^{|y|-1} \frac{\Delta}{i^\alpha} = O(1)$$

as $|y| \to \infty$. □

A.6 Variance of the potential when $\alpha > 1$

Lemma A.5. *Assume $\alpha > 1$. Then for almost every $\omega \in \mathbf{\Omega}$, as $|y| \to \infty$ we have that*

$$\sum_{O_n \prec x \preceq y} \mathrm{Var}[\log(\rho_x)] = O(1).$$

Proof. It follows from (A.1) that for all $t \in [0, 1]$,

$$\sum_{O_\infty \prec x \preceq y} \mathrm{Var}(\log \rho_x) = \sum_{i=1}^{|y|} \left\{\psi'\left(\frac{(i-1)^\alpha}{2\Delta} + \frac{1}{2}\right) + \psi'\left(\frac{i^\alpha}{2\Delta}\right)\right\}.$$

Since both arguments in the derivatives of the digamma function blow up as $i \to \infty$, it follows from (A.4) that as $i \to \infty$,

$$\psi'\left(\frac{(i-1)^\alpha}{2\Delta} + \frac{1}{2}\right) + \psi'\left(\frac{i^\alpha}{2\Delta}\right) = \frac{4\Delta}{i^\alpha}(1 + o(1)).$$

This implies that

$$\sum_{O_\infty \prec x \preceq y} \mathrm{Var}(\log \rho_x) \sim \sum_{i=1}^{|y|} \frac{4\Delta}{i^\alpha} = O(1)$$

as $|y| \to \infty$. □

References

[1] R. Abraham, J-F. Delmas, and P. Hoscheit, *A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces*, Electron. J. Probab. **18(14)** (2013), 1–21. MR3035742

[2] E. Aidékon, *Transient random walks in random environment on a galton–watson tree*, Probability Theory and Related Fields **142** (2008), 525–559. MR2438700

[3] J. Akahori, A. Collevecchio, and M. Takei, *Phase transitions for edge-reinforced random walks on the half-line*, Electronic Communications in Probability **24** (2019). MR3978688

[4] D. Aldous, *Asymptotic fringe distributions for general families of random trees*, Ann. Appl. Probab. **1** (1991), no. 2, 228–266. MR1102319

[5] D. Aldous, *The continuum random tree. I*, Ann. Probab. **19** (1991), no. 1, 1–28. MR1085326

[6] D. Aldous, *The continuum random tree. II. An overview*, Stochastic analysis (Durham, 1990), London Math. Soc. Lecture Note Ser., vol. 167, Cambridge Univ. Press, Cambridge, 1991, pp. 23–70. MR1166406

[7] D. Aldous, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR1207226

[8] G. Andriopoulos, *Invariance principles for random walks in random environment on trees*, Electron. J. Probab. **26** (2021), Paper No. 115, 38. MR4309973

[9] E. Archer, *Infinite stable looptrees*, Electronic Journal of Probability **25** (2020). MR4059189

[10] E Archer, *Random walks on decorated Galton-Watson trees*, arXiv preprint arXiv:2011.07266 (2020).

[11] S. Athreya, M. Eckhoff, and A. Winter, *Brownian motion on* $\mathbb{R}$*-trees*, Transactions of the American Mathematical Society **365** (2013), no. 6, 3115–3150. MR3034461

[12] S. Athreya, W. Löhr, and A. Winter, *The gap between gromov-vague and gromov–hausdorff-vague topology*, Stochastic Processes and their Applications **126** (2016), no. 9, 2527–2553. MR3522292

[13] S. Athreya, W. Löhr, and A. Winter, *Invariance principle for variable speed random walks on trees*, Ann. Probab. **45** (2017), no. 2, 625–667. MR3630284

[14] S. Bacallado, *Bayesian analysis of variable-order, reversible Markov chains*, Ann. Statist. **39** (2011), no. 2, 838–864. MR2816340

[15] S. Bacallado, J. Chodera, and V. Pande, *Bayesian comparison of markov models of molecular dynamics with detailed balance constraint*, The Journal of chemical physics **131** (2009), no. 4, 07B622.

[16] S. Bacallado, V. Pande, S. Favaro, and L. Trippa, *Bayesian regularization of the length of memory in reversible sequences*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **78** (2016), no. 4, 933–946. MR3534357

[17] S. Bacallado, C. Sabot, and P. Tarrès, *The*-edge-reinforced random walk*, arXiv preprint arXiv:2102.08984 (2021). MR3420510

[18] M. Barlow, *Diffusions on fractals*, Lectures on probability theory and statistics (Saint-Flour, 1995), Lecture Notes in Math., vol. 1690, Springer, Berlin, 1998, pp. 1–121. MR1668115

[19] M. Barlow and T. Kumagai, *Random walk on the incipient infinite cluster on trees*, Illinois J. Math. **50** (2006), no. 1-4, 33–65. MR2247823

[20] R. Bass and K. Burdzy, *Stochastic bifurcation models*, Ann. Probab. **27** (1999), no. 1, 50–108. MR1681142

[21] J. Bertoin, *An extension of Pitman's theorem for spectrally positive Lévy processes*, Ann. Probab. **20** (1992), no. 3, 1464–1483. MR1175272

[22] J. Bertoin, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996. MR1406564

[23] P. Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999, A Wiley-Interscience Publication. MR1700749

[24] N. Bingham, C. Goldie, and J. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1989. MR1015093

[25] T. Brox, *A one-dimensional diffusion process in a Wiener medium*, Ann. Probab. **14** (1986), no. 4, 1206–1218. MR0866343

[26] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001. MR1835418

[27] P. Carmona, *The mean velocity of a brownian motion in a random lévy potential*, The Annals of Probability (1997), 1774–1788. MR1487435

[28] A. Collevecchio, *Limit theorems for reinforced random walks on certain trees*, Probab. Theory Related Fields **136** (2006), no. 1, 81–101. MR2240783

[29] D. Coppersmith and P. Diaconis, *Random walk with reinforcement*, Unpublished manuscript (1986).

[30] D. Croydon, *An introduction to stochastic processes associated with resistance forms and their scaling limits*, RIMS Kokyuroku 2030, no. 1 (2017). MR3865663

[31] D. Croydon, *Scaling limits of stochastic processes associated with resistance forms*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018), no. 4, 1939–1968. MR3865663

[32] D. Croydon and T. Kumagai, *Random walks on galton-watson trees with infinite variance offspring distribution conditioned to survive*, Electronic Journal of Probability **13** (2008), 1419–1441. MR2438812

[33] N. Curien and I. Kortchemski, *Random stable looptrees*, Electron. J. Probab. **19(108)** (2014), 1–35. MR3286462

[34] B. Davis and S. Volkov, *Continuous time vertex-reinforced jump processes*, Probability theory and related fields **123** (2002), no. 2, 281–300. MR1900324

[35] P. Diaconis and D. Freedman, *de finetti's theorem for markov chains*, The Annals of Probability (1980), 115–130. MR0556418

[36] P. Diaconis and S. Rolles, *Bayesian analysis for reversible Markov chains*, Ann. Statist. **34** (2006), no. 3, 1270–1292. MR2278358

[37] T. Duquesne, *A limit theorem for the contour process of conditioned Galton-Watson trees*, Ann. Probab. **31** (2003), no. 2, 996–1027. MR1964956

[38] T. Duquesne, *A limit theorem for the contour process of conditioned Galton-Watson trees*, Ann. Probab. **31** (2003), no. 2, 996–1027. MR1964956

[39] T. Duquesne, *Continuum random trees and branching processes with immigration*, Stochastic Process. Appl. **119** (2009), no. 1, 99–129. MR2485021

[40] T. Duquesne and J-F. Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque (2002), no. 281, vi+147. MR1954248

[41] T. Duquesne and J-F. Le Gall, *Probabilistic and fractal aspects of Lévy trees*, Probab. Theory Related Fields **131** (2005), no. 4, 553–603. MR2147221

[42] T. Duquesne and J-F. Le Gall, *The Hausdorff measure of stable trees*, ALEA Lat. Am. J. Probab. Math. Stat. **1** (2006), 393–415. MR2291942

[43] T. Duquesne and G. Wang, *Exceptionally small balls in stable trees*, arXiv preprint arXiv:1111.3465 (2011). MR3269345

[44] S. Evans, J. Pitman, and A. Winter, *Rayleigh processes, real trees, and root growth with re-grafting*, Probab. Theory Related Fields **134** (2006), no. 1, 81–126. MR2221786

[45] C. Goldschmidt and B. Haas, *Behavior near the extinction time in self-similar fragmentations. I. The stable case*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 2, 338–368. MR2667702

[46] P. Hall and C. Heyde, *Martingale limit theory and its application*, Academic press, 2014. MR0624435

[47] S. Janson and J-F. Marckert, *Convergence of discrete snakes*, J. Theoret. Probab. **18** (2005), no. 3, 615–647. MR2167644

[48] H. Kesten, *The limit distribution of Sinaĭ's random walk in random environment*, Phys. A **138** (1986), no. 1-2, 299–309. MR0865247

[49] H. Kesten, *Subdiffusive behavior of random walk on a random cluster*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 4, 425–487. MR0871905

[50] I. Kortchemski, *Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees*, Probability Theory and Related Fields **168** (2017), no. 1, 1–40. MR3651047

[51] I. Kortchemski and L. Richier, *Condensation in critical Cauchy Bienaymé–Galton–Watson trees*, The Annals of Applied Probability **29** (2019), no. 3, 1837–1877. MR3914558

[52] T. Kumagai, *Heat kernel estimates and parabolic Harnack inequalities on graphs and resistance forms*, Publ. Res. Inst. Math. Sci. **40** (2004), no. 3, 793–818. MR2074701

[53] J-F. Le Gall, *Random real trees*, Ann. Fac. Sci. Toulouse Math. (6) **15** (2006), no. 1, 35–62. MR2225746

[54] J-F. Le Gall and Y. Le Jan, *Branching processes in Lévy processes: the exploration process*, Ann. Probab. **26** (1998), no. 1, 213–252. MR1617047

[55] D. Levin, Y. Peres, and E. Wilmer, *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. MR3726904

[56] T. Lupu, C. Sabot, and P. Tarrès, *Fine mesh limit of the vrjp in dimension one and bass–burdzy flow*, Probability Theory and Related Fields **177** (2020), no. 1, 55–90. MR4095013

[57] C. Marzouk, *Scaling limits of discrete snakes with stable branching*, Annales de l’Institut Henri Poincaré, Probabilités et Statistiques **56** (2020), no. 1, 502–523. MR4058997

[58] Pertti Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, Cambridge, 1995, Fractals and rectifiability. MR1333890

[59] F. Merkl and S. Rolles, *A random environment for linearly edge-reinforced random walks on infinite graphs*, Probability theory and related fields **138** (2007), no. 1-2, 157–176. MR2288067

[60] J. Neveu, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 199–207. MR850756

[61] C. Pacheco, *From the sinai’s walk to the brox diffusion using bilinear forms*, arXiv preprint arXiv:1605.02826 (2016).

[62] R. Pemantle, *Phase transition in reinforced random walk and RWRE on trees*, Ann. Probab. **16** (1988), no. 3, 1229–1241. MR942765

[63] J. Pitman, *Combinatorial stochastic processes: Ecole d’eté de probabilités de saint-flour xxxii-2002*, Springer, 2006. MR2245368

[64] C. Sabot and P. Tarrès, *Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model*, J. Eur. Math. Soc. (JEMS) **17** (2015), no. 9, 2353–2378. MR3420510

[65] C. Sabot and L. Tournier, *Random walks in Dirichlet environment: an overview*, Ann. Fac. Sci. Toulouse Math. (6) **26** (2017), no. 2, 463–509. MR3640900

[66] P. Seignourel, *Discrete schemes for processes in random media*, Probability theory and related fields **118** (2000), no. 3, 293–322. MR1800534

[67] Y. Sinaĭ, *The limit behavior of a one-dimensional random walk in a random environment*, Teor. Veroyatnost. i Primenen. **27** (1982), no. 2, 247–258. MR657919

[68] R. Slack, *A branching process with mean one and possibly infinite variance*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **9** (1968), 139–145. MR228077

[69] M. Takei, *Almost sure behavior of linearly edge-reinforced random walks on the half-line*, Electron. J. Probab. **26** (2021), Paper No. 104, 18. MR4290506

[70] M. Takeshima, *Behavior of 1-dimensional reinforced random walk*, Osaka Journal of Mathematics **37** (2000), no. 2, 355–372. MR1772837

[71] P. Tetali, *Random walks and the effective resistance of networks*, J. Theoret. Probab. **4** (1991), no. 1, 101–109. MR1088395

[72] A. van der Vaart and J. Wellner, *Weak convergence and empirical processes*, Springer Series in Statistics, Springer-Verlag, New York, 1996, With applications to statistics. MR1385671

Acknowledgments. We would like to thank David Croydon for introducing us to this topic, and Andrea Collevecchio for useful comments on the manuscript.