

SCALING LIMITS FOR SIMPLE RANDOM WALKS ON RANDOM ORDERED GRAPH TREES

D. A. CROYDON,* *University of Warwick*

Abstract

Consider a family of random ordered graph trees $(T_n)_{n \geq 1}$, where T_n has n vertices. It has previously been established that if the associated search-depth processes converge to the normalised Brownian excursion when rescaled appropriately as $n \rightarrow \infty$, then the simple random walks on the graph trees have the Brownian motion on the Brownian continuum random tree as their scaling limit. Here, this result is extended to demonstrate the existence of a diffusion scaling limit whenever the volume measure on the limiting real tree is nonatomic, supported on the leaves of the limiting tree, and satisfies a polynomial lower bound for the volume of balls. Furthermore, as an application of this generalisation, it is established that the simple random walks on a family of Galton–Watson trees with a critical infinite variance offspring distribution, conditioned on the total number of offspring, can be rescaled to converge to the Brownian motion on a related α -stable tree.

Keywords: Branching process; random graph tree; random walk; scaling limit; α -stable tree

2010 Mathematics Subject Classification: Primary 60K37

Secondary 60G99; 60J15; 60J80

1. Introduction

If $(T_n)_{n \geq 1}$ is a family of random ordered graph trees, where T_n has n vertices for each n , and the rescaled graph trees $n^{-1/2}T_n$ converge suitably to the Brownian continuum random tree, then the associated simple random walks can be rescaled to converge to a diffusion limit, namely the Brownian motion on the Brownian continuum random tree; see [7, Theorem 1.2]. (Note that, in [7], the Brownian continuum random tree was referred to simply as the continuum random tree. The ‘Brownian’ is included here to make clear the distinction between this specific random real tree and the other random real trees that will feature in the discussion.) The collection of random graphs to which this result applies includes the case when T_n is a Galton–Watson tree with an offspring distribution, $(p_i)_{i \geq 0}$ say, that is, aperiodic (in particular, it is not supported on a proper subgroup of \mathbb{Z}), critical (mean 1), and has finite variance, conditioned on its total progeny being equal to n . Motivated by the problem of extending the methods of [7] to deal with the case when the finite variance assumption is dropped, in this paper we prove convergence results for the simple random walks on a much broader class of graph trees. Although the overall structure of the argument used here closely matches that of [7], in that paper several rather precise properties of the Brownian continuum random tree were applied, meaning that extensions to more general limiting trees could not be made immediately. Here, new techniques are introduced to allow us to remove some of the restrictions of [7], and in particular prove a

Received 24 April 2008; revision received 15 March 2010.

* Postal address: Department of Statistics, University of Warwick, Coventry CV4 7AL, UK.

Email address: d.a.croydon@warwick.ac.uk

scaling limit result for the simple random walks on critical Galton–Watson trees with infinite variance offspring distributions.

Denote the Brownian continuum random tree and its canonical measure by (\mathcal{T}, μ) (see [1] for background). Let ρ be a distinguished vertex of \mathcal{T} and call it the root of \mathcal{T} . Suppose that $\mathcal{T}(k)$ is the minimal subtree of \mathcal{T} spanning ρ and a k -sample of μ -random vertices; such sets can be constructed so that $\mathcal{T}(k) \subseteq \mathcal{T}(k + 1)$, and also we can assume that $\bigcup_k \mathcal{T}(k)$ is dense in \mathcal{T} , because μ has full support, almost surely. The set $\mathcal{T}(k)$ consists of a finite number of finite length line segments and, therefore, we can well define the one-dimensional Hausdorff measure on $\mathcal{T}(k)$ normalised to be a probability measure, $\lambda^{(k)}$ say. In [7], the fact that $\lambda^{(k)}$ converges weakly to μ as probability measures on \mathcal{T} was used in a time-change argument to show that certain Brownian motions on the subtrees $\mathcal{T}(k)$ converge to the Brownian motion on the Brownian continuum random tree (see [7, Lemma 3.1]). Whilst this assumption is known to hold for the Brownian continuum random tree, it has not been established for the related α -stable trees, $\alpha \in (1, 2)$, which are the scaling limits of infinite variance Galton–Watson trees (see [19, Section 4] for a brief introduction to α -stable trees and this convergence result). By considering alternative Brownian motions on $\mathcal{T}(k)$, we are able to show that the assumption that $\lambda^{(k)}$ converges to μ is redundant (Proposition 2.1 is the relevant convergence result). To proceed as we do, however, we are required to check that the time-change additive functionals we consider satisfy a tightness property (Lemma 2.5), and our proof of this relies on the technical result that the local times of the Brownian motion on the limiting tree diverge uniformly (Lemma 2.3).

A second fact about the Brownian continuum random tree applied in [7] is that it has no vertex of degree greater than three, by which we mean that the number of connected components of $\mathcal{T} \setminus \{\sigma\}$ is no greater than three for any $\sigma \in \mathcal{T}$. This property was important because it meant that the graph trees $T_n(k)$ constructed analogously to $\mathcal{T}(k)$ eventually had to have the same ‘tree shape’ as $\mathcal{T}(k)$ as n became large. However, α -stable trees with $\alpha \in (1, 2)$ admit infinite branch points (see [13, Theorem 4.6]), and so the same argument does not apply in general. To overcome this problem, in Section 3 we establish a method to demonstrate the convergence of Brownian motions on a sequence of finite-branched trees that converge in a specific way to the subtrees $\mathcal{T}(k)$ when the limiting tree \mathcal{T} has branch points of arbitrary degree. Our argument involves considering approximations to the processes of interest that jump over branch points, so that the precise geometry of the trees at branch points is not seen by the approximations.

Specifically, the limit space \mathbb{T}^* that we study in this paper is the collection of pairs (\mathcal{T}, μ) , where $\mathcal{T} = (\mathcal{T}, d_{\mathcal{T}})$ is a compact real tree (see [13, Definition 2.1] for example) and μ is a nonatomic Borel probability measure on \mathcal{T} that satisfies

$$\liminf_{r \rightarrow 0} \inf_{\sigma \in \mathcal{T}} \frac{\mu(B(\sigma, r))}{r^\kappa} > 0 \tag{1.1}$$

for some $\kappa > 0$, where $B(\sigma, r)$ is the open \mathcal{T} -ball of radius r (with respect to the metric $d_{\mathcal{T}}$) centred at $\sigma \in \mathcal{T}$. Furthermore, for $(\mathcal{T}, \mu) \in \mathbb{T}^*$, we assume that μ is supported on the leaves of \mathcal{T} , where by saying that $\sigma \in \mathcal{T}$ is a leaf of \mathcal{T} we mean that $\mathcal{T} \setminus \{\sigma\}$ is connected. We note that the Brownian continuum random tree is an element of \mathbb{T}^* almost surely, with (1.1) being satisfied whenever $\kappa > 2$; see [8, Theorem 1.2]. Furthermore, we will later check that all the above properties are also almost surely satisfied by α -stable trees for $\alpha \in (1, 2)$, with (1.1) holding for any choice of $\kappa > \alpha/(\alpha - 1)$.

The main result of this paper is the following theorem, which is stated within the framework of [7], so that ℓ^1 is the Banach space of infinite sequences of real numbers equipped with

the metric d_{ℓ^1} induced by the norm $\sum_{i \geq 1} |x(i)|$ for $x \in \ell^1$; $\mathcal{K}(\ell^1)$ is the space of compact subsets of ℓ^1 equipped with the Hausdorff topology; $\mathcal{M}_1(\ell^1)$ is the space of Borel probability measures on ℓ^1 equipped with the topology of weak convergence; and $\mathcal{M}_1(C([0, 1], \ell^1))$ is the space of Borel probability measures on $C([0, 1], \ell^1)$, also equipped with the topology of weak convergence. The set \mathcal{W} is the collection of continuous functions $w : [0, 1] \rightarrow \mathbb{R}_+$ such that $w(t) > 0$ if and only if $t \in (0, 1)$. For an excursion $w \in \mathcal{W}$, \mathcal{T}_w is the rooted real tree associated with w , μ_w is the natural measure on \mathcal{T}_w , and $\mathbb{P}_\rho^{\mathcal{T}_w, \mu_w}$ is the law of the Brownian motion on (\mathcal{T}_w, μ_w) started from the root $\rho = \rho(\mathcal{T}_w)$ (see the end of Section 2 for details). For an ordered graph tree T_n with n vertices, we write μ_n to represent the uniform probability measure on the vertices of T_n , and denote by $\mathbb{P}_\rho^{T_n}$ the law of the discrete-time simple random walk on T_n , started from the root, $\rho = \rho(T_n)$, of T_n . The search-depth process w_n of T_n is defined in Section 4.

Theorem 1.1. *Let $(\alpha_n)_{n \geq 1}$ be a positive divergent sequence such that $\alpha_n = o(n)$. If $(T_n)_{n \geq 1}$ is a sequence of ordered graph trees whose search-depth functions $(w_n)_{n \geq 1}$ satisfy*

$$\alpha_n^{-1} w_n \rightarrow w$$

in $C([0, 1], \mathbb{R}_+)$ for some $w \in \mathcal{W}$ with $(\mathcal{T}_w, \mu_w) \in \mathbb{T}^$, then there exists, for each n , an isometric embedding $(\tilde{T}_n, \tilde{\mu}_n, \tilde{\mathbb{P}}_\rho^{T_n})$ of the triple $(T_n, \mu_n, \mathbb{P}_\rho^{T_n})$ into ℓ^1 such that*

$$(\alpha_n^{-1} \tilde{T}_n, \tilde{\mu}_n(\alpha_n \cdot), \mathbb{P}_\rho^{T_n}(\{f \in C([0, 1], \ell^1) : (\alpha_n^{-1} f(t\alpha_n))_{t \in [0, 1]} \in \cdot\}))$$

converges to $(\tilde{\mathcal{T}}, \tilde{\mu}, \tilde{\mathbb{P}}_\rho^{\mathcal{T}, \mu})$ in the space $\mathcal{K}(\ell^1) \times \mathcal{M}_1(\ell^1) \times \mathcal{M}_1(C([0, 1], \ell^1))$, where the law $\tilde{\mathbb{P}}_\rho^{T_n}$ is extended to an element of $\mathcal{M}_1(C([0, 1], \ell^1))$ by linear interpolation of discrete-time processes, and $(\tilde{\mathcal{T}}, \tilde{\mu}, \tilde{\mathbb{P}}_\rho^{\mathcal{T}, \mu})$ is an isometric embedding of $(\mathcal{T}_w, \mu_w, \mathbb{P}_\rho^{\mathcal{T}_w, \mu_w})$ into ℓ^1 .

As an extension to [7, Theorem 1.2], a random version of this result can be formulated by applying the fact that the isometrically embedded triple $(\tilde{\mathcal{T}}, \tilde{\mu}, \tilde{\mathbb{P}}_\rho^{\mathcal{T}, \mu})$ can be constructed as a measurable function of a pair (w, u) , where w is the relevant excursion and u is an element of $[0, 1]^{\mathbb{N}}$. In particular, suppose that (W, U) is a $(\mathcal{W} \times [0, 1]^{\mathbb{N}})$ -valued random variable, built on a complete probability space with probability measure \mathbb{P} , such that W is a random excursion satisfying $\mathbb{P}((\mathcal{T}_W, \mu_W) \in \mathbb{T}^*) = 1$, and $U = (U_i)_{i \geq 1}$ is an independent sequence of independent $U[0, 1]$ random variables. We can define a probability law Pr on $\mathcal{K}(\ell^1) \times \mathcal{M}_1(\ell^1) \times C([0, 1], \ell^1)$ that satisfies

$$\text{Pr}(A \times B \times C) = \int_{C([0, 1], \mathbb{R}_+) \times [0, 1]^{\mathbb{N}}} \mathbb{P}((W, U) \in (dw, du)) \mathbf{1}_{\{\tilde{\mathcal{T}} \in A, \tilde{\mu} \in B\}} \tilde{\mathbb{P}}_\rho^{\mathcal{T}, \mu}(C) \quad (1.2)$$

for every measurable $A \subseteq \mathcal{K}(\ell^1)$, $B \subseteq \mathcal{M}_1(\ell^1)$, and $C \subseteq C([0, 1], \ell^1)$. In the discrete case, as in [7], let $(T_n)_{n \geq 1}$ be a sequence of random ordered graph trees with corresponding search-depth functions $(W_n)_{n \geq 1}$, and suppose that these are built on our underlying probability space independently of the random variable U . We can assume that $(\tilde{T}_n, \tilde{\mu}_n, \tilde{\mathbb{P}}_\rho^{T_n})$ are constructed measurably from (W_n, U) . Moreover, there exists a unique probability law Pr_n on $\mathcal{K}(\ell^1) \times \mathcal{M}_1(\ell^1) \times C([0, 1], \ell^1)$ that satisfies

$$\text{Pr}_n(A \times B \times C) = \int_{C([0, 1], \mathbb{R}_+) \times [0, 1]^{\mathbb{N}}} \mathbb{P}((W_n, U) \in (dw, du)) \mathbf{1}_{\{\tilde{T}_n \in A, \tilde{\mu}_n \in B\}} \tilde{\mathbb{P}}_\rho^{T_n}(C) \quad (1.3)$$

for every measurable $A \subseteq \mathcal{K}(\ell^1)$, $B \subseteq \mathcal{M}_1(\ell^1)$, and $C \subseteq C([0, 1], \ell^1)$. In the following result,

the rescaling operator Θ_n is defined on $\mathcal{K}(\ell^1) \times \mathcal{M}_1(\ell^1) \times C([0, 1], \ell^1)$, so that if $(\tilde{K}, \tilde{v}, \tilde{f})$ is an element of this space then

$$\Theta_n(\tilde{K}, \tilde{v}, \tilde{f}) := (\alpha_n^{-1}\tilde{K}, \tilde{v}(\alpha_n \cdot), (\alpha_n^{-1}\tilde{f}(t\alpha_n))_{t \in [0,1]}).$$

Theorem 1.2. *Let $(\alpha_n)_{n \geq 1}$ be a positive divergent sequence such that $\alpha_n = o(n)$. Suppose that $(T_n)_{n \geq 1}$ is a sequence of random ordered graph trees whose rescaled search-depth functions $(\alpha_n^{-1}W_n)_{n \geq 1}$ converge in distribution to W in $C([0, 1], \mathbb{R}_+)$, where W is a random excursion satisfying $P((\mathcal{T}_W, \mu_W) \in \mathbb{T}^*) = 1$. If \Pr and $(\Pr_n)_{n \geq 1}$ are the unique probability measures satisfying (1.2) and (1.3), respectively, then*

$$\Pr_n \circ \Theta_n^{-1} \rightarrow \Pr$$

weakly as measures on the space $\mathcal{K}(\ell^1) \times \mathcal{M}_1(\ell^1) \times C([0, 1], \ell^1)$.

In addition to the above convergence results for the laws of the simple random walks on the graph trees $(T_n)_{n \geq 1}$, let us remark that, in the settings of Theorem 1.1 or 1.2, it is possible to proceed as in [9, Section 7.2] to deduce related local limit theorems demonstrating that the associated discrete transition densities can be rescaled to converge in an appropriate space to the transition densities of the Brownian motion on the limiting space.

To prove the results stated above, we commence, in Section 2, by proving some general results about Brownian motion and the corresponding local times on real trees. In Section 2, we also describe the procedures we use for embedding real trees into ℓ^1 , and the connection between real trees and excursions. The heart of the paper is Section 3, which is where we establish some key results about the convergence of Brownian motions on real trees with a finite number of branches embedded into ℓ^1 . In Section 4, which contains the proofs of Theorems 1.1 and 1.2, we explain how the argument of [7] can be reworked to apply in our more general setting. Finally, in Section 5, we describe the application of our results to Galton–Watson trees with infinite variance offspring distributions and the related α -stable trees.

2. Brownian motion and local times on real trees

Suppose that $\mathcal{T} = (\mathcal{T}, d_{\mathcal{T}})$ is a compact real tree and that μ is a finite Borel measure on \mathcal{T} with full support. To avoid trivialities, assume that \mathcal{T} contains more than one point. It is possible (see [7, Proposition 2.2]) to construct a strong Markov process

$$X^{\mathcal{T}, \mu} = ((X_t^{\mathcal{T}, \mu})_{t \geq 0}, P_{\sigma}^{\mathcal{T}, \mu}, \sigma \in \mathcal{T}),$$

with continuous sample paths that is reversible with respect to its invariant measure μ and satisfies the following properties.

- (i) For $\sigma_1, \sigma_2 \in \mathcal{T}$, $\sigma_1 \neq \sigma_2$, we have

$$P_{\sigma}^{\mathcal{T}, \mu}(h(\sigma_1) < h(\sigma_2)) = \frac{d_{\mathcal{T}}(b^{\mathcal{T}}(\sigma, \sigma_1, \sigma_2), \sigma_2)}{d_{\mathcal{T}}(\sigma_1, \sigma_2)} \quad \text{for all } \sigma \in \mathcal{T},$$

where $h(\sigma) := \inf\{t > 0: X_t^{\mathcal{T}, \mu} = \sigma\}$ is the hitting time of $\sigma \in \mathcal{T}$, and $b^{\mathcal{T}}(\sigma, \sigma_1, \sigma_2)$ is the unique branch point of σ, σ_1 , and σ_2 in \mathcal{T} . In particular, if $[[\sigma, \sigma_1]]$, $[[\sigma_1, \sigma_2]]$, and $[[\sigma_2, \sigma]]$ are the unique injective paths between the relevant pairs of vertices, then $b^{\mathcal{T}}(\sigma, \sigma_1, \sigma_2)$ is the unique point in the set $[[\sigma, \sigma_1]] \cap [[\sigma_1, \sigma_2]] \cap [[\sigma_2, \sigma]]$.

- (ii) For $\sigma_1, \sigma_2 \in \mathcal{T}$, the mean occupation measure for the process started at σ_1 and killed on hitting σ_2 has density $2d_{\mathcal{T}}(b^{\mathcal{T}}(\sigma, \sigma_1, \sigma_2), \sigma_2)\mu(d\sigma)$, so that

$$E_{\sigma_1}^{\mathcal{T}, \mu} \int_0^{h(\sigma_2)} f(X_s) ds = 2 \int_{\mathcal{T}} f(\sigma) d_{\mathcal{T}}(b^{\mathcal{T}}(\sigma, \sigma_1, \sigma_2), \sigma_2)\mu(d\sigma)$$

for every continuous bounded function $f : \mathcal{T} \rightarrow \mathbb{R}$.

In the terminology of [1, Section 5.2], $X^{\mathcal{T}, \mu}$ is Brownian motion on (\mathcal{T}, μ) , and is in fact uniquely determined by these properties. Moreover, that $X^{\mathcal{T}, \mu}$ admits jointly measurable local times $(L_t(\sigma))_{\sigma \in \mathcal{T}, t \geq 0}$ can be checked as in the proof of [8, Lemma 8.2]. In the arguments of subsequent sections we will require further that $(L_t(\sigma))_{\sigma \in \mathcal{T}, t \geq 0}$ is jointly continuous in t and σ , and we will demonstrate that this is the case whenever μ satisfies a polynomial lower bound of the form of (1.1). In the proof of this result, we apply the two following properties that (1.1) implies.

Lemma 2.1. *Suppose that \mathcal{T} is a compact real tree and that μ is a finite Borel measure on \mathcal{T} that satisfies (1.1) for some $\kappa > 0$.*

- (a) *If $N(\mathcal{T}, r)$ is the smallest number of balls of radius r needed to cover \mathcal{T} , then*

$$\limsup_{r \rightarrow 0} r^{\kappa} N(\mathcal{T}, r) < \infty.$$

- (b) *The Markov process $X^{\mathcal{T}, \mu}$ admits a transition density $(p_t(\sigma, \sigma'))_{\sigma, \sigma' \in \mathcal{T}, t > 0}$ that satisfies*

$$\limsup_{t \rightarrow 0} t^{\kappa/(\kappa+1)} \sup_{\sigma, \sigma' \in \mathcal{T}} p_t(\sigma, \sigma') < \infty.$$

Proof. The proof of part (a) is elementary. Part (b) can be obtained by applying a general heat kernel bound of the type proved in [18, Proposition 4.1] or [6, Proposition 5] for example.

Lemma 2.2. *If \mathcal{T} is a compact real tree and μ is a finite Borel measure on \mathcal{T} that satisfies (1.1) for some $\kappa > 0$, then the local times $(L_t(\sigma))_{\sigma \in \mathcal{T}, t \geq 0}$ of $X^{\mathcal{T}, \mu}$ are jointly continuous in t and σ , $P_{\sigma'}^{\mathcal{T}, \mu}$ -almost surely ($P_{\sigma'}^{\mathcal{T}, \mu}$ -a.s.), for every $\sigma' \in \mathcal{T}$.*

Proof. Given the estimates of Lemma 2.1, the proof is identical to that of [7, Lemma 2.5]. In particular, it is easily checked from Lemma 2.1(b) that the 1-potential density $u(\sigma, \sigma') := \int_0^{\infty} e^{-t} p_t(\sigma, \sigma') dt$ is finite for all $\sigma, \sigma' \in \mathcal{T}$. As a consequence, by applying [20, Theorem 1], we see that the continuity of local times is equivalent to the continuity of the centred Gaussian process $(G(\sigma))_{\sigma \in \mathcal{T}}$ with covariances given by $(u(\sigma, \sigma'))_{\sigma, \sigma' \in \mathcal{T}}$. By [11, Theorem 2.1], for the latter process to be continuous, it is enough that the integral $\int_0^1 \sqrt{\ln N(\mathcal{T}, r)} dr$ is finite, which in view of Lemma 2.1(a) is clearly the case.

The local times of $X^{\mathcal{T}, \mu}$ will be used in a time-change argument that depends on their uniform divergence, which we now prove. Note that, by definition, the Brownian motion on (\mathcal{T}, μ) satisfies $E_{\sigma}^{\mathcal{T}, \mu} h(\sigma') \leq 2 \text{diam}(\mathcal{T})\mu(\mathcal{T}) < \infty$ for every $\sigma, \sigma' \in \mathcal{T}$, where $\text{diam}(\mathcal{T})$ is the diameter of the metric space $(\mathcal{T}, d_{\mathcal{T}})$.

Lemma 2.3. *Suppose that \mathcal{T} is a compact real tree and that μ is a finite Borel measure on \mathcal{T} that satisfies (1.1) for some $\kappa > 0$. For every $\sigma \in \mathcal{T}$, $P_{\sigma}^{\mathcal{T}, \mu}$ -a.s., we have*

$$\liminf_{t \rightarrow \infty} \inf_{\sigma' \in \mathcal{T}} L_t(\sigma') = \infty.$$

Proof. Fix $\sigma, \sigma' \in \mathcal{T}$. By [20, Lemma 3.6] we have $E_{\sigma'}^{\mathcal{J}, \mu} \int_0^\infty e^{-t} d_t L_t(\sigma') > 0$. Hence, there is a strictly positive $P_{\sigma'}^{\mathcal{J}, \mu}$ -probability that $L_t(\sigma') > 0$ for large t . Applying the joint continuity of the local times, it follows that there exist $r = r(\sigma') > 0, \varepsilon = \varepsilon(\sigma') > 0$, and $t_0 = t_0(\sigma') < \infty$ such that

$$P_{\sigma'}^{\mathcal{J}, \mu} \left(\inf_{\sigma'' \in B(\sigma', r)} L_{t_0}(\sigma'') > \varepsilon \right) > 0. \tag{2.1}$$

Now, set $h(\sigma', t_0, \sigma) := \inf\{t > t_0 + h(\sigma') : X_t^{\mathcal{J}, \mu} = \sigma\}$. Applying the observation made above this lemma about the finite moments of hitting times, the strong Markov property, and (2.1), it is easy to check that $h(\sigma', t_0, \sigma)$ is finite, $P_\sigma^{\mathcal{J}, \mu}$ -a.s., and also that

$$P_\sigma^{\mathcal{J}, \mu} \left(\inf_{\sigma'' \in B(\sigma', r)} L_{h(\sigma', t_0, \sigma)}(\sigma'') > \varepsilon \right) > 0. \tag{2.2}$$

Observe that the additivity of local times and the strong Markov property implies that

$$\liminf_{t \rightarrow \infty} \inf_{\sigma'' \in B(\sigma', r)} L_t(\sigma'') \geq \sum_{i=1}^\infty \xi_i,$$

where, under $P_\sigma^{\mathcal{J}, \mu}$, the $(\xi_i)_{i=1}^\infty$ are independent copies of $\inf_{\sigma'' \in B(\sigma', r)} L_{h(\sigma', t_0, \sigma)}(\sigma'')$. The strong law of large numbers lets it be deduced from (2.2) that the right-hand side of the above inequality is infinite, $P_\sigma^{\mathcal{J}, \mu}$ -a.s., which proves the uniform divergence of local times uniformly over $B(\sigma', r)$, $P_\sigma^{\mathcal{J}, \mu}$ -a.s.

To extend the conclusion of the previous paragraph, note that $B(\sigma', r(\sigma'))_{\sigma' \in \mathcal{T}}$ is an open cover for \mathcal{T} . Thus, by the compactness of \mathcal{T} , it admits a finite subcover, $B(\sigma_i, r(\sigma_i))_{i=1}^N$, and, clearly,

$$\liminf_{t \rightarrow \infty} \inf_{\sigma' \in \mathcal{T}} L_t(\sigma') = \min_{i=1, \dots, N} \lim_{t \rightarrow \infty} \inf_{\sigma' \in B(\sigma_i, r(\sigma_i))} L_t(\sigma').$$

Since, by our above argument, the right-hand side of this expression is infinite, $P_\sigma^{\mathcal{J}, \mu}$ -a.s., the proof is complete.

Following [7], we now show how $X^{\mathcal{J}, \mu}$ can be approximated by a family of Brownian motions on subtrees of \mathcal{T} with a finite number of branches. Henceforth, we suppose that the real tree \mathcal{T} has a distinguished vertex $\rho \in \mathcal{T}$ called the root and consider a dense sequence of vertices $(\sigma_i)_{i=1}^\infty$ in \mathcal{T} . Without loss of generality, we assume that the $(\sigma_i)_{i=1}^\infty$ are distinct and that $\sigma_i \neq \rho$ for any i . For each $k \geq 1$, define a subset of \mathcal{T} by

$$\mathcal{T}(k) := \bigcup_{i=1}^k [[\rho, \sigma_i]], \tag{2.3}$$

where, as above, for $\sigma, \sigma' \in \mathcal{T}$, $[[\sigma, \sigma']]$ is the unique injective path in \mathcal{T} from σ to σ' . Clearly, $\mathcal{T}(k)$ is a compact real tree when endowed with the appropriate restriction of $d_{\mathcal{T}}$, and we set its root to be ρ , which is contained in $\mathcal{T}(k)$ by construction. The natural projection $\phi_{\mathcal{T}, \mathcal{T}(k)}$ from \mathcal{T} to $\mathcal{T}(k)$ is obtained by setting $\phi_{\mathcal{T}, \mathcal{T}(k)}(\sigma)$ to be the unique point in $\mathcal{T}(k)$ satisfying

$$d_{\mathcal{T}}(\sigma, \phi_{\mathcal{T}, \mathcal{T}(k)}(\sigma)) = \inf_{\sigma' \in \mathcal{T}(k)} d_{\mathcal{T}}(\sigma, \sigma'). \tag{2.4}$$

It is elementary to check that $\phi_{\mathcal{T}, \mathcal{T}(k)}$ is continuous and $\sup_{x \in \mathcal{T}} d_{\mathcal{T}}(x, \phi_{\mathcal{T}, \mathcal{T}(k)}(x)) \rightarrow 0$ (cf. [7, Lemma 2.4]). Consequently, $\mu^{(k)} := \mu \circ \phi_{\mathcal{T}, \mathcal{T}(k)}^{-1}$ defines a Borel probability measure on $\mathcal{T}(k)$ with full support for each k , and $\mu^{(k)} \rightarrow \mu$ weakly as probability measures on \mathcal{T} . Since $\mathcal{T}(k)$ is a compact real tree containing more than one point and $\mu^{(k)}$ is a Borel probability measure on $\mathcal{T}(k)$ with full support, we can construct the Brownian motion $X^{\mathcal{T}(k), \mu^{(k)}}$ on $(\mathcal{T}(k), \mu^{(k)})$. For these processes, we are able to deduce the following convergence result. Since it is a relatively simple adaptation of [7, Lemma 3.1], we only sketch the proof.

Proposition 2.1. *If (\mathcal{T}, μ) and $\{(\mathcal{T}(k), \mu^{(k)})\}_{k=1}^{\infty}$ are as described above, then*

$$P_{\rho}^{\mathcal{T}(k), \mu^{(k)}} \rightarrow P_{\rho}^{\mathcal{T}, \mu}$$

weakly as probability measures on $C(\mathbb{R}_+, \mathcal{T})$.

Proof. Applying the weak convergence of $\mu^{(k)}$ to μ and the joint continuity of the local times of $X^{\mathcal{T}, \mu}$ (see Lemma 2.2), we obtain, for every $t \geq 0$, $P_{\rho}^{\mathcal{T}, \mu}$ -a.s.,

$$\tilde{A}_t^{(k)} := \int_{\mathcal{T}(k)} L_t(\sigma) \mu^{(k)}(d\sigma) \rightarrow t.$$

Moreover, an elementary monotonicity argument yields this convergence result uniformly on compact intervals. As a consequence of this, $\tilde{\tau}^{(k)}(t) := \inf\{s : \tilde{A}_s^{(k)} > t\} \rightarrow t$ uniformly on compacts, $P_{\rho}^{\mathcal{T}, \mu}$ -a.s. Now, the trace theorem for Dirichlet forms (see [15, Theorem 6.2.1] for example) allows us to check that the law of $(X_{\tilde{\tau}^{(k)}(t)}^{\mathcal{T}, \mu})_{t \geq 0}$ under $P_{\rho}^{\mathcal{T}, \mu}$ is precisely $P_{\rho}^{\mathcal{T}(k), \mu^{(k)}}$ (cf. [7, Lemma 2.6]), and, hence, the result follows.

We continue by presenting a characterisation of $X^{\mathcal{T}(k), \mu^{(k)}}$ as a time change of another Brownian motion on $\mathcal{T}(k)$. For $k \geq 1$, let $\lambda^{(k)}$ be the one-dimensional Hausdorff measure on $\mathcal{T}(k)$ normalised to have total mass equal to 1. Since $\mathcal{T}(k)$ consists of a finite number of line segments, $\lambda^{(k)}$ is a Borel probability measure on $\mathcal{T}(k)$ with full support. Consequently, the Brownian motion $X^{\mathcal{T}(k), \lambda^{(k)}}$ on $(\mathcal{T}(k), \lambda^{(k)})$ exists. Furthermore, it is elementary to check that $\lambda^{(k)}$ satisfies (1.1) with $\kappa = 1$ and, therefore, we can apply Lemma 2.2 to deduce that $X^{\mathcal{T}(k), \lambda^{(k)}}$ admits jointly continuous local times $(L_t^{(k)}(\sigma))_{\sigma \in \mathcal{T}(k), t \geq 0}$. As in [7], define a continuous additive functional $\hat{A}^{(k)} = (\hat{A}_t^{(k)})_{t \geq 0}$ by setting

$$\hat{A}_t^{(k)} := \int_{\mathcal{T}(k)} L_t^{(k)}(\sigma) \mu^{(k)}(d\sigma), \tag{2.5}$$

and its inverse by

$$\hat{\tau}^{(k)}(t) := \inf\{s : \hat{A}_s^{(k)} > t\}. \tag{2.6}$$

As with the time change employed in the proof of the previous result, the following lemma is a relatively straightforward consequence of the trace theorem for Dirichlet forms (see [15, Theorem 6.2.1] for the trace theorem and [7, Lemma 2.6] for a similar application), and so will be stated without proof.

Lemma 2.4. *Suppose that (\mathcal{T}, μ) and $\{(\mathcal{T}(k), \mu^{(k)})\}_{k=1}^{\infty}$ are as described above. If the process $X^{\mathcal{T}(k), \lambda^{(k)}}$ has law $P_{\rho}^{\mathcal{T}(k), \lambda^{(k)}}$ then the process*

$$(X_{\hat{\tau}^{(k)}(t)}^{\mathcal{T}(k), \lambda^{(k)}})_{t \geq 0}$$

has law $P_{\rho}^{\mathcal{T}(k), \mu^{(k)}}$.

We now deduce some simple path properties of $\hat{A}^{(k)}$ that will be useful to us later.

Lemma 2.5. *If (\mathcal{T}, μ) and $\{(\mathcal{T}(k), \mu^{(k)})\}_{k=1}^\infty$ are as described above, then, for every $k \geq 1$, $P_\rho^{\mathcal{T}(k), \lambda^{(k)}}$ -a.s., the functions $\hat{A}^{(k)}$ are continuous and strictly increasing. Moreover, for every $t_0 \geq 0$,*

$$\lim_{t \rightarrow \infty} \limsup_{k \rightarrow \infty} P_\rho^{\mathcal{T}(k), \lambda^{(k)}}(\hat{A}_{t_0}^{(k)} > t) = 0. \tag{2.7}$$

Proof. The continuity of $\hat{A}^{(k)}$ follows from the continuity of $(L_t^{(k)}(\sigma))_{\sigma \in \mathcal{T}(k), t \geq 0}$, which was noted above Lemma 2.4. Hence, it remains to show that $\hat{A}_t^{(k)}$ strictly increases in t and satisfies (2.7). We start by showing that $(\hat{A}_t^{(k)})_{t \geq 0}$ is strictly increasing. The following argument holds $P_\rho^{\mathcal{T}(k), \lambda^{(k)}}$ -a.s. Let $s < t$. Then, clearly,

$$\int_{\mathcal{T}(k)} (L_t^{(k)}(x) - L_s^{(k)}(x)) \lambda^{(k)}(dx) = t - s > 0.$$

Hence, there exist an $\varepsilon > 0$ and a nonempty open set $A \subseteq \mathcal{T}(k)$ such that $L_t^{(k)}(x) - L_s^{(k)}(x) > \varepsilon$ for $x \in A$. Since $\mu^{(k)}$ has full support, it charges every nonempty open set and so we must therefore have $\hat{A}_s^{(k)} \leq \hat{A}_t^{(k)} - \varepsilon \mu^{(k)}(A) < \hat{A}_t^{(k)}$, which proves the desired result.

We now prove the tightness result of (2.7). Use the local times of $X^{\mathcal{T}, \mu}$ to define an additive functional $A^{(k)} = (A_t^{(k)})_{t \geq 0}$ by setting

$$A_t^{(k)} := \int_{\mathcal{T}} L_t(\sigma) \lambda^{(k)}(d\sigma),$$

and its inverse $\tau^{(k)}$ similarly to the definition of $\hat{\tau}^{(k)}$ in (2.6). By again applying the trace theorem for Dirichlet forms, it can be deduced that the law of the process $(X_{\tau^{(k)}(t)}^{\mathcal{T}, \mu})_{t \geq 0}$ under $P_\rho^{\mathcal{T}, \mu}$ is the same as that of the process $X^{\mathcal{T}(k), \lambda^{(k)}}$ under $P_\rho^{\mathcal{T}(k), \lambda^{(k)}}$. Similarly to [7, Lemma 3.4], it follows that, under $P_\rho^{\mathcal{T}, \mu}$, the two-parameter process $(L_{\tau^{(k)}(t)}(\sigma))_{\sigma \in \mathcal{T}(k), t \geq 0}$ has the same distribution as $(L_t^{(k)}(\sigma))_{\sigma \in \mathcal{T}(k), t \geq 0}$ under $P_\rho^{\mathcal{T}(k), \lambda^{(k)}}$. Consequently, to complete the proof, it will suffice to demonstrate that, for every $t_0 \geq 0$, $P_\rho^{\mathcal{T}, \mu}$ -a.s.,

$$\limsup_{k \rightarrow \infty} \int_{\mathcal{T}} L_{\tau^{(k)}(t_0)}(\sigma) \mu^{(k)}(d\sigma) < \infty.$$

Recalling that $\mu^{(k)}$ converges weakly to μ , we find that the left-hand side of the above expression is bounded above by

$$\int_{\mathcal{T}} L_{\sup_k \tau^{(k)}(t_0)}(\sigma) \mu(d\sigma) = \sup_k \tau^{(k)}(t_0),$$

and this supremum is finite whenever $\inf_k A_t^{(k)}$ diverges as $t \rightarrow \infty$. Applying the uniform divergence of local times proved in Lemma 2.3 and the fact that $\lambda^{(k)}$ is by definition a probability measure, this result holds $P_\rho^{\mathcal{T}, \mu}$ -a.s., as required.

As in [7], to embed \mathcal{T} into ℓ^1 , we use the sequential embedding of [2, Section 2.2]. In particular, given a sequence $(\mathcal{T}(k))_{k \geq 1}$ as above, it is possible to construct a distance-preserving map $\psi : (\mathcal{T}, d_{\mathcal{T}}) \rightarrow (\ell^1, d_{\ell^1})$ that satisfies $\psi(\rho) = 0$ and

$$\pi_k(\psi(\sigma)) = \psi(\phi_{\mathcal{T}, \mathcal{T}(k)}(\sigma)) \tag{2.8}$$

for every $\sigma \in \mathcal{T}$ and $k \geq 1$, where π_k is the projection map on ℓ^1 defined by setting $\pi_k(x(1), x(2), \dots) = (x(1), \dots, x(k), 0, \dots)$. Such a map is determined uniquely by

insisting that $\psi(\mathcal{T}) \subseteq \{(x(1), x(2), \dots) \in \ell^1 : x(i) \geq 0, i = 1, 2, \dots\}$. We will denote the ℓ^1 -embedded versions of the objects $\mathcal{T}, \mu, P_\rho^{\mathcal{T}, \mu}, \dots$ by $\tilde{\mathcal{T}}, \tilde{\mu}, \tilde{P}_\rho^{\mathcal{T}, \mu}, \dots$, respectively.

To complete this section, we present the well-known relation between continuous excursions and real trees, and define a collection of pairs of excursions and sequences that will be of interest later in this paper. Let \mathcal{W} be defined as in the introduction. For $w \in \mathcal{W}$, define a distance on $[0, 1]$ by setting

$$d_w(s, t) := w(s) + w(t) - 2 \inf\{w(r) : r \in [s \wedge t, s \vee t]\}, \tag{2.9}$$

and then use the equivalence $s \sim_w t$ if and only if $d_w(s, t) = 0$, to define $\mathcal{T}_w := [0, 1] / \sim_w$. Denoting the canonical projection (with respect to ‘ \sim_w ’) from $[0, 1]$ to \mathcal{T}_w by \hat{w} , it is possible to check that $d_{\mathcal{T}_w}(\hat{w}(s), \hat{w}(t)) := d_w(s, t)$ defines a metric on \mathcal{T}_w , and also that with this metric \mathcal{T}_w is a compact real tree (see [13, Theorem 2.1]). The root of the tree \mathcal{T}_w is defined to be the equivalence class $\hat{w}(0)$, and is denoted by ρ_w . A Borel probability measure on \mathcal{T}_w with full support can be constructed by setting $\mu_w := \lambda^{[0,1]} \circ \hat{w}^{-1}$, where $\lambda^{[0,1]}$ is the usual one-dimensional Lebesgue measure on $[0, 1]$. Furthermore, given a pair (w, u) , where $w \in \mathcal{W}$ and $u = (u_i)_{i \geq 1} \in [0, 1]^{\mathbb{N}}$, we define a sequence of vertices $(\sigma_i)_{i \geq 1}$ by setting $\sigma_i = \hat{w}(u_i)$ for each i . This allows us to construct a sequence of subtrees $\mathcal{T}_{w,u}(k)$ of \mathcal{T}_w as in (2.3). Note that we will usually suppress the dependence on w and u from the notation for all of these objects when it is clear which excursion and sequence is being considered.

Definition 2.1. The set Γ is the collection of pairs $(w, u) \in \mathcal{W} \times [0, 1]^{\mathbb{N}}$ such that μ satisfies (1.1) for some $\kappa > 0$, the sequence $(u_i)_{i \geq 1}$ is dense in $[0, 1]$, and the vertices $(\sigma_i)_{i \geq 1}$ are a dense collection of leaves of \mathcal{T} , distinct and not equal to ρ for any i .

3. ℓ^1 convergence

For $x = (x(1), x(2), \dots) \in \ell^1$, define $[[0, x]]_{sp}$ as in [2] to be the union of line segments connecting 0 to $(x(1), 0, 0, \dots)$ to $(x(1), x(2), 0, 0, \dots), \dots$. Fix $k \geq 1$, and suppose that we are given distinct $x^{(1)}, \dots, x^{(k)} \in \ell^1 \setminus \{0\}$. Write $\mathbf{x} = (x^{(1)}, \dots, x^{(k)})$, and set

$$\mathcal{T}^{\mathbf{x}} := \bigcup_{i=1}^k [[0, x^{(i)}]]_{sp}, \tag{3.1}$$

which is a compact real tree. We assume that every $x \in \{0, x^{(1)}, \dots, x^{(k)}\}$ is a leaf of $\mathcal{T}^{\mathbf{x}}$. Define $\lambda^{\mathbf{x}}$ to be the one-dimensional Hausdorff measure on $\mathcal{T}^{\mathbf{x}}$; in this section it will be convenient not to normalise this to be a probability measure. Denote the law of the Brownian motion $X^{\mathbf{x}}$ on $(\mathcal{T}^{\mathbf{x}}, \lambda^{\mathbf{x}})$ started from 0 by $P_0^{\mathbf{x}}$. Given a Borel probability measure ν on $\mathcal{T}^{\mathbf{x}}$, let $A^{\mathbf{x}, \nu}$ be the additive functional defined by

$$A_t^{\mathbf{x}, \nu} := \int_{\mathcal{T}^{\mathbf{x}}} L_t^{\mathbf{x}}(x) \nu(dx),$$

where $(L_t^{\mathbf{x}}(x))_{x \in \mathcal{T}^{\mathbf{x}}, t \geq 0}$ are the local times of $X^{\mathbf{x}}$, which exist and are jointly continuous, $P_0^{\mathbf{x}}$ -a.s., because $\lambda^{\mathbf{x}}$ satisfies (1.1) for $\kappa = 1$ and, therefore, Lemma 2.2 applies.

The aim of this section is to show that if we have a sequence $(\mathbf{x}_n)_{n \geq 1}$, where $\mathbf{x}_n = (x_n^{(1)}, \dots, x_n^{(k)}) \in (\ell^1)^k$, that satisfies $\mathbf{x}_n \rightarrow \mathbf{x}$, then the Brownian motions $X^{\mathbf{x}_n}$ on $(\mathcal{T}^{\mathbf{x}_n}, \lambda^{\mathbf{x}_n})$ started from 0 converge in distribution to the Brownian motion $X^{\mathbf{x}}$ started from 0. Note that we construct $\mathcal{T}^{\mathbf{x}_n}$ from \mathbf{x}_n similarly to the definition of $\mathcal{T}^{\mathbf{x}}$ in (3.1) and $\lambda^{\mathbf{x}_n}$ is the one-dimensional

Hausdorff measure on \mathcal{T}^{x_n} ; the assumption that $x_n \rightarrow x$ means that we can define the law $P_0^{x_n}$ of the Brownian motion on $(\mathcal{T}^{x_n}, \lambda^{x_n})$ started from 0 as in the previous section, at least for large n . Moreover, simultaneously with this convergence, we will prove that if $(\nu_n)_{n \geq 1}$ is a sequence of Borel probability measures, where ν_n is supported on \mathcal{T}^{x_n} , such that $\nu_n \rightarrow \nu$ weakly as probability measures on ℓ^1 , then the related additive functionals A^{x_n, ν_n} , defined by

$$A_t^{x_n, \nu_n} := \int_{\mathcal{T}^{x_n}} L_t^{x_n}(x) \nu_n(dx),$$

converge in distribution to $A^{x, \nu}$ in $C(\mathbb{R}_+, \mathbb{R}_+)$. To deduce the existence and continuity of the local times $(L_t^x(x))_{x \in \mathcal{T}^x, t \geq 0}$ of X^x , we again apply Lemma 2.2.

On line segments, the processes X^{x_n} and X^x look like standard one-dimensional Brownian motions. However, the structure of \mathcal{T}^{x_n} and \mathcal{T}^x can vary at branch points, and so we must be careful about analysing the processes close to these. In our arguments, we will approximate X^{x_n} and X^x by processes that avoid the branch points of the trees. Define the finite set of ‘vertices’ of \mathcal{T}^x by

$$\mathcal{B}^x := \{b^x(x, x', x'') : x, x', x'' \in \{0, x^{(1)}, \dots, x^{(k)}\}\},$$

where $b^x(x, x', x'')$ is the branch point of x, x' , and x'' in \mathcal{T}^x . Note that \mathcal{B}^x contains the set $\{0, x^{(1)}, \dots, x^{(k)}\}$. An ε -neighbourhood of \mathcal{B}^x in ℓ^1 is given by

$$\mathcal{B}_\varepsilon^x := \bigcup_{x \in \mathcal{B}^x} B_{\ell^1}(x, \varepsilon),$$

where $B_{\ell^1}(x, \varepsilon)$ is the open ball in (ℓ^1, d_{ℓ^1}) of radius ε centred at x . Now, fix a strictly positive constant $\varepsilon_1 < \varepsilon_0$, where

$$\varepsilon_0 := \frac{1}{2} \inf_{\{x, x' \in \mathcal{B}^x : x \neq x'\}} d_{\ell^1}(x, x'). \tag{3.2}$$

Set $v_0 = 0$ and, for $i \geq 0$, let

$$\varsigma_i := \inf\{t \geq v_i : X_t^x \notin \mathcal{B}_{\varepsilon_1}^x\}, \quad v_{i+1} := \inf\{t \geq \varsigma_i : X_t^x \in \overline{\mathcal{B}_{\varepsilon_1/2}^x}\}.$$

Define

$$A_t^{x, \varepsilon_1} := \int_0^t \mathbf{1}_{\{s \in I\}} ds,$$

where

$$I := \mathbb{R} \setminus \bigcup_{i=0}^\infty (v_i, \varsigma_i), \tag{3.3}$$

and its inverse $\tau^{x, \varepsilon_1}(t) := \inf\{s : A_s^{x, \varepsilon_1} > t\}$. Finally, let X^{x, ε_1} be the process defined by

$$X_t^{x, \varepsilon_1} := X_{\tau^{x, \varepsilon_1}(t)}^x,$$

which takes values in the space $D(\mathbb{R}_+, \ell^1)$ of càdlàg paths in ℓ^1 . To prove that X^{x, ε_1} approximates X^x well for small ε_1 , we will apply the following (rather crude) bound on the expectations of the local times $(L_t^x(x))_{x \in \mathcal{T}^x, t \geq 0}$.

Lemma 3.1. *The local times of X^x satisfy*

$$\sup_{x \in \mathcal{T}^x} E_0^x L_t^x(x) \leq \frac{t + 4\lambda^x(\mathcal{T}^x)e^t}{\lambda^x(\mathcal{T}^x)}.$$

Proof. By [4, Section V.3.28] we have $P_0^x(|L_t^x(x) - L_t^x(x')| > 2\delta) \leq 2e^t e^{-\delta}$ for every $x, x' \in \mathcal{T}^x$. Integrating this inequality implies that $E_0^x(|L_t^x(x) - L_t^x(x')|) \leq 4e^t$. Thus,

$$E_0^x(L_t^x(x))\lambda^x(\mathcal{T}^x) \leq E_0^x\left(\int_{\mathcal{T}^x} L_t^x(x')\lambda^x(dx')\right) + \int_{\mathcal{T}^x} E_0^x(|L_t^x(x) - L_t^x(x')|)\lambda^x(dx') \leq t + 4e^t \lambda^x(\mathcal{T}^x).$$

Lemma 3.2. For $t_0, \varepsilon > 0$,

$$\lim_{\varepsilon_1 \rightarrow 0} P_0^x\left(\sup_{t \in [0, t_0]} d_{\ell^1}(X_t^x, X_t^{x, \varepsilon_1}) > \varepsilon\right) = 0.$$

Proof. By construction, $X_t^{x, \varepsilon_1} = X_{\tau^{x, \varepsilon_1}(t)}^x$. Hence, applying the continuity of X^x and the definition of τ^{x, ε_1} as the inverse of A^{x, ε_1} , the result will follow if we can show that

$$\lim_{\varepsilon_1 \rightarrow 0} P_0^x\left(\sup_{t \in [0, t_0+1]} |t - A_t^{x, \varepsilon_1}| > \varepsilon\right) = 0. \tag{3.4}$$

To prove this, first note that if $X_0^x = 0$ then

$$\sup_{t \in [0, t_0+1]} |t - A_t^{x, \varepsilon_1}| \leq \int_0^{t_0+1} \mathbf{1}_{\{X_s \in \mathcal{B}_{\varepsilon_1}^x\}} ds = \int_{\mathcal{B}_{\varepsilon_1}^x} L_{t_0+1}^x(x)\lambda^x(dx).$$

Thus, applying Markov’s inequality, Fubini’s theorem, and Lemma 3.1, it follows that

$$P_0^x\left(\sup_{t \in [0, t_0+1]} |t - A_t^{x, \varepsilon_1}| > \varepsilon\right) \leq \varepsilon^{-1} \int_{\mathcal{B}_{\varepsilon_1}^x} E_0^x(L_{t_0+1}^x(x))\lambda^x(dx) \leq c\lambda^x(\mathcal{B}_{\varepsilon_1}^x), \tag{3.5}$$

where c is a constant that does not depend on ε_1 . Since λ^x is nonatomic and \mathcal{B}^x is a finite set, the result follows.

We now prove a similar result for X^{x_n} that is uniform in n . Set $v_0^n = 0$ and, for $i \geq 0$, let

$$\zeta_i^n := \inf\{t \geq v_i^n : X_t^{x_n} \notin \mathcal{B}_{\varepsilon_1}^x\}, \tag{3.6}$$

$$v_{i+1}^n := \inf\{t \geq \zeta_i^n : X_t^{x_n} \in \overline{\mathcal{B}_{\varepsilon_1/2}^x}\}. \tag{3.7}$$

Note that although $\mathcal{B}_{\varepsilon_1/2}^x$ and $\mathcal{B}_{\varepsilon_1}^x$ do not depend on n , the above quantities will be well defined and finite for large n ; when they are not, simply set $\zeta_0^n = 0$ and $\zeta_i^n = v_i^n = \infty$ for $i \geq 1$. Define A^{x_n, ε_1} , $\tau^{x_n, \varepsilon_1}$, and X^{x_n, ε_1} from these stopping times analogously to the definitions of A^{x, ε_1} , τ^{x, ε_1} , and X^{x, ε_1} , respectively.

Lemma 3.3. If $x_n \rightarrow x$ then, for $t_0, \varepsilon > 0$,

$$\lim_{\varepsilon_1 \rightarrow 0} \limsup_{n \rightarrow \infty} P_0^{x_n}\left(\sup_{t \in [0, t_0]} d_{\ell^1}(X_t^{x_n}, X_t^{x_n, \varepsilon_1}) > \varepsilon\right) = 0.$$

Proof. Similarly to the bound in (3.5), it is possible to deduce that

$$\limsup_{n \rightarrow \infty} P_0^{x_n}\left(\sup_{t \in [0, t_0+1]} |t - A_t^{x_n, \varepsilon_1}| > \varepsilon\right) \leq \limsup_{n \rightarrow \infty} c\lambda^{x_n}(\mathcal{B}_{\varepsilon_1}^x) = c\lambda^x(\mathcal{B}_{\varepsilon_1}^x), \tag{3.8}$$

where c is a constant that does not depend on n or ε_1 . The result will follow from this if we

can show that the sequence $(P_0^{x_n})_{n \geq 1}$ is tight. To prove this, first observe that it is elementary to check that there exist constants $c_1, c_2 \in (0, \infty)$ such that

$$c_1 r \leq \liminf_{n \rightarrow \infty} \inf_{x \in \mathcal{T}^{x_n}} \lambda^{x_n}(B_{\ell^1}(x, r)) \leq \limsup_{n \rightarrow \infty} \sup_{x \in \mathcal{T}^{x_n}} \lambda^{x_n}(B_{\ell^1}(x, r)) \leq c_2 r$$

for every $r \in (0, 1]$. By applying the argument of [18, Lemma 4.2], this implies that

$$\limsup_{n \rightarrow \infty} \sup_{x \in \mathcal{T}^{x_n}} P_x^{x_n}(\inf\{s : d_{\ell^1}(x, X_s^{x_n}) > r\} < t) \leq c_3 e^{-c_4 r^2/t}$$

for every $r \in (0, 1], t \in (0, t_1]$, and some constants $c_3, c_4, t_1 \in (0, \infty)$. Consequently,

$$\lim_{t \rightarrow 0} \limsup_{n \rightarrow \infty} t^{-1} \sup_{x \in \mathcal{T}^{x_n}} P_x^{x_n}(\inf\{s : d_{\ell^1}(x, X_s^{x_n}) > r\} < t) = 0$$

for any $r > 0$, which implies the tightness of $(P_0^{x_n})_{n \geq 1}$, as desired (cf. the corollary to Theorem 7.4 of [3]).

We now construct an approximation for the additive functional $A^{x, \nu}$. First, formulate a local time L^{x, ε_1} for the process X^{x, ε_1} by setting

$$L_t^{x, \varepsilon_1}(x) := \int_0^{\tau^{x, \varepsilon_1}(t)} \mathbf{1}_{\{s \in I\}} d_s L_s^x(x) \tag{3.9}$$

for $x \in \mathcal{T}^x$ and $t \geq 0$, where I is as defined in (3.3), then let $A^{x, \nu, \varepsilon_1}$ be defined by

$$A_t^{x, \nu, \varepsilon_1} := \int_{\mathcal{T}_{\varepsilon_1}^x} L_t^{x, \varepsilon_1}(x) \nu(dx) + \sum_{x \in \mathcal{B}^x} \nu(B_{\ell^1}(x, \varepsilon_1)) \sup_{y \in \partial B_{\ell^1}(x, \varepsilon_1) \cap \mathcal{T}^x} L_t^{x, \varepsilon_1}(y), \tag{3.10}$$

where $\mathcal{T}_{\varepsilon_1}^x := \mathcal{T}^x \setminus \mathcal{B}_{\varepsilon_1}^x$. That $A^{x, \nu, \varepsilon_1}$ is uniformly close to $A^{x, \nu}$ is confirmed by the following lemma.

Lemma 3.4. For $t_0, \varepsilon > 0$,

$$\lim_{\varepsilon_1 \rightarrow 0} P_0^x \left(\sup_{t \in [0, t_0]} |A_t^{x, \nu} - A_t^{x, \nu, \varepsilon_1}| > \varepsilon \right) = 0.$$

Proof. Since $\{s : X_s \in \mathcal{T}_{\varepsilon_1}^x\} \subseteq I$, the definition of L^{x, ε_1} in (3.9) implies that $L_t^{x, \varepsilon_1}(x) = L_{\tau^{x, \varepsilon_1}(t)}^x(x)$ for every $x \in \mathcal{T}_{\varepsilon_1}^x$ and $t \geq 0$. Consequently, we can check that

$$\sup_{t \in [0, t_0]} |A_t^{x, \nu} - A_t^{x, \nu, \varepsilon_1}| \leq \sup_{t \in [0, t_0]} \sup_{\{x, y \in \mathcal{T}^x : d_{\ell^1}(x, y) \leq 2\varepsilon_1\}} |L_t^x(x) - L_{\tau^{x, \varepsilon_1}(t)}^x(y)|,$$

where we apply the fact that ν is a probability measure. The P_0^x -a.s. joint continuity of the local times L^x allows us to deduce the result from (3.4).

For large n , we obtain objects analogous to L^{x, ε_1} and $A^{x, \nu, \varepsilon_1}$ by setting

$$L_t^{x_n, \varepsilon_1}(x) := \int_0^{\tau^{x_n, \varepsilon_1}(t)} \mathbf{1}_{\{s \in I_n\}} d_s L_s^{x_n}(x)$$

for $x \in \mathcal{T}^{x_n}$ and $t \geq 0$, where $I_n = \mathbb{R} \setminus \bigcup_{i=0}^{\infty} (\nu_i^n, \varsigma_i^n)$ is defined from the stopping times introduced in (3.6) and (3.7), then letting $A^{x_n, \nu_n, \varepsilon_1}$ be defined by

$$A_t^{x_n, \nu_n, \varepsilon_1} := \int_{\mathcal{T}_{\varepsilon_1}^{x_n}} L_t^{x_n, \varepsilon_1}(x) \nu_n(dx) + \sum_{x \in \mathcal{B}^{x_n}} \nu_n(B_{\ell^1}(x, \varepsilon_1)) \sup_{y \in \partial B_{\ell^1}(x, \varepsilon_1) \cap \mathcal{T}^{x_n}} L_t^{x_n, \varepsilon_1}(y), \tag{3.11}$$

where $\mathcal{T}_{\varepsilon_1}^{x_n} := \mathcal{T}^{x_n} \setminus \mathcal{B}_{\varepsilon_1}^{x_n}$ and the summands where $B_{\ell^1}(x, \varepsilon_1) \cap \mathcal{T}^{x_n} = \emptyset$ are assumed to be

equal to 0. To check that $A^{x_n, v_n, \varepsilon_1}$ is close to A^{x_n, v_n} uniformly in n , we will apply the following tightness result for the local times $(L^{x_n})_{n \geq 1}$.

Lemma 3.5. *If $x_n \rightarrow x$ then, for $t_0, \varepsilon > 0$,*

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P_0^{x_n} \left(\sup_{\{x, y \in \mathcal{T}^{x_n} : d_{\ell^1}(x, y) \leq \delta\}} \sup_{\{s, t \in [0, t_0] : |s-t| \leq \delta\}} |L_s^{x_n}(x) - L_t^{x_n}(y)| > \varepsilon \right) = 0. \tag{3.12}$$

Proof. First note that any two vertices $x, y \in \mathcal{T}^{x_n}$ are contained in a set of the form $[0, x_n^{(i)}]^{x_n}$ or $[[x_n^{(i)}, x_n^{(j)}]^{x_n}$ for some $i, j \in \{1, \dots, k\}, i \neq j$, where, for $x', y' \in \mathcal{T}^{x_n}$, we write $[[x', y']^{x_n}$ to represent the path from x' to y' in \mathcal{T}^{x_n} . Secondly, writing $b^{x_n}(x, x', x'')$ to represent the branch point of x, x' , and x'' in \mathcal{T}^{x_n} , applying the assumption $x_n \rightarrow x$, and Equation (5) of [2], it is possible to check that there exists a finite integer n_0 such that

$$\eta(n) := \sup_{i, j=1, \dots, k} d_{\ell^1}(b^{x_n}(0, x_n^{(i)}, x_n^{(j)}), b^x(0, x^{(i)}, x^{(j)})) < \frac{\varepsilon_0}{2}$$

for $n \geq n_0$, where ε_0 is defined in (3.2). This implies that $d_{\ell^1}(0, x_n^{(i)}) > \varepsilon_0$ and also that $d_{\ell^1}(b^{x_n}(0, x_n^{(i)}, x_n^{(j)}), x_n^{(i)}) > \varepsilon_0$ for every $i \neq j$, whenever $n \geq n_0$. For $n \geq n_0$, it is an elementary exercise to deduce from these two facts the existence of a collection of paths $([[a_i, b_i]]^{x_n})_{i \in \mathcal{I}_n}$, where $d_{\ell^1}(a_i, b_i) = \varepsilon_0$ and $d_{\ell^1}(b^{x_n}(0, a_i, b_i), a_i) \in \{0, \varepsilon_0/2, \varepsilon_0\}$, such that, if $x, y \in \mathcal{T}^{x_n}$ and $d_{\ell^1}(x, y) < \varepsilon_0/2$, then x and y are both contained in a single set of the form $[[a_i, b_i]]^{x_n}$ for some $i \in \mathcal{I}_n$. Moreover, the collection $([[a_i, b_i]]^{x_n})_{i \in \mathcal{I}_n}$ can be chosen in such a way that $\#\mathcal{I}_n$ is uniformly bounded in n .

Suppose now that $n \geq n_0$ is large enough so that a cover of the form described above exists and $\delta < \varepsilon_0/2$. For $i \in \mathcal{I}_n$, define

$$A_i^t := \int_{[[a_i, b_i]]^{x_n}} L_t^{x_n}(x) \lambda^{x_n}(dx)$$

and $\tau^i(t) := \inf\{s : A_s^i > t\}$. Similarly to Lemma 2.4, the law of the process X^i , where $X_t^i := X_{\tau^i(t)}^{x_n}$, under $P_0^{x_n}$ is equal to the law of the Brownian motion on the measured compact real tree $([[a_i, b_i]]^{x_n}, \lambda^{x_n}([a_i, b_i]^{x_n} \cap \cdot))$ started from $b^{x_n}(0, a_i, b_i)$. Moreover, the local times of X^i are given by $L_t^i(x) = L_{\tau^i(t)}^{x_n}(x)$ for $x \in [[a_i, b_i]]^{x_n}$ and $t \geq 0$ (cf. [7, Lemma 3.4]). Now, by our choice of the intervals $([[a_i, b_i]]^{x_n})_{i \in \mathcal{I}_n}$ we have

$$\begin{aligned} & \sup_{\{x, y \in \mathcal{T}^{x_n} : d_{\ell^1}(x, y) \leq \delta\}} \sup_{\{s, t \in [0, t_0] : |s-t| \leq \delta\}} |L_s^{x_n}(x) - L_t^{x_n}(y)| \\ & \leq \sum_{i \in \mathcal{I}_n} \sup_{\{x, y \in [[a_i, b_i]]^{x_n} : d_{\ell^1}(x, y) \leq \delta\}} \sup_{\{s, t \in [0, t_0] \cap \overline{\tau^i(\mathbb{R}_+)}\} : |s-t| \leq \delta} |L_s^{x_n}(x) - L_t^{x_n}(y)|, \end{aligned}$$

where the condition $s, t \in \overline{\tau^i(\mathbb{R}_+)}$ is justified by the observation that, for $x \in [[a_i, b_i]]^{x_n}$, the local time $L_t^{x_n}(x)$ increases only on the set $\{t \geq 0 : X_t = x\} \subseteq \overline{\tau^i(\mathbb{R}_+)}$. Furthermore, simple continuity arguments allow us to replace $\overline{\tau^i(\mathbb{R}_+)}$ by $\tau^i(\mathbb{R}_+)$. Since $|A_s^i - A_t^i| \leq |s - t|$, we have $|\tau^i(s) - \tau^i(t)| \geq |s - t|$, and, consequently, we obtain from this bound

$$\begin{aligned} & \sup_{\{x, y \in \mathcal{T}^{x_n} : d_{\ell^1}(x, y) \leq \delta\}} \sup_{\{s, t \in [0, t_0] : |s-t| \leq \delta\}} |L_s^{x_n}(x) - L_t^{x_n}(y)| \\ & \leq \sum_{i \in \mathcal{I}_n} \sup_{\{x, y \in [[a_i, b_i]]^{x_n} : d_{\ell^1}(x, y) \leq \delta\}} \sup_{\{s, t \in [0, t_0] : |s-t| \leq \delta\}} |L_s^i(x) - L_t^i(y)|, \end{aligned}$$

where we apply the representation of $(L_t^i(x))_{x \in [[a_i, b_i]]^{x_n}, t \geq 0}$ noted above. All the measured real trees $\{([a_i, b_i]]^{x_n}, \lambda^{x_n}([a_i, b_i]]^{x_n} \cap \cdot)\}_{i \in \mathcal{I}_n}$ are equivalent and, by construction, the Brownian motion X^{x_n} first hits $[a_i, b_i]^{x_n}$ at $b^{x_n}(0, a_i, b_i)$; hence, it follows that the probability on the left-hand side of (3.12) is bounded above by

$$\# \mathcal{I}_n \sup_{x \in (0, \varepsilon_0/2)} P_x^{[0, \varepsilon_0]} \left(\sup_{\{y, z \in [0, \varepsilon_0]: |y-z| \leq \delta\}} \sup_{\{s, t \in [0, t_0]: |s-t| \leq \delta\}} |L_s^{[0, \varepsilon_0]}(y) - L_t^{[0, \varepsilon_0]}(z)| > \varepsilon (\# \mathcal{I}_n)^{-1} \right),$$

where $P_x^{[0, \varepsilon_0]}$ is the law of the Brownian motion on the real interval $[0, \varepsilon_0]$ equipped with the one-dimensional Hausdorff measure and the $(L_t^{[0, \varepsilon_0]}(x))_{x \in [0, \varepsilon_0], t \geq 0}$ are the local times of this process. Recalling that $\# \mathcal{I}_n$ is uniformly bounded in n , Lemma 2.2 allows it to be deduced that the above expression decays to 0 as first $n \rightarrow \infty$ and then $\delta \rightarrow 0$.

By following a similar proof to that of Lemma 3.4, combining the previous lemma with (3.8) allows it to be deduced that $A^{x_n, \nu_n, \varepsilon_1}$ does indeed approximate A^{x_n, ν_n} , as desired.

Lemma 3.6. *If $x_n \rightarrow x$ then, for $t_0, \varepsilon > 0$,*

$$\lim_{\varepsilon_1 \rightarrow 0} \limsup_{n \rightarrow \infty} P_0^{x_n} \left(\sup_{t \in [0, t_0]} |A_t^{x_n, \nu_n} - A_t^{x_n, \nu_n, \varepsilon_1}| > \varepsilon \right) = 0.$$

Our next step is to demonstrate that X^{x_n, ε_1} converges to X^{x, ε_1} and $A^{x_n, \nu_n, \varepsilon_1}$ is close to $A^{x, \nu, \varepsilon_1}$, which we do by applying a particular sample path construction of the processes from their excursions between jump times. Let us start by introducing some further notation. Set $\mathcal{T}_{\varepsilon_1/2}^x := \mathcal{T}^x \setminus \mathcal{B}_{\varepsilon_1/2}^x$, which consists of a finite number of connected components, each a closed line segment with endpoints in the finite set $\partial \mathcal{T}_{\varepsilon_1/2}^x$. For $x \in \partial \mathcal{T}_{\varepsilon_1/2}^x$, write

$$N_{\varepsilon_1}^x(x) := \{y \in \partial \mathcal{T}_{\varepsilon_1}^x : [[x, y]]^x \cap \partial \mathcal{T}_{\varepsilon_1}^x = \{y\}\}$$

to represent the collection of ‘nearest neighbours’ of x in $\partial \mathcal{T}_{\varepsilon_1}^x$, where $[[x, y]]^x$ is the path from x to y in \mathcal{T}^x . For a pictorial representation of the definitions at a typical branch point, see Figure 1. For $x \in \mathcal{T}_{\varepsilon_1/2}^x$, we will denote by $C_{\varepsilon_1}^x(x)$ the connected component of $\mathcal{T}_{\varepsilon_1/2}^x$ containing x .

Let $(\zeta_i)_{i \geq 0}$ be the jump times of X^{x, ε_1} ; by convention we set $\zeta_0 = 0$. Applying the definition of Brownian motion on a dendrite, conditional on $X_{\zeta_i}^{x, \varepsilon_1}$, the path segment $(X_{\zeta_i+t}^{x, \varepsilon_1})_{t \in [0, \zeta_{i+1}-\zeta_i]}$ is distributed precisely as a standard one-dimensional Brownian motion on the line segment $C_{\varepsilon_1}^x(X_{\zeta_i}^{x, \varepsilon_1})$, started from $X_{\zeta_i}^{x, \varepsilon_1}$ and run until it hits $\partial \mathcal{T}_{\varepsilon_1/2}^x$. The local times

$$(L_{\zeta_i+t}^{x, \varepsilon_1}(x))_{x \in C_{\varepsilon_1}^x(X_{\zeta_i}^{x, \varepsilon_1}), t \in [0, \zeta_{i+1}-\zeta_i]}$$

are distributed exactly as the local times of the same one-dimensional Brownian motion; outside $C_{\varepsilon_1}^x(X_{\zeta_i}^{x, \varepsilon_1})$, the local times L^{x, ε_1} do not increase on the time interval $[\zeta_i, \zeta_{i+1})$. Moreover, the strong Markov property implies that, conditional on $X_{\zeta_i}^{x, \varepsilon_1}, (X_{\zeta_i+t}^{x, \varepsilon_1}, (L_{\zeta_i+t}^{x, \varepsilon_1}(x))_{x \in \mathcal{T}^x})_{t \in [0, \zeta_{i+1}-\zeta_i]}$ is independent of the σ -algebra generated by $(X_t^{x, \varepsilon_1})_{t \leq \zeta_i}$. At discontinuities, the process satisfies the following transition law:

$$p^{x, \varepsilon_1}(x, y) := P_0^x(X_{\zeta_i}^{x, \varepsilon_1} = y \mid X_{\zeta_i^-}^{x, \varepsilon_1} = x) = \begin{cases} \frac{1 + \# N_{\varepsilon_1}^x(x)}{2\# N_{\varepsilon_1}^x(x)} & \text{if } y \in C_{\varepsilon_1}^x(x), \\ \frac{1}{2\# N_{\varepsilon_1}^x(x)} & \text{otherwise,} \end{cases}$$

for $x \in \partial \mathcal{T}_{\varepsilon_1/2}^x$ and $y \in N_{\varepsilon_1}^x(x)$. As a consequence of this description, we can construct

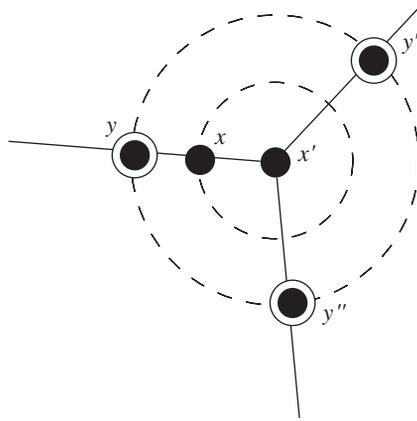


FIGURE 1: Example of $x \in \partial\mathcal{T}_{\varepsilon_1/2}^x$, $x' \in \mathcal{B}^x$, and $N_{\varepsilon_1}^x(x) = \{y, y', y''\}$.

$(X^{x,\varepsilon_1}, L^{x,\varepsilon_1})$ from a countable collection

$$\{(\alpha^{x,i})_{x \in \partial\mathcal{T}_{\varepsilon_1/2}^x}, (\beta^{y,i}, \gamma^{y,i}, \xi^{y,i})_{y \in \partial\mathcal{T}_{\varepsilon_1}^x}\}_{i \geq 0} \tag{3.13}$$

of random variables built on an underlying probability space with probability measure P that satisfy the following properties.

- The random variables $(\alpha^{x,i})_{x \in \partial\mathcal{T}_{\varepsilon_1/2}^x, i \geq 0}$ are independent. The random variable $\alpha^{x,i}$ is $N_{\varepsilon_1}^x(x)$ -valued and distributed according to the law determined by $p^{x,\varepsilon_1}(x, \cdot)$.
- The triples $\{(\beta^{y,i}, \gamma^{y,i}, \xi^{y,i})\}_{y \in \partial\mathcal{T}_{\varepsilon_1}^x, i \geq 0}$ are independent of each other and of the collection $(\alpha^{x,i})_{x \in \partial\mathcal{T}_{\varepsilon_1/2}^x, i \geq 0}$. The process $\beta^{y,i} = (\beta_t^{y,i})_{t \geq 0}$ is a Brownian motion on the line segment $C_{\varepsilon_1}^x(y)$ (equipped with the appropriate restriction of λ^x) started from y ; $\gamma^{y,i} := \inf\{s : \beta_s^{y,i} \in \partial\mathcal{T}_{\varepsilon_1/2}^x\}$; and $\xi^{y,i} = (\xi_t^{y,i}(x))_{x \in C_{\varepsilon_1}^x(y), t \geq 0}$ are the jointly continuous local times of $\beta^{y,i}$.

We now define a càdlàg process $\tilde{X}^{x,\varepsilon_1}$, its times of discontinuities $(\tilde{\zeta}_i)_{i \geq 1}$, and its local times $\tilde{L}^{x,\varepsilon_1}$ from these random variables. First, set $\tilde{X}_0^{x,\varepsilon_1}$ to be the unique point in \mathcal{T}^x satisfying $d_{\ell_1}(0, \tilde{X}_0^{x,\varepsilon_1}) = \varepsilon_1$ (such a point is well defined, because we are assuming that 0 is a leaf of \mathcal{T}^x and ε_1 is less than ε_0 , as defined in (3.2)), $\tilde{\zeta}_0 = 0$, and $\tilde{L}_0^{x,\varepsilon_1}(x) = 0$ for every $x \in \mathcal{T}^x$. Given $(\tilde{X}_t^{x,\varepsilon_1}, (\tilde{L}_t^{x,\varepsilon_1}(x))_{x \in \mathcal{T}^x})_{t \in [0, \tilde{\zeta}_i]}$, define

$$\begin{aligned} \tilde{\zeta}_{i+1} &:= \tilde{\zeta}_i + \gamma_{\tilde{\zeta}_i}^{\tilde{X}_{\tilde{\zeta}_i}^{x,\varepsilon_1}, i}, \\ (\tilde{X}_t^{x,\varepsilon_1})_{t \in (\tilde{\zeta}_i, \tilde{\zeta}_{i+1})} &:= (\beta_{t-\tilde{\zeta}_i}^{\tilde{X}_{\tilde{\zeta}_i}^{x,\varepsilon_1}, i})_{t \in (\tilde{\zeta}_i, \tilde{\zeta}_{i+1})}, \quad \tilde{X}_{\tilde{\zeta}_{i+1}}^{x,\varepsilon_1} := \alpha_{\tilde{\zeta}_{i+1}}^{\tilde{X}_{\tilde{\zeta}_i}^{x,\varepsilon_1}, i}, \end{aligned}$$

and also

$$(\tilde{L}_t^{x,\varepsilon_1}(x))_{t \in (\tilde{\zeta}_i, \tilde{\zeta}_{i+1})} := \begin{cases} (\tilde{L}_{\tilde{\zeta}_i}^{x,\varepsilon_1}(x) + \xi_{t-\tilde{\zeta}_i}^{\tilde{X}_{\tilde{\zeta}_i}^{x,\varepsilon_1}, i}(x))_{t \in (\tilde{\zeta}_i, \tilde{\zeta}_{i+1})} & \text{if } x \in C_{\varepsilon_1}^x(\tilde{X}_{\tilde{\zeta}_i}^{x,\varepsilon_1}), \\ (\tilde{L}_{\tilde{\zeta}_i}^{x,\varepsilon_1}(x))_{t \in (\tilde{\zeta}_i, \tilde{\zeta}_{i+1})} & \text{otherwise.} \end{cases}$$

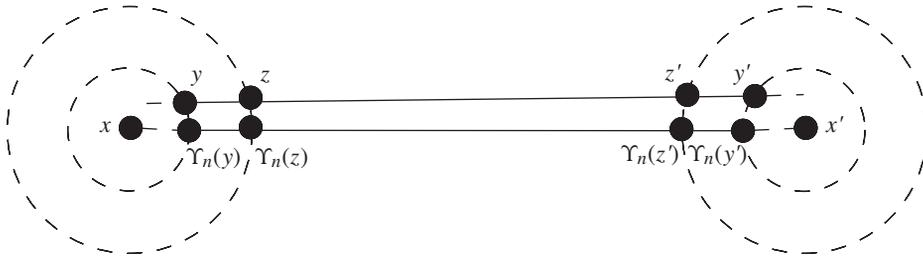


FIGURE 2: Example of a homeomorphism between line segments.

Applying the above description of $(X^{x,\varepsilon_1}, L^{x,\varepsilon_1})$, it is straightforward to check that the law of $(\tilde{X}^{x,\varepsilon_1}, \tilde{L}^{x,\varepsilon_1})$ under P is identical to the law of $(X^{x,\varepsilon_1}, L^{x,\varepsilon_1})$ under P_0^x .

We continue by describing the structure of $\mathcal{T}_{\varepsilon_1/2}^{x_n}$. Define $\eta(n)$ as in the proof of Lemma 3.5. The assumption $x_n \rightarrow x$ implies that there exists a finite integer n_0 such that $\eta(n) < \varepsilon_1/2$ for $n \geq n_0$, which implies in turn that $\mathcal{T}_{\varepsilon_1/2}^{x_n}$ is homeomorphic to $\mathcal{T}_{\varepsilon_1/2}^x$ for $n \geq n_0$. In particular, if we suppose that x and x' are neighbours in \mathcal{B}^x , by which we mean that $x, x' \in \mathcal{B}^x$, $x \neq x'$ and $[[x, x']]^x \cap \mathcal{B}^x = \{x, x'\}$, then there exists a unique connected component of $\mathcal{T}_{\varepsilon_1/2}^{x_n}$ which is a closed line segment with endpoints in $\partial B_{\varepsilon_1/2}(x, \varepsilon_1/2)$ and $\partial B_{\varepsilon_1/2}(x', \varepsilon_1/2)$; moreover, every connected component of $\mathcal{T}_{\varepsilon_1/2}^{x_n}$ can be represented in this way. We define a homeomorphism from $\mathcal{T}_{\varepsilon_1/2}^{x_n}$ to $\mathcal{T}_{\varepsilon_1/2}^x$ that maps the endpoint of such a line segment, y say, in $\partial B_{\varepsilon_1/2}(x, \varepsilon_1/2)$ to the unique point $\Upsilon_n(y)$ in $[[x, x']]^x$ that satisfies $d_{\ell^1}(x, \Upsilon_n(y)) = \varepsilon_1/2$; maps the point of the line segment at a distance $\varepsilon_1/2$ from y , z say, to the unique point $\Upsilon_n(z)$ in $[[x, x']]^x$ that satisfies $d_{\ell^1}(x, \Upsilon_n(z)) = \varepsilon_1$; and Υ_n is extended by linear interpolation on the line segments between points for which we have not already defined it. Figure 2 depicts a typical configuration on a line segment.

We now assume that $\eta(n) < \varepsilon_1/2$. The transition law at the discontinuities $(\zeta_i^n)_{i \geq 1}$ of X^{x_n, ε_1} is given by

$$p^{x_n, \varepsilon_1}(x, y) := P_0^{x_n}(X_{\zeta_i^n}^{x_n, \varepsilon_1} = y \mid X_{\zeta_i^{n-}}^{x_n, \varepsilon_1} = x),$$

where $x \in \partial \mathcal{T}_{\varepsilon_1/2}^{x_n}$ and $y \in N_{\varepsilon_1}^{x_n}(x) := \Upsilon_n^{-1}(N_{\varepsilon_1}^x(\Upsilon_n(x)))$. We use this to define a càdlàg process $\tilde{X}^{x_n, \varepsilon_1}$ for each n from a countable collection

$$\{(\alpha^{n,x,i})_{x \in \partial \mathcal{T}_{\varepsilon_1/2}^{x_n}}, (\beta^{n,y,i}, \gamma^{n,y,i}, \xi^{n,y,i})_{y \in \partial \mathcal{T}_{\varepsilon_1}^{x_n}}\}_{i \geq 0}$$

of random variables built on our underlying probability space with probability measure P that satisfy the following properties, where we apply the notation $C_{\varepsilon_1}^{x_n}(x)$ to represent the connected component of $\mathcal{T}_{\varepsilon_1/2}^{x_n}$ containing $x \in \mathcal{T}_{\varepsilon_1/2}^{x_n}$.

- The random variables $(\alpha^{n,x,i})_{x \in \partial \mathcal{T}_{\varepsilon_1/2}^{x_n}, i \geq 0}$ are independent. The random variable $\alpha^{n,x,i}$ is $N_{\varepsilon_1}^{x_n}(x)$ -valued and distributed according to $p^{x_n, \varepsilon_1}(x, \cdot)$.
- The triples $\{(\beta^{n,y,i}, \gamma^{n,y,i}, \xi^{n,y,i})\}_{y \in \partial \mathcal{T}_{\varepsilon_1}^{x_n}, i \geq 0}$ are independent of each other and of the collection $(\alpha^{n,x,i})_{x \in \partial \mathcal{T}_{\varepsilon_1/2}^{x_n}, i \geq 0}$. The process $\beta^{n,y,i} = (\beta_t^{n,y,i})_{t \geq 0}$ is a Brownian motion on the line segment $C_{\varepsilon_1}^{x_n}(y)$ (equipped with the appropriate restriction of λ^{x_n}) started

from y ; $\gamma^{n,y,i} := \inf\{s : \beta_s^{n,y,i} \in \partial\mathcal{T}_{\varepsilon_1/2}^{x_n}\}$; and $\xi^{n,y,i} = (\xi_t^{n,y,i}(x))_{x \in C_{\varepsilon_1}^{x_n}(y), t \geq 0}$ are the jointly continuous local times of $\beta^{n,y,i}$.

Constructing $(\tilde{X}^{x_n, \varepsilon_1}, \tilde{L}^{x_n, \varepsilon_1})$ from these random variables, similarly to the definition of

$$(\tilde{X}^{x, \varepsilon_1}, \tilde{L}^{x, \varepsilon_1}),$$

results in a process whose law under P is equal to that of $(X^{x_n, \varepsilon_1}, L^{x_n, \varepsilon_1})$ under $P_0^{x_n}$. The following simple lemma is crucial in proving that X^{x_n, ε_1} converges to X^{x, ε_1} and $A^{x_n, \nu_n, \varepsilon_1}$ is close to $A^{x, \nu, \varepsilon_1}$, as we do in the subsequent result. Note that the topology on the countable collections with the same index sets we consider is that generated by uniform convergence of finite subcollections (with respect to the appropriate product topology).

Lemma 3.7. *If $x_n \rightarrow x$, the countable collection of random variables consisting of*

$$(\Upsilon_n(\alpha^n, \Upsilon_n^{-1}(x), i))_{x \in \partial\mathcal{T}_{\varepsilon_1/2}^x, i \geq 0}$$

and

$$(\beta^{n, \Upsilon_n^{-1}(y), i}, \gamma^{n, \Upsilon_n^{-1}(y), i}, (\xi^{n, \Upsilon_n^{-1}(y), i}(\Upsilon_n^{-1}(x))))_{x \in C_{\varepsilon_1}^x(y), y \in \partial\mathcal{T}_{\varepsilon_1}^x, i \geq 0}$$

converges in distribution to the collection in (3.13).

Proof. To deduce that $\Upsilon_n(\alpha^n, \Upsilon_n^{-1}(x), i)$ converges in distribution to $\alpha^{x,i}$, it suffices to check that

$$p^{x_n, \varepsilon_1}(\Upsilon_n^{-1}(x), \Upsilon_n^{-1}(y)) \rightarrow p^{x, \varepsilon_1}(x, y)$$

for each $x \in \partial\mathcal{T}_{\varepsilon_1/2}^x$ and $y \in N_{\varepsilon_1}^x(x)$. The proof that this is true involves an elementary electrical network (or harmonic analysis) calculation and is omitted (that the Brownian motion on a real tree can be constructed in terms of electrical resistance networks is guaranteed by [7, Proposition 2.2], and a detailed study of harmonic analysis on such spaces appears in [17]).

The convergence of the triple

$$(\beta^{n, \Upsilon_n^{-1}(y), i}, \xi^{n, \Upsilon_n^{-1}(y), i}, (\gamma^{n, \Upsilon_n^{-1}(y), i}(\Upsilon_n^{-1}(x))))_{x \in C_{\varepsilon_1}^x(y)}$$

to $(\beta^{y,i}, \gamma^{y,i}, \xi^{y,i})$ is a simple consequence of the fact that

$$((B_t)_{t \geq 0}, h(-a_n, b_n), (L_t(x))_{x \in \mathbb{R}, t \geq 0}) \rightarrow ((B_t)_{t \geq 0}, h(-a, b), (L_t(x))_{x \in \mathbb{R}, t \geq 0})$$

in distribution whenever $a_n \rightarrow a$ and $b_n \rightarrow b$ for some $a, b > 0$, where B is a standard Brownian motion in \mathbb{R} started from 0, $h(-a, b)$ is the hitting time of $\{-a, b\}$ by B , and L are the jointly continuous local times of B . Note that to map this result into our setting, we apply the fact that $\sup_{x \in \mathcal{T}^{x_n}} d_{\ell^1}(\Upsilon_n(x), x) \rightarrow 0$ as $n \rightarrow \infty$.

Lemma 3.8. *Suppose that $x_n \rightarrow x$. If F is a continuous bounded function on $D(\mathbb{R}_+, \ell^1)$ then*

$$\lim_{n \rightarrow \infty} E_0^{x_n}(F(X^{x_n, \varepsilon_1})) = E_0^x(F(X^{x, \varepsilon_1})). \tag{3.14}$$

Moreover, if $\nu_n \rightarrow \nu$ weakly as probability measures on ℓ^1 and F is a continuous bounded function on $D(\mathbb{R}_+, \ell^1) \times C(\mathbb{R}_+, \mathbb{R}_+)$, then

$$\lim_{\varepsilon_1 \rightarrow 0} \limsup_{n \rightarrow \infty} |E_0^{x_n}(F(X^{x_n, \varepsilon_1}, A^{x_n, \nu_n, \varepsilon_1})) - E_0^x(F(X^{x, \varepsilon_1}, A^{x, \nu, \varepsilon_1}))| = 0. \tag{3.15}$$

Proof. By separability, it is possible to assume that all the relevant random variables are constructed in such a way that the convergence of Lemma 3.7 holds P-a.s. It will follow easily from this that $\tilde{X}^{x_n, \varepsilon_1} \rightarrow \tilde{X}^{x, \varepsilon_1}$ in $D(\mathbb{R}_+, \ell^1)$, P-a.s., if we can show that the number of discontinuities that $\tilde{X}^{x, \varepsilon_1}$ admits in any finite time interval is finite. By continuity, there exists an $\varepsilon > 0$ such that $P(\inf_{y \in \partial \mathcal{T}_{\varepsilon_1}^x} \gamma^{y, i} > \varepsilon) > 0$. Thus, the strong law of large numbers implies that

$$\lim_{i \rightarrow \infty} \tilde{\zeta}_i \geq \sum_{i=0}^{\infty} \inf_{y \in \partial \mathcal{T}_{\varepsilon_1}^x} \gamma^{y, i} = \infty, \tag{3.16}$$

P-a.s., and, hence, there can indeed only be a finite number of discontinuities of $\tilde{X}^{x, \varepsilon_1}$ in any finite time interval. Due to the equivalence of the laws of $\tilde{X}^{x_n, \varepsilon_1}$, $\tilde{X}^{x, \varepsilon_1}$ and X^{x_n, ε_1} , X^{x, ε_1} under the appropriate probability measures, this yields the convergence result in (3.14).

Let us now suppose that we have a realisation of random variables such that the convergence of Lemma 3.7 occurs and (3.16) holds. Under these conditions, it is possible to check that, for any $t_0 > 0$,

$$\varepsilon(n) := \sup_{t \in [0, t_0]} \sup_{x \in \mathcal{T}_{\varepsilon_1}^x} |\tilde{L}_t^{x_n, \varepsilon_1}(\Upsilon_n^{-1}(x)) - \tilde{L}_t^{x, \varepsilon_1}(x)| \rightarrow 0 \tag{3.17}$$

as $n \rightarrow \infty$. Moreover, we can define a function $(\tilde{L}_t^{x, \varepsilon_1}(x))_{x \in \mathcal{T}^x \cup \mathcal{B}_{\varepsilon_1}^x, t \geq 0}$ that is jointly continuous and agrees with $\tilde{L}_t^{x, \varepsilon_1}(x)$ for $x \in \mathcal{T}_{\varepsilon_1}^x$. If $x \in B_{\ell^1}(y, \varepsilon_1)$ for some $y \in \mathcal{B}^x$ then we set

$$\tilde{L}_t^{x, \varepsilon_1}(x) := \sum_{z \in \partial B_{\ell^1}(y, \varepsilon_1) \cap \mathcal{T}^x} d_{\ell^1}(x, z)^{-1} \tilde{L}_t^{x, \varepsilon_1}(z) \Big/ \sum_{z \in \partial B_{\ell^1}(y, \varepsilon_1) \cap \mathcal{T}^x} d_{\ell^1}(x, z)^{-1},$$

so that $\tilde{L}_t^{x, \varepsilon_1}(x)$ is a weighted average of the points in \mathcal{T}^x on the boundary of $B_{\ell^1}(y, \varepsilon_1)$. Furthermore, for large n , let $\phi_n : \mathcal{T}^{x_n} \cup \mathcal{B}_{\varepsilon_1}^{x_n} \rightarrow \mathcal{T}^x \cup \mathcal{B}_{\varepsilon_1}^x$ be defined by setting $\phi_n(x) = \Upsilon_n(x)$ on $\mathcal{T}_{\varepsilon_1}^{x_n}$ and $\phi_n(x) = x$ otherwise, and note that $\nu_n \circ \phi_n^{-1} \rightarrow \nu$ weakly as probability measures on $\mathcal{T}^x \cup \mathcal{B}_{\varepsilon_1}^x$. If we construct $\tilde{A}_t^{x_n, \nu_n, \varepsilon_1}$ and $\tilde{A}_t^{x, \nu, \varepsilon_1}$ from $\tilde{L}_t^{x_n, \varepsilon_1}$ and $\tilde{L}_t^{x, \varepsilon_1}$ analogously to the definitions of $A_t^{x_n, \nu_n, \varepsilon_1}$ and $A_t^{x, \nu, \varepsilon_1}$ in (3.10) and (3.11), respectively, then we can apply the above notation to deduce that, for any $0 \leq t_1 \leq \dots \leq t_m \leq t_0$,

$$\begin{aligned} & \sup_{t \in \{t_1, \dots, t_m\}} |\tilde{A}_t^{x_n, \nu_n, \varepsilon_1} - \tilde{A}_t^{x, \nu, \varepsilon_1}| \\ & \leq \varepsilon(n) + 2 \sup_{t \in [0, t_0]} \sup_{\{x, y \in \mathcal{T}^x \cup \mathcal{B}_{\varepsilon_1}^x : d_{\ell^1}(x, y) \leq 2\varepsilon_1\}} |\tilde{L}_t^{x, \varepsilon_1}(x) - \tilde{L}_t^{x, \varepsilon_1}(y)| \\ & \quad + \sup_{t \in \{t_1, \dots, t_m\}} \left| \int_{\mathcal{T}^x \cup \mathcal{B}_{\varepsilon_1}^x} \tilde{L}_t^{x, \varepsilon_1}(x) \nu_n \circ \phi_n^{-1}(dx) - \int_{\mathcal{T}^x \cup \mathcal{B}_{\varepsilon_1}^x} \tilde{L}_t^{x, \varepsilon_1}(x) \nu(dx) \right|. \end{aligned}$$

By (3.17), the first term in the upper bound decays to 0 as $n \rightarrow \infty$. The third term also converges to 0, since $\nu_n \circ \phi_n^{-1} \rightarrow \nu$. It follows from this and the definition of $\tilde{L}^{x, \varepsilon_1}$ that, for any $\varepsilon > 0$,

$$\begin{aligned} & P\left(\limsup_{n \rightarrow \infty} \sup_{t \in \{t_1, \dots, t_m\}} |\tilde{A}_t^{x_n, \nu_n, \varepsilon_1} - \tilde{A}_t^{x, \nu, \varepsilon_1}| > \varepsilon\right) \\ & \leq P_0^x \left(2 \sup_{t \in [0, t_0]} \sup_{\{x, y \in \mathcal{T}_{\varepsilon_1}^x : d_{\ell^1}(x, y) \leq 4\varepsilon_1\}} |\tilde{L}_t^{x, \varepsilon_1}(x) - \tilde{L}_t^{x, \varepsilon_1}(y)| > \varepsilon \right) \\ & = P_0^x \left(2 \sup_{t \in [0, t_0]} \sup_{\{x, y \in \mathcal{T}_{\varepsilon_1}^x : d_{\ell^1}(x, y) \leq 4\varepsilon_1\}} |L_{\tau^{x, \varepsilon_1}(t)}^x(x) - L_{\tau^{x, \varepsilon_1}(t)}^x(y)| > \varepsilon \right), \end{aligned}$$

where the equality follows from the observation made in the proof of Lemma 3.4 that $L_t^{x,\varepsilon_1}(x) = L_{\tau_{x,\varepsilon_1}(t)}^x(x)$ for $x \in \mathcal{T}_{\varepsilon_1}^x$. By applying the P_0^x -a.s. continuity of the local times L^x and (3.4), this implies that

$$\lim_{\varepsilon_1 \rightarrow 0} P\left(\limsup_{n \rightarrow \infty} \sup_{t \in \{t_1, \dots, t_m\}} |\tilde{A}_t^{x_n, \nu_n, \varepsilon_1} - \tilde{A}_t^{x, \nu, \varepsilon_1}| > \varepsilon\right) = 0.$$

In conjunction with the result of the opening paragraph of the proof and Lemma 3.6, this will imply the convergence in (3.15) once the tightness of $(A^{x_n, \nu_n})_{n \geq 1}$ in $C(\mathbb{R}_+, \mathbb{R}_+)$ has been checked. However, this is an easy corollary of Lemma 3.5, and so the proof is complete.

The main result of this section, which we can now state precisely, is an immediate consequence of Lemmas 3.2, 3.3, 3.4, 3.6, and 3.8.

Proposition 3.1. *Suppose that $x_n \rightarrow x$ and $\nu_n \rightarrow \nu$ weakly as probability measures on ℓ^1 . If F is a continuous bounded function on $C(\mathbb{R}_+, \ell^1) \times C(\mathbb{R}_+, \mathbb{R}_+)$ then*

$$\lim_{n \rightarrow \infty} E_0^{x_n}(F(X^{x_n}, A^{x_n, \nu_n})) = E_0^x(F(X^x, A^{x, \nu})).$$

4. Simple random walk convergence

The results of the two previous sections allow us to prove our main scaling limit theorems by relatively straightforward adaptations of the proofs in [7]. Since most of the objects we study here were also considered in [7], we will be brief in introducing them, and refer the reader to [7] for further details.

Let $(T_n)_{n \geq 1}$ be a collection of ordered graph trees such that T_n has n vertices for each n . The root of T_n will be denoted by $\rho = \rho(T_n)$. Define the depth-first search $\hat{w}_n : \{0, \dots, 2n\} \rightarrow T_n$ as in [2, Section 2.6] (\hat{w}_n is extended from the definition there by setting $\hat{w}_n(0) = \hat{w}_n(2n) = \rho$), and suppose that the search-depth process $(w_n(t))_{t \in [0, 1]}$ is the function satisfying

$$w_n\left(\frac{i}{2n}\right) := d_{T_n}(\rho, \hat{w}_n(i)), \quad 0 \leq i \leq 2n,$$

where d_{T_n} is the graph distance on T_n , and which is linear between these values, so that w_n takes values in $C([0, 1], \mathbb{R}_+)$. The uniform probability measure on the vertices of T_n will be written as μ_n . For each n , if we construct a function $\gamma_n : [0, 1] \rightarrow [0, 1]$ by setting

$$\gamma_n(t) := \begin{cases} \lfloor 2nt \rfloor / 2n & \text{if } w_n(\lfloor 2nt \rfloor / 2n) \geq w_n(\lceil 2nt \rceil / 2n), \\ \lceil 2nt \rceil / 2n & \text{otherwise,} \end{cases}$$

then it is the case that

$$\mu_n = \lambda^{[0, 1]} \circ (2n\gamma_n)^{-1} \circ \hat{w}_n^{-1}, \tag{4.1}$$

where, as previously, $\lambda^{[0, 1]}$ is the one-dimensional Lebesgue measure on $[0, 1]$ (this result can be checked by arguing along the lines of [2, Lemma 12]). It will be useful later to note that γ_n satisfies $\sup_{t \in [0, 1]} |\gamma_n(t) - t| \leq (2n)^{-1}$.

For a sequence $(u_k^n)_{k \geq 1} \in [0, 1]^{\mathbb{N}}$, define $(\sigma_k^n)_{k \geq 1} \in T_n^{\mathbb{N}}$ by $\sigma_k^n := \hat{w}_n(2n\gamma_n(u_k^n))$. Set $T_n(k)$ to be the minimal graph tree spanning $\{\rho, \sigma_1^n, \dots, \sigma_k^n\}$. The measure projection of μ_n onto $T_n(k)$ is denoted by

$$\mu_n^{(k)} := \mu_n \circ \phi_{T_n, T_n(k)}^{-1},$$

where the projection operator $\phi_{T_n, T_n(k)}$ is defined on graph trees analogously to the projection operator for real trees (see Section 2).

We will repeatedly apply the following assumption throughout the remainder of the section. The set Γ was introduced in Definition 2.1.

Assumption 4.1. For each n , the sequence $(u_k^n)_{k \geq 1}$ is dense in $[0, 1]$. Furthermore, there exists a divergent sequence $(\alpha_n)_{n \geq 1}$ such that $\alpha_n = o(n)$ and

$$(\alpha_n^{-1} w_n, u^n) \rightarrow (w, u),$$

in $C([0, 1], \mathbb{R}_+) \times [0, 1]^{\mathbb{N}}$, for some $(w, u) \in \Gamma$.

As in Section 2 for real trees, we use the sequential construction of [2, Section 2.2] to isometrically embed the vertices of T_n into ℓ^1 from the vertex sequence $(\sigma_k^n)_{k \geq 1}$. Observe that, under Assumption 4.1, because we are assuming that $(u_k^n)_{k \geq 1}$ is dense in $[0, 1]$, the sequence $(\sigma_k^n)_{k \geq 1}$ will contain all the vertices of T_n , and so this procedure does result in an isometric embedding for T_n . We will denote by ψ_n the unique distance-preserving map from the vertices of T_n into the set $\{(x(1), x(2), \dots) \in \ell^1 : x(i) \geq 0, i = 1, 2, \dots\} \subseteq \ell^1$ that satisfies $\psi_n(\rho) = 0$ and

$$\pi_k(\psi_n(\sigma)) = \psi_n(\phi_{T_n, T_n(k)}(\sigma)) \tag{4.2}$$

for every $\sigma \in T_n$ and $k \geq 1$, where π_k is the projection operator defined below (2.8). We write $\tilde{T}_n, \tilde{\mu}_n, \dots$ to represent the ℓ^1 -embedded versions of objects.

Let us now introduce the discrete-time processes that we will consider. Suppose that $X^n = (X_m^n)_{m \geq 0}$ is the discrete-time simple random walk on T_n started from ρ , and denote its law by $P_\rho^{T_n}$. Set $X^{n,k} = \phi_{T_n, T_n(k)}(X^n)$, and let $J^{n,k}$ be the associated jump chain, so that $J^{n,k}$ is the simple random walk on the vertices of $T_n(k)$ started from ρ . If $(A_m^{n,k})_{m \geq 0}$ are the jump times of $X^{n,k}$, i.e. $A_0^{n,k} = 0$ and, for $m \geq 1$,

$$A_m^{n,k} := \min\{l \geq A_{m-1}^{n,k} : X_l^n \in T_n(k) \setminus \{X_{A_{m-1}^{n,k}}^{n,k}\}\},$$

and $(\tau^{n,k}(m))_{m \geq 0}$ is the discrete inverse of $A^{n,k}$, i.e. $\tau^{n,k}(m) := \max\{l : A_l^{n,k} \leq m\}$, then we can recover $X^{n,k}$ from $J^{n,k}$ through the identity

$$X_m^{n,k} = J_{\tau^{n,k}(m)}^{n,k}. \tag{4.3}$$

The local times of $J^{n,k}$ are determined by

$$L_m^{n,k}(\sigma) := \frac{2}{\text{deg}_{n,k}(\sigma)} \sum_{l=0}^m \mathbf{1}_\sigma(J_l^{n,k})$$

for σ a vertex in $T_n(k)$, where $\text{deg}_{n,k}(\sigma)$ is the usual graph degree of σ in $T_n(k)$, and we use these to define an additive functional, $(\hat{A}_m^{n,k})_{m \geq 0}$, by setting $\hat{A}_0^{n,k} = 0$, and, for $m \geq 1$,

$$\hat{A}_m^{n,k} := n \int_{T_n(k)} L_{m-1}^{n,k}(\sigma) \mu_n^{(k)}(d\sigma).$$

The discrete-time inverse of $\hat{A}^{n,k}$ is given by $\hat{\tau}^{n,k}(m) := \max\{l : \hat{A}_l^{n,k} \leq m\}$, and we use this to define a time-changed version of $J^{n,k}$, denoted by $(\hat{X}_m^{n,k})_{m \geq 0}$, by setting

$$\hat{X}_m^{n,k} := J_{\hat{\tau}^{n,k}(m)}^{n,k}. \tag{4.4}$$

As in [7], to show that $\hat{X}^{n,k}$ and $X^{n,k}$ are close, we will prove a tightness result for $A^{n,k}$ and $\hat{A}^{n,k}$. We start by proving a continuity result for the local times $L^{n,k}$. Coinciding with the notation of [7], we define $\Lambda_n^{(k)} := \alpha_n^{-1} \# E(T_n(k))$, where $E(T_n(k))$ is the edge set of $T_n(k)$.

Lemma 4.1. Fix $k \in \mathbb{N}$ and $t_0 > 0$. Under Assumption 4.1,

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P_\rho^{T_n} \left(\alpha_n^{-1} \sup_{\{\sigma, \sigma' \in T_n(k) : d_{T_n}(\sigma, \sigma') \leq \delta \alpha_n\}} \sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} |L_m^{n,k}(\sigma) - L_m^{n,k}(\sigma')| > \varepsilon \right) = 0.$$

Proof. By assumption, $\alpha_n^{-1} d_{T_n}(\rho, \sigma_i^n) \rightarrow d_{\mathcal{T}}(\rho, \sigma_i) > 0$ for each $i \geq 1$. Hence, there exists a constant $L > 0$ such that

$$\inf_{i=1, \dots, k} \alpha_n^{-1} d_{T_n}(\rho, \sigma_i^n) \geq 2L \tag{4.5}$$

for large n . For the remainder of the proof, we will suppose that n is large enough so that this bound holds. If $\sigma \in T_n(k)$ then σ is on the path between ρ and σ_i^n for some $i \in \{1, \dots, k\}$. Consequently, (4.5) implies that there exists an injective path in $T_n(k)$ of length at least $\lfloor \alpha_n L \rfloor$ with endpoint σ . By considering the random walk observed on this path and the other vertices adjacent to σ , we can deduce that, for $t > 0$,

$$P_\rho^{T_n} \left(\alpha_n^{-1} L_{t_0 \alpha_n^2 \Lambda_n^{(k)}}^{n,k}(\sigma) \geq t \right) \leq P_0^{\Gamma(\lfloor \alpha_n L \rfloor, \deg_{n,k}(\sigma)-1)} \left(\alpha_n^{-1} L_{t_0 \alpha_n^2 \Lambda_n^{(k)}}^{\Gamma(\lfloor \alpha_n L \rfloor, \deg_{n,k}(\sigma)-1)} \geq t \right),$$

where $\Gamma(i, D)$ is a graph tree consisting of a path of length i emanating from a vertex 0 along with D other vertices each attached to 0 by a single edge, and $(L_m^{\Gamma(i, D)})_{m \geq 0}$ is the local time at 0 of the simple random walk on $\Gamma(i, D)$. Using a strong Markov argument, it is possible to check that $L_m^{\Gamma(i, D)}$ is stochastically dominated by $\sum_{j=1}^{L_m^{\Gamma(i, 0)}} \xi_j^D$, where the $(\xi_j^D)_{j \geq 1}$ are independent geometric, parameter $D/(D+1)$, random variables, independent of the random walk on $\Gamma(i, 0)$. Therefore,

$$\begin{aligned} P_\rho^{T_n} \left(\alpha_n^{-1} L_{t_0 \alpha_n^2 \Lambda_n^{(k)}}^{n,k}(\sigma) \geq t \right) &\leq \sup_{D \leq k} P_0^{\Gamma(\lfloor \alpha_n L \rfloor, D)} \left(\alpha_n^{-1} L_{t_0 \alpha_n^2 \Lambda_n^{(k)}}^{\Gamma(\lfloor \alpha_n L \rfloor, D)} \geq t \right) \\ &\leq (t \alpha_n)^{-4} \sup_{D \leq k} E_0^{\Gamma(\lfloor \alpha_n L \rfloor, 0)} \left(\sum_{i_1, i_2, i_3, i_4=1}^{L_{t_0 \alpha_n^2 \Lambda_n^{(k)}}^{\Gamma(\lfloor \alpha_n L \rfloor, 0)}} \xi_{i_1}^D \xi_{i_2}^D \xi_{i_3}^D \xi_{i_4}^D \right) \\ &\leq c_1 (t \alpha_n)^{-4} E_0^{\Gamma(\lfloor \alpha_n L \rfloor, 0)} \left(\left(L_{t_0 \alpha_n^2 \Lambda_n^{(k)}}^{\Gamma(\lfloor \alpha_n L \rfloor, 0)} \right)^4 \right) \\ &\leq c_2 t^{-4}, \end{aligned} \tag{4.6}$$

where c_1 and c_2 are constants that do not depend on σ, n , or t . The final inequality here is an application of [7, Lemma B.2] and the easily checked fact that $\Lambda_n^{(k)}$ is uniformly bounded in n .

Given the above bound, it is possible to deduce that, for fixed $\varepsilon > 0$,

$$\sup_{\{\sigma, \sigma' \in T_n(k) : d_{T_n}(\sigma, \sigma') \leq \delta \alpha_n\}} P_\rho^{T_n} \left(\alpha_n^{-1} \sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} |L_m^{n,k}(\sigma) - L_m^{n,k}(\sigma')| > \varepsilon \right) \leq c_3 \delta^2 \tag{4.7}$$

for every $n \geq n_0$ and $\delta \in (0, 1)$, where n_0 is a suitably large integer, by following the same argument as used to prove [7, Lemma 4.5] (which itself is an adaptation of a result appearing in [5]), inserting α_n in place of the scaling factor $n^{1/2}$ where appropriate. More specifically,

fix $\sigma \neq \sigma' \in T_n(k)$ such that $d_{T_n}(\sigma, \sigma') \leq \delta\alpha_n$ (note that in what follows we may assume that $\delta\alpha_n \geq 1$, else the left-hand side of (4.7) is clearly 0). Conditional on the event where the jump chain $J^{n,k}$ hits σ before σ' occurs, we have, by a simple calculation,

$$\begin{aligned} & \sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} \left| L_m^{n,k}(\sigma) - L_m^{n,k}(\sigma') + 2 \sum_{i=1}^{\deg_{n,k}(\sigma) L_m^{n,k}(\sigma)/2} \eta_i \right| \\ & \leq \sup \left\{ 2N_i \deg_{n,k}(\sigma')^{-1} : i \leq \deg_{n,k}(\sigma) L_{\lfloor t_0 \alpha_n^2 \Lambda_n^{(k)} \rfloor}^{n,k}(\sigma)/2 \right\}, \end{aligned} \tag{4.8}$$

where N_i is the number of visits by $J^{n,k}$ to σ' between the i th and $(i + 1)$ th visits to σ , and $\eta_i := N_i \deg_{n,k}(\sigma')^{-1} - \deg_{n,k}(\sigma)^{-1}$ is a centred random variable (for the precise distribution of N_i and estimates of the moments of η_i , see [7, Section B.2]). Now, since $(\sum_{i=1}^m \eta_i)_{m \geq 1}$ is a martingale with respect to the filtration $(\mathcal{F}_m)_{m \geq 1}$, where \mathcal{F}_m is the σ -algebra generated by $J^{n,k}$ up to the $(m + 1)$ th hitting time of σ , and $L := \deg_{n,k}(\sigma) L_{\lfloor t_0 \alpha_n^2 \Lambda_n^{(k)} \rfloor}^{n,k}(\sigma)/2$ is a stopping time for this martingale, Doob’s martingale norm inequality implies that

$$P_\rho^{T_n} \left(2\alpha_n^{-1} \sup_{m \leq L} \left| \sum_{i=1}^m \eta_i \right| > \varepsilon \right) \leq c_4 \alpha_n^{-4} E_\rho^{T_n} \left(\left(\sum_{i=1}^L \eta_i \right)^4 \right).$$

An upper bound for the right-hand side in terms of the moments of L and η_i can be computed by following the steps that lead to [5, Equation (1.29)], which is an analogous bound for a simple random walk on the line. In particular, we can check from (4.6) and the estimates for the moments of η_i of the form $|E_\rho^{T_n}(\eta_i^p)| \leq c(\delta\alpha_n)^{p-1}$ appearing in the appendix of [7] that

$$P_\rho^{T_n} \left(2\alpha_n^{-1} \sup_{m \leq L} \left| \sum_{i=1}^m \eta_i \right| > \varepsilon \right) \leq c_5 \delta^2, \tag{4.9}$$

uniformly in n, σ , and σ' . Moreover, it also holds that

$$\begin{aligned} P_\rho^{T_n} \left(\alpha_n^{-1} \sup_{i \leq L} 2N_i \deg_{n,k}(\sigma')^{-1} > \varepsilon \right) & \leq E_\rho^{T_n}(L) P_\rho^{T_n} (2\alpha_n^{-1} N_1 \deg_{n,k}(\sigma')^{-1} > \varepsilon) \\ & \leq c_6 \alpha_n^{-3} (1 + E_\rho^{T_n}(|\eta_1|^4)) \\ & \leq c_7 \delta^3, \end{aligned} \tag{4.10}$$

uniformly in n, σ , or σ' . Piecing together (4.8), (4.9), and (4.10), we obtain the estimate (4.7), as desired.

Subsequently, by considering, for each $\sigma \in T_n(k)$, the behaviour of the local times on the paths from σ to the at most $k + 1$ leaves of $T_n(k)$, we can apply a standard maximal inequality (for example, the extension of [3, Theorem 10.3], suggested as [3, Problem 10.1]) to deduce from this result that

$$\sup_{\sigma \in T_n(k)} P_\rho^{T_n} \left(\alpha_n^{-1} \sup_{\{\sigma' \in T_n(k) : d_{T_n}(\sigma, \sigma') \leq \delta\alpha_n\}} \sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} |L_m^{n,k}(\sigma) - L_m^{n,k}(\sigma')| > \varepsilon \right) \leq c_8 \delta^2$$

for every $n \geq n_0$ and $\delta \in (0, 1)$ (cf. [7, Equation (39)]).

Finally, note that, for each n and δ , we can choose a set $A_n(\delta)$ of at most $c_9\delta^{-1}$ vertices of $T_n(k)$ (where c_9 is independent of n and δ) such that the sets $\{\sigma' : d_{T_n}(\sigma, \sigma') \leq \delta\alpha_n\}$, $\sigma \in A_n(\delta)$, cover $T_n(k)$. Therefore,

$$\begin{aligned} & \mathbb{P}_\rho^{T_n} \left(\alpha_n^{-1} \sup_{\{\sigma, \sigma' \in T_n(k) : d_{T_n}(\sigma, \sigma') \leq \delta\alpha_n\}} \sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} |L_m^{n,k}(\sigma) - L_m^{n,k}(\sigma')| > \varepsilon \right) \\ & \leq \sum_{\sigma \in A_n(\delta)} 2 \mathbb{P}_\rho^{T_n} \left(\alpha_n^{-1} \sup_{\{\sigma' \in T_n(k) : d_{T_n}(\sigma, \sigma') \leq 2\delta\alpha_n\}} \sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} |L_m^{n,k}(\sigma) - L_m^{n,k}(\sigma')| > \frac{\varepsilon}{2} \right) \\ & \leq c_{10}\delta \end{aligned}$$

(cf. [7, Lemma 4.6]), from which the lemma follows.

We now show that the rescaled jump chains $\tilde{J}^{n,k} := \psi_n(J^{n,k})$ and additive functionals $\hat{A}^{n,k}$ converge. Henceforth, we extend $\tilde{J}^{n,k}$ and $\hat{A}^{n,k}$ to continuous-time processes by linear interpolation in ℓ^1 and \mathbb{R}_+ , respectively. We recall the definition of $\mathcal{T}(k)$ from (2.3) and also that the process $\tilde{X}^{\mathcal{T}(k), \lambda^{(k)}}$ is the ℓ^1 -embedding of $X^{\mathcal{T}(k), \lambda^{(k)}}$ (as introduced below Proposition 2.1). The definition of $\hat{A}^{(k)}$ appears in (2.5).

Lemma 4.2. *Under Assumption 4.1, the joint laws of the pairs*

$$\left(\alpha_n^{-1} \tilde{J}_{t\alpha_n^2 \Lambda_n^{(k)}}^{n,k}, (n\alpha_n)^{-1} \hat{A}_{t\alpha_n^2 \Lambda_n^{(k)}}^{n,k} \right)_{t \geq 0}$$

under $\mathbb{P}_\rho^{T_n}$ converge to the joint law of

$$\left(\tilde{X}_t^{\mathcal{T}(k), \lambda^{(k)}}, \hat{A}_t^{(k)} \right)_{t \geq 0}$$

under $\mathbb{P}_\rho^{\mathcal{T}(k), \lambda^{(k)}}$ weakly as probability measures on the space $C(\mathbb{R}_+, \ell^1) \times C(\mathbb{R}_+, \mathbb{R}_+)$.

Proof. Fix $k \geq 1$. Consider the vectors $\mathbf{x}_n := (\alpha_n^{-1}\psi_n(\sigma_1^n), \dots, \alpha_n^{-1}\psi_n(\sigma_k^n))$ and $\mathbf{x} := (\psi(\sigma_1), \dots, \psi(\sigma_k))$. Under Assumption 4.1, it is possible to check that $\mathbf{x}_n \rightarrow \mathbf{x}$. In fact, this is true for any k , and so we can write

$$(\alpha_n^{-1}\psi_n(\hat{w}_n(2n\gamma_n(u_i^n))))_{i \geq 1} \rightarrow (\psi(\hat{w}(u_i)))_{i \geq 1}$$

as sequences in ℓ^1 . Since $(u_i)_{i \geq 1}$ is dense in $[0, 1]$ and $\psi \circ \hat{w}$ is continuous, if we can show that $(\alpha_n^{-1}\psi_n \circ \hat{w}_n \circ (2n\gamma_n))_{n \geq 1}$ is tight then we will obtain

$$\alpha_n^{-1}\psi_n \circ \hat{w}_n \circ (2n\gamma_n) \rightarrow \psi \circ \hat{w} \tag{4.11}$$

in $C([0, 1], \ell^1)$. The necessary tightness can be proved as follows:

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\substack{s, t \in [0, 1] \\ |s-t| \leq \delta}} d_{\ell^1}(\alpha_n^{-1}\psi_n(\hat{w}_n(2n\gamma_n(s))), \alpha_n^{-1}\psi_n(\hat{w}_n(2n\gamma_n(t)))) \\ & = \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\substack{s, t \in [0, 1] \\ |s-t| \leq \delta}} d_{\alpha_n^{-1}w_n}(\gamma_n(s), \gamma_n(t)) \\ & = \lim_{\delta \rightarrow 0} \sup_{\substack{s, t \in [0, 1] \\ |s-t| \leq \delta}} d_w(s, t) \\ & = 0, \end{aligned}$$

where $d_{\alpha_n^{-1}w_n}$ is a distance on $[0, 1]$ defined similarly to the distance d_w introduced in (2.9).

The second equality is a result of Assumption 4.1, and the final equality holds because w is a continuous function. Now, by (4.1) we can write

$$\tilde{\mu}_n = \lambda^{[0,1]} \circ (2n\gamma_n)^{-1} \circ \hat{w}_n^{-1} \circ \psi_n^{-1},$$

and so (4.11) implies that $\tilde{\mu}_n(\alpha_n \cdot)$ converges to $\tilde{\mu} = \lambda^{[0,1]} \circ \hat{w}^{-1} \circ \psi^{-1}$ weakly as probability measures on ℓ^1 . Thus, because $\tilde{\mu}^{(k)} = \tilde{\mu} \circ \pi_k^{-1}$ and $\tilde{\mu}_n^{(k)} = \tilde{\mu}_n \circ \pi_k^{-1}$ (these expressions follow from (2.8) and (4.2), respectively), we obtain

$$\tilde{\mu}_n^{(k)}(\alpha_n \cdot) \rightarrow \tilde{\mu}^{(k)}$$

weakly as probability measures on ℓ^1 for each $k \geq 1$.

The vector and measure convergence of the previous paragraph allows us to apply Proposition 3.1 to deduce that the law of

$$(X_{t\lambda^{x_n}(\mathcal{T}^{x_n})}^{x_n}, A_{t\lambda^{x_n}(\mathcal{T}^{x_n})}^{x_n, v_n})_{t \geq 0}$$

under $P_0^{x_n}$ converges to the law of

$$(X_{t\lambda^x(\mathcal{T}^x)}^x, A_{t\lambda^x(\mathcal{T}^x)}^{x, v})_{t \geq 0}$$

under P_0^x , weakly as probability measures on the space $C(\mathbb{R}_+, \ell^1) \times C(\mathbb{R}_+, \mathbb{R}_+)$, where $v_n := \tilde{\mu}_n^{(k)}(\alpha_n \cdot)$ and $v := \tilde{\mu}^{(k)}$. We can readily check that the distribution of the limit is equal to the distribution of $(\tilde{X}^{\mathcal{T}^{(k)}, \lambda^{(k)}}, \hat{A}^{(k)})$ under $P_\rho^{\mathcal{T}^{(k)}, \lambda^{(k)}}$. To complete the proof, we will use a coupling argument to show how the pairs $(\tilde{J}^{n,k}, \hat{A}^{n,k})$ and (X^{x_n}, A^{x_n, v_n}) can be related.

The defining properties of Brownian motion on a dendrite imply that, under $P_0^{x_n}$, the process $(X_{h^n(m)}^{x_n})_{m \geq 0}$, where $h^n(0) := 0$ and

$$h^n(m) := \inf\{t \geq h^n(m-1) : d_{\ell^1}(X_t^{x_n}, X_{h^n(m-1)}^{x_n}) = \alpha_n^{-1}\},$$

has the same law as $(\alpha_n^{-1} \tilde{J}_m^{n,k})_{m \geq 0}$ under $P_\rho^{T_n}$. In view of this fact, for the remainder of the proof, we will abuse notation slightly by identifying $J^{n,k}$ with the process $(\psi_n^{-1}(\alpha_n X_{h^n(m)}^{x_n}))_{m \geq 0}$, and $L^{n,k}$ with its local times. Furthermore, by considering the path segments of X^{x_n} between the hitting times $(h^n(m))_{m \geq 0}$, it is possible to check that the rescaled increments $\alpha_n^{-2}(h^n(m) - h^n(m-1))$ are independently and identically distributed as the hitting time of $\{\pm 1\}$ by a standard Brownian motion in \mathbb{R} started from 0, which is a random variable with mean 1 and finite fourth moments. After applying a standard martingale estimate (see [16, Proposition 7.16] for example), it readily follows that, for any $t_0, \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P_0^{x_n} \left(\sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} \left| h^n(m) - \frac{m}{\alpha_n^2} \right| > \varepsilon \right) = 0 \tag{4.12}$$

(cf. the proof of [7, Lemma 4.2]). Consequently, since the sequence $(P_0^{x_n})_{n \geq 1}$ is tight and $\lambda^{x_n}(\mathcal{T}^{x_n}) = \Lambda_n^{(k)}$, we are able to deduce that, for $t_0, \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P_0^{x_n} \left(\sup_{t \in [0, t_0]} \left| \alpha_n^{-1} \tilde{J}_{t\alpha_n^2 \Lambda_n^{(k)}}^{n,k} - X_{t\lambda^{x_n}(\mathcal{T}^{x_n})}^{x_n} \right| > \varepsilon \right) = 0. \tag{4.13}$$

For the related additive functionals, to prove that

$$\lim_{n \rightarrow \infty} \mathbb{P}_0^{x_n} \left(\sup_{t \in [0, t_0]} \left| (n\alpha_n)^{-1} \hat{A}_{t\alpha_n^2 \Lambda_n^{(k)}}^{n,k} - A_{t\lambda^{x_n}(\mathcal{T}^{x_n})}^{x_n, v_n} \right| > \varepsilon \right) = 0, \tag{4.14}$$

by applying the tightness results of Lemmas 3.5 and 4.1, it will suffice to prove that

$$\lim_{n \rightarrow \infty} \sup_{\sigma \in T_n(k)} \mathbb{P}_0^{x_n} \left(\sup_{t \in [0, t_0]} \left| \alpha_n^{-1} L_{t\alpha_n^2 \Lambda_n^{(k)}}^{n,k}(\sigma) - L_{t\lambda^{x_n}(\mathcal{T}^{x_n})}^{x_n}(\alpha_n^{-1} \psi_n(\sigma)) \right| > \varepsilon \right) = 0, \tag{4.15}$$

where the definition of $L^{n,k}$ is extended to continuous time by linear interpolation. To demonstrate that this is the case requires a simple adaptation of [7, Lemma 4.7]. In particular, for $\sigma \in T_n(k)$, denote by $(\varsigma_i)_{i \geq 1}$ the hitting times of σ by $J^{n,k}$, and define

$$\eta_i := L_{h^n(\varsigma_i+1)}^{x_n}(\alpha_n^{-1} \psi_n(\sigma)) - L_{h^n(\varsigma_i)}^{x_n}(\alpha_n^{-1} \psi_n(\sigma)),$$

so that the $(\eta_i)_{i=1}^\infty$ is a sequence of independent random variables, each distributed as

$$\frac{2Z}{\alpha_n \text{deg}_{n,k}(\sigma)},$$

where Z represents the local time at 0 of a standard Brownian motion in \mathbb{R} started from 0, evaluated at the hitting time of $\{\pm 1\}$. By conditioning on $L_{t_0 \alpha_n^2 \Lambda_n^{(k)}}^{n,k}(\sigma)$, applying the fact that Z is a random variable with mean 1 and finite fourth moments and recalling (4.6), it is possible to check that

$$\lim_{n \rightarrow \infty} \sup_{\sigma \in T_n(k)} \mathbb{P}_0^{x_n} \left(\sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} \left| \eta_1 + \dots + \eta_{\text{deg}_{n,k}(\sigma) L_m^{n,k}(\sigma)/2} - \alpha_n^{-1} L_m^{n,k}(\sigma) \right| > \varepsilon \right) = 0 \tag{4.16}$$

(cf. [7, Equation (40)]). Now, if $J_m^{n,k} = \sigma$ then

$$\eta_1 + \dots + \eta_{\text{deg}_{n,k}(\sigma) L_m^{n,k}(\sigma)/2} = L_{h^n(m+1)}^{x_n}(\alpha_n^{-1} \psi_n(\sigma)),$$

otherwise the sum is equal to $L_{h^n(m)}^{x_n}(\alpha_n^{-1} \psi_n(\sigma))$. From this and the fact that, conditional on $J_m^{n,k} = \sigma$, $L_{h^n(m+1)}^{x_n}(\alpha_n^{-1} \psi_n(\sigma)) - L_{h^n(m)}^{x_n}(\alpha_n^{-1} \psi_n(\sigma))$ is distributed as $2Z/\alpha_n \text{deg}_{n,k}(\sigma)$, it is readily deduced that (4.16) also holds when $\alpha_n^{-1} L_m^{n,k}(\sigma)$ is replaced by the continuous local time $L_{h^n(m)}^{x_n}(\alpha_n^{-1} \psi_n(\sigma))$. These observations, together with the hitting time estimate of (4.12) and the local time tightness results of Lemmas 3.5 and 4.1, imply (4.15), which thereby establishes (4.14). Finally, combining (4.13) and (4.14) with the convergence result of the first part of the proof yields the lemma.

Lemma 4.3. *If Assumption 4.1 holds then, for $t_0, \varepsilon > 0$,*

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}_\rho^{T_n} \left((n\alpha_n)^{-1} \sup_{m \leq t_0 \alpha_n^2 \Lambda_n^{(k)}} |A_m^{n,k} - \hat{A}_m^{n,k}| > \varepsilon \right) = 0.$$

Proof. The proof of this result is an adaptation of that used to obtain [7, Proposition 5.2]. In particular, for $m \in \mathbb{N}$, $|A_m^{n,k} - \hat{A}_m^{n,k}|$ is bounded above by

$$\left| \sum_{l=0}^{m-1} \left(A_{l+1}^{n,k} - A_l^{n,k} - \frac{2n\mu_n^{(k)}(\{J_l^{n,k}\}) - 2 + \text{deg}_{n,k}(J_l^{n,k})}{\text{deg}_{n,k}(J_l^{n,k})} \right) \right| + \sum_{l=0}^{m-1} \frac{|2 - \text{deg}_{n,k}(J_l^{n,k})|}{\text{deg}_{n,k}(J_l^{n,k})}.$$

Clearly, the second term is no greater than m , and so its supremum over $m \leq t_0 \alpha_n^2 \Delta_n^{(k)}$ converges to 0 when multiplied by $(n\alpha_n)^{-1}$ (recall that $\alpha_n = o(n)$ by assumption). We now deal with the first term, which we will denote by $\Sigma(m)$. Since, conditional on knowing $J^{n,k}$, the expected value of $A_{l+1}^{n,k} - A_l^{n,k}$ is precisely

$$\frac{2n\mu_n^{(k)}(\{J_l^{n,k}\}) - 2 + \deg_{n,k}(J_l^{n,k})}{\deg_{n,k}(J_l^{n,k})}$$

and its expected square is bounded by

$$36n^2(\deg_{n,k}(J_l^{n,k}) + \Delta_n^{(k)}) \frac{\mu_n^{(k)}(\{J_l^{n,k}\})^2}{\deg_{n,k}(J_l^{n,k})},$$

where $\Delta_n^{(k)} := \sup_{\sigma \in T_n} d_{T_n}(\sigma, \phi_{T_n, T_n(k)}(\sigma))$ (these are elementary simple random walk estimates; see [7, Lemma B.3]), we can use Kolmogorov’s maximum inequality (see Lemma 4.15 of [16]) to deduce that, for $\varepsilon > 0$,

$$\begin{aligned} & \mathbb{P}_\rho^{T_n} \left((n\alpha_n)^{-1} \sup_{m \leq t_0 \alpha_n^2 \Delta_n^{(k)}} \Sigma(m) > \varepsilon \mid J^{n,k} \right) \\ & \leq \frac{1}{n^2 \alpha_n^2 \varepsilon^2} \sum_{l=0}^{\lfloor t_0 \alpha_n^2 \Delta_n^{(k)} \rfloor - 1} 36n^2 (\deg_{n,k}(J_l^{n,k}) + \Delta_n^{(k)}) \frac{\mu_n^{(k)}(\{J_l^{n,k}\})^2}{\deg_{n,k}(J_l^{n,k})} \\ & \leq \frac{18}{n \alpha_n^2 \varepsilon^2} \hat{A}^{n,k}_{\lfloor t_0 \alpha_n^2 \Delta_n^{(k)} \rfloor} \left(\max_{\sigma \in T_n(k)} \deg_{n,k}(\sigma) + \Delta_n^{(k)} \right). \end{aligned}$$

Hence, if we can show that

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \alpha_n^{-1} \left(\max_{\sigma \in T_n(k)} \deg_{n,k}(\sigma) + \Delta_n^{(k)} \right) = 0, \tag{4.17}$$

and also that

$$\lim_{t \rightarrow \infty} \limsup_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}_\rho^{T_n} \left((n\alpha_n)^{-1} \hat{A}^{n,k}_{t_0 \alpha_n^2 \Delta_n^{(k)}} > t \right) = 0, \tag{4.18}$$

then the lemma will follow. That

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \alpha_n^{-1} \Delta_n^{(k)} = 0 \tag{4.19}$$

is a straightforward consequence of Assumption 4.1 (cf. [7, Lemma 2.7]). Moreover, we clearly have $\max_{\sigma \in T_n(k)} \deg_{n,k}(x) \leq k + 1$, and so (4.17) holds. The distributional convergence result of Lemma 4.2 and the tail bound of (2.7) together imply (4.18), which completes the proof.

We can now prove the convergence of simple random walks. In the statement of the result and the proof, discrete-time processes are extended to continuous time by linear interpolation in the appropriate spaces.

Proposition 4.1. *Under Assumption 4.1, the laws of*

$$(\alpha_n^{-1} \tilde{X}_{t n \alpha_n}^n)_{t \geq 0}$$

under $\mathbb{P}_\rho^{T_n}$ converge to $\tilde{\mathbb{P}}_\rho^{\mathcal{J}, \mu}$ weakly as probability measures on the space $C(\mathbb{R}_+, \ell^1)$.

Proof. By Lemma 2.5, $\hat{A}^{(k)}$ is $\mathbb{P}_\rho^{\mathcal{J}(k), \lambda^{(k)}}$ -a.s. continuous and strictly increasing. Consequently, Lemma 4.2 implies that the joint laws of the pairs

$$(\alpha_n^{-1} \tilde{J}_{t\alpha_n^2 \Lambda_n}^{n,k}, (\alpha_n^2 \Lambda_n^{(k)})^{-1} \hat{\tau}_{t\alpha_n}^{n,k})_{t \geq 0}$$

under $\mathbb{P}_\rho^{T_n}$ converge to the law of

$$(\tilde{X}_t^{\mathcal{J}(k), \lambda^{(k)}}, \hat{\tau}_t^{(k)})_{t \geq 0}$$

under $\mathbb{P}_\rho^{\mathcal{J}(k), \lambda^{(k)}}$ weakly as probability measures on the space $C(\mathbb{R}_+, \ell^1) \times C(\mathbb{R}_+, \mathbb{R}_+)$. Hence, we see from the continuous mapping theorem that the laws under $\mathbb{P}_\rho^{T_n}$ of $(\alpha_n^{-1} \psi_n(\hat{X}_{t\alpha_n}^{n,k}))_{t \geq 0}$ converge to $\tilde{\mathbb{P}}_\rho^{\mathcal{J}(k), \mu^{(k)}}$ weakly as probability measures on the space $C(\mathbb{R}_+, \ell^1)$ for each $k \geq 1$, where we apply the representations of $\hat{X}^{n,k}$ and $X^{\mathcal{J}(k), \mu^{(k)}}$ from (4.4) and Lemma 2.4, respectively. Moreover, it is immediate from Proposition 2.1 and the continuity of ψ that $\tilde{\mathbb{P}}_\rho^{\mathcal{J}(k), \mu^{(k)}}$ converges to $\tilde{\mathbb{P}}_\rho^{\mathcal{J}, \mu}$ as $k \rightarrow \infty$. The proposition will follow from these convergence results by applying [3, Theorem 3.2] for example, if we can demonstrate the following tightness condition: for every $t_0, \varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}_\rho^{T_n} \left(\sup_{m \leq t_0 n \alpha_n} \alpha_n^{-1} d_{T_n}(X_m^n, \hat{X}_m^{n,k}) > \varepsilon \right) = 0.$$

To prove this, first observe that, by construction, $d_{T_n}(X_m^n, X_m^{n,k}) \leq \Delta_n^{(k)}$ for every $m \geq 0$, where $\Delta_n^{(k)}$ was defined in the proof of Lemma 4.3. Thus, (4.19) implies that it will actually suffice to prove that

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}_\rho^{T_n} \left(\sup_{m \leq t_0 n \alpha_n} \alpha_n^{-1} d_{T_n}(X_m^{n,k}, \hat{X}_m^{n,k}) > \varepsilon \right) = 0.$$

Applying the representations of $X^{n,k}$ and $\hat{X}^{n,k}$ in terms of the jump chain from (4.3) and (4.4), respectively, for any $\delta > 0$, the probability on the left-hand side of this expression is bounded above by $p_1(n, k) + p_2(n, k)$, where

$$p_1(n, k) := \mathbb{P}_\rho^{T_n} \left(\sup_{\substack{s, t \in [0, t_0 + \delta] \\ |s-t| \leq \delta}} \alpha_n^{-1} d_{\ell^1}(\tilde{J}_{\hat{\tau}^{n,k}(s\alpha_n)}^{n,k}, \tilde{J}_{\hat{\tau}^{n,k}(t\alpha_n)}^{n,k}) > \varepsilon \right),$$

$$p_2(n, k) := \mathbb{P}_\rho^{T_n} (\tau^{n,k}(t\alpha_n) \notin [\hat{\tau}^{n,k}((t - \delta)\alpha_n \vee 0), \hat{\tau}^{n,k}((t + \delta)\alpha_n)]) \text{ for some } t \in [0, t_0].$$

It is elementary to check from Lemma 4.3 that

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} p_2(n, k) = 0$$

for any $\delta > 0$. Furthermore, the convergence results above yield

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} p_1(n, k) = \mathbb{P}_\rho^{\mathcal{J}, \mu} \left(\sup_{\substack{s, t \in [0, t_0 + \delta] \\ |s-t| \leq \delta}} d_{\mathcal{J}}(X_s^{\mathcal{J}, \mu}, X_t^{\mathcal{J}, \mu}) > \varepsilon \right),$$

which can be made arbitrarily small by letting $\delta \rightarrow 0$, because $X^{\mathcal{J}, \mu}$ is a diffusion under $\mathbb{P}_\rho^{\mathcal{J}, \mu}$. This completes the proof.

We can now complete the proof of our main results.

Proof of Theorem 1.1. Let $w \in \mathcal{W}$ satisfy $(\mathcal{T}, \mu) \in \mathbb{T}^*$, and let $U = (U_i)_{i \geq 1}$ be an independent sequence of $U[0, 1]$ random variables. Since the Lebesgue measure $\lambda^{[0,1]}$ on $[0, 1]$ has full support, and the measure μ is nonatomic, has full support, and is supported on the leaves of \mathcal{T} , it is clear that $(U_i)_{i \geq 1}$ is dense in $[0, 1]$, and the vertices $(\sigma_i)_{i \geq 1} = (\hat{w}(U_i))_{i \geq 1}$ are a dense collection of leaves of \mathcal{T} , distinct and not equal to ρ for any i , for almost every realisation of U . In particular, there exists a $u \in [0, 1]^{\mathbb{N}}$ such that $(w, u) \in \Gamma$. Set $u^n = u$ for each n , so that, under the assumptions of the theorem, we have

$$(\alpha_n^{-1}w_n, u^n) \rightarrow (w, u)$$

in $C([0, 1], \mathbb{R}_+) \times C(\mathbb{R}_+, \mathbb{R}_+)$ for some $(w, u) \in \Gamma$, which is Assumption 4.1. Applying Proposition 4.1, this implies that

$$\tilde{P}_\rho^{T_n} (\{f \in C([0, 1], \ell^1) : (\alpha_n^{-1}f(tn\alpha_n))_{t \in [0,1]} \in \cdot\}) \rightarrow \tilde{P}_\rho^{\mathcal{T}, \mu}$$

weakly as probability measures on $C([0, 1], \ell^1)$. The convergence $\tilde{\mu}_n(\alpha_n \cdot) \rightarrow \tilde{\mu}$ in $\mathcal{M}(\ell^1)$ was established in the proof of Lemma 4.2. That $\alpha_n^{-1}\tilde{T}_n(k)$ converges to $\tilde{T}(k)$ in $\mathcal{K}(\ell^1)$ is a straightforward consequence of the convergence $\mathbf{x}_n \rightarrow \mathbf{x}$ in $(\ell^1)^k$, where the vertices \mathbf{x}_n and \mathbf{x} are defined as in the proof of Lemma 4.2. To extend this to the result that $\alpha_n^{-1}\tilde{T}_n \rightarrow \tilde{T}$, we apply the tightness result of (4.19) and the fact that $\sup_{\sigma \in \mathcal{T}} d_{\mathcal{T}}(\sigma, \phi_{\mathcal{T}, \mathcal{T}(k)}(\sigma)) \rightarrow 0$, which was noted below (2.4).

Proof of Theorem 1.2. We start our proof, which is an adaptation of [7, Section 8], by outlining the construction of the measure Pr . First, by following the proof of Lemma 4.2, it is possible to check that, for each $k \geq 1$, the map $(w, u) \mapsto (\mathbf{x}, \nu)$, where $\mathbf{x} := (\psi(\sigma_1), \dots, \psi(\sigma_k))$ and $\nu := \tilde{\mu}^{(k)}$, is continuous on Γ . By Lemma 2.4, Lemma 2.5, and Proposition 3.1, this implies that $(w, u) \mapsto (\tilde{T}(k), \tilde{\mu}^{(k)}, \tilde{P}_\rho^{\mathcal{T}(k), \mu^{(k)}})$ is also continuous on Γ . Consequently, we find from Proposition 2.1 that $(w, u) \mapsto (\tilde{T}, \tilde{\mu}, \tilde{P}_\rho^{\mathcal{T}, \mu})$ is a measurable map on Γ . Given this, and noting that, under the assumptions of the theorem, $(W, U) \in \Gamma$, P-a.s., checking the existence of a unique probability measure satisfying (1.2) is straightforward. The construction of Pr_n is similar, but easier. Finally, to check that $\text{Pr}_n \circ \Theta_n^{-1} \rightarrow \text{Pr}$, we consider a Skorokhod-type coupling of random variables. In particular, since the relevant spaces are separable, if $(\alpha_n^{-1}W_n)_{n \geq 1}$ converges in distribution to W then there exists a probability space upon which random variables $(W_n^*, U_n^*)_{n \geq 1}$ and (W^*, U^*) are defined in such a way that (W_n^*, U_n^*) has the distribution of (W_n, U) for each n , (W^*, U^*) has the distribution of (W, U) , and $(\alpha_n^{-1}W_n^*, U_n^*) \rightarrow (W^*, U^*)$ almost surely. Defining all the related objects on this probability space, then exactly as in the previous proof we are able to deduce from Proposition 4.1 that $\Theta_n(\tilde{T}_n, \tilde{\mu}_n, \tilde{P}_\rho^{T_n}) \rightarrow (\tilde{T}, \tilde{\mu}, \tilde{P}_\rho^{\mathcal{T}, \mu})$ almost surely, and the result easily follows.

5. Application to α -stable trees

In this section we describe the application of our results to Galton–Watson trees with a possibly infinite offspring distribution variance. Suppose that ξ is a nonnegative integer-valued random variable that is aperiodic, has mean 1, and is in the domain of attraction of a stable law with index $\alpha \in (1, 2)$, by which we mean that there exists a sequence $a_n \uparrow \infty$ such that

$$\frac{\xi[n] - n}{a_n} \xrightarrow{D} \Xi,$$

where $\xi[n]$ is the sum of n independent copies of ξ and the limit random variable satisfies $E(e^{-\lambda \Xi}) = e^{-\lambda^\alpha}$. If $(T_n)_{n \geq 1}$ is a family of random trees such that T_n is a Galton–Watson tree with offspring distribution ξ , conditioned on the total progeny being equal to n , then it is known (see [12, Theorem 3.1]) that the associated rescaled search-depth processes $(n^{-1}a_n W_n)_{n \geq 1}$ converge in distribution to a random excursion W with law $N_\alpha^{(1)}$, say. The corresponding random real tree \mathcal{T} is the α -stable tree conditioned to have total mass equal to 1, and we denote its law by $\Theta_\alpha^{(1)}$ (see also [19, Section 4]). Consequently, to allow us to apply Theorem 1.2 to the sequence $(T_n)_{n \geq 1}$ with $\alpha_n = n\alpha_n^{-1}$, it remains to check that the $\Theta_\alpha^{(1)}$ -almost everywhere (a.e.) realisation of μ is nonatomic, supported on the leaves of \mathcal{T} , and satisfies (1.1) for some $\kappa > 0$. In fact, rather than investigating μ under the conditioned measure $\Theta_\alpha^{(1)}$, by rescaling, it will suffice to check that the required properties hold under the unconditioned ‘excursion’ measure Θ_α of the α -stable tree (see [19, Definition 4.2]), as is done in the following proposition. For a rooted real tree, we use the notation $\chi(\mathcal{T}) := \sup\{d_{\mathcal{T}}(\rho, \sigma) : \sigma \in \mathcal{T}\}$ to represent its height.

Proposition 5.1. *Let $\alpha \in (1, 2)$. For Θ_α -a.e. realisation of \mathcal{T} , μ is nonatomic, supported on the leaves of \mathcal{T} , and satisfies*

$$\liminf_{r \rightarrow 0} \frac{\inf_{\sigma \in \mathcal{T}} \mu(B(\sigma, r))}{r^{\alpha/(\alpha-1)} (\ln r^{-1})^{-\alpha/(\alpha-1)}} > 0. \tag{5.1}$$

In particular, (5.1) implies (1.1) for any $\kappa > \alpha/(\alpha - 1)$.

Proof. That μ is nonatomic and supported on the leaves of \mathcal{T} for Θ_α -a.e. realisation of \mathcal{T} follows from results of [13]. Thus, it remains to check (5.1).

Fix a compact rooted real tree \mathcal{T} and $r > 0$. Following [13], denote by $(\mathcal{T}^{(i),o})_{i \in \mathcal{I}}$ the connected components of the open set $\{\sigma \in \mathcal{T} : d_{\mathcal{T}}(\rho, \sigma) > r\}$. Define $\mathcal{T}^{(i)} := \mathcal{T}^{(i),o} \cup \{\sigma_i\}$, where σ_i is the common ancestor in level r of every $\sigma \in \mathcal{T}^{(i),o}$, so that $\mathcal{T}^{(i)}$ is a compact rooted real tree, and we set its root to be σ_i . The trees $(\mathcal{T}^{(i)})_{i \in \mathcal{I}}$ are the subtrees of \mathcal{T} originating from level r . If $\chi(\mathcal{T}^{(i)}) \geq \delta$ then we say that $\mathcal{T}^{(i)}$ hits level $r + \delta$. The number of subtrees of \mathcal{T} originating from level r that hit level $r + \delta$ will be written as $Z(r, \delta)$.

For integers $n, k \geq 0$, define $(\mathcal{T}^{n,k,(i)})_{i=1}^{Z(k2^{-n}, 2^{-n})}$ to be the collection of subtrees of \mathcal{T} originating at level $k2^{-n}$ that hit level $(k + 1)2^{-n}$. If we assume that $\chi(\mathcal{T}) \geq 2^{-n}$ then it is elementary to check that, for fixed $r > 0$,

$$\inf_{\sigma \in B(\rho, r)} \mu(B(\sigma, 3 \cdot 2^{-n})) \geq \inf_{\substack{0 \leq k \leq 2^n r \\ 1 \leq i \leq Z(k2^{-n}, 2^{-n})}} \mu(\mathcal{T}^{n,k,(i)}(2^{-n})),$$

where $\mathcal{T}^{n,k,(i)}(2^{-n})$ is the ball in $\mathcal{T}^{n,k,(i)}$ of radius 2^{-n} centred at the root of $\mathcal{T}^{n,k,(i)}$, which implies that

$$\begin{aligned} \Theta_\alpha \left(\inf_{\sigma \in B(\rho, r)} \mu(B(\sigma, 3 \cdot 2^{-n})) \leq x, \chi(\mathcal{T}) \geq 2^{-n} \right) \\ \leq \sum_{0 \leq k \leq 2^n r} \Theta_\alpha \left(\inf_{1 \leq i \leq Z(k2^{-n}, 2^{-n})} \mu(\mathcal{T}^{n,k,(i)}(2^{-n})) \leq x, \chi(\mathcal{T}) \geq (k + 1)2^{-n} \right). \end{aligned} \tag{5.2}$$

Note that if $\chi(\mathcal{T}) < (k + 1)2^{-n}$ for some $k \geq 0$ then $Z(k2^{-n}, 2^{-n}) = 0$ and the infimum appearing in the k th summand is infinite. Hence, the summands are not decreased by including the statement $\chi(\mathcal{T}) \geq (k + 1)2^{-n}$, as we do.

The branching property of Lévy trees (see [13, Theorem 4.2]) implies that, under the measure $\Theta_\alpha(\cdot \mid \chi(\mathcal{T}) \geq k2^{-n})$, conditional on $Z(k2^{-n}, 2^{-n})$, the trees in the collection $(\mathcal{T}^{n,k,(i)})_{i=1}^{Z(k2^{-n}, 2^{-n})}$ are distributed as an independent sample chosen from the law

$$\Theta_\alpha(\cdot \mid \chi(\mathcal{T}) \geq 2^{-n}).$$

Hence, writing $\tilde{\Theta}_\alpha := \Theta_\alpha(\cdot \mid \chi(\mathcal{T}) \geq k2^{-n})$ and $Z := Z(k2^{-n}, 2^{-n})$, when $k \geq 1$, the k th summand of (5.2) satisfies

$$\begin{aligned} &\Theta_\alpha\left(\inf_{1 \leq i \leq Z} \mu(\mathcal{T}^{n,k,(i)}(2^{-n})) \leq x, \chi(\mathcal{T}) \geq (k+1)2^{-n}\right) \\ &= \tilde{\Theta}_\alpha\left(\tilde{\Theta}_\alpha\left(\inf_{1 \leq i \leq Z} \mu(\mathcal{T}^{n,k,(i)}(2^{-n})) \leq x, \chi(\mathcal{T}) \geq (k+1)2^{-n} \mid Z\right)\right) \\ &\quad \times \Theta_\alpha(\chi(\mathcal{T}) \geq k2^{-n}) \\ &= \tilde{\Theta}_\alpha\left(\mathbf{1}_{\{Z \neq 0\}} \tilde{\Theta}_\alpha\left(\inf_{1 \leq i \leq Z} \mu(\mathcal{T}^{n,k,(i)}(2^{-n})) \leq x \mid Z\right)\right) \Theta_\alpha(\chi(\mathcal{T}) \geq k2^{-n}) \\ &\leq \tilde{\Theta}_\alpha(Z \Theta_\alpha(\mu(B(\rho, 2^{-n})) \leq x \mid \chi(\mathcal{T}) \geq 2^{-n})) \Theta_\alpha(\chi(\mathcal{T}) \geq k2^{-n}) \\ &= \Theta_\alpha(Z) \Theta_\alpha(\mu(B(\rho, 2^{-n})) \leq x \mid \chi(\mathcal{T}) \geq 2^{-n}). \end{aligned}$$

Moreover, we have $\Theta_\alpha(Z) = \Theta_\alpha(\chi(\mathcal{T}) \geq 2^{-n})$ (cf. proof of [13, Lemma 5.4]), and so summing over k yields

$$\begin{aligned} &\Theta_\alpha\left(\inf_{\sigma \in B(\rho,r)} \mu(B(\sigma, 3 \cdot 2^{-n})) \leq x, \chi(\mathcal{T}) \geq 2^{-n}\right) \\ &\leq (2^n r + 1) \Theta_\alpha(\mu(B(\rho, 2^{-n})) \leq x, \chi(\mathcal{T}) \geq 2^{-n}) \\ &\leq c_1 (2^n r + 1) 2^{n/(\alpha-1)} e^{-c_2 2^{-n} x^{-(\alpha-1)/\alpha}}, \end{aligned}$$

where c_1 and c_2 are constants not depending on n, r , or x , and the second inequality is a result of Equation (5.8) of [14]. Taking c_3 suitably small, this implies that, for any $R > 0$,

$$\sum_{n=0}^\infty \Theta_\alpha\left(\inf_{\sigma \in B(\rho,r)} \mu(B(\sigma, 3 \cdot 2^{-n})) \leq c_3 (n2^n)^{-\alpha/(\alpha-1)}, \chi(\mathcal{T}) \geq R\right) < \infty.$$

The result follows by applying the Borel–Cantelli lemma, and then letting $r \rightarrow \infty$ and $R \rightarrow 0$.

To complete this section, we note that the above proposition allows the heat kernel estimate of Lemma 2.1(b) to be improved in the following way in the α -stable case. By comparison with results appearing in [10] for random walks on infinite variance Galton–Watson trees conditioned to survive, we expect that the polynomial term in the following lemma is the best possible. For an analogous estimate in the Brownian case, see [8].

Corollary 5.1. *Let $\alpha \in (1, 2)$. For Θ_α -a.e. realisation of \mathcal{T} , the Brownian motion $X^{\mathcal{T},\mu}$ admits a transition density $(p_t(\sigma, \sigma'))_{\sigma, \sigma' \in \mathcal{T}, t > 0}$ that satisfies*

$$\limsup_{t \rightarrow 0} \frac{\sup_{\sigma, \sigma' \in \mathcal{T}} p_t(\sigma, \sigma')}{t^{-\alpha/(2\alpha-1)} (\ln t^{-1})^{\alpha/(2\alpha-1)}} < \infty.$$

References

- [1] ALDOUS, D. (1991). The continuum random tree. II. An overview. In *Stochastic Analysis* (Durham, 1990; *London Math. Soc. Lecture Note Ser.* **167**), Cambridge University Press, pp. 23–70.
- [2] ALDOUS, D. (1993). The continuum random tree. III. *Ann. Prob.* **21**, 248–289.
- [3] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.
- [4] BLUMENTHAL, R. M. AND GETTOOR, R. K. (1968). *Markov Processes and Potential Theory* (Pure Appl. Math. **29**). Academic Press, New York.
- [5] BORODIN, A. N. (1981). The asymptotic behavior of local times of recurrent random walks with finite variance. *Teor. Veroyat. Primen.* **26**, 769–783.
- [6] CROYDON, D. A. (2007). Heat kernel fluctuations for a resistance form with non-uniform volume growth. *Proc. London Math. Soc.* **94**, 672–694.
- [7] CROYDON, D. A. (2008). Convergence of simple random walks on random discrete trees to Brownian motion on the continuum random tree. *Ann. Inst. H. Poincaré Prob. Statist.* **44**, 987–1019.
- [8] CROYDON, D. A. (2008). Volume growth and heat kernel estimates for the continuum random tree. *Prob. Theory Relat. Fields* **140**, 207–238.
- [9] CROYDON, D. A. AND HAMBLY, B. M. (2008). Local limit theorems for sequences of simple random walks on graphs. *Potential Analysis* **29**, 351–389.
- [10] CROYDON, D. AND KUMAGAI, T. (2008). Random walks on Galton–Watson trees with infinite variance offspring distribution conditioned to survive. *Electron. J. Prob.* **13**, 1419–1441.
- [11] DUDLEY, R. M. (1973). Sample functions of the Gaussian process. *Ann. Prob.* **1**, 66–103.
- [12] DUQUESNE, T. (2003). A limit theorem for the contour process of conditioned Galton–Watson trees. *Ann. Prob.* **31**, 996–1027.
- [13] DUQUESNE, T. AND LE GALL, J.-F. (2005). Probabilistic and fractal aspects of Lévy trees. *Prob. Theory Relat. Fields* **131**, 553–603.
- [14] DUQUESNE, T. AND LE GALL, J.-F. (2006). The Hausdorff measure of stable trees. *ALEA* **1**, 393–415.
- [15] FUKUSHIMA, M., ŌSHIMA, Y. AND TAKEDA, M. (1994). *Dirichlet Forms and Symmetric Markov Processes* (de Gruyter Stud. Math. **19**). Walter de Gruyter, Berlin.
- [16] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.
- [17] KIGAMI, J. (1995). Harmonic calculus on limits of networks and its application to dendrites. *J. Funct. Anal.* **128**, 48–86.
- [18] KUMAGAI, T. (2004). Heat kernel estimates and parabolic Harnack inequalities on graphs and resistance forms. *Publ. Res. Inst. Math. Sci.* **40**, 793–818.
- [19] LE GALL, J.-F. (2006). Random real trees. *Ann. Fac. Sci. Toulouse Math.* **15**, 35–62.
- [20] MARCUS, M. B. AND ROSEN, J. (1992). Sample path properties of the local times of strongly symmetric Markov processes via Gaussian processes. *Ann. Prob.* **20**, 1603–1684.