The Annals of Probability
2017, Vol. 45, No. 1, 4–55
DOI: 10.1214/15-AOP1030

SUBSEQUENTIAL SCALING LIMITS OF SIMPLE RANDOM WALK ON THE TWO-DIMENSIONAL UNIFORM SPANNING TREE

BY M. T. BARLOW[1], D. A. CROYDON[2] AND T. KUMAGAI[3]

University of British Columbia, University of Warwick and Kyoto University

The first main result of this paper is that the law of the (rescaled) two-dimensional uniform spanning tree is tight in a space whose elements are measured, rooted real trees continuously embedded into Euclidean space. Various properties of the intrinsic metrics, measures and embeddings of the subsequential limits in this space are obtained, with it being proved in particular that the Hausdorff dimension of any limit in its intrinsic metric is almost surely equal to $8/5$. In addition, the tightness result is applied to deduce that the annealed law of the simple random walk on the two-dimensional uniform spanning tree is tight under a suitable rescaling. For the limiting processes, which are diffusions on random real trees embedded into Euclidean space, detailed transition density estimates are derived.

1. Introduction. The study of uniform spanning trees (USTs) has a long history; in the 1840s Kirchhoff used them in his classic paper [34] on electrical resistance. Much of the recent theory in the probability literature is based on the discovery that paths in the UST have the same law as loop erased random walks. Using this connection, algorithms to construct the UST from random walks have been given in [5, 14, 49]. See [12] for a survey of the properties of the UST, and a description of Wilson's algorithm, which will be important for this article, and [38] for a survey of the properties of the loop erased random walk (LERW). We also remark that USTs can be considered as a boundary case of the random cluster model; see [29].

In [48], Schramm studied the scaling limit of the UST in $\mathbb{Z}^2$, and this led him to introduce the SLE process. In [41], it was proved that the LERW in $\mathbb{Z}^2$ has SLE_2 as its scaling limit, and this connection was used in [10, 45] to improve earlier results of Kenyon [31] on the growth function of two-dimensional LERW. In [11], this good control on the length of LERW paths, combined with Wilson's algorithm, was used to obtain volume growth and resistance estimates for the two-dimensional UST $\mathcal{U}$. Using the connection between random walks and electrical

Received July 2014; revised April 2015.

[1]Supported in part by NSERC (Canada).
[2]Supported in part by EPSRC First Grant EP/K029657/1.
[3]Supported in part by the JSPS Grant-in-Aid for Scientific Research (A) 25247007.
MSC2010 subject classifications. 60D05, 60G57, 60J60, 60J67, 60K37.
Key words and phrases. Uniform spanning tree, loop-erased random walk, random walk, scaling limit, continuum random tree.

resistance, and the methods of [9, 37], these bounds then led to heat kernel bounds for $\mathcal{U}$.

In this paper, we study scaling limits of $\mathcal{U}$, as well as the random walk on it. While very significant progress in this direction was made on the first topic in [2, 48], those papers are focused on the topological properties of the scaling limit as a subset of $\mathbb{R}^2$. Here, we work in a framework that allows us to describe properties of the joint scaling limit of the corresponding intrinsic metric, uniform measure and simple random walk.

We begin by introducing our main notation. Throughout this article, $\mathcal{U}$ will represent the uniform spanning tree on $\mathbb{Z}^2$, and $\mathbf{P}$ the probability measure on the probability space on which this is built. As proved in [47], $\mathcal{U}$ is the local limit of the uniform spanning tree on $[-n,n]^2 \cap \mathbb{Z}^2$ (equipped with nearest-neighbour bonds) as $n \to \infty$. We note that $\mathcal{U}$ is $\mathbf{P}$-a.s. indeed a spanning tree of $\mathbb{Z}^2$, that is, it is a graph with vertex set $\mathbb{Z}^2$, and any two of its vertices are connected by a unique path in $\mathcal{U}$. We will denote by $d_{\mathcal{U}}$ the intrinsic (shortest path) metric on the graph $\mathcal{U}$, and $\mu_{\mathcal{U}}$ the uniform measure on $\mathcal{U}$ (i.e., the measure which places a unit mass at each vertex).

To describe the scaling limit of the metric measure space $(\mathcal{U}, d_{\mathcal{U}}, \mu_{\mathcal{U}})$, we work with a Gromov–Hausdorff-type topology of the kind that has proved useful for studying real trees. (See [15] for an introduction to the classical theory, and [25] for its application to real trees.) In particular, we will build on the notions of Gromov–Hausdorff–Prohorov topology of [1, 25, 46], and the topology for spatial trees of [23] (cf. the spectral Gromov–Hausdorff topology of [22]). We extend the metric space $(\mathcal{U}, d_{\mathcal{U}})$ to a complete and locally compact real tree by adding unit line segments along edges. The measure $\mu_{\mathcal{U}}$ is then viewed as a locally finite (atomic) Borel measure on this space. To retain information about $\mathcal{U}$ in the Euclidean topology, we consider $(\mathcal{U}, d_{\mathcal{U}})$ as a spatial tree, that is, as an abstract real tree embedded into $\mathbb{R}^2$ via a continuous map $\phi_{\mathcal{U}} : \mathcal{U} \to \mathbb{R}^2$, which we take in our example to be just the identity on vertices, with linear interpolation along edges. In addition, we will suppose the space $(\mathcal{U}, d_{\mathcal{U}})$ is rooted at the origin of $\mathbb{Z}^2$. Thus, we define a random quintuplet $(\mathcal{U}, d_{\mathcal{U}}, \mu_{\mathcal{U}}, \phi_{\mathcal{U}}, 0)$, and our first result (Theorem 1.1 below) is that the law of this object is tight under rescaling in the appropriate space of "measured, rooted spatial trees." The principal advantage of working in this topology is that it allows us to preserve information about the intrinsic metric $d_{\mathcal{U}}$ and measure $\mu_{\mathcal{U}}$; these parts of the picture were missing from the earlier scaling results of [2, 48].

The final ingredient we need in order to state our first main result comes from the growth function for LERW in $\mathbb{Z}^2$. This is the function $G_2(r) = \mathbb{E}|L_r|$, where $|L_r|$ is the length of a LERW run from 0 until it first exits the ball of radius r. In particular, from the results in [40, 45] we have (see [8], Corollary 3.15) that there exist constants $c_1, c_2 \in (0,\infty)$ such that

$$c_1 r^{\kappa} \le G_2(r) \le c_2 r^{\kappa}, \tag{1.1}$$

where the growth exponent $\kappa := 5/4$. This exponent plays a key role in the comparison of the intrinsic and Euclidean metrics on the UST. We remark that in [11], where the key result of [40] was not available, the heat kernel estimates on $\mathcal{U}$ take on a more complicated form involving the function G_2 and functions derived from it.

THEOREM 1.1. *If $\mathbf{P}_\delta$ is the law of the measured, rooted spatial tree $(\mathcal{U}, \delta^\kappa d_{\mathcal{U}}, \delta^2 \mu_{\mathcal{U}}, \delta\phi_{\mathcal{U}}, 0)$ under $\mathbf{P}$, then the collection $(\mathbf{P}_\delta)_{\delta\in(0,1)}$ is tight.*

As already noted, this theorem extends the results of [2, 48] to include scaling of the intrinsic metric and uniform measure. We further note that the tightness in [2, 48] was essentially a finite-dimensional statement, since it described the shape in Euclidean space of the tree spanning a finite number of points, while the result above establishes tightness for the entire space.

REMARK 1.2. To extend the above theorem to a full convergence result, and establish that the scaling limit satisfies the obvious scale invariance properties, it would be sufficient to characterise the limit uniquely from a suitable finite-dimensional convergence result. We expect that such a characterisation will be possible once it is known that two-dimensional loop-erased random walk converges as a process. Proving this is an open problem, but see [3, 42, 43] for recent progress on proving the convergence of LERW to the SLE_2 curve in its "natural parameterisation."

The tightness in Theorem 1.1 implies the existence of subsequential scaling limits for the collection $(\mathbf{P}_\delta)_{\delta\in(0,1)}$ of laws on measured, rooted spatial trees as $\delta \to 0$. The following theorem gives a number of properties of these limits. We note that (a)(ii) translates part of [2], Theorem 1.2, into our setting, and the topological aspects of (c)(i) and (c)(ii) are a restatement of parts of [48], Theorem 1.6. [In particular, the set $\phi_{\mathcal{T}}(\mathcal{T}^o)$ that appears in the statement of our result is identical to Schramm's notion of the "trunk" for the UST scaling limit; see Lemma 5.7.] We do not expect the powers of logarithms and log-logarithms in (1.3) and (1.4) to be optimal. We write $\deg_{\mathcal{T}}(x)$ for the degree of a point x in a real tree $\mathcal{T}$, that is, the number of connected components of $\mathcal{T} \setminus \{x\}$, $|A|$ to represent the cardinality of a subset $A \subseteq \mathcal{T}$, and $\mathcal{L}$ to represent Lebesgue measure on $\mathbb{R}^2$.

THEOREM 1.3. *If $\tilde{\mathbf{P}}$ is a subsequential limit of $(\mathbf{P}_\delta)_{\delta\in(0,1)}$, then for $\tilde{\mathbf{P}}$-a.e. measured, rooted spatial tree $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \phi_{\mathcal{T}}, \rho_{\mathcal{T}})$ it holds that:*

(a) (i) *the Hausdorff dimension of the complete and locally compact real tree $(\mathcal{T}, d_{\mathcal{T}})$ is given by*

$$d_f := \frac{2}{\kappa} = \frac{8}{5}; \tag{1.2}$$

(ii) $(\mathcal{T}, d_{\mathcal{T}})$ *has precisely one end at infinity* [*i.e., there exists a unique isometric embedding of* $\mathbb{R}_+$ *into* $(\mathcal{T}, d_{\mathcal{T}})$ *that maps* 0 *to* $\rho_{\mathcal{T}}$];

(b) (i) *the locally finite Borel measure* $\mu_{\mathcal{T}}$ *on* $(\mathcal{T}, d_{\mathcal{T}})$ *is nonatomic and supported on the leaves of* $\mathcal{T}$, *that is*, $\mu_{\mathcal{T}}(\mathcal{T}^o) = 0$, *where* $\mathcal{T}^o := \mathcal{T} \setminus \{x \in \mathcal{T} : \deg_{\mathcal{T}}(x) = 1\}$;

(ii) *given* $R > 0$, *there exists a random* $r_0(\mathcal{T}) > 0$ *and deterministic* $c_1, c_2 \in (0, \infty)$ *such that*

$$c_1 r^{d_f} (\log r^{-1})^{-80} \le \mu_{\mathcal{T}}(B_{\mathcal{T}}(x, r)) \le c_2 r^{d_f} (\log r^{-1})^{80}, \tag{1.3}$$

for every $x \in B_{\mathcal{T}}(\rho_{\mathcal{T}}, R)$ *and* $r \in (0, r_0(\mathcal{T}))$, *where* $B_{\mathcal{T}}(x, r)$ *is the open ball centred at* x *with radius* r *in* $(\mathcal{T}, d_{\mathcal{T}})$;

(iii) *there exists a random* $r_0(\mathcal{T}) > 0$ *and deterministic* $c_1, c_2 \in (0, \infty)$ *such that*

$$c_1 r^{d_f} (\log\log r^{-1})^{-9} \le \mu_{\mathcal{T}}(B_{\mathcal{T}}(\rho_{\mathcal{T}}, r)) \le c_2 r^{d_f} (\log\log r^{-1})^{3}, \tag{1.4}$$

for every $r \in (0, r_0(\mathcal{T}))$;

(c) (i) *the restriction of the continuous map* $\phi_{\mathcal{T}} : \mathcal{T} \to \mathbb{R}^2$ *to* $\mathcal{T}^o$ *is a homeomorphism between* $\mathcal{T}^o$ (*equipped with the topology induced by the metric* $d_{\mathcal{T}}$) *and its image* $\phi_{\mathcal{T}}(\mathcal{T}^o)$ (*equipped with the Euclidean topology*), *the latter of which is dense in* $\mathbb{R}^2$;

(ii) $\max_{x \in \mathcal{T}} \deg_{\mathcal{T}}(x) = 3 = \max_{x \in \mathbb{R}^2} |\phi_{\mathcal{T}}^{-1}(x)|$;

(iii) $\mu_{\mathcal{T}} = \mathcal{L} \circ \phi_{\mathcal{T}}$.

The second topic of this paper is the scaling limit of the simple random walk (SRW) on the two-dimensional UST. For a given realisation of the graph $\mathcal{U}$, the SRW on $\mathcal{U}$ is the discrete time Markov process $X^{\mathcal{U}} = ((X_n)_{n \ge 0}, (P_x^{\mathcal{U}})_{x \in \mathbb{Z}^2})$ which at each time step jumps from its current location to a uniformly chosen neighbour in $\mathcal{U}$ (considered as a graph); see Figure 1. For $x \in \mathbb{Z}^2$, the law $P_x^{\mathcal{U}}$ is called the *quenched* law of the simple random walk on $\mathcal{U}$ started at x. Since 0 is always an element of $\mathcal{U}$, we can define the *annealed* or *averaged* law $\mathbb{P}$ as the semi-direct product of the environment law $\mathbf{P}$ and the quenched law $P_0^{\mathcal{U}}$ by setting

$$\mathbb{P}(\cdot) := \int P_0^{\mathcal{U}}(\cdot)\, d\mathbf{P}. \tag{1.5}$$

It is this measure for which we will deduce scaling behaviour.

Techniques for deriving the scaling limits of random walks on the kinds of trees generated by critical branching processes have previously been developed in [16, 20, 21]; see also [36], Chapter 7, for a survey. In the present work, we adapt these to prove a general result of the following form; see Theorem 6.1 below for details and some additional technical conditions. If we have a sequence of graph trees (T_n), $n \ge 1$, each equipped with its intrinsic metric d_{T_n}, a measure μ_{T_n}, an embedding $\phi_{T_n} : T_n \to \mathbb{R}^2$ and a distinguished root vertex ρ_n, for which there exist null sequences $(a_n)_{n\ge1}$, $(b_n)_{n\ge1}$, $(c_n)_{n\ge1}$ with $b_n = o(a_n)$

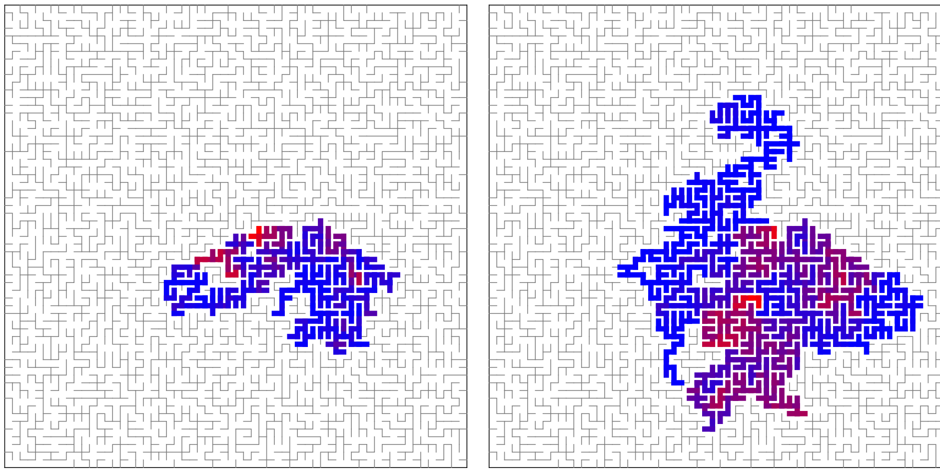

FIG. 1. *The range of a realisation of the simple random walk on uniform spanning tree on a* 60×60 *box* (*with wired boundary conditions*), *shown after* 5000 *and* 50,000 *steps*. *From most to least crossed edges*, *colours blend from red to blue*.

such that $(T_n, a_n d_{T_n}, b_n \mu_{T_n}, c_n \phi_{T_n}, \rho_{T_n}) \to (\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \phi_{\mathcal{T}}, \rho_{\mathcal{T}})$ in the space of measured, rooted spatial trees, then the corresponding rescaled random walks $(c_n \phi_{T_n}(X^{T_n}_{t/a_n b_n}))_{t \geq 0}$ converge in distribution. Further, the limiting process can be written as $(\phi_{\mathcal{T}}(X^{\mathcal{T}}_t))_{t\geq 0}$, where $X^{\mathcal{T}} = ((X^{\mathcal{T}}_t)_{t \geq 0}, (P^{\mathcal{T}}_x)_{x \in \mathcal{T}})$ is the canonical Brownian motion on $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}})$, as constructed in [6], for example, cf. [32]. (We give a brief introduction to Brownian motion on measured real trees at the start of Section 6.)

Combining Theorems 1.1 and 6.1 we obtain the following theorem, which establishes the existence of subsequential scaling limits for the annealed law of the simple random walk on $\mathcal{U}$. Given the volume estimates (1.3), the general results of [18] yield sub-diffusive transition density bounds for the limiting diffusion. These demonstrate that, uniformly over bounded regions of space, the transition density in question has at most logarithmic fluctuations from the leading order polynomial terms in both the on-diagonal and exponential off-diagonal decay parts. In Section 7, we also deduce pointwise on-diagonal estimates with only log-logarithmic fluctuations (cf. the discrete result of [11], Theorem 4.5(a)), as well as annealed on-diagonal polynomial bounds. We note the similarity between these results and the transition density estimates for the Brownian continuum random tree given in [19].

THEOREM 1.4. *If* $(\mathbf{P}_{\delta_i})_{i \geq 1}$ *is a convergent sequence with limit* $\tilde{\mathbf{P}}$, *then the following statements hold*:

(a) *The annealed law of* $(\phi_{\mathcal{T}}(X_t^{\mathcal{T}}))_{t\geq 0}$, *where* $X^{\mathcal{T}}$ *is Brownian motion on* $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}})$ *started from* $\rho_{\mathcal{T}}$, *that is*,

$$\tilde{\mathbb{P}}(\cdot) := \int P_{\rho_{\mathcal{T}}}^{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1}(\cdot)\, d\tilde{\mathbf{P}}, \tag{1.6}$$

is a well-defined probability measure on $C(\mathbb{R}_+, \mathbb{R}^2)$.

(b) *If* $\mathbb{P}_\delta$ *is defined to be the law of* $(\delta X^{\mathcal{U}}_{\delta^{-\kappa d_w} t})_{t\geq 0}$ *under* $\mathbb{P}$, *where the walk dimension* d_w *of* $\mathcal{U}$ *is defined by*

$$d_w := 1 + d_f = \tfrac{13}{5},$$

then $(\mathbb{P}_{\delta_i})_{i\geq 1}$ *converges to* $\tilde{\mathbb{P}}$.

(c) $\tilde{\mathbf{P}}$-*a.s., the process* $X^{\mathcal{T}}$ *is recurrent and admits a jointly continuous transition density* $(p_t^{\mathcal{T}}(x,y))_{x,y\in\mathcal{T},t>0}$. *Moreover, it* $\tilde{\mathbf{P}}$-*a.s. holds that, for any* $R > 0$, *there exist random constants* $c_i(\mathcal{T})$ *and* $t_0(\mathcal{T}) \in (0,\infty)$ *and deterministic constants* $\theta_1, \theta_2, \theta_3, \theta_4 \in (0,\infty)$ (*not depending on* R) *such that*

$$\begin{aligned} p_t^{\mathcal{T}}(x,y) &\leq c_1(\mathcal{T}) t^{-d_f/d_w} \ell(t^{-1})^{\theta_1} \\ &\quad \times \exp\left\{-c_2(\mathcal{T})\left(\frac{d_{\mathcal{T}}(x,y)^{d_w}}{t}\right)^{1/(d_w-1)} \ell(d_{\mathcal{T}}(x,y)/t)^{-\theta_2}\right\}, \\ p_t^{\mathcal{T}}(x,y) &\geq c_3(\mathcal{T}) t^{-d_f/d_w} \ell(t^{-1})^{-\theta_3} \\ &\quad \times \exp\left\{-c_4(\mathcal{T})\left(\frac{d_{\mathcal{T}}(x,y)^{d_w}}{t}\right)^{1/(d_w-1)} \ell(d_{\mathcal{T}}(x,y)/t)^{\theta_4}\right\}, \end{aligned}$$

for all $x, y \in B_{\mathcal{T}}(\rho_{\mathcal{T}}, R)$, $t \in (0, t_0(\mathcal{T}))$, *where* $\ell(x) := 1 \vee \log x$.

REMARK 1.5. If follows that for $\tilde{\mathbf{P}}$-a.e. realisation of $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \phi_{\mathcal{T}}, \rho_{\mathcal{T}})$, we have that

$$-\lim_{t\to 0} \frac{2\log p_t^{\mathcal{T}}(x,x)}{\log t} = \frac{2d_f}{1+d_f} = \frac{16}{13} \qquad \text{for every } x \in \mathcal{T}.$$

Using the language of diffusions on fractals, this means that the spectral dimension of the limiting tree is $\tilde{\mathbf{P}}$-a.s. equal to $16/13$, which is the same as for the discrete model (see [11]).

The remainder of this article is organised as follows. In Section 2, we prove some key estimates for $\mathcal{U}$, which enable us to compare distances in the Euclidean and intrinsic metrics on this set. These allow us to extend some of the volume estimates of [11]. In Section 3, we introduce our topology for measured, rooted spatial trees, and in Section 4 we prove tightness in this topology for the rescaled trees. The properties of limiting trees are studied in Section 5. Following this, we turn our attention to the simple random walk on $\mathcal{U}$, establishing in Section 6 a general convergence result for simple random walks on measured, rooted spatial trees and applying this to the two-dimensional UST. In addition, we explain how this convergence result can be applied to branching random walks and trees without embeddings. In Section 7, we then derive the transition density estimates for the limiting diffusion.

We write c or c_i for constants in $(0, \infty)$; these will be universal and nonrandom, but may change in value from line to line. We use the notation $c_i(\mathcal{T})$ for (random) constants which depend on the tree $\mathcal{T}$.

2. UST estimates. In this section, we obtain estimates for the two-dimensional UST $\mathcal{U}$, which improve those in [11]. Our arguments will depend heavily on Wilson's algorithm, which gives the construction of $\mathcal{U}$ in terms of LERW. In particular, we can construct $\mathcal{U}$ by first running an infinite loop-erased random walk from 0 to ∞ (for details of this see [45]), and then, sequentially running through vertices $x \in \mathbb{Z}^2 \setminus \{0\}$, adding a loop-erased random walk path from $x \in \mathbb{Z}^2$ to the part of the tree already created. We remark that $\mathcal{U}$ is a one-ended tree; see [12].

We will consider three metrics on $\mathcal{U}$, which we now introduce. We define d_E to be the Euclidean metric on $\mathbb{Z}^2$, and write $B_E(x, r) = \{y : d_E(x, y) \leq r\}$ for balls in this metric. For $x \in \mathbb{Z}^2$, we let $\gamma(x, y)$ be the unique path in $\mathcal{U}$ between x and y. We define the intrinsic (shortest path) metric $d_{\mathcal{U}}$ by setting $d_{\mathcal{U}}(x, y) := |\gamma(x, y)|$, that is, the number of edges on the path $\gamma(x, y)$, and write $B_{\mathcal{U}}(x, r)$ for balls in this metric. Finally, it will also be helpful to use a modification of a metric introduced by Schramm in [48], given by

$$d_{\mathcal{U}}^S(x, y) := \operatorname{diam}(\gamma(x, y)), \tag{2.1}$$

where the right-hand side refers to the diameter of $\gamma(x, y)$ in the metric d_E.

We begin by recalling the comparison between $B_{\mathcal{U}}(0, r^{1/\kappa})$ and $B_E(0, r)$ and the estimates on the size of $|B_{\mathcal{U}}(0, r)|$ from [11].

THEOREM 2.1 (See [11], Theorems 1.1, 1.2). (a) *There exist c_1, c_2 such that for every $r \geq 1$ and $\lambda \geq 1$,*

$$\mathbf{P}\big(B_{\mathcal{U}}(0, \lambda^{-1} r^{\kappa}) \not\subset B_E(0, r)\big) \leq c_1 e^{-c_2 \lambda^{2/3}},$$

$$\mathbf{P}\big(B_E(0, r) \not\subset B_{\mathcal{U}}(0, \lambda r^{\kappa})\big) \leq c_1 \lambda^{-1/5}.$$

(b) *There exist c_1, c_2 such that for every $r \geq 1$ and $\lambda \geq 1$,*

$$\mathbf{P}\big(|B_{\mathcal{U}}(0, r)| \geq \lambda r^{d_f}\big) \leq c_1 e^{-c_2 \lambda^{1/3}},$$

$$\mathbf{P}\big(|B_{\mathcal{U}}(0, r)| \leq \lambda^{-1} r^{d_f}\big) \leq c_1 e^{-c_2 \lambda^{1/9}}.$$

Since the law of $\mathcal{U}$ is translation invariant, the above result also holds for $B_{\mathcal{U}}(x, r)$ for any $x \in \mathbb{Z}^d$. However, we wish to have these bounds (for suitable r, n)

for every $x \in B_E(0,n)$; obtaining such uniform estimates is one of the main goals of this section. If we use a simple union bound, as, for example, in [11], (4.47), we obtain an error estimate of the form $n^2 \exp(-\lambda^c)$, which is only small when $\lambda \gg (\log n)^{1/c}$. To improve this, for a suitable $\delta = \delta(\lambda) > 0$ we choose a δ-cover D of $B_E(0,n)$ with $|D| \le c\delta^{-2}$. (Recall that a subset A of $\mathbb{Z}^2$ is called a λ-*cover* if every point of $\mathbb{Z}^2$ is within distance λ of a point of A.) We then obtain good behaviour of $B_{\mathcal{U}}(x,r)$ for all $x \in D$, except on a set of probability $|D|\exp(-\lambda^c)$. Using a "filling in lemma" (see Lemma 2.3 below), together with some additional bounds, we are able to extend this good behaviour to $B_{\mathcal{U}}(y,r)$ for all $y \in B_E(0,n)$; see Proposition 2.10 below for a uniform version of part (b) in particular. An example of the kind of additional result that we need is that if $d_E(x,y) = 3r$ then every path in $\mathcal{U}$ between $B_E(x,r)$ and $B_E(y,r)$ is of length at least cr^κ, except on a set of trees of small probability; note that [10], Theorem 1.2, shows that with high probability the unique infinite self avoiding path in $\mathcal{U}$ started from x takes at least cr^κ steps to escape a Euclidean ball of radius r, but again this does not readily extend to a uniform bound and so further work is required.

We proceed by introducing some further notation and results from [11]. Let $\gamma_x = \gamma(x,\infty)$ be the unique infinite self avoiding path in $\mathcal{U}$ started at x; by Wilson's algorithm γ_x has the law of the loop-erased random walk from x to ∞. Write $\gamma_x[i]$ for the ith point on γ_x, and let $\tau_{y,r} = \tau_{y,r}(\gamma_x) = \min\{i : \gamma_x[i] \notin B_E(y,r)\}$. Whenever we use notation such as $\gamma_x[\tau_{y,r}]$, the exit time $\tau_{y,r}$ will always be for the path γ_x. We define the segment of the path γ_x between its ith and jth points by $\gamma_x[i,j] = (\gamma_x[i], \gamma_x[i+1], \dots, \gamma_x[j])$, and define $\gamma_x[i,\infty)$ in a similar fashion. For such paths, the following was established in [11].

LEMMA 2.2 (See [11], Lemma 2.4). *There exists c_1 such that for every $r \ge 1$ and $k \ge 2$,*

$$\mathbf{P}\big(\gamma_x[\tau_{x,kr},\infty) \cap B_E(x,r) \neq \emptyset\big) \le c_1 k^{-1}.$$

We next give the filling in lemma that we will use several times. This is a small extension of [11], Proposition 3.2. Note that [10], Proposition 6.2, shows that the function $G(r)$ considered in [11] is comparable with the function $G_2(r)$ appearing in (1.1).

LEMMA 2.3. *There exist constants $c_1, c_2 \in (0,\infty)$ such that for each $\delta \le 1$ the following holds. Let $r \ge 1$, and U_0 be a fixed tree in $\mathbb{Z}^2$ with the property that $d_E(x,U_0) \le \delta r$ for each $x \in B_E(0,r)$. Let $\mathcal{U}$ be the random spanning tree in $\mathbb{Z}^2$ obtained by running Wilson's algorithm with root U_0 (i.e., starting from the tree U_0). Then there exists an event G such that $\mathbf{P}(G^c) \le c_1 e^{-c_2\delta^{-1/3}}$, and on G we have that for all $x \in B_E(0,r/2)$,*

$$d_{\mathcal{U}}(x,U_0) \le (\delta^{1/2} r)^\kappa; \qquad d^S_{\mathcal{U}}(x,U_0) \le \delta^{1/2} r; \qquad \gamma(x,U_0) \subset B_E(0,r).$$

PROOF. Except for the bound involving $d^S_{\mathcal{U}}$ this is proved in [11]. [Note that the hypothesis there that U_0 connects 0 to $B_E(0,2r)^c$ is unnecessary.] The proof of the $d^S_{\mathcal{U}}$ bound is similar. □

The following sequence of lemmas will improve the results in [11] on the comparison of the metrics $d_{\mathcal{U}}$, d_E and $d^S_{\mathcal{U}}$. Fix for now $r,k\geq 1$ and $x\in\mathbb{Z}^2$, and choose points z_j on γ_x so that $z_0=x$ and $z_j=\gamma_x[s_j]$, where $s_j=\min\{i: d_E(\gamma_x[i],\{z_0,\dots,z_{j-1}\})\geq r/k\}$. Let $N=N(r,k)=\max\{j: s_j\leq \tau_{x,r}(\gamma_x)\}$. Moreover, define a collection of disjoint balls $\mathcal{B}_{r,k}=\{B_j=B_E(z_j,r/3k), j=1,\dots,N(r,k)\}$. These depend on the path γ_x, and when we need to recall this we will write $\mathcal{B}_{r,k}(\gamma_x)$. Let $a=1+k^{-1/8}$, and set

$$F_1(x,r,k)=\{\gamma_x[\tau_{x,ar},\infty] \text{ hits fewer than } k^{1/2} \text{ of } B_1,\dots,B_{N(r,k)}\}.$$

LEMMA 2.4. *There exist constants $c_1,c_2\in(0,\infty)$ such that, if $r,k\geq 1$ and $x\in\mathbb{Z}^2$, then*

$$\mathbf{P}(F_1(x,r,k)^c)\leq c_1 e^{-c_2 k^{1/8}}.$$

PROOF (see [11], Lemma 3.7). Write $\tau_s=\tau_{x,s}(\gamma_x)$, and let $b=e^{k^{1/8}}\geq 2$. Then by Lemma 2.2,

$$\mathbf{P}(\gamma_x[\tau_{br},\infty]\cap B_E(x,r)\neq\varnothing)\leq cb^{-1}=ce^{-k^{1/8}}. \tag{2.2}$$

If $\gamma_x[\tau_{ar},\infty]$ hits more than $k^{1/2}$ balls from the family $\mathcal{B}_{r,k}(\gamma_x)$, then either γ_x hits $B_E(0,r)$ after time τ_{br}, or $\gamma_x[\tau_{ar},\tau_{br}]$ hits more than $k^{1/2}$ balls. Given (2.2), it is therefore sufficient to prove that

$$\mathbf{P}(\gamma_x[\tau_{ar},\tau_{br}] \text{ hits more than } k^{1/2} \text{ balls})\leq c_1 e^{-c_2 k^{1/8}}. \tag{2.3}$$

Let S be a simple random walk on $\mathbb{Z}^2$ started at x, L' be the loop-erasure of $S[0,\tau_{x,4br}(S)]$, and $L''=L'[\tau_{x,ar}(L'),\tau_{x,br}(L')]$. Then by [45], Corollary 4.5, in order to prove (2.3), it is sufficient to prove that

$$\mathbf{P}(L'' \text{ hits more than } k^{1/2} \text{ balls in } \mathcal{B}_{r,k}(L'))\leq c_1 e^{-c_2 k^{1/8}}.$$

Define stopping times for S by letting $T_0=\tau_{x,ar}(S)$ and for $j\geq 1$, setting $R_j=\min\{n\geq T_{j-1}: S_n\in B_E(x,r)\}$ and $T_j=\min\{n\geq R_j: S_n\notin B_E(x,ar)\}$. Note that the balls in $\mathcal{B}_{r,k}(L')$ can only be hit by S in the intervals $[R_j,T_j]$ for $j\geq 1$. Let $M=\min\{j: R_j\geq \tau_{x,4br}(S)\}$. Then, by the result of [39], Exercise 1.6.8, $\mathbf{P}(M=j+1|M>j)\geq c(\log(ar)-\log r)/(\log(4br)-\log r)\geq ck^{-2/8}$. Hence, $\mathbf{P}(M\geq k^{3/8})\leq c_1\exp(-c_2k^{1/8})$. Now, for each $j\geq 1$, let L_j be the loop-erasure of $S[0,T_j]$, α_j be the first exit by L_j from $B_E(x,ar)$, and β_j be the number of steps in L_j. If L'' hits more than $k^{1/2}$ balls in $\mathcal{B}_{r,k}(L')$, then there must exist some $j\leq M$ such that $L_j[\alpha_j,\beta_j]$ hits more than $k^{1/2}$ of the balls in the collection $\mathcal{B}_{r,k}(L_j)$. Hence, if $M\leq k^{3/8}$ and L'' hits more than $k^{1/2}$ balls in $\mathcal{B}_{r,k}(L')$,

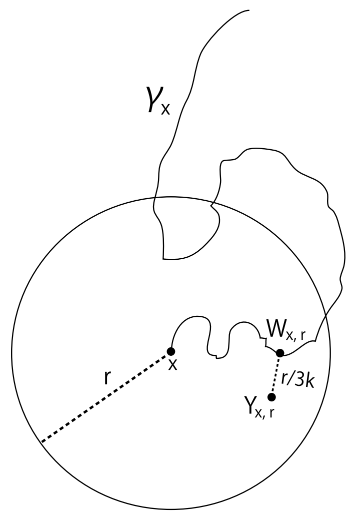

FIG. 2. *A sample of* $A_1(x,r,k)$ *in Lemma* 2.5.

then S must hit more than $k^{1/8}$ balls in $\mathcal{B}_{r,k}(L_j)$ in one of the intervals $[R_j, T_j]$, without hitting the path $L_j[0,\alpha_j]$. However, by Beurling's estimate (see [39], Lemma 2.5.3, e.g.), the probability of this event is less than $c_1 \exp(-c_2 k^{1/8})$. Combining these estimates completes the proof. □

Our next lemma shows that if D_0 is δr-cover of $B_E(x,2r)$, then with high probability we can find points $Y_{x,r}$ and $W_{x,r}$ which are close to the boundary of $B_E(x,r)$ and to each other, and such that $Y_{x,r} \in D_0$ and $W_{x,r} \in \gamma_x \cap B_E(x,r)$. (See Figure 2.) In the proof, we refer to the event $F_2(x,r,k) = \{8k^{-1/4} r^\kappa \le \tau_{x,r}(\gamma_x) \le k^{1/4} r^\kappa\}$. From [10], Theorems 5.8, 6.1, we have

$$\mathbf{P}\big(F_2(x,r,k)^c\big) \le c_1 \exp\big(-c_2 k^{1/6}\big). \tag{2.4}$$

LEMMA 2.5. *Let* $r \ge 1$, $k \ge 2$, $x \in \mathbb{Z}^2$, *and* $D_0 \subset \mathbb{Z}^2$ *satisfy* $B_E(x,2r) \subset \bigcup_{y \in D_0} B_E(y, r/18k)$. *Then there exists an event* $A_1 = A_1(x,r,k)$, *defined in* (2.8) *below, which satisfies*

$$\mathbf{P}\big(A_1^c\big) \le e^{-k^{1/8}}, \tag{2.5}$$

and on $A_1(x,r,k)$ *there exists* $T \le \tau_{x,r}(\gamma_x)$ *such that, writing* $W_{x,r} = \gamma_x(T)$:

(a) $k^{-1/4} r^\kappa \le T \le k^{1/4} r^\kappa$;
(b) $a^{-2} r \le d_E(x, W_{x,r}) \le r$;
(c) *there exists* $Y_{x,r} \in D_0$ *such that* $d_E(Y_{x,r}, W_{x,r}) \le r/3k$, $d_{\mathcal{U}}^S(Y_{x,r}, W_{x,r}) \le 2r/3k$ *and also* $d_{\mathcal{U}}(Y_{x,r}, W_{x,r}) \le c_1 (r/k)^\kappa$.

PROOF. Fix $k \geq 1$ and recall that $a = 1 + k^{-1/8}$. Suppose that the event

$$F_1(x, r/a, k) \cap F_2(x, r/a^2, k) \cap F_2(x, r, k) \tag{2.6}$$

occurs. Write $\tau_s = \tau_{x,s}(\gamma_x)$, $T_1 = \tau_{r/a^2}$, and $T_2 = \tau_{r/a}$. Let $J_0 = J_0(\omega)$ be the set of j such that $z_j \in \gamma_x[T_1, T_2]$ and $B_E(z_j, r/3ak) \subset B_E(x, r/a) \setminus B_E(x, r/a^2)$. Then $|J_0| \geq ck^{7/8}$. Since $F_1(x, r/a, k)$ holds, at most $k^{1/2}$ of the balls $(B_E(z_j, r/3ak)$, $j \in J_0)$ are hit by $\gamma_x[\tau_r, \infty]$. So if $J = J(\omega)$ is the set of $j \in J_0$ such that $B_E(z_j, r/3ak) \cap \gamma_x[\tau_r, \infty] = \varnothing$, then $|J| \geq k^{7/8} - k^{1/2} \geq ck^{3/4}$. For each $j \in J$, we can find a point $y_j \in D_0$ with $d_E(y_j, z_j) \leq r/18k$. Hence, $B_E(y_j, r/18k) \cap \gamma_x[T_1, T_2] \neq \varnothing$, while $B_E(y_j, r/9k) \cap \gamma_x[\tau_r, \infty] = \varnothing$. Note that $B_E(y_j, r/9k)$ may however intersect the path γ_x in the interval $[T_2, \tau_r]$.

For the remainder of the proof, it will be helpful to regard γ_x as a fixed deterministic path which satisfies the conditions in (2.6). For each $j \in J$, let X^j be a SRW started at y_j and run until it hits γ_x, and let L^j be the loop-erasure of X^j. Let

$$H_j = \{X^j \text{ hits } \gamma_x \text{ before it exits } B_E(z_j, r/3ak), |L^j| \leq c_0(r/3k)^\kappa\}.$$

By [11], Theorem 2.2, we have [taking $D = \mathbb{Z}^2 \setminus \gamma_x$ and $D' = D \cap B_E(z_j, r/3ak)$],

$$\mathbf{P}(|L^j \cap B_E(z_j, r/3ak)| > \lambda(r/k)^\kappa) \leq c_1 \exp(-c_2\lambda).$$

So, by Beurling's estimate (see [39], Lemma 2.5.3, e.g.), we can choose c_0 so that there exists $p > 0$ such that $\mathbf{P}(H_j) \geq p$.

Recall now the implementation of Wilson's algorithm using "stacks" (see [49]). For each j, assume we have stack variables $\xi_{x,i}$ for $x \in B_E(z_j, r/3ak)$. We use these to make a random walk path X^j started at y_j and run either it hits γ_x or leaves $B_E(z_j, r/3ak)$. Thus, the event H_j is measurable with respect to $\sigma(\xi_{x,i}, i \geq 1, x \in B_E(z_j, r/3ak))$. We now consider the y_j one at a time, and continue until either we obtain a success, or we have tried $k^{3/4}$ of the points y_j. Since these events are independent, if H is the event that we obtain a success, then

$$\mathbf{P}(H^c) \leq (1-p)^{k^{3/4}} \leq c_1 \exp(-c_2 k^{3/4}). \tag{2.7}$$

If H occurs, with a success for y_j, set $Y = Y_{x,r} = y_j$, let $W = W_{x,r}$ be the point where X^j hits γ_x, and let T be such that $\gamma_x(T) = W$. We take

$$A_1 = A_1(x, r, k) = H \cap F_1(x, r/a, k) \cap F_2(x, r/a^2, k) \cap F_2(x, r, k). \tag{2.8}$$

By Lemma 2.4, (2.4) and (2.7), we have the upper bound (2.5) on $\mathbf{P}(A_1^c)$.

Finally, suppose that $A_1(x, r, k)$ occurs. By construction, we have $d_E(Y, W) \leq r/3k$, and since the path X^j lies inside $B_E(z_j, r/3k)$ we also have $d_{\mathcal{U}}^S(Y, W) \leq 2r/3k$. The definition of the event H_j gives that $d_{\mathcal{U}}(Y, W) \leq c(r/k)^\kappa$. Since X^j hits γ_x inside $B_E(z_j, r/3ak)$, we must have $W \in B_E(x, r) \setminus B_E(x, a^{-2}r)$. Moreover, because $j \in J$, $T \leq \tau_r(\gamma_x)$, so since $F_2(x, k, r)$ holds we have $T \leq k^{1/4} r^\kappa$.

Since $B_E(z_j, r/3ak) \cap B_E(x, r/a^2) = \varnothing$, we must also have $T \geq \tau_{r/a^2}$, and so $T \geq 8k^{-1/4}(r/a^2)^\kappa \geq k^{-1/4}r^\kappa$. $\square$

The next lemma allows us to compare $d_{\mathcal{U}}^S$ and $d_{\mathcal{U}}$ on a large family of paths in a ball.

LEMMA 2.6. *Let $r \geq 1$, $k \geq 8$, and $x \in \mathbb{Z}^2$. Set $M_1 = e^{k^{1/8}/27}$, $M_2 = e^{k^{1/8}/3}$, $R_i = rM_i$, and let $D_0 \subset B_E(x, 2R_2)$ satisfy $|D_0| \leq ck^2M_2^2$, $|D_0 \cap B_E(x, 2R_1)| \leq ck^2M_1^2$, $B_E(x, 2R_2) \subset \bigcup_{y \in D_0} B_E(y, r/18k)$. Write $D_1 = D_0 \cap B_E(x, 2R_1)$. Then there exist constants b_1, b_2 and an event $A_2 = A_2(x, r, k)$ with*

$$\mathbf{P}(A_2^c) \leq c\exp(-k^{1/8}/4), \tag{2.9}$$

such that on A_2 the following holds for every $y \in D_1$:

(a) $\gamma_y[\tau_{x,R_2}, \infty] \cap B_E(x, 4R_1) = \varnothing$;
(b) *if $x_1, x_2 \in \gamma_y[0, \tau_{x,R_2}]$ and $d_{\mathcal{U}}^S(x_1, x_2) > b_2 r$, then $d_{\mathcal{U}}(x_1, x_2) \geq \frac{1}{2}k^{-1/4}r^\kappa$;*
(c) *if $x_1, x_2 \in \gamma_y[0, \tau_{x,R_2}]$ and $d_{\mathcal{U}}^S(x_1, x_2) < b_1 r$, then $d_{\mathcal{U}}(x_1, x_2) \leq 2k^{1/4}r^\kappa$.*

PROOF. For $y \in D_1$, let $F_3(y, r, k) = \{\gamma_y[\tau_{x,R_2}, \infty] \cap B_E(x, 4R_1) = \varnothing\}$. By Lemma 2.2, we have $\mathbf{P}(F_3^c) \leq cM_1/M_2$. Now set

$$A_2 = \Big(\bigcap_{y \in D_0} A_1(y, r, k)\Big) \cap \Big(\bigcap_{y \in D_1} F_3(y, r, k)\Big),$$

where $A_1(y, r, k)$ is the event defined by (2.8). From (2.5), we note that

$$\mathbf{P}(A_2^c) \leq ck^2M_2^2e^{-k^{1/8}} + cM_1^2k^2M_1M_2^{-1} \leq c\exp(-k^{1/8}/4).$$

Now suppose that A_2 holds, and let $y \in D_1$. It is immediate that (a) holds. Write $W_0 = Y_0 = y$, and let $Y_1 = Y_{Y_0,r}$ and $W_1 = W_{Y_0,r}$ be the points given by the event $A_1(Y_0, r, k)$. Similarly write Y_{j+1} and W_{j+1} for the points given by the event $A_1(Y_j, k, r)$ for $j \geq 1$, and continue until we have for some $N = N_y$ that $W_N \notin B_E(x, 3R_2/2)$. Note that both $d_{\mathcal{U}}^S$ and $d_{\mathcal{U}}$ are monotone on the path γ_y, in the sense that if $x_1, x_2 \in \gamma_y$ and $x_3 \in \gamma(x_1, x_2)$ then for $\rho = d_{\mathcal{U}}^S$ or $\rho = d_{\mathcal{U}}$ then $\rho(x_1, x_3) \leq \rho(x_1, x_2)$. This is immediate for $d_{\mathcal{U}}$ and easily proved from the definition of $d_{\mathcal{U}}^S$.

The construction of the (Y_j, W_j) gives that

$$\frac{r}{a^2} \leq d_{\mathcal{U}}^S(Y_j, W_{j+1}) \leq r, \qquad d_{\mathcal{U}}^S(Y_j, W_j) \leq \frac{2r}{k},$$
$$k^{-1/4}r^\kappa \leq d_{\mathcal{U}}(Y_j, W_{j+1}) \leq k^{1/4}r^\kappa, \qquad d_{\mathcal{U}}(Y_j, W_j) \leq c(r/k)^\kappa.$$

Thus, we have

$$d^S_{\mathcal{U}}(W_j, W_{j+1}) \le d^S_{\mathcal{U}}(W_j, Y_j) + d^S_{\mathcal{U}}(Y_j, W_{j+1}) \le \frac{2r}{k} + r = \frac{1}{2} b_2 r,$$

$$d^S_{\mathcal{U}}(W_j, W_{j+1}) \ge d^S_{\mathcal{U}}(Y_j, W_{j+1}) - d^S_{\mathcal{U}}(W_j, Y_j) \ge r/a^2 - \frac{2r}{k} = b_1 r.$$

Here, we have used the equations above to define b_1 and b_2. Similarly, we have

$$d_{\mathcal{U}}(W_j, W_{j+1}) \le d_{\mathcal{U}}(Y_j, W_{j+1}) \le k^{1/4} r^{\kappa},$$

$$d_{\mathcal{U}}(W_j, W_{j+1}) \ge d_{\mathcal{U}}(Y_j, W_{j+1}) - d_{\mathcal{U}}(Y_j, W_j) \ge k^{-1/4} r^{\kappa} - c(r/k)^{\kappa} \ge \tfrac{1}{2} k^{-1/4} r^{\kappa}.$$

Let $x_1, x_2 \in \gamma_y[0, \tau_{x,3R_2/2}]$. We can assume that $x_1 \in \gamma(y, x_2)$. Let $j = \min\{i : W_i \in \gamma(x_1, \infty)\}$. If $x_2 \in \gamma(x_1, W_{j+1})$, then $d^S_{\mathcal{U}}(x_1, x_2) \le d^S_{\mathcal{U}}(W_{j-1}, W_j) + d^S_{\mathcal{U}}(W_j, W_{j+1}) \le b_2 r$. So if $d^S_{\mathcal{U}}(x_1, x_2) > b_2 r$, then both W_j and W_{j+1} are on the path $\gamma(x_1, x_2)$, and so $d_{\mathcal{U}}(x_1, x_2) \ge \frac{1}{2} k^{-1/4} r^{\kappa}$, proving (b). Similarly, if both W_j and W_{j+1} are on the path $\gamma(x_1, x_2)$, then we have $d^S_{\mathcal{U}}(x_1, x_2) \ge b_1 r$. So if $d^S_{\mathcal{U}}(x_1, x_2) < b_1 r$, then $W_{j+1} \in \gamma(x_2, \infty)$, and hence $d_{\mathcal{U}}(x_1, x_2) \le 2k^{1/4} r^{\kappa}$. □

We now extend this result to all paths γ_x in a ball.

LEMMA 2.7. *Let $r \ge 1$, $k \ge 8$, $x_0 \in \mathbb{Z}^2$, M_i, R_i, and b_1, b_2 be as in Lemma 2.6. Then there exist constants b_3, b_4 (depending on k) and an event $A_3 = A_3(x_0, r, k)$ with*

$$\mathbf{P}(A_3^c) \le c_1 \exp(-c_2 k^{1/8}), \tag{2.10}$$

such that on A_3 the following holds for every $x \in B_E(x_0, R_1)$:

(a) $\gamma_x[\tau_{x_0,R_2}, \infty] \cap B_E(x_0, 4R_1) = \varnothing$;
(b) *If $x_1, x_2 \in \gamma_x[0, \tau_{x_0,R_2}]$ and $d^S_{\mathcal{U}}(x_1, x_2) > b_3 r$, then $d_{\mathcal{U}}(x_1, x_2) \ge \frac{1}{2} k^{-1/4} r^{\kappa}$;*
(c) *If $x_1, x_2 \in \gamma_x[0, \tau_{x_0,R_2}]$ and $d^S_{\mathcal{U}}(x_1, x_2) < b_1 r$, then $d_{\mathcal{U}}(x_1, x_2) \le b_4 k^{1/4} r^{\kappa}$;*
(d) *If $x_1, x_2 \in B_E(x_0, R_1)$ and $d^S_{\mathcal{U}}(x_1, x_2) > 2b_3 r$, then $d_{\mathcal{U}}(x_1, x_2) \ge \frac{1}{2} k^{-1/4} r^{\kappa}$;*
(e) *If $x_1, x_2 \in B_E(x_0, R_1)$ and $d^S_{\mathcal{U}}(x_1, x_2) < b_1 r$, then $d_{\mathcal{U}}(x_1, x_2) \le 2b_4 k^{1/4} r^{\kappa}$.*

PROOF. We begin by choosing a set D_0 which satisfies the conditions of Lemma 2.6. Let $A_2(x_0, r, k)$ be the event defined in that lemma, and let U_0 be the random tree obtained by applying Wilson's algorithm with initial points in $D_1 = D_0 \cap B(x_0, 2R_1)$. Let $z \in B_E(x_0, R_1)$. We now apply the filling in Lemma 2.3 to $B_E(z, r)$ taking $\delta = 1/18k$. Let $G(z)$ be the "good" event given by the lemma; we have

$$\mathbf{P}(G(z)^c) \le c \exp(-ck^{1/3}). \tag{2.11}$$

Now choose z_i, $i = 1, \ldots, N$ so that $N \le cM_1^2$ and $B_E(x_0, R_1) \subset \bigcup_i B_E(z_i, r/4)$, and let $A_3 = A_2(x_0, r, k) \cap (\bigcap_{i=1}^N G(z_i))$. The bound (2.10) then follows from (2.9) and (2.11).

Let $x \in B(x_0, R_1)$, and let W_x be the point where γ_x first hits the tree U_0. Since $G(z_i)$ holds for some z_i with $d_E(x, z_i) \le r/4$, we have by Lemma 2.3 that $d^S_{\mathcal{U}}(x, W_x) \le ck^{-1/2}r$, $d_{\mathcal{U}}(x, W_x) \le c(k^{-1/2}r)^\kappa$. Since $U_0 = \bigcup_{y \in D_1} \gamma_y$ there must exist a $y \in D_1$ such that $W_x \in \gamma_y$. Let W_j be the points given in the proof of Lemma 2.6. By property Lemma 2.6(a), we have that γ_y does not return to $B_E(x_0, 4R_1)$ after leaving $B_E(x_0, R_2)$ and, therefore, there exists j such that $W_x \in \gamma(W_{j-1}, W_j)$. (We take $W_x = W_j$ if W_x is one of the points W_i.) Note also that property (a) of γ_x follows from the same property for γ_y.

Let x_1, x_2 be on the path $\gamma_x[0, \tau_{x_0,R_2}]$; we can assume that $x_1 \in \gamma(x, x_2)$. If $x_1 \in \gamma(W_x, \infty)$ then both x_1 and x_2 are in γ_y, and so properties (b) and (c) follow from Lemma 2.6. So suppose that $x_1 \in \gamma(x, W_x)$. If $x_2 \in \gamma(x, W_{j+1})$, then

$$\begin{aligned} d^S_{\mathcal{U}}(x_1, x_2) &\le d^S_{\mathcal{U}}(x, W_x) + d^S_{\mathcal{U}}(W_x, W_{j+1}) \\ &\le d^S_{\mathcal{U}}(x, W_x) + d^S_{\mathcal{U}}(W_{j-1}, W_j) + d^S_{\mathcal{U}}(W_j, W_{j+1}) \\ &\le ck^{-1/2}r + b_2 r \le (c + b_2)r = b_3 r. \end{aligned}$$

So if $d^S_{\mathcal{U}}(x_1, x_2) > b_3 r$, then $x_2 \in \gamma(W_{j+1}, \infty)$, and hence $d_{\mathcal{U}}(x_1, x_2) \ge d_{\mathcal{U}}(W_j, W_{j+1}) \ge \frac{1}{2}k^{-1/4}r^\kappa$. Similarly, if $x_2 \in \gamma(W_{j+1}, \infty)$, then $d^S_{\mathcal{U}}(x_1, x_2) \ge d^S_{\mathcal{U}}(W_j, W_{j+1}) \ge b_1 r$. So if $d^S_{\mathcal{U}}(x_1, x_2) < b_1 r$, then $x_2 \in \gamma(x, W_{j+1})$, and so

$$\begin{aligned} d_{\mathcal{U}}(x_1, x_2) &\le d_{\mathcal{U}}(x, W_x) + d_{\mathcal{U}}(W_{j-1}, W_j) + d_{\mathcal{U}}(W_j, W_{j+1}) \\ &\le c\big(k^{-1/2}r\big)^\kappa + 2k^{1/4}r^\kappa \le b_4 k^{1/4} r^\kappa. \end{aligned}$$

This proves properties (b) and (c) of γ_x.

Finally, let $x_1, x_2 \in B_E(x_0, R_1)$, and let W be the point where γ_{x_1} and γ_{x_2} meet. If $d^S_{\mathcal{U}}(x_1, x_2) > 2b_3 r$ and $W \in \gamma_{x_1}[0, \tau_{x_0,R_2}] \cap \gamma_{x_2}[0, \tau_{x_0,R_2}]$, then we have $\max_i d^S_{\mathcal{U}}(x_i, W) > b_3 r$, and so $d_{\mathcal{U}}(x_1, x_2) \ge \max_i d_{\mathcal{U}}(x_i, W) \ge \frac{1}{2}k^{-1/4}r^\kappa$. If, on the other hand, $d^S_{\mathcal{U}}(x_1, x_2) > 2b_3 r$ and $W \notin \gamma_{x_i}[0, \tau_{x_0,R_2}]$ for either $i = 1$ or $i = 2$, then set $W' = \gamma_{x_i}(\tau_{x_0,R_2})$ for the relevant i. Note that $W' \in \gamma(x_1, x_2) \cap \gamma_{x_i}[0, \tau_{x_0,R_2}]$ and $d^S_{\mathcal{U}}(x_i, W') \ge b_3 r$, and so $d_{\mathcal{U}}(x_1, x_2) \ge d_{\mathcal{U}}(x_i, W') \ge \frac{1}{2}k^{-1/4}r^\kappa$ in this case as well. Similarly, if $d^S_{\mathcal{U}}(x_1, x_2) < b_1 r$, then necessarily we have $W \in \gamma_{x_1}[0, \tau_{x_0,R_2}] \cap \gamma_{x_2}[0, \tau_{x_0,R_2}]$ and $\max_i d^S_{\mathcal{U}}(x_i, W) < b_1 r$, which implies $d_{\mathcal{U}}(x_1, x_2) \le 2b_4 k^{1/4} r^\kappa$. This proves properties (d) and (e). □

Note that there is a gap between the conditions (d) and (e) above. We could fill this by a direct calculation, but instead we will handle this in the next result by varying r.

PROPOSITION 2.8. *Let* $r \geq 1$, $\lambda \geq \lambda_0$ (*where* λ_0 *is a large, finite constant*), $x_0 \in \mathbb{Z}^2$, *and* $R = re^{c_1\lambda^{1/2}}$. *There exists an event* A_4 *with* $\mathbf{P}(A_4^c) \leq c\exp(-c_2\lambda^{1/2})$ *such that on* A_4, *for all* $x, y \in B_E(x_0, R)$,

$$\lambda^{-1} d_{\mathcal{U}}^S(x,y)^\kappa \leq d_{\mathcal{U}}(x,y) \leq \lambda d_{\mathcal{U}}^S(x,y)^\kappa \qquad \text{if } r \leq d_{\mathcal{U}}^S(x,y) \leq R,$$

$$d_{\mathcal{U}}(x,y) \leq \lambda r^\kappa \qquad \text{if } d_{\mathcal{U}}^S(x,y) \leq r, \tag{2.12}$$

$$d_{\mathcal{U}}(x,y) \geq \lambda^{-1} R^\kappa \qquad \text{if } d_{\mathcal{U}}^S(x,y) \geq R.$$

PROOF. Choose $k = c\lambda^4$, let m be such that $2^{m-1} < \exp(k^{1/8}/27) \leq 2^m$, and define $A_4 = \bigcap_{i=0}^m A_3(x_0, 2^i r, k)$. Then $\mathbf{P}(A_4^c) \leq \exp(-ck^{1/8}) \leq \exp(-c'\lambda^{1/2})$. Now let $x, y \in B_E(x_0, R)$, and suppose $r' = d_{\mathcal{U}}^S(x,y) \leq R$. Then choosing the largest $i \in \{0, 1, \dots, m\}$ so that $r' \geq 2b_3 2^i r$, we have $d_{\mathcal{U}}(x_1, x_2) \geq c\lambda^{-1}(2^i r)^\kappa \geq c\lambda^{-1}(r')^\kappa$. Similarly, we have $d_{\mathcal{U}}(x_1, x_2) \leq c\lambda(r')^\kappa$. Replacing $c\lambda$ by λ this gives (2.12), and the other two inequalities follow. □

One consequence of the above proposition is the following approximation result, which shows that if a set of points is an $r/18k^2$-cover in the Euclidean metric, then it is also a cover with respect to the metrics $d_{\mathcal{U}}^S$ and $d_{\mathcal{U}}$.

PROPOSITION 2.9. *Let* $r \geq k \geq 1$. *Define* $R_1 := re^{k^{1/32}}$, $R_2 := re^{k^{1/16}}$, *and suppose* $D_2 \subseteq \mathbb{Z}^2$ *satisfies*

$$B_E(0, 6R_2) \subseteq \bigcup_{x \in D_2} B_E(x, r/18k^2). \tag{2.13}$$

Then there exists an event $A_5 = A_5(r,k)$ *such that* $\mathbf{P}(A_5^c) \leq c_1 e^{-c_2 k^{1/16}}$ *and on* A_5 *the following holds*:

$$\max_{x \in B_E(0,R_1)} d_{\mathcal{U}}^S(x, D_2) \leq \frac{2r}{k}, \tag{2.14}$$

$$\max_{x \in B_E(0,R_1)} d_{\mathcal{U}}(x, D_2) \leq \frac{4r^\kappa}{k^{1/4}}. \tag{2.15}$$

PROOF. First, choose a subset $D_2' \subseteq D_2$ such that (2.13) holds when D_2 is replaced by D_2' and also $|D_2'| \leq ck^4 e^{2k^{1/16}}$. Set $A'(r,k) := \bigcap_{x \in D_2'} A_1(x, r/k, k)$, where A_1 is defined in the statement of Lemma 2.5. From that result, we know that

$$\mathbf{P}(A'^c) \leq ck^4 e^{2k^{1/16}} \mathbf{P}(A_1(0, r/k, k)) \leq ce^{-ck^{1/8}}. \tag{2.16}$$

Moreover, if A' holds, then for $x \in B_E(0, 2R_1) \cap D_2'$ we can define $(W_j, Y_j)_{j=0}^N$ similarly to the proof of Lemma 2.6. In particular, set $W_0 = Y_0 = x$, and let W_j, Y_j

be given by the event $A_1(Y_{j-1}, r/k, k)$, up to $j = N := \inf\{m : d_E(x, W_m) > 2R_2\}$. By construction, it follows that

$$\begin{aligned}\max_{z\in\gamma_x(0,\tau_{x,R_2})} d^S_{\mathcal{U}}(z, D_2) &\le \max_{j=1,\dots,N} d^S_{\mathcal{U}}(W_{j-1}, W_j)\\ &\le \max_{j=1,\dots,N} d^S_{\mathcal{U}}(Y_{j-1}, W_j) \le \frac{r}{k}.\end{aligned}\tag{2.17}$$

Next, choose $D''_2 \subseteq D'_2 \cap B_E(0, 2R_1)$ such that $B_E(0, R_1) \subseteq \bigcup_{x\in D''_2} B_E(x, r/18k^2)$ and $|D''_2| \le ck^4 e^{2k^{1/32}}$. Set $A''(r, k) := A'(r, k) \cap (\bigcap_{x\in D''_2} B(x, r, k))$, where $B(x, r, k) := \{\gamma_x(\tau_{x,R_2}, \infty) \cap B_E(0, 2R_1) = \varnothing\}$. By applying Lemma 2.2 in conjunction with (2.16), we obtain

$$\mathbf{P}(A''^c) \le ce^{-ck^{1/8}} + ck^4 e^{2k^{1/32}} \mathbf{P}\big(\gamma_0(\tau_{0,R_2}, \infty) \cap B_E(0, 4R_1) \ne \varnothing\big) \le c_1 e^{-c_2 k^{1/16}}.$$

Define U_0 to be the subtree of $\mathcal{U}$ spanned by D''_2 and suppose A'' holds. If $x \in U_0 \cap B_E(0, 2R_1)$, then it must be the case that $x \in \gamma_y(0, \tau_{y,R_2})$ for some $y \in D''_2$. Hence, by (2.17), it holds that $\max_{x\in U_0\cap B_E(0,2R_1)} d^S_{\mathcal{U}}(x, D_2) \le r/k$. Now, by applying Lemma 2.3 with root U_0, it is possible to deduce

$$\mathbf{P}\Big(\max_{x\in B_E(0,R_1)} d^S_{\mathcal{U}}(x, U_0) > \frac{r}{k}\Big) \le Ce^{-ce^{k^{1/32}}}.$$

So, if A''' is defined to be the event that both A'' and $\max_{x\in B_E(0,R_1)} d^S_{\mathcal{U}}(x, U_0) \le r/k$ hold, then we have $\mathbf{P}(A'''^c) \le c_1 e^{-c_2 k^{1/16}}$ and also (2.14) holds on A'''.

To complete the proof, we will use Proposition 2.8 with (x_0, r, λ) given by $(0, 2r/k, k)$ to compare the relevant distances. Since $R = 2rk^{-1}e^{c_1 k^{1/2}} \ge 2R_1$ for large k, we find that with probability exceeding $1 - ce^{-c_2 k^{1/2}}$ it is the case that

$$\max_{\substack{x,y\in B_E(0,2R_1):\\ d^S_{\mathcal{U}}(x,y)\le 2r/k}} d_{\mathcal{U}}(x, y) \le k\Big(\frac{2r}{k}\Big)^{\kappa} \le \frac{4r^\kappa}{k^{1/4}}.$$

Note that if A''' and the above inequality both hold, then so does (2.15). Hence, in conjunction with the conclusion of the previous paragraph, this completes the proof. □

We can now improve the volume estimates of [11]. Recall from (1.2) that $d_f = 2/\kappa = 8/5$, and define, for $\lambda, n \ge 1$,

$$\begin{aligned}\tilde{A}(\lambda, n) := \{\omega : \lambda^{-1} R^{d_f} \le |B_{\mathcal{U}}(x, R)| \le \lambda R^{d_f}\\ \text{for all } x \in B_E(0, n), R \in [e^{-\lambda^{1/40}} n^\kappa, n^\kappa]\}.\end{aligned}$$

The following result extends a fundamental estimate of [11]; the key improvement is that the upper bound does not depend on n (once n is suitably large). Although we do not need to do so here, we note that the same approach can also be used to obtain a similar improvement of the resistance estimates in [11].

PROPOSITION 2.10. *There exist constants* $c_1, c_2 \in (0,\infty)$ *such that*

$$\mathbf{P}\big(\tilde{A}(\lambda, n)^c\big) \le c_1 \exp\big(-c_2\lambda^{1/80}\big) \qquad \textit{for all } n \ge e^{\lambda^{1/16}}.$$

PROOF. Let $k = \lambda$, $r = ne^{-\lambda^{1/32}}$, and let $R_1 = n$, $R_2 = re^{k^{1/6}}$ and D_2 be as in Proposition 2.9, with $|D_2| \le ck^4 e^{2k^{1/16}}$. Set $m_0 := \inf\{m : k^m \ge e^{k^{1/32}}\}$. Let $A_5(r,k)$ be the event given in the statement of Proposition 2.9, and

$$E(r,k) := \bigcap_{x\in D_2} \bigcap_{m=1}^{m_0+1} \{k^{-1}(rk^m)^\kappa \le |B_{\mathcal{U}}(x,(rk^m)^\kappa)| \le k(rk^m)^\kappa\}.$$

A simple union bound allows us to deduce from Theorem 2.1(b) that

$$\mathbf{P}\big(E(r,k)^c\big) \le Ck^4 e^{2k^{1/16}} k^{1/32} ce^{-k^{1/9}} \le Ce^{-ck^{1/9}}.$$

Consequently, we have $\mathbf{P}(E(r,k)^c \cup A_5(r,k)^c) \le c\exp(-c\lambda^{1/16})$.

Suppose that $E(r,k) \cap A_5(r,k)$ holds. Let $x \in B_E(0,n)$, and $s \in [rk^3, n]$. Choose $m \in \{3, \dots, m_0+1\}$ such that $s \in [rk^m, rk^{m+1})$. Since $A_5(r,k)$ holds, there exists $y \in D_2$ with $d_{\mathcal{U}}(x,y) \le 4r^\kappa/k^{1/4}$. Hence,

$$\begin{aligned}|B_{\mathcal{U}}(x,s^\kappa)| &\le |B_{\mathcal{U}}(y,(rk^{m+1})^\kappa + 4r^\kappa/k^{1/4})| \le |B_{\mathcal{U}}(y,(rk^{m+2})^\kappa)| \le k(rk^{m+2})^2 \\ &\le k^5 s^2.\end{aligned}$$

Similarly, $|B_{\mathcal{U}}(x,s^\kappa)| \ge k^{-5}s^2$. Since $(rk^3)^\kappa \le n^\kappa \exp(-\lambda^{1/40})$ it follows that $E(r,k) \cap A_5(r,k) \subset \tilde{A}(\lambda^5, n)$, which completes the proof of the proposition. □

From this, we can prove the following distributional measure bounds, which will be used in the proof of Theorem 1.3(b)(ii).

COROLLARY 2.11. *Given* $R > 0$, *there exist constants* $c_1, \dots, c_7 \in (0,\infty)$ (*depending on* R) *such that for every* $r \in (0, c_7)$,

$$\begin{aligned}&\limsup_{\delta\to 0} \mathbf{P}\Big(\delta^2 \min_{x\in B_E(0,\delta^{-1}R)} \mu_{\mathcal{U}}(B_{\mathcal{U}}(x,\delta^{-\kappa}r)) \le c_1 r^{d_f}(\log r^{-1})^{-80}\Big) \\ &\qquad \le c_2 r^{c_3},\end{aligned} \tag{2.18}$$

$$\begin{aligned}&\limsup_{\delta\to 0} \mathbf{P}\Big(\delta^2 \max_{x\in B_E(0,\delta^{-1}R)} \mu_{\mathcal{U}}(B_{\mathcal{U}}(x,\delta^{-\kappa}r)) \ge c_4 r^{d_f}(\log r^{-1})^{80}\Big) \\ &\qquad \le c_5 r^{c_6}.\end{aligned} \tag{2.19}$$

PROOF. We just prove (2.19); the proof of (2.18) is similar. Fix $R \ge 1$, and suppose $r \in (0,1)$, $\delta \in (0,1)$. Define $n := \delta^{-1}R$ and $\lambda := (\log(R^\kappa/r))^{80}$. Since $\delta^{-\kappa} r \in [e^{-\lambda^{1/40}} n^\kappa, n^\kappa]$, we have that, on $\tilde{A}(\lambda, n)$,

$$\min_{x\in B_E(0,\delta^{-1}R)} \mu_{\mathcal{U}}(B_{\mathcal{U}}(x,\delta^{-\kappa}r)) \ge \lambda^{-1}\delta^{-2}r^{d_f} \ge c_1\delta^{-2}r^{d_f}(\log r^{-1})^{-80}.$$

Hence, by Proposition 2.10, the left-hand side of (2.19) is bounded above by $Ce^{-c\lambda^{1/80}}$. $\square$

Let $N_{\mathcal{U}}(r,s)$ the minimum number of $d_{\mathcal{U}}$-balls of radius s required to cover $B_{\mathcal{U}}(0,r)$. Another consequence of Proposition 2.10 is the following bound on $N_{\mathcal{U}}(r, r/\lambda)$.

LEMMA 2.12. *There exist constants* $c_1, c_2, c_3, \lambda_0 \in (0,\infty)$ *such that, for* $r \geq e^{\kappa(\log\lambda)^{41/16}}$ *and* $\lambda \geq \lambda_0$,

$$\mathbf{P}\big(N_{\mathcal{U}}(r,r/\lambda) \geq c_1(\log\lambda)^{107}\lambda^{d_f}\big) \leq c_2 e^{-c_3(\log\lambda)^{41/80}}.$$

PROOF. Let $\theta \geq 1$ be such that $2\lambda \leq \theta^{-1}\exp(\theta^{1/40})$. By Theorem 2.1(a), we have that

$$\mathbf{P}\big(B_{\mathcal{U}}(0,r) \not\subset B_E\big(0,\theta^{1/\kappa}r^{1/\kappa}\big)\big) \leq e^{-c\theta^{2/3}}.$$

Now it is straightforward to check that one can cover $B_{\mathcal{U}}(0,r)$ by balls $B_{\mathcal{U}}(z_i, r/\lambda)$, $i = 1,\dots,M$, such that $B_{\mathcal{U}}(z_i, r/2\lambda)$ are disjoint and $z_i \in B_{\mathcal{U}}(0,r)$. Moreover, it is necessarily the case that $M \geq N_{\mathcal{U}}(r,r/\lambda)$. Setting $n^\kappa = \theta r$, if $\tilde{A}(\theta,n)$ holds and $B_{\mathcal{U}}(0,r) \subset B_E(0,n)$ then we have $|B_{\mathcal{U}}(0,r)| \leq (\theta r)^{d_f}$ and $|B_{\mathcal{U}}(z_i, r/2\lambda)| \geq c\theta^{-1}(r/\lambda)^{d_f}$ for each i. Thus, we deduce from Proposition 2.10 that

$$\mathbf{P}\big(N_{\mathcal{U}}(r,r/\lambda) \geq c\theta^{1+d_f}\lambda^{d_f}\big) \leq c\exp\big(-c\theta^{1/80}\big). \tag{2.20}$$

Taking $\theta = (\log\lambda)^{41}$ completes the proof. $\square$

REMARK 2.13. Taking $\theta = \lambda$ in (2.20) gives the bound, for $r \geq e^{\kappa\lambda^{1/16}}$ and λ large,

$$\mathbf{P}\big(N_{\mathcal{U}}(r,r/\lambda) \geq c\lambda^{1+2d_f}\big) \leq c\exp\big(-c\lambda^{1/80}\big).$$

3. Topology for UST scaling limit. In this section, we introduce the topology on measured, rooted spatial trees for which we prove tightness for the law of the rescaled UST. This topology is finer than that considered in [2, 48], since it incorporates the full convergence of real trees embedded into Euclidean space, rather than merely the shape of subsets spanning a finite number of vertices. This point will be important when it comes to the proof of Theorem 1.4.

We define $\mathbb{T}$ to be the collection of quintuplets of the form

$$\underline{\mathcal{T}} = (\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \phi_{\mathcal{T}}, \rho_{\mathcal{T}}),$$

where: $(\mathcal{T}, d_{\mathcal{T}})$ is a complete and locally compact real tree (see [44], Definition 1.1, e.g.); $\mu_{\mathcal{T}}$ is a locally finite Borel measure on $(\mathcal{T}, d_{\mathcal{T}})$; $\phi_{\mathcal{T}}$ is a continuous map from $(\mathcal{T}, d_{\mathcal{T}})$ into a separable metric space (M, d_M); and $\rho_{\mathcal{T}}$ is a distinguished vertex in $\mathcal{T}$. [Usually the image space (M, d_M) we consider is $\mathbb{R}^2$

equipped with the Euclidean distance, though we will also consider other image spaces at certain places in our arguments.] We call such a quintuplet a *measured, rooted, spatial tree*. Let $\mathbb{T}_c$ be the subset of $\mathbb{T}$ for which $(\mathcal{T}, d_{\mathcal{T}})$ is compact. We will say that two elements of $\mathbb{T}$, $\underline{\mathcal{T}}$ and $\underline{\mathcal{T}}'$ say, are equivalent if there exists an isometry $\pi : (\mathcal{T}, d_{\mathcal{T}}) \to (\mathcal{T}', d'_{\mathcal{T}})$ for which $\mu_{\mathcal{T}} \circ \pi^{-1} = \mu'_{\mathcal{T}}$, $\phi_{\mathcal{T}} = \phi'_{\mathcal{T}} \circ \pi$ and also $\pi(\rho_{\mathcal{T}}) = \rho'_{\mathcal{T}}$.

In order to introduce a topology on $\mathbb{T}$, we will start by defining a topology on $\mathbb{T}_c$. In particular, for two elements of $\mathbb{T}_c$, we set $\Delta_c(\underline{\mathcal{T}}, \underline{\mathcal{T}}')$ to be equal to

$$\inf_{\substack{Z,\psi,\psi',\mathcal{C}:\\ (\rho_{\mathcal{T}},\rho'_{\mathcal{T}})\in\mathcal{C}}} \Big\{ d_P^Z(\mu_{\mathcal{T}} \circ \psi^{-1}, \mu'_{\mathcal{T}} \circ \psi'^{-1}) + \sup_{(x,x')\in\mathcal{C}} \big(d_Z(\psi(x), \psi'(x')) + d_M(\phi_{\mathcal{T}}(x), \phi'_{\mathcal{T}}(x')) \big) \Big\}, \tag{3.1}$$

where the infimum is taken over all metric spaces $Z = (Z, d_Z)$, isometric embeddings $\psi : (\mathcal{T}, d_{\mathcal{T}}) \to Z$, $\psi' : (\mathcal{T}', d'_{\mathcal{T}}) \to Z$, and correspondences $\mathcal{C}$ between $\mathcal{T}$ and $\mathcal{T}'$, and we define d_P^Z to be the Prohorov distance between finite Borel measures on Z. Note that, by a correspondence $\mathcal{C}$ between $\mathcal{T}$ and $\mathcal{T}'$, we mean a subset of $\mathcal{T} \times \mathcal{T}'$ such that for every $x \in \mathcal{T}$ there exists at least one $x' \in \mathcal{T}'$ such that $(x, x') \in \mathcal{C}$ and conversely for every $x' \in \mathcal{T}'$ there exists at least one $x \in \mathcal{T}$ such that $(x, x') \in \mathcal{C}$.

PROPOSITION 3.1. *The function Δ_c defines a metric on the equivalence classes of $\mathbb{T}_c$. Moreover, the resulting metric space is separable.*

PROOF. The proof of this result is almost identical to that of [22], Lemma 2.1, taking, in the notation of that paper, $I = \{1\}$ and $q_1(x, y) := \phi_{\mathcal{T}}(x)$. The main change is that when considering a correspondence between $\mathcal{T}$ and $\mathcal{T}'$, one has to require that the pair of roots $(\rho_{\mathcal{T}}, \rho'_{\mathcal{T}})$ is included, and, when selecting the points x_i, x'_i as in [22], one should take $x_1 = \rho_{\mathcal{T}}$ and $x'_1 = \rho'_{\mathcal{T}}$. A second change is that in the proof of separability, rather than approximating by metric spaces with a finite number of vertices, one should approximate by real trees formed of a finite number of line segments; however, making these changes is routine and we omit the details. □

REMARK 3.2. Even if (M, d_M) is assumed to be complete, the space of equivalence classes of $\mathbb{T}_c$ is not complete with respect to the metric Δ_c in general. Indeed, suppose $(M, d_M) = (\mathbb{R}^2, d_E^{(2)})$ and consider $([0, 1], d_E^{(1)}, \mathcal{L}, f, 0) \in \mathbb{T}_c$, where $d_E^{(d)}$ is the d-dimensional Euclidean distance, $\mathcal{L}$ is Lebesgue measure on $[0, 1]$, and $f : [0, 1] \to \mathbb{R}^2$ is any continuous nonconstant function. If we replace $d_E^{(1)}$ by $\varepsilon d_E^{(1)}$, then the sequence of elements in $\mathbb{T}_c$ that we obtain is Cauchy as $\varepsilon \to 0$, but does not have a limit in $\mathbb{T}_c$. One way to ensure completeness would be to restrict to a subset of $\mathbb{T}_c$ for which the functions $\phi_{\mathcal{T}}$ satisfy an equi-continuity condition.

To extend Δ_c to a metric on the equivalence classes of $\mathbb{T}$, we consider bounded restrictions of elements of $\mathbb{T}$ (cf. [1]). Thus, for $\underline{\mathcal{T}} \in \mathbb{T}$, let $\underline{\mathcal{T}}^{(r)} = (\mathcal{T}^{(r)}, d_{\mathcal{T}}^{(r)}, \mu_{\mathcal{T}}^{(r)}, \phi_{\mathcal{T}}^{(r)}, \rho_{\mathcal{T}}^{(r)})$ be obtained by taking: $\mathcal{T}^{(r)}$ to be the closed ball in $(\mathcal{T}, d_{\mathcal{T}})$ of radius r centred at $\rho_{\mathcal{T}}$; $d_{\mathcal{T}}^{(r)}$ $\mu_{\mathcal{T}}^{(r)}$ and $\phi_{\mathcal{T}}^{(r)}$ to be the restriction of $d_{\mathcal{T}}$, $\mu_{\mathcal{T}}$ and $\phi_{\mathcal{T}}$, respectively, to $\mathcal{T}^{(r)}$, and $\rho_{\mathcal{T}}^{(r)}$ to be equal to $\rho_{\mathcal{T}}$. As in [1], the fact that $(\mathcal{T}, d_{\mathcal{T}})$ is a real tree, and therefore a length space, means we can apply the Hopf–Rinow theorem (which implies that all closed, bounded subsets of a complete and locally compact length space are compact) to establish that $\underline{\mathcal{T}}^{(r)}$ is an element of $\mathbb{T}_c$. Furthermore, as in [1], Lemma 2.8, we can check the regularity of this restriction with respect to the metric Δ_c.

LEMMA 3.3. *For any two elements of* $\mathbb{T}$, $\underline{\mathcal{T}}$ *and* $\underline{\mathcal{T}}'$, *the function* $r \mapsto \Delta_c(\underline{\mathcal{T}}^{(r)}, \underline{\mathcal{T}}'^{(r)})$ *is cadlag.*

PROOF. By considering the natural embedding of $\mathcal{T}^{(r)}$ into $\mathcal{T}^{(r+\varepsilon)}$, along with the correspondence consisting of pairs (x, x') such that x is the closest point in $\mathcal{T}^{(r)}$ to $x' \in \mathcal{T}^{(r+\varepsilon)}$, we have, as in [1], Lemma 5.2, that

$$\Delta_c(\underline{\mathcal{T}}^{(r)}, \underline{\mathcal{T}}^{(r+\varepsilon)}) \leq \mu_{\mathcal{T}}(\mathcal{T}^{(r+\varepsilon)} \setminus \mathcal{T}^{(r)}) + \varepsilon + \sup_{\substack{x,x' \in \mathcal{T}^{(r+\varepsilon)}: \\ d_{\mathcal{T}}(x,x') \leq \varepsilon}} d_M(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(x'));$$

given this, the proof is a straightforward adaption of the proof of [1], Lemma 2.8. □

This result allows us to well define a function Δ on $\mathbb{T}^2$ by setting

$$\Delta(\underline{\mathcal{T}}, \underline{\mathcal{T}}') := \int_0^\infty e^{-r}(1 \wedge \Delta_c(\underline{\mathcal{T}}^{(r)}, \underline{\mathcal{T}}'^{(r)}))\, dr. \tag{3.2}$$

PROPOSITION 3.4. *The function* Δ *defines a metric on the equivalence classes of* $\mathbb{T}$. *Moreover, the resulting metric space is separable.*

PROOF. Again, the proof is similar to the corresponding result in [1]. Positivity, finiteness and symmetry of Δ are clear. Moreover, the triangle inequality is easy to check from the definition and the fact that the triangle inequality holds for Δ_c. So, to establish that Δ is a metric, it remains to prove positive definiteness. To this end, suppose that $\underline{\mathcal{T}}$ and $\underline{\mathcal{T}}'$ are such that the expression at (3.2) is equal to zero. From Lemma 3.3, it follows that $\Delta_c(\underline{\mathcal{T}}^{(r)}, \underline{\mathcal{T}}'^{(r)}) = 0$ for every $r > 0$. Consequently, for each r, there exists an isometry $\pi_r : (\mathcal{T}^{(r)}, d_{\mathcal{T}}^{(r)}) \to (\mathcal{T}'^{(r)}, d_{\mathcal{T}}'^{(r)})$ such that $\mu_{\mathcal{T}}^{(r)} \circ \pi_r^{-1} = \mu_{\mathcal{T}}'^{(r)}$, $\phi_{\mathcal{T}}^{(r)} = \phi_{\mathcal{T}}'^{(r)} \circ \pi_r$ and also $\pi_r(\rho_{\mathcal{T}}^{(r)}) = \rho_{\mathcal{T}}'^{(r)}$. For $n, k \geq 1$, let

$(x_i^{n,k})_{i=1}^{N(n,k)}$ be a finite k^{-1}-cover of $\mathcal{T}^{(n)}$ containing the root $\rho_{\mathcal{T}}$ (such a collection exists as a result of the compactness of $\mathcal{T}^{(n)}$). Since π_r is an isometry, we have that $(\pi_m(x_i^{n,k}))_{m\geq n}$ is a bounded sequence for each $n, k \geq 1$ and $1 \leq i \leq N(n,k)$, and so has a convergent subsequence. By a diagonal procedure, one can thus find a subsequence $(m_j)_{j\geq 1}$ such that $\pi(x_i^{n,k}) = \lim_{j\to\infty} \pi_{m_j}(x_i^{n,k})$ exists for every $n, k \geq 1$ and $1 \leq i \leq N(n,k)$. From this construction, we obtain that π is distance-preserving on $\{x_i^{n,k} : n, k \geq 1, 1 \leq i \leq N(n,k)\}$ and, since the latter set is dense in $\mathcal{T}$, we can extend it to a distance-preserving map on $\mathcal{T}$. Clearly, by reversing the roles of $\mathcal{T}$ and $\mathcal{T}'$, it is also possible to find a distance-preserving map from $\mathcal{T}'$ to $\mathcal{T}$. Hence, π must be an isometry. Moreover, it is clear that this map is root-preserving, that is, $\pi(\rho_{\mathcal{T}}) = \rho'_{\mathcal{T}}$. To check that it is measure-preserving, that is, $\mu_{\mathcal{T}} \circ \pi^{-1} = \mu'_{\mathcal{T}}$, one can follow an identical argument to that applied in the proof of [1], Proposition 5.3, based on considering approximations to the measures $\mu_{\mathcal{T}}^{(n)}$ and $\mu_{\mathcal{T}}^{\prime(n)}$ supported on $(x_i^{n,k})_{i=1}^{N(n,k)}$ and $(\pi(x_i^{n,k}))_{i=1}^{N(n,k)}$, respectively. Finally, we note that the continuity of $\phi'_{\mathcal{T}}$ implies

$$\phi'_{\mathcal{T}}(\pi(x_i^{n,k})) = \lim_{j\to\infty} \phi_{\mathcal{T}}^{\prime(m_j)} \circ \pi_{m_j}(x_i^{n,k}) = \lim_{j\to\infty} \phi_{\mathcal{T}}^{(m_j)}(x_i^{n,k}) = \phi_{\mathcal{T}}(x_i^{n,k}).$$

Since $\phi_{\mathcal{T}}$ is also continuous, it follows that $\phi_{\mathcal{T}} = \phi'_{\mathcal{T}} \circ \pi$. Hence, we have shown that $\underline{\mathcal{T}}$ and $\underline{\mathcal{T}}'$ are equivalent, and so Δ is indeed a metric on the equivalence classes of $\mathbb{T}$.

For separability, we first note that $\Delta(\underline{\mathcal{T}}, \underline{\mathcal{T}}^{(r)}) \leq e^{-r}$, and so $\mathbb{T}_c$ is dense in $(\mathbb{T}, \Delta)$. Since $(\mathbb{T}_c, \Delta_c)$ is separable, it will thus be sufficient to check that convergence in $(\mathbb{T}_c, \Delta_c)$ implies convergence in $(\mathbb{T}, \Delta)$ (cf. [1], Proposition 2.10). So let us start by supposing that we have a sequence $\underline{\mathcal{T}}_n$ that converges to $\underline{\mathcal{T}}$ in $(\mathbb{T}_c, \Delta_c)$. In particular, we can find a sequence of metric spaces Z_n, isometric embeddings $\psi_n : \mathcal{T} \to Z_n$, $\psi'_n : \mathcal{T}_n \to Z_n$ and correspondences $\mathcal{C}_n$ between $\mathcal{T}$ and $\mathcal{T}_n$ containing $(\rho_{\mathcal{T}}, \rho_{\mathcal{T}_n})$ such that

$$\begin{aligned} & d_P^{Z_n}(\mu_{\mathcal{T}} \circ \psi_n^{-1}, \mu_{\mathcal{T}_n} \circ \psi_n^{\prime -1}) \\ & \quad + \sup_{(x,x')\in\mathcal{C}_n} \big(d_{Z_n}(\psi_n(x), \psi'_n(x')) + d_M(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}_n}(x'))\big) \\ & \quad < \varepsilon_n, \end{aligned} \tag{3.3}$$

where $\varepsilon_n \to 0$. Now, define $\psi_n^{(r)}$ to be the restriction of ψ_n to $\mathcal{T}^{(r)}$, $\psi_n^{\prime(r)}$ to be the restriction of ψ'_n to $\mathcal{T}_n^{(r)}$, and $\mathcal{C}_n^{(r)}$ to be the collection of pairs (x, x') such that: either $x \in \mathcal{T}^{(r)}$ and x' is the closest point in $\mathcal{T}_n^{(r)}$ to an element $x'' \in \mathcal{T}_n$ such that $(x, x'') \in \mathcal{C}_n$; or $x' \in \mathcal{T}_n^{(r)}$ and x is the closest point in $\mathcal{T}^{(r)}$ to an element $x'' \in \mathcal{T}$ such that $(x'', x') \in \mathcal{C}_n$. Note that $\psi_n^{(r)}$ and $\psi_n^{\prime(r)}$ are isometric embeddings of $\mathcal{T}^{(r)}$ and $\mathcal{T}_n^{(r)}$, respectively, into Z_n, and that $\mathcal{C}_n^{(r)}$ is a correspondence between $\mathcal{T}^{(r)}$ and $\mathcal{T}_n^{(r)}$ such that $(\rho_{\mathcal{T}}^{(r)}, \rho_{\mathcal{T}_n}^{(r)}) \in \mathcal{C}_n^{(r)}$. If we suppose that $x \in \mathcal{T}^{(r)}$ and x' is the closest point in $\mathcal{T}_n^{(r)}$ to an element $x'' \in \mathcal{T}_n$ such that $(x, x'') \in \mathcal{C}_n$, then

$$d_{\mathcal{T}_n}(\rho_{\mathcal{T}_n}, x'') \leq d_{Z_n}(\psi_n'(\rho_{\mathcal{T}_n}), \psi_n(\rho_{\mathcal{T}})) + d_{Z_n}(\psi_n(\rho_{\mathcal{T}}), \psi_n(x)) + d_{Z_n}(\psi_n(x), \psi_n'(x'')),$$

which is bounded above by $r + 2\varepsilon_n$. It follows that $d_{\mathcal{T}_n}(x', x'') < 2\varepsilon_n$ and, therefore, also $d_{Z_n}(\psi_n(x), \psi_n'(x')) < 3\varepsilon_n$. A similar argument applies to the case when $x' \in \mathcal{T}_n^{(r)}$ and x is the closest point in $\mathcal{T}^{(r)}$ to an element $x'' \in \mathcal{T}$ such that $(x'', x') \in \mathcal{C}_n$. Consequently, we obtain that

$$\sup_{(x,x')\in\mathcal{C}_n^{(r)}} d_{Z_n}(\psi_n^{(r)}(x), \psi_n'^{(r)}(x')) < 3\varepsilon_n. \tag{3.4}$$

From this, one can proceed as in the proof of [1], Proposition 2.10, to deduce that

$$d_P^{Z_n}(\mu_{\mathcal{T}}^{(r)} \circ (\psi_n^{(r)})^{-1}, \mu_{\mathcal{T}_n}^{(r)} \circ (\psi_n'^{(r)})^{-1}) < \varepsilon_n + \mu_{\mathcal{T}}(\mathcal{T}^{(r+4\varepsilon_n)} \setminus \mathcal{T}^{(r-4\varepsilon_n)}).$$

Moreover, it is also elementary to deduce from (3.3) and (3.4) that

$$\sup_{(x,x')\in\mathcal{C}_n^{(r)}} |\phi_{\mathcal{T}}^{(r)}(x) - \phi_{\mathcal{T}_n}^{(r)}(x')| \leq \varepsilon_n + \sup_{\substack{(x,x')\in\mathcal{T}^{(r+4\varepsilon_n)}:\\ d_{\mathcal{T}}(x,x')<4\varepsilon_n}} d_M(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(x')).$$

Hence, we have established that

$$\begin{aligned}\Delta_c(\underline{\mathcal{T}}_n^{(r)}, \underline{\mathcal{T}}^{(r)}) \leq{}& 5\varepsilon_n + \mu_{\mathcal{T}}(\mathcal{T}^{(r+4\varepsilon_n)} \setminus \mathcal{T}^{(r-4\varepsilon_n)}) \\ &+ \sup_{\substack{(x,x')\in\mathcal{T}^{(r+4\varepsilon_n)}:\\ d_{\mathcal{T}}(x,x')<4\varepsilon_n}} d_M(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(x')).\end{aligned} \tag{3.5}$$

Since $\mu_{\mathcal{T}}$ is a finite measure, this expression must converge to zero for all but at most a countable number of values of r. Thus, dominated convergence implies that $\Delta(\underline{\mathcal{T}}_n, \underline{\mathcal{T}}) \to 0$, as desired. □

Next, under the additional assumption that (M, d_M) is proper (i.e., every closed ball in M is compact), we provide a sufficient condition for a subset $\mathcal{A}$ of $\mathbb{T}$ to be relatively compact with respect to the topology induced by Δ. This extends the corresponding result of [1], Theorem 2.11, to include the spatial embedding.

LEMMA 3.5. *Suppose (M, d_M) is proper. Let $\mathcal{A}$ be a subset of $\mathbb{T}$ such that, for every $r > 0$:*

(i) *for every $\varepsilon > 0$, there exists a finite integer $N(r, \varepsilon)$ such that for any element $\underline{\mathcal{T}}$ of $\mathcal{A}$ there is an ε-cover of $\mathcal{T}^{(r)}$ of cardinality less than $N(r, \varepsilon)$;*

(ii) *it holds that*

$$\sup_{\underline{\mathcal{T}}\in\mathcal{A}} \mu_{\mathcal{T}}(\mathcal{T}^{(r)}) < \infty;$$

(iii) $\{\phi_{\mathcal{T}}(\rho_{\mathcal{T}}) : \underline{\mathcal{T}} \in \mathcal{A}\}$ *is a bounded subset of* M, *and for every* $\varepsilon > 0$, *there exists a* $\delta = \delta(r, \varepsilon) > 0$ *such that*

$$\sup_{\underline{\mathcal{T}} \in \mathcal{A}} \sup_{\substack{x, y \in \mathcal{T}^{(r)}: \\ d_{\mathcal{T}}(x,y) \leq \delta}} d_M\big(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y)\big) < \varepsilon.$$

Then $\mathcal{A}$ *is relatively compact.*

PROOF. We follow closely the proof [1], Theorem 2.11. Suppose that $\underline{\mathcal{T}}_n$ is a sequence in a set $\mathcal{A} \subseteq \mathbb{T}$ that is assumed to satisfy the properties listed in the statement of the lemma. We can then define U to be a countable index set such that $\{x_u^n : u \in U\}$ is dense in $\mathcal{T}_n$ for each n (we further assume that $0 \in U$ and $x_0^n = \rho_{\mathcal{T}_n}$), and also introduce an abstract space $\mathcal{T}' := \{x_u : u \in U\}$ such that, for some subsequence $(n_i)_{i \geq 1}$,

$$d_{\mathcal{T}_{n_i}}\big(x_u^{n_i}, x_v^{n_i}\big) \to d_{\mathcal{T}}(x_u, x_v) \tag{3.6}$$

for each pair of indices $u, v \in U$, where the right-hand side may be taken as a definition of the function $d_{\mathcal{T}} : \mathcal{T}' \times \mathcal{T}' \to \mathbb{R}_+$. In fact, $d_{\mathcal{T}}$ is a quasi-metric on $\mathcal{T}'$, and so, with a slight abuse of notation, we obtain a metric space $(\mathcal{T}', d_{\mathcal{T}})$ by identifying points that are a $d_{\mathcal{T}}$-distance of zero apart. Moreover, the argument of [1] gives us that the completion $(\mathcal{T}, d_{\mathcal{T}})$ of this metric space is locally compact, and identifies $\rho_{\mathcal{T}} := x_0$ as the root for the space. It also describes how to construct a corresponding locally finite Borel measure on $\mathcal{T}$, which we will call $\mu_{\mathcal{T}}$. Now, from property (iii) and (3.6), it is easy to see that $\phi_{\mathcal{T}_{n_i}}(x_u^{n_i})$ is bounded for each u, and so a diagonal procedure yields that, by taking a further subsequence if necessary, $\phi_{\mathcal{T}_{n_i}}(x_u^{n_i}) \to \phi_{\mathcal{T}}(x_u)$, for each $u \in U$, where, similarly to the definition of $d_{\mathcal{T}}$, the right-hand side provides a definition of $\phi_{\mathcal{T}}(x_u)$ [that this function is well-defined on $\mathcal{T}'$ is readily checked from (iii) and (3.6)]. Moreover, it is not difficult to check that

$$\sup_{\substack{x, y \in \mathcal{T}'^{(r)}: \\ d_{\mathcal{T}}(x,y) \leq \delta(r,\varepsilon)}} d_M\big(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y)\big) < \varepsilon,$$

and so the function can be extended continuously to the whole of $\mathcal{T}$. In particular, we have so far constructed $\underline{\mathcal{T}}$, and to check this is an element of $\mathbb{T}$, it remains to show that $(\mathcal{T}, d_{\mathcal{T}})$ is a real tree. However, in [1], Lemma 2.7, it is shown that $(\mathcal{T}, d_{\mathcal{T}})$ is a length space, and so it is connected. Moreover, the four-point condition for the metric for $(\mathcal{T}, d_{\mathcal{T}})$ follows from the four-point condition that must hold for $(\mathcal{T}_n, d_{\mathcal{T}_n})$ (see [26], (2.1)). It follows that $(\mathcal{T}, d_{\mathcal{T}})$ must be a real tree, as desired.

It remains to show that $\underline{\mathcal{T}}_{n_i} \to \underline{\mathcal{T}}$ in $(\mathbb{T}, \Delta)$. For this it is sufficient to show that $\underline{\mathcal{T}}_{n_i}^{(r)} \to \underline{\mathcal{T}}^{(r)}$ in $(\mathbb{T}_c, \Delta_c)$, at least whenever $\mu_{\mathcal{T}}(\partial B_{\mathcal{T}}(\rho_{\mathcal{T}}, r)) = 0$. Again, this may be accomplished by following the argument of [1], which involves introducing finite subsets $U_{k,l} \subset U$ such that $\{x_u^{n_i} : u \in U_{k,l}\}$ and $\{x_u : u \in U_{k,l}\}$ suitably well-approximate $\mathcal{T}_{n_i}^{(r)}$ and $\mathcal{T}^{(r)}$, respectively. Moreover, a consideration of the correspondence between these finite sets given by $(x_u^{n_i}, x_u)$, $u \in U_{k,l}$, allows it to be deduced in our case that

$$\lim_{i\to\infty} \Delta_c(\underline{\mathcal{T}}_{n_i}^{(r)}, \underline{\mathcal{T}}^{(r)}) \leq 2 \sup_{i\geq 1} \sup_{\substack{x,y\in\mathcal{T}_{n_i}^{(r+\delta)}:\\ d_{\mathcal{T}_{n_i}}(x,y)\leq\delta}} d_M(\phi_{\mathcal{T}_{n_i}}(x), \phi_{\mathcal{T}_{n_i}}(y)) + 2 \sup_{\substack{x,y\in\mathcal{T}^{(r+\delta)}:\\ d_{\mathcal{T}}(x,y)\leq\delta}} d_M(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y)),$$

for any $\delta > 0$ [cf. the extra term involving the continuity of $\phi_{\mathcal{T}}$ in (3.5)]. Since the right-hand side can be made arbitrarily small by suitable choice of δ, this completes the proof. □

REMARK 3.6. The restriction to real trees for $(\mathbb{T}_c, \Delta_c)$ has actually been unnecessary in this section so far, and so the same topology could be extended to the setting where the metric space part of an element—$(\mathcal{T}, d_{\mathcal{T}})$—is simply assumed to be a compact metric space. Similarly, for the topology, $(\mathbb{T}, \Delta)$, it would have been enough to assume that the metric space part of an element is a locally compact length space (cf. [1]). In both cases, the restriction to the case where the metric space is a real tree would then simply be the restriction to a closed subset of the relevant topology (cf. [25], Lemma 4.22).

To conclude this section, we present two consequences of convergence in $(\mathbb{T}_c, \Delta_c)$, again assuming that (M, d_M) is proper. First, we prove convergence of the push-forward measures. In what follows, $B_X(x, r)$ is the open ball in the metric space $X = (X, d_X)$ with radius r centred at x.

LEMMA 3.7. *Suppose (M, d_M) is proper. If $\underline{\mathcal{T}}_n \to \underline{\mathcal{T}}$ in $(\mathbb{T}_c, \Delta_c)$, then*

$$\mu_{\mathcal{T}_n} \circ \phi_{\mathcal{T}_n}^{-1} \to \mu_{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1} \tag{3.7}$$

weakly as Borel measures on (M, d_M).

PROOF. Note first that if $\underline{\mathcal{T}}_n \to \underline{\mathcal{T}}$ in $(\mathbb{T}_c, \Delta_c)$ then for each n we can find a measurable function $f_n : \mathcal{T}_n \to \mathcal{T}$ such that $\mu_{\mathcal{T}_n} \circ f_n^{-1} \to \mu_{\mathcal{T}}$ weakly as measures on $\mathcal{T}$, and also

$$\sup_{x\in\mathcal{T}_n} d_M(\phi_{\mathcal{T}}(f_n(x)), \phi_{\mathcal{T}_n}(x)) \to 0. \tag{3.8}$$

Indeed, let $Z_n, \psi_n, \psi_n', \mathcal{C}_n$ be defined as in the proof of Proposition 3.4, that is, so that (3.3) holds. Let $(x_i^n)_{i=1}^{N(n)}$ be a ε_n-cover of $\mathcal{T}$. Set $A_1^n := B_{Z_n}(\psi_n(x_1^n), 2\varepsilon_n)$ and $A_i^n := B_{Z_n}(\psi_n(x_i^n), 2\varepsilon_n) \setminus A_{i-1}^n$ for $i = 2, \dots, N(n)$. Then the sets A_i^n, $i =$

$1, \dots, N(n)$, are disjoint and their union contains all those points in Z_n within a distance ε_n of $\psi_n(\mathcal{T})$. In particular, they cover $\psi'_n(\mathcal{T}_n)$, so one can define a (measurable) map $f_n : \mathcal{T}_n \to \mathcal{T}$ by setting $f_n(x) := x_i^n$ if $\psi'_n(x) \in A_i^n$. For this map, we have

$$d_P^{\mathcal{T}}(\mu_{\mathcal{T}_n} \circ f_n^{-1}, \mu_{\mathcal{T}}) \leq d_P^{Z_n}(\mu_{\mathcal{T}_n} \circ f_n^{-1} \circ \psi_n^{-1}, \mu_{\mathcal{T}_n} \circ \psi_n'^{-1}) + d_P^{Z_n}(\mu_{\mathcal{T}_n} \circ \psi_n'^{-1}, \mu_{\mathcal{T}} \circ \psi_n^{-1}), \tag{3.9}$$

where $d_P^{\mathcal{T}}$ is the Prohorov distance between measures on $\mathcal{T}$. By (3.3), the second term in (3.9) is bounded above by ε_n. Moreover, by definition we have that $d_{Z_n}(\psi_n(f_n(x)), \psi'_n(x))$ is strictly less than $2\varepsilon_n$ for all $x \in \mathcal{T}_n$, and so the first term is bounded above by $2\varepsilon_n$. This confirms that $\mu_{\mathcal{T}_n} \circ f_n^{-1} \to \mu_{\mathcal{T}}$. Next, observe that if $f_n(x) = x_i^n$ and $(x', x) \in \mathcal{C}_n$, then

$$\begin{aligned} d_M(\phi_{\mathcal{T}}(f_n(x)), \phi_{\mathcal{T}_n}(x)) &\leq \varepsilon_n + d_M(\phi_{\mathcal{T}}(x_i^n), \phi_{\mathcal{T}}(x')) \\ &\leq \varepsilon_n + \sup_{\substack{(x,x') \in \mathcal{T}: \\ d_{\mathcal{T}}(x,x') < 3\varepsilon_n}} d_M(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(x')). \end{aligned}$$

By the continuity of $\phi_{\mathcal{T}}$, this upper bound converges to zero as $n \to \infty$, and we have thereby established (3.8). As a consequence, if $g : M \to \mathbb{R}$ is continuous and compactly supported, then

$$\begin{aligned} &|\mu_{\mathcal{T}_n} \circ \phi_{\mathcal{T}_n}^{-1}(g) - \mu_{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1}(g)| \\ &\quad \leq \int_{\mathcal{T}_n} |g(\phi_{\mathcal{T}_n}(x)) - g(\phi_{\mathcal{T}}(f_n(x)))| \mu_{\mathcal{T}_n}(dx) \\ &\qquad + \left| \int_{\mathcal{T}} g(\phi_{\mathcal{T}}(x)) \mu_{\mathcal{T}_n} \circ f_n^{-1}(dx) - \int_{\mathcal{T}} g(\phi_{\mathcal{T}}(x)) \mu_{\mathcal{T}}(dx) \right| \\ &\quad \to 0, \end{aligned}$$

where the convergence of the first term in the upper bound to zero follows from (3.8) [and the fact that $\mu_{\mathcal{T}_n}(\mathcal{T}_n) \to \mu_{\mathcal{T}}(\mathcal{T}) < \infty$, as follows from $\mu_{\mathcal{T}_n} \circ f_n^{-1} \to \mu_{\mathcal{T}}$], and the convergence of the second term to zero also follows from $\mu_{\mathcal{T}_n} \circ f_n^{-1} \to \mu_{\mathcal{T}}$. This establishes that $\mu_{\mathcal{T}_n} \circ \phi_{\mathcal{T}_n}^{-1}$ converges vaguely to $\mu_{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1}$. Finally, since the masses of the measures in the sequence converge to the mass of the limit, which is finite, it also demonstrates weak convergence. □

REMARK 3.8. It is not difficult to extend the above proof to deduce that the conclusion of (3.7) holds in the sense of vague convergence of measures whenever $\underline{\mathcal{T}}_n \to \underline{\mathcal{T}}$ in $(\mathbb{T}, \Delta)$, and in addition we have the following condition which prevents an explosion of mass in a bounded region of the proper space (M, d_M): for each $r \in (0, \infty)$, there exists an $R < \infty$ such that

$$\phi_{\mathcal{T}_n}^{-1}(B_M(\rho_M, r)) \subseteq B_{\mathcal{T}_n}(\rho_{\mathcal{T}_n}, R) \qquad \text{for all } n, \tag{3.10}$$

where ρ_M is a distinguished point in M. We will apply a probabilistic version of such an argument to prove Lemma 5.2.

Our second result is that convergence in $\mathbb{T}_c$ with respect to Δ_c implies convergence in a generalisation of the topology for path ensembles considered by Schramm in [48]. In that paper, the space M considered was the one-point compactification of $\mathbb{R}^2$, $\mathbb{S}^2$ say. This result will be used when we wish to transfer the results of [48] to our setting. Given a metric space X, write $\mathcal{H}(X)$ for the Hausdorff space of compact subsets of X. We write $\gamma_{\mathcal{T}}(x,y)$ for the unique path between x and y in $\mathcal{T}$ (including its endpoints).

LEMMA 3.9. *If we define*

$$\mathfrak{T}(\underline{\mathcal{T}}) := \{(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y), \phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y))) : x, y \in \mathcal{T}\},$$

then the convergence $\underline{\mathcal{T}}_n \to \underline{\mathcal{T}}$ in $(\mathbb{T}_c, \Delta_c)$ implies that $\mathfrak{T}(\underline{\mathcal{T}}_n) \to \mathfrak{T}(\underline{\mathcal{T}})$ in $\mathcal{H}(M \times M \times \mathcal{H}(M))$.

PROOF. Suppose that $\underline{\mathcal{T}}_n \to \underline{\mathcal{T}}$ holds in $(\mathbb{T}_c, \Delta_c)$, and that $Z_n, \psi_n, \psi_n', \mathcal{C}_n$ are defined as in the proof of Proposition 3.4, so that (3.3) holds. We claim that if $(x, x_n), (y, y_n) \in \mathcal{C}_n$, then

$$\begin{aligned} &d_H^M(\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y)), \phi_{\mathcal{T}_n}(\gamma_{\mathcal{T}_n}(x_n,y_n))) \\ &\quad \le \eta_n := \varepsilon_n + \sup_{\substack{(z,z')\in\mathcal{T}: \\ d_{\mathcal{T}}(z,z')<5\varepsilon_n}} d_M(\phi_{\mathcal{T}}(z), \phi_{\mathcal{T}}(z')), \end{aligned} \tag{3.11}$$

where d_H^M is the Hausdorff distance between subsets of M. To prove this, we start by considering $z \in \gamma_{\mathcal{T}}(x,y)$, and defining z_n to be any element of $\mathcal{T}_n$ such that $(z, z_n) \in \mathcal{C}_n$. By applying (3.3) and the fact that the metric $d_{\mathcal{T}}$ is additive along paths, we obtain

$$\begin{aligned} d_{\mathcal{T}_n}(x_n, z_n) + d_{\mathcal{T}_n}(z_n, y_n) &< d_{\mathcal{T}}(x,z) + d_{\mathcal{T}}(z,y) + 4\varepsilon_n = d_{\mathcal{T}}(x,y) + 4\varepsilon_n \\ &< d_{\mathcal{T}_n}(x_n, y_n) + 6\varepsilon_n. \end{aligned}$$

It follows that z_n is within a distance of $3\varepsilon_n$ (with respect to $d_{\mathcal{T}_n}$) of $\gamma_{\mathcal{T}_n}(x_n, y_n)$. Now, if we let z_n' be the closest point in $\gamma_{\mathcal{T}_n}(x_n, y_n)$ to z_n, and z_n'' be such that $(z_n'', z_n') \in \mathcal{C}_n$, then it is the case that $d_{\mathcal{T}}(z, z_n'') < d_{\mathcal{T}_n}(z_n, z_n') + 2\varepsilon_n < 5\varepsilon_n$, and so $d_M(\phi_{\mathcal{T}}(z), \phi_{\mathcal{T}_n}(z_n')) < \varepsilon_n + d_M(\phi_{\mathcal{T}}(z), \phi_{\mathcal{T}}(z_n'')) \le \eta_n$. Thus $\phi_{\mathcal{T}}(z)$ is within a d_M-distance η_n of $\phi_{\mathcal{T}_n}(\gamma_{\mathcal{T}_n}(x_n, y_n))$. A similar argument yields that any point of $\phi_{\mathcal{T}_n}(\gamma_{\mathcal{T}_n}(x_n, y_n))$ is within a d_M-distance η_n of $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y))$. This establishes (3.11), from which the result follows. □

REMARK 3.10. As with Lemma 3.7, this result is readily extended to the noncompact case when (M, d_M) is proper. Indeed, under this assumption, if $\mathcal{T}_n \to \mathcal{T}$ in $(\mathbb{T}, \Delta)$ and (3.10) holds, then $\mathfrak{T}(\mathcal{T}_n) \to \mathfrak{T}(\mathcal{T})$ in $\mathcal{H}(\dot{M} \times \dot{M} \times \mathcal{H}(\dot{M}))$, where $\dot{M}$ is defined to be the one-point compactification of M. A probabilistic version of this argument will be used to prove Lemma 5.5.

REMARK 3.11. While we do not need the result, we note that a similar argument can be used to relate convergence in our topology to convergence in the topology of [2]. This topology is similar to that of Schramm, but it incorporates convergence of the shape of subtrees spanning an arbitrary finite number of vertices, rather than just two.

4. Tightness of UST law under rescaling. The aim of this section is to prove Theorem 1.1, that is, to establish that the law of the UST, considered as a measured, rooted, spatial tree, is tight under rescaling. The key estimates for this purpose were already established in Section 2. As discussed in the Introduction, here we extend $\mathcal{U}$ to a (locally compact) real tree by adding line segments of unit length along its edges, and define $\phi_{\mathcal{U}} : \mathcal{U} \to \mathbb{R}^2$ to be the identity map on vertices with linear interpolation along edges. Throughout this section, we suppose that the image space (M, d_M) introduced in Section 3 is $\mathbb{R}^2$ equipped with the Euclidean distance.

LEMMA 4.1. *For every $r > 1$, $\varepsilon \in (0, \varepsilon_0)$, it holds that*

$$\lim_{N\to\infty} \liminf_{\delta\to 0} \mathbf{P}\big(\textit{there exists a } \delta^{-\kappa}\varepsilon\textit{-cover for } B_{\mathcal{U}}(0, \delta^{-\kappa} r) \textit{ of cardinality} \le N\big) = 1. \tag{4.1}$$

PROOF. Recalling the notation $N_{\mathcal{U}}$ introduced above Lemma 2.12. we have that the probability in (4.1) is at least $\mathbf{P}(N_{\mathcal{U}}(\delta^{-\kappa} r, \delta^{-\kappa}\varepsilon) \le N)$. Let $\theta = \theta(N)$ be such that $c\theta^{1+d_f}(r/\varepsilon)^{d_f} = N$; then by (2.20) we have that $\limsup_{\delta\to 0} \mathbf{P}(N_{\mathcal{U}}(\delta^{-\kappa} r, \delta^{-\kappa}\varepsilon) \ge N) \le c\exp(-\theta^{1/80})$, and since $\lim_{N\to\infty}\theta(N) = \infty$, this proves the result. □

LEMMA 4.2. *For every $r < \infty$, it holds that*

$$\lim_{\lambda\to\infty} \liminf_{\delta\to 0} \mathbf{P}\big(\delta^2 \mu_{\mathcal{U}}\big(B_{\mathcal{U}}(0, \delta^{-\kappa} r)\big) \le \lambda\big) = 1.$$

PROOF. This is a simple consequence of Theorem 2.1(b). □

LEMMA 4.3. *For every $\varepsilon > 0$, $r < \infty$, it holds that*

$$\lim_{\eta\to 0} \liminf_{\delta\to 0} \mathbf{P}\Big(\max_{\substack{x,y\in B_{\mathcal{U}}(0,\delta^{-\kappa} r):\\ d_{\mathcal{U}}(x,y)\le \delta^{-\kappa}\eta}} \big|\phi_{\mathcal{U}}(x) - \phi_{\mathcal{U}}(y)\big| \le \delta^{-1}\varepsilon\Big) = 1.$$

PROOF. Since $|\phi_{\mathcal{U}}(x) - \phi_{\mathcal{U}}(y)| \leq d^S_{\mathcal{U}}(x, y)$ it is sufficient to prove that

$$\lim_{\eta\to 0} \liminf_{\delta\to 0} \mathbf{P}\Big(\max_{\substack{x,y\in B_{\mathcal{U}}(0,c_1\delta^{-\kappa}r):\\ d_{\mathcal{U}}(x,y)\leq c_2\delta^{-\kappa}\eta}} d^S_{\mathcal{U}}(x, y) > \delta^{-1}\varepsilon\Big) = 0. \tag{4.2}$$

Let $r' = \delta^{-1}\varepsilon$, and set $A_*(\lambda) := \{B_{\mathcal{U}}(0, c_1\delta^{-\kappa}r) \subset B_E(0, r'e^{c_1\lambda^{1/2}})\}$, where c_1 is the constant of Proposition 2.8. By Theorem 2.1(a),

$$\mathbf{P}\big(\{B_{\mathcal{U}}(0, c_1\delta^{-\kappa}r) \subset B_E(0, (\lambda r)^{1/\kappa}\delta^{-1})\}^c\big) \leq c_2e^{-c_3\lambda^{2/3}} \qquad \forall\delta^\kappa \leq \lambda r, \lambda \geq c_1,$$

and in addition $(\lambda r)^{1/\kappa}\delta^{-1} \leq r'e^{c_1\lambda^{1/2}}$ for λ large. Thus, $\mathbf{P}(A_*(\lambda)^c) \leq c_2e^{-c_3\lambda^{2/3}}$ for all $\delta^\kappa \leq \lambda r$, $\lambda \geq \lambda_1$, where λ_1 is some large, finite constant. Next, let A_4 be as in Proposition 2.8 [taking (x_0, r, λ) in that result to be $(0, r', \lambda)$ in our current parameterisation], so that $\mathbf{P}(A_4^c) \leq c_4 \exp(-c_5\lambda^{1/2})$ for all $\delta \leq \varepsilon$, $\lambda \geq \lambda_0$. Clearly, it is enough to consider the event of (4.2) on $A_*(\lambda) \cap A_4$. On $A_*(\lambda) \cap A_4$, if $x, y \in B_{\mathcal{U}}(0, c_1\delta^{-\kappa}r)$ satisfy $d^S_{\mathcal{U}}(x, y) > \delta^{-1}\varepsilon = r'$, then by Proposition 2.8, $d_{\mathcal{U}}(x, y) \geq \lambda^{-1}d^S_{\mathcal{U}}(x, y)^\kappa \geq \lambda^{-1}\varepsilon^\kappa\delta^{-\kappa}$. Thus, by taking $\eta < \lambda^{-1}\varepsilon^\kappa$, $d_{\mathcal{U}}(x, y) > \eta\delta^{-\kappa}$, so (4.2) is proved. □

PROOF OF THEOREM 1.1. This is clear given the pre-compactness result of Lemma 3.5, and Lemmas 4.1–4.3. □

5. Properties of limit measures. In this section, we establish properties of the limit measure of the UST and will prove Theorem 1.3. Throughout, we fix a sequence $\delta_n \to 0$ such that $(\mathbf{P}_{\delta_n})_{n\geq 1}$ converges weakly [as measures on $(\mathbb{T}, \Delta)$], and write $\underline{\mathcal{U}}_{\delta_n} = (\mathcal{U}, \delta_n^\kappa d_{\mathcal{U}}, \delta_n^2\mu_{\mathcal{U}}, \delta_n\phi_{\mathcal{U}}, 0)$. Letting $\tilde{\mathbf{P}}$ be the relevant limiting law, we denote by $\underline{\mathcal{T}} = (\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \phi_{\mathcal{T}}, \rho_{\mathcal{T}})$ a random variable with law $\tilde{\mathbf{P}}$. Again, we take the image space (M, d_M) of Section 3 to be $\mathbb{R}^2$ equipped with the Euclidean distance. In many of the arguments, the following coupling result will be useful.

LEMMA 5.1. *There exist realisations of $(\underline{\mathcal{U}}_{\delta_n})_{n\geq 1}$ and $\underline{\mathcal{T}}$ built on the same probability space, with probability measure $\mathbf{P}^*$ say, such that: for some subsequence $(n_i)_{i\geq 1}$ and divergent sequence $(r_j)_{j\geq 1}$ it holds that, $\mathbf{P}^*$-a.s.,*

$$D_{i,j} := \Delta_c\big(\underline{\mathcal{U}}^{(r_j)}_{\delta_{n_i}}, \underline{\mathcal{T}}^{(r_j)}\big) \to 0 \tag{5.1}$$

as $i \to \infty$, for every $j \geq 1$.

PROOF. Recall that by the definition of $\tilde{\mathbf{P}}$ we have that $\underline{\mathcal{U}}_{\delta_n} \to \underline{\mathcal{T}}$ in distribution (where the laws of random variables on the left-hand side are considered under $\mathbf{P}$, and those on the right under $\tilde{\mathbf{P}}$). Thus, since the space $(\mathbb{T}, \Delta)$ is separable (see Proposition 3.4), we can suppose that we have versions of the random variables built on a common probability space, with probability measure $\mathbf{P}^*$, such that the convergence holds $\mathbf{P}^*$-a.s. From the definition of Δ and Fubini's theorem, it follows that $\int_0^\infty e^{-r}(1 \wedge \mathbf{E}^*(\Delta_c(\underline{\mathcal{U}}_{\delta_n}^{(r)}, \underline{\mathcal{T}}^{(r)})))\,dr \to 0$. Some standard analysis now yields that there exists a subsequence $(n_i)_{i\geq 1}$ such that for Lebesgue almost-every r, $\mathbf{E}^*(\Delta_c(\underline{\mathcal{U}}_{\delta_{n_i}}^{(r)}, \underline{\mathcal{T}}^{(r)})) \to 0$. In turn, letting $(r_j)_{j\geq 1}$ be a divergent sequence such that the above holds for every r_j, a straightforward diagonalisation argument yields the result. □

We now show that the push-forward of $\mu_{\mathcal{T}}$ by $\phi_{\mathcal{T}}$ is $\tilde{\mathbf{P}}$-a.s. equal to Lebesgue measure on $\mathbb{R}^2$.

LEMMA 5.2. *$\tilde{\mathbf{P}}$-a.s., it holds that $\mu_{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1} = \mathcal{L}$.*

PROOF. We first note that since $\delta^2\mu_{\mathcal{U}} \circ \phi_{\mathcal{U}}^{-1}(\delta^{-1}\cdot) \to \mathcal{L}$ for any realisation of the UST, it will suffice to show that

$$\delta_n^2\mu_{\mathcal{U}} \circ \phi_{\mathcal{U}}^{-1}(\delta_n^{-1}\cdot) \to \mu_{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1} \tag{5.2}$$

in distribution with respect to the topology of vague convergence of probability measures on $\mathbb{R}^2$. For this, it will be enough to establish that, for any continuous, positive, compactly supported function f,

$$\delta_n^2 \int_{\mathbb{R}^2} f(\delta_n x)\mu_{\mathcal{U}} \circ \phi_{\mathcal{U}}^{-1}(dx) \to \int_{\mathbb{R}^2} f(x)\mu_{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1}(dx) \tag{5.3}$$

in distribution (see [30], Theorem 16.16, e.g.).

Applying the coupling of Lemma 5.1 in conjunction with Lemma 3.7, we obtain

$$\delta_{n_i}^2\mu_{\mathcal{U}}(\phi_{\mathcal{U}}^{-1}(\delta_{n_i}^{-1}\cdot) \cap B_{\mathcal{U}}(0, \delta_{n_i}^{-\kappa} r_j)) \to \mu_{\mathcal{T}}(\phi_{\mathcal{T}}^{-1}(\cdot) \cap \mathcal{T}^{(r_j)}) \tag{5.4}$$

weakly as measures on $\mathbb{R}^2$ as $i \to \infty$, for every r_j, $\mathbf{P}^*$-a.s. In particular, this confirms that, for every r_j, the above convergence holds in distribution (under the convention that the laws of random variables on the left-hand side are considered under $\mathbf{P}$, and those on the right under $\tilde{\mathbf{P}}$). By monotonicity, we also clearly have $\tilde{\mathbf{P}}$-a.s. that, for any positive measurable f,

$$\int_{\mathbb{R}^2} f(x)\mu_{\mathcal{T}}(\phi_{\mathcal{T}}^{-1}(\cdot) \cap \mathcal{T}^{(r)})(dx) \to \int_{\mathbb{R}^2} f(x)\mu_{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1}(dx), \tag{5.5}$$

as $r \to \infty$.

As a consequence of (5.4) and (5.5), to establish the convergence at (5.3) along the subsequence $(n_i)_{i\geq 1}$, it is sufficient to show that $\mu_{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1}$ is locally finite and also that, for any continuous, positive, compactly supported function f,

$$\lim_{j\to\infty} \limsup_{i\to\infty} \delta_{n_i}^2 \left| \mathbf{E}\left(\int_{B_{\mathcal{U}}(0, \delta_{n_i}^{-\kappa} r_j)^c} f(\delta_{n_i}\phi_{\mathcal{U}}(x))\mu_{\mathcal{U}}(dx) \right) \right| = 0 \tag{5.6}$$

(cf. [13], Theorem 3.2). To show that the latter is true, first choose r such that the support of f is contained within $\overline{B}_E(0,r)$ (where we write $\overline{A}$ to represent the closure of a set A), and define $A(i,j)$ to be the event that

$$\phi_{\mathcal{U}}^{-1}\big(\overline{B}_E\big(0,\delta_{n_i}^{-1}r\big)\big) \subseteq B_{\mathcal{U}}\big(0,\delta_{n_i}^{-\kappa}r_j\big) \tag{5.7}$$

[i.e., similar to the inclusion at (3.10)]. It is then the case that the expression within the limits on the left-hand side of (5.6) is equal to

$$\delta_{n_i}^2\left|\mathbf{E}\left(\int_{B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j)^c} f\big(\delta_{n_i}\phi_{\mathcal{U}}(x)\big)\mu_{\mathcal{U}}(dx)\mathbf{1}_{A(i,j)^c}\right)\right|,$$

which is bounded above by $\sup_{x\in\mathbb{R}^2} f(x)\delta_{n_i}^2\mu_{\mathcal{U}}(\overline{B}_E(0,\delta_{n_i}^{-1}r))\mathbf{P}(A(i,j)^c) \leq c\mathbf{P}(A(i,j)^c)$ for some finite constant c. Consequently, since

$$\lim_{j\to\infty}\limsup_{i\to\infty}\mathbf{P}\big(A(i,j)^c\big)=0 \tag{5.8}$$

by Theorem 2.1(a), we have proved (5.6), as desired.

To check that $\mu_{\mathcal{T}}\circ\phi_{\mathcal{T}}^{-1}$ is locally finite, we will show that, for every $r>0$,

$$\lim_{R\to\infty}\tilde{\mathbf{P}}\big(\phi_{\mathcal{T}}^{-1}\big(\overline{B}_E(0,r)\big)\nsubseteq\mathcal{T}^{(R)}\big)=0. \tag{5.9}$$

Suppose that (5.9) is not true for some $r'>0$, with the limit instead being equal to $\varepsilon>0$. It is then the case that for every R, there exists an R' such that

$$\tilde{\mathbf{P}}\big(\phi_{\mathcal{T}}(x)\in\overline{B}_E\big(0,r'\big)\text{ for some } x\in\mathcal{T}^{(R')}\setminus\mathcal{T}^{(R)}\big)\geq\varepsilon/2. \tag{5.10}$$

Next, let us suppose that the sequences $(n_i)_{i\geq1}$ and $(r_j)_{j\geq1}$ are given by Lemma 5.1, and $D_{i,j}$, as defined by (5.1), is bounded strictly above by δ. We can then find a correspondence $\mathcal{C}_{i,j}\subseteq B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j)\times\mathcal{T}^{(r_j)}$ such that $|\delta_{n_i}^{\kappa}d_{\mathcal{U}}(0,x)-d_{\mathcal{T}}(\rho_{\mathcal{T}},x')|<2\delta$ and also $|\delta_{n_i}\phi_{\mathcal{U}}(x)-\phi_{\mathcal{T}}(x')|\leq\delta$ for every $(x,x')\in\mathcal{C}_{i,j}$. It is then easy to check that if $r_j>R'$ and the event within the probability at (5.10) holds, then so does the event that $\phi_{\mathcal{U}}(x)\in\overline{B}_E(0,\delta_{n_i}^{-1}(r'+\delta))$ for some $x\in B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j)\setminus B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}(R-2\delta))$, which is a subset of $\{\phi_{\mathcal{U}}^{-1}(\overline{B}_E(0,\delta_{n_i}^{-1}(r'+\delta)))\nsubseteq B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}(R-2\delta))\}$. Since we know that $\mathbf{P}^*(D_{i,j}>\delta)\to0$ as $i\to\infty$, it follows that

$$\liminf_{i\to\infty}\mathbf{P}\big(\phi_{\mathcal{U}}^{-1}\big(\overline{B}_E\big(0,\delta_{n_i}^{-1}\big(r'+\delta\big)\big)\big)\nsubseteq B_{\mathcal{U}}\big(0,\delta_{n_i}^{-\kappa}(R-2\delta)\big)\big)\geq\varepsilon/2.$$

However, replacing $r'+\delta$ by r and $R-2\delta$ by r_j for suitably large j, we see that this contradicts the statement at (5.8). Consequently, (5.9) must actually be true.

What we have proved already is enough to yield the lemma. We do note, though, that for any subsequence $(n_i)_{i\geq1}$, we could have applied the same argument to find a sub-subsequence $(n_{i_j})_{j\geq1}$ along which the convergence at (5.3) holds. Since the limit is identical for any such sub-subsequence, it must be the case that the full sequence also converges to this limit, which thereby establishes (5.2). □

Next, similar to (2.1), define a "Schramm-metric" on $\mathcal{T}$ by setting, for $x, y \in \mathcal{T}$,

$$d^S_{\mathcal{T}}(x, y) := \operatorname{diam}\bigl(\phi_{\mathcal{T}}\bigl(\gamma_{\mathcal{T}}(x, y)\bigr)\bigr),$$

where the diameter is in the Euclidean metric. It follows immediately from the continuity of $\phi_{\mathcal{T}}$ that $d^S_{\mathcal{T}}$ takes values in $[0, \infty)$, and it is easy to verify from the definition that $d^S_{\mathcal{T}}$ is symmetric and satisfies the triangle inequality. In the next two lemmas, we show that $d^S_{\mathcal{T}}$ is a metric on $\mathcal{T}$, and that it gives the same topology as $d_{\mathcal{T}}$.

LEMMA 5.3. *For every* $r, \eta > 0$, *we have*

$$\lim_{\varepsilon\to 0} \tilde{\mathbf{P}}\Bigl(\inf_{\substack{x,y\in B_{\mathcal{T}}(\rho_{\mathcal{T}},r):\\ d_{\mathcal{T}}(x,y)\geq\eta}} d^S_{\mathcal{T}}(x, y) < \varepsilon\Bigr) = 0. \tag{5.11}$$

PROOF. We start by proving the discrete analogue of the result: for every $r, \eta > 0$,

$$\lim_{\varepsilon\to 0} \limsup_{\delta\to 0} \mathbf{P}\Bigl(\inf_{\substack{x,y\in B_{\mathcal{U}}(0,\delta^{-\kappa}r):\\ d_{\mathcal{U}}(x,y)\geq\delta^{-\kappa}\eta}} d^S_{\mathcal{U}}(x, y) < \delta^{-1}\varepsilon\Bigr) = 0. \tag{5.12}$$

We argue similar to the proof of Lemma 4.3. Again, it is enough to consider the event in $A_*(\lambda) \cap A_4$. On $A_*(\lambda) \cap A_4$, if $x, y \in B_{\mathcal{U}}(0, c_1\delta^{-\kappa}r)$ satisfy $d^S_{\mathcal{U}}(x, y) \leq \delta^{-1}\varepsilon$, then by Proposition 2.8, $d_{\mathcal{U}}(x, y) \leq \lambda\varepsilon^\kappa\delta^{-\kappa}$. Thus, by taking ε small enough so that $\lambda\varepsilon^\kappa < \eta$, we have $d_{\mathcal{U}}(x, y) < \eta\delta^{-\kappa}$. This proves (5.12) with r replaced by $c_1 r$, and a simple reparameterisation yields the result.

To transfer to the continuous setting, let us suppose that the sequences $(n_i)_{i\geq 1}$ and $(r_j)_{j\geq 1}$ are defined by Lemma 5.1 and that the Δ_c distance between $\underline{\mathcal{U}}^{(r_j)}_{\delta_{n_i}}$ and $\underline{\mathcal{T}}^{(r_j)}$, again denoted by $D_{i,j}$, is bounded strictly above by δ. Similarly to (3.11), we can then find a correspondence $\mathcal{C}_{i,j} \subseteq B_{\mathcal{U}}(0, \delta_{n_i}^{-\kappa} r_j) \times \mathcal{T}^{(r_j)}$ such that for every $(x, x'), (y, y') \in \mathcal{C}_{i,j}$, we have

$$d^{\mathbb{R}^2}_H\bigl(\delta_{n_i}\phi_{\mathcal{U}}\bigl(\gamma(x, y)\bigr), \phi_{\mathcal{T}}\bigl(\gamma_{\mathcal{T}}(x', y')\bigr)\bigr) \leq \delta + \sup_{(z,z')\in\mathcal{T}^{(r_j)}:d_{\mathcal{T}}(z,z')<5\delta} \bigl|\phi_{\mathcal{T}}(z) - \phi_{\mathcal{T}}(z')\bigr|,$$

where $d^{\mathbb{R}^2}_H$ is the Hausdorff distance on $\mathbb{R}^2$, and so

$$\bigl|\delta_{n_i} d^S_{\mathcal{U}}(x, y) - d^S_{\mathcal{T}}(x', y')\bigr| \leq 2\delta + 2 \sup_{\substack{(z,z')\in\mathcal{T}^{(r_j)}:\\ d_{\mathcal{T}}(z,z')<5\delta}} \bigl|\phi_{\mathcal{T}}(z) - \phi_{\mathcal{T}}(z')\bigr|.$$

We can further assume that $|\delta^\kappa_{n_i} d_{\mathcal{U}}(x, y) - d_{\mathcal{T}}(x', y')| \leq 2\delta$ for every (x, x'), $(y, y') \in \mathcal{C}_{i,j}$. Next, fix $r, \eta > 0$ and select j so that $r_j > r$. It then holds that the probability on the left-hand side of (5.11) is bounded above by

$$\mathbf{P}\Big(\inf_{\substack{x,y\in B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j):\\ d_{\mathcal{U}}(x,y)\geq\delta_{n_i}^{-\kappa}(\eta-2\delta)}} d_{\mathcal{U}}^S(x,y)<2\delta_{n_i}^{-1}\varepsilon\Big) \\ +\tilde{\mathbf{P}}\Big(2\delta+2\sup_{\substack{(z,z')\in\mathcal{T}^{(r_j)}:\\ d_{\mathcal{T}}(z,z')<5\delta}}\big|\phi_{\mathcal{T}}(z)-\phi_{\mathcal{T}}(z')\big|>\varepsilon\Big)+\mathbf{P}^*(D_{i,j}>\delta),$$

where $\mathbf{P}^*$ is the coupling measure defined in the statement of Lemma 5.1. Now, by our choice of subsequence $(n_i)_{i\geq 1}$, the final expression converges to zero as $i\to\infty$, for any value of $\delta>0$. Since $\phi_{\mathcal{T}}$ is continuous, the second term converges to zero as $\delta\to 0$, for any value of $\varepsilon>0$. Hence, we can conclude

$$\tilde{\mathbf{P}}\Big(\inf_{\substack{x,y\in B_{\mathcal{T}}(\rho_{\mathcal{T}},r):\\ d_{\mathcal{T}}(x,y)\geq\eta}} d_{\mathcal{T}}^S(x,y)<\varepsilon\Big)\leq\limsup_{i\to\infty}\mathbf{P}\Big(\inf_{\substack{x,y\in B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j):\\ d_{\mathcal{U}}(x,y)\geq\delta_{n_i}^{-\kappa}\eta/2}} d_{\mathcal{U}}^S(x,y)<2\delta_{n_i}^{-1}\varepsilon\Big).$$

Since the upper bound converges to zero as $\varepsilon\to 0$ by (5.12), this completes the proof. □

LEMMA 5.4. *$\tilde{\mathbf{P}}$-a.s., $d_{\mathcal{T}}^S$ is a metric on $\mathcal{T}$, and the identity map from $(\mathcal{T},d_{\mathcal{T}})$ to $(\mathcal{T},d_{\mathcal{T}}^S)$ is a homeomorphism.*

PROOF. To establish that $d_{\mathcal{T}}^S$ is a metric, it remains to check that it is positive definite. For this, we note $\tilde{\mathbf{P}}(d_{\mathcal{T}}^S(x,y)=0$ for some $x,y\in\mathcal{T}$ with $d_{\mathcal{T}}(x,y)>0)$ is equal to

$$\lim_{r\to\infty}\lim_{\eta\to 0}\lim_{\varepsilon\to 0}\tilde{\mathbf{P}}\big(d_{\mathcal{T}}^S(x,y)<\varepsilon \text{ for some } x,y\in B_{\mathcal{T}}(\rho_{\mathcal{T}},r)\text{ with } d_{\mathcal{T}}(x,y)\geq\eta\big),$$

which in turn is equal to zero by Lemma 5.3. Next, we check that the identity map from $(\mathcal{T},d_{\mathcal{T}})$ to $(\mathcal{T},d_{\mathcal{T}}^S)$ is a homeomorphism. Clearly, it is a bijection. Moreover, its continuity follows from the continuity of $\phi_{\mathcal{T}}$. For the continuity of the inverse, we start by noting that a simple Borel–Cantelli argument yields that, $\tilde{\mathbf{P}}$-a.s., for every $\eta>0$, there exists a $\varepsilon_\eta>0$ such that $\inf_{x,y\in B_{\mathcal{T}}(\rho_{\mathcal{T}},r):d_{\mathcal{T}}(x,y)\geq\eta}d_{\mathcal{T}}^S(x,y)>\varepsilon_\eta$. In particular, this implies that if $x,y\in B_{\mathcal{T}}(\rho_{\mathcal{T}},r)$ and $d_{\mathcal{T}}^S(x,y)\leq\varepsilon_\eta$, then $d_{\mathcal{T}}(x,y)<\eta$. Hence, the identity map from $(\mathcal{T},d_{\mathcal{T}}^S)$ to $(\mathcal{T},d_{\mathcal{T}})$ is continuous, as desired. □

In order to transfer results from [48], we now show that the push-forward of $\tilde{\mathbf{P}}$ by the map $\mathfrak{T}$ introduced in Lemma 3.9 gives precisely a subsequential limiting measure as considered in the latter paper. In particular, in [48], Schramm studied properties of the subsequential limits as $\delta\to 0$ of the laws of $\mathfrak{T}(\underline{\mathcal{U}}_\delta)$, viewed as probability measures on the space $\mathcal{H}(\mathbb{S}^2\times\mathbb{S}^2\times\mathcal{H}(\mathbb{S}^2))$ (with $\mathbb{S}^2$ the one-point compactification of $\mathbb{R}^2$). Whilst the space $\mathcal{H}(\mathbb{S}^2 \times \mathbb{S}^2 \times \mathcal{H}(\mathbb{S}^2))$ is compact, and so it is immediate that the laws of $(\mathfrak{T}(\underline{\mathcal{U}}_\delta))_{\delta>0}$ are tight and admit such subsequential limits, the next result shows that along the subsequence $(\delta_n)_{n\geq1}$ we actually have convergence, with the limit being the law of $\mathfrak{T}(\underline{\mathcal{T}})$ under $\tilde{\mathbf{P}}$.

LEMMA 5.5. *The laws of* $(\mathfrak{T}(\underline{\mathcal{U}}_{\delta_n}))_{n\geq1}$ *under* $\mathbf{P}$ *converge to the law of* $\mathfrak{T}(\underline{\mathcal{T}})$ *under* $\tilde{\mathbf{P}}$, *weakly as probability measures on* $\mathcal{H}(\mathbb{S}^2 \times \mathbb{S}^2 \times \mathcal{H}(\mathbb{S}^2))$.

PROOF. We again consider the coupling of Lemma 5.1. Together with Lemma 3.9, this gives that there exists a divergent sequence $(r_j)_{j\geq1}$ such that, for every r_j, $\mathbf{P}^*$-a.s., $\mathfrak{T}(\underline{\mathcal{U}}^{(r_j)}_{\delta_{n_i}}) \to \mathfrak{T}(\underline{\mathcal{T}}^{(r_j)})$ in $\mathcal{H}(\mathbb{R}^2 \times \mathbb{R}^2 \times \mathcal{H}(\mathbb{R}^2))$, and thus also in $\mathcal{H}(\mathbb{S}^2 \times \mathbb{S}^2 \times \mathcal{H}(\mathbb{S}^2))$.

Let $d_{\mathbb{S}^2}$ be the usual metric on $\mathbb{S}^2$. Set $\delta_r := \sup_{x,y\in\mathbb{S}^2\setminus B_E(0,r)} d_{\mathbb{S}^2}(x, y)$, and note that $\delta_r \to 0$ as $r \to \infty$. Let also

$$d_{\mathbb{S}^2\times\mathbb{S}^2\times\mathcal{H}(\mathbb{S}^2)}\big((x, y, A), (x', y', A')\big) := d_{\mathbb{S}^2}(x, x') + d_{\mathbb{S}^2}(y, y') + d_H^{\mathbb{S}^2}(A, A'),$$

where $d_H^{\mathbb{S}^2}$ is the Hausdorff distance on $\mathcal{H}(\mathbb{S}^2)$. Now, suppose that i and j are indices such that the event $A(i, j)$ holds, where $A(i, j)$ is defined as in the proof of Lemma 5.2 [see the definition at (5.7) in particular]. Denoting by $d_H^{\mathbb{S}^2\times\mathbb{S}^2\times\mathcal{H}(\mathbb{S}^2)}$ the Hausdorff distance on $\mathcal{H}(\mathbb{S}^2 \times \mathbb{S}^2 \times \mathcal{H}(\mathbb{S}^2))$, we claim that on $A(i, j)$,

$$d_H^{\mathbb{S}^2\times\mathbb{S}^2\times\mathcal{H}(\mathbb{S}^2)}\big(\mathfrak{T}(\underline{\mathcal{U}}^{(r_j)}_{\delta_{n_i}}), \mathfrak{T}(\underline{\mathcal{U}}_{\delta_{n_i}})\big) < 3\delta_r,$$

and similarly, if $\phi_{\mathcal{T}}^{-1}(\overline{B}_E(0, r)) \subseteq \mathcal{T}^{(R)}$, then

$$d_H^{\mathbb{S}^2\times\mathbb{S}^2\times\mathcal{H}(\mathbb{S}^2)}\big(\mathfrak{T}(\underline{\mathcal{T}}^{(R)}), \mathfrak{T}(\underline{\mathcal{T}})\big) < 3\delta_r.$$

Since the two statements can be proved in the same way, let us consider only the latter. We need to show that if $x \in \mathcal{T} \setminus \mathcal{T}^{(R)}$ and $y \in \mathcal{T}$ then $(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y), \phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x, y)))$ is within a distance of δ_r of $\mathfrak{T}(\underline{\mathcal{T}}^{(R)})$ with respect to the metric $d_{\mathbb{S}^2\times\mathbb{S}^2\times\mathcal{H}(\mathbb{S}^2)}$. First, define x_0, y_0 to be the closest point of $\mathcal{T}^{(R)}$ to x, y, respectively, so that the triple $(\phi_{\mathcal{T}}(x_0), \phi_{\mathcal{T}}(y_0), \phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x_0, y_0)))$ is an element of $\mathfrak{T}(\underline{\mathcal{T}}^{(R)})$. By definition, we have that $\gamma_{\mathcal{T}}(x_0, x) \setminus \{x_0\}$ is a subset of $\mathcal{T} \setminus \mathcal{T}^{(R)}$, and so its image under $\phi_{\mathcal{T}}$ must fall outside of $\overline{B}_E(0, r)$. A similar observation holds in the case that $y \notin \mathcal{T}^{(R)}$. It follows that $d_{\mathbb{S}^2\times\mathbb{S}^2\times\mathcal{H}(\mathbb{S}^2)}((\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y), \phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x, y))), (\phi_{\mathcal{T}}(x_0), \phi_{\mathcal{T}}(y_0), \phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x_0, y_0)))) < 3\delta_r$, as desired.

Given (5.8), (5.9) and the conclusions of the previous two paragraphs, it is not difficult to show that $\mathfrak{T}(\underline{\mathcal{U}}_{\delta_{n_i}})$ converges to $\mathfrak{T}(\underline{\mathcal{T}})$ in distribution. The full convergence result can be obtained from this by applying a subsequence argument as in the proof of Lemma 5.2. □

As a consequence of the previous lemma, we immediately inherit a number of results from [48].

LEMMA 5.6 (see [48], Theorem 1.6, Corollary 10.4). *For $\tilde{\mathbf{P}}$-a.e. realisation of $\underline{\mathcal{T}}$, the following properties are satisfied*:

(a) *For every $(a,b,\omega)\in\mathfrak{T}(\underline{\mathcal{T}})$, if $a\neq b$, then ω is a simple path, that is, homeomorphic to* $[0,1]$. *If $a=b$, then ω is a single point or homeomorphic to a circle.*

(b) *Considered as a subset of $\mathbb{S}^2$,*

$$\text{trunk} := \bigcup_{(a,b,\omega)\in\mathfrak{T}(\underline{\mathcal{T}})} \omega \Big\backslash \{a,b\} \tag{5.13}$$

is a dense topological tree.

(c) *For each $x\in$ trunk, there are at most three connected components of* trunk $\setminus\{x\}$.

(d) *The Hausdorff dimension of* trunk *is in* $(1,2)$.

Note that, by construction, the set trunk defined at (5.13) is actually a subset of $\mathbb{R}^2$, and is also a dense topological tree when considered a subset of this space. In the following lemma, we show further that trunk is topologically equivalent to the set $\mathcal{T}^o$ introduced in the statement of Theorem 1.3. For the proof of this result, we observe that $\mathcal{T}^o$ can equivalently be defined by

$$\mathcal{T}^o = \bigcup_{x,y\in\mathcal{T}} \gamma_{\mathcal{T}}(x,y)\Big\backslash\{x,y\}. \tag{5.14}$$

Define, for $x,y\in$ trunk, $d^S_{\text{trunk}}(x,y) := \operatorname{diam}(\gamma_{\text{trunk}}(x,y))$, where $\gamma_{\text{trunk}}(x,y)$ is the unique path between x and y in trunk, and the diameter is taken with regards to the Euclidean metric. We remark that although the metric d^S_{trunk} behaves quite differently to the Euclidean one, the topologies these two metrics induce on trunk are the same; see the proof of Theorem 1.3 for details (cf. [48], Remark 10.15).

LEMMA 5.7. *$\tilde{\mathbf{P}}$-a.s., $\phi_{\mathcal{T}}$ is an isometry from $(\mathcal{T}^o, d^S_{\mathcal{T}})$ to* $(\text{trunk}, d^S_{\text{trunk}})$.

PROOF. We start the proof by establishing that

$$\phi_{\mathcal{T}}\big(\gamma_{\mathcal{T}}(x,y)\setminus\{x,y\}\big) = \phi_{\mathcal{T}}\big(\gamma_{\mathcal{T}}(x,y)\big)\setminus\big\{\phi_{\mathcal{T}}(x),\phi_{\mathcal{T}}(y)\big\} \tag{5.15}$$

for every $x,y\in\mathcal{T}$. The inclusion $\supseteq$ is easy, and so we work toward showing $\subseteq$. Let $z\in\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y)\setminus\{x,y\})$ for some $x\neq y$, and suppose that it is also the case that $z=\phi_{\mathcal{T}}(x)$. By assumption, we know that $z=\phi_{\mathcal{T}}(x')$ for some $x'\in\gamma_{\mathcal{T}}(x,y)\setminus\{x,y\}$. Now, by Lemma 5.5, because $\phi_{\mathcal{T}}(x')=\phi_{\mathcal{T}}(x)$ we can apply Lemma 5.6(a) to deduce that $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,x'))$ is either a single point or homeomorphic to a circle. Actually, since $d_{\mathcal{T}}(x,x')>0$, Lemma 5.4 tells us that $d^S_{\mathcal{T}}(x,x')>0$, and so it must be the latter option that holds true. We continue to consider two cases. First, if $\phi_{\mathcal{T}}(y)\neq\phi_{\mathcal{T}}(x)$, then Lemma 5.6(a) tells us that $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y))$ must be a simple path. However, the circle $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,x'))$ is a subset of $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y))$, and so we arrive at a contradiction. Second, if $\phi_{\mathcal{T}}(y)=\phi_{\mathcal{T}}(x)$, then one can again apply

Lemma 5.4 to choose $x'' \in \gamma_{\mathcal{T}}(x,y)$ such that $\phi_{\mathcal{T}}(x'') \neq \phi_{\mathcal{T}}(x)$. Clearly, we have that either $z \in \phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,x'') \setminus \{x,x''\})$ or $z \in \phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x'',y) \setminus \{x'',y\})$. Since $x'' \notin \{x,y\}$ and $\phi_{\mathcal{T}}(x'') \notin \{\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y)\}$, the situation reduces to the first case, and yields another contradiction. Hence, we cannot have $z = \phi_{\mathcal{T}}(x)$. Similarly, $z \neq \phi_{\mathcal{T}}(y)$, so the claim at (5.15) is proved.

Now, from (5.13), (5.14) and (5.15), it is clear that

$$\begin{aligned}\phi_{\mathcal{T}}(\mathcal{T}^o) &= \bigcup_{x,y\in\mathcal{T}} \phi_{\mathcal{T}}\big(\gamma_{\mathcal{T}}(x,y) \setminus \{x,y\}\big) = \bigcup_{x,y\in\mathcal{T}} \phi_{\mathcal{T}}\big(\gamma_{\mathcal{T}}(x,y)\big) \setminus \{\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y)\}\\ &= \text{trunk},\end{aligned}$$

and so the map $\phi_{\mathcal{T}} : \mathcal{T}^o \to$ trunk is a surjection. To complete the proof, we will again use the fact that, for every $x, y \in \mathcal{T}$ with $\phi_{\mathcal{T}}(x) \neq \phi_{\mathcal{T}}(y)$, $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y))$ is a simple path, and note that the proof of this result in [48] includes showing that the endpoints of this path are $\phi_{\mathcal{T}}(x)$ and $\phi_{\mathcal{T}}(y)$. In particular, if $x, y \in \mathcal{T}^o$ are such that $\phi_{\mathcal{T}}(x) \neq \phi_{\mathcal{T}}(y)$, then we know that the simple path $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y))$ from $\phi_{\mathcal{T}}(x)$ to $\phi_{\mathcal{T}}(y)$ is contained in trunk. Recalling that trunk is a topological tree, which implies there is be a unique path $\gamma_{\text{trunk}}(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y))$ between $\phi_{\mathcal{T}}(x)$ and $\phi_{\mathcal{T}}(y)$ within this set, it must be the case that $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y)) = \gamma_{\text{trunk}}(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y))$. On the other hand, if $x, y \in \mathcal{T}^o$ are such that $\phi_{\mathcal{T}}(x) = \phi_{\mathcal{T}}(y)$, then, by Lemma 5.6(a), it must hold that $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y)) = \{\phi_{\mathcal{T}}(x)\}$, where we note that we can exclude the possibility that $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y))$ is homeomorphic to a circle, since trunk is a topological tree and cannot contain such a subset. Hence we obtain that $\phi_{\mathcal{T}}(\gamma_{\mathcal{T}}(x,y)) = \gamma_{\text{trunk}}(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y))$ in this case, also. Consequently, $d^S_{\mathcal{T}}(x,y) = d^S_{\text{trunk}}(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y))$ for every $x, y \in \mathcal{T}^o$. This confirms that $\phi_{\mathcal{T}}$ is an isometry, as desired. □

Before we complete the proof of Theorem 1.3, we mention another property of the trunk that will be needed. This is that the trunk can be used to reconstruct the *dual trunk*, that is, the (subsequential) scaling limit of the dual graph of the UST (see [48], Remarks 10.13 and 10.14). More precisely, for any two points $x, y \in \mathbb{S}^2 \setminus$ trunk, there exists a unique path in $\mathbb{S}^2 \setminus$ trunk between them. Denote this path by $\gamma_{\text{trunk}^\dagger}(x,y)$, and set $\text{trunk}^\dagger := \bigcup_{x,y\in\mathbb{S}^2\setminus\text{trunk}} \gamma_{\text{trunk}^\dagger}(x,y) \setminus \{x,y\}$. This is the dual trunk, which is distributed identically to trunk.

PROOF OF THEOREM 1.3. It readily follows from Theorem 1.1 and the unboundedness of $(\mathcal{U}, d_{\mathcal{U}})$ that the diameter of $(\mathcal{T}, d_{\mathcal{T}})$ is infinite, and so it has at least one end at infinity. Thus, to complete the proof of part (a)(ii), it will suffice to show that there can be no more than one end at infinity. To this end, note that, for any $r > 0$,

$$\tilde{\mathbf{P}}\big((\mathcal{T}, d_{\mathcal{T}}) \text{ has} \geq 2 \text{ ends at infinity}\big) \leq \lim_{R\to\infty} \tilde{\mathbf{P}}\big(C_{\mathcal{T}}(r,R) \geq 2\big),$$

where $C_{\mathcal{T}}(r, R)$ is the event that there exist $x, y \notin \mathcal{T}^{(R)}$ such that $\gamma_{\mathcal{T}}(x, y) \cap \mathcal{T}^{(r)} \neq \varnothing$. By applying the coupling of Lemma 5.1, it is possible to bound the inner probability by $\limsup_{\delta\to 0} \mathbf{P}(C_{\mathcal{U}}(2\delta^{-\kappa} r, \delta^{-\kappa} R/2))$, where $C_{\mathcal{U}}(r, R)$ is defined similarly to $C_{\mathcal{T}}(r, R)$, with $\mathcal{T}$ replaced by $\mathcal{U}$. (Since we have already presented similar coupling arguments in the proofs of Lemmas 5.2, 5.3 and 5.5, we omit the details.) Now, for $\lambda > 0$,

$$\begin{aligned}
&\mathbf{P}\big(C_{\mathcal{U}}(2\delta^{-\kappa} r, \delta^{-\kappa} R/2)\big) \\
&\quad \leq \mathbf{P}\big(C^E_{\mathcal{U}}(\lambda\delta^{-1} r^{1/\kappa}, \delta^{-1} R^{1/\kappa}/\lambda)\big) + \mathbf{P}\big(B_{\mathcal{U}}(0, 2\delta^{-\kappa} r) \nsubseteq B_E(0, \lambda\delta^{-1} r^{1/\kappa})\big) \\
&\qquad + \mathbf{P}\big(B_E(0, \lambda^{-1}\delta^{-1} R^{1/\kappa}) \nsubseteq B_{\mathcal{U}}(0, \delta^{-\kappa} R/2)\big),
\end{aligned}$$

where $C^E_{\mathcal{U}}(r, R)$ is the event that there exist $x, y \notin B_E(0, R)$ such that $\gamma(x, y) \cap B_E(0, r) \neq \varnothing$. Hence, from Theorem 2.1(a) and [2], we obtain, for $R \geq \lambda \geq 2$ and $R \geq \lambda^{2\kappa} r$,

$$\mathbf{P}\big(C_{\mathcal{U}}(2\delta^{-\kappa} r, \delta^{-\kappa} R/2)\big) \leq c_1 \left(\frac{\lambda^2 r^{1/\kappa}}{R^{1/\kappa}}\right)^{c_2} + c_3 \lambda^{-1/6}.$$

Taking $\lambda = R^{1/3\kappa}$, this converges to 0 as $R \to \infty$. It follows that the $\tilde{\mathbf{P}}$-probability of $(\mathcal{T}, d_{\mathcal{T}})$ having ≥ 2 ends at infinity is zero, as desired.

For part (b)(i), we begin by noting that $A \subseteq \phi_{\mathcal{T}}^{-1}(\phi_{\mathcal{T}}(A))$ for any set A. Consequently, from Lemma 5.2, we obtain $\mu_{\mathcal{T}}(\{x\}) \leq \mathcal{L}(\{\phi_{\mathcal{T}}(x)\}) = 0$ for every $x \in \mathcal{T}$. Moreover, by Lemma 5.7, $\mu_{\mathcal{T}}(\mathcal{T}^o) \leq \mathcal{L}(\phi_{\mathcal{T}}(\mathcal{T}^o)) = \mathcal{L}(\text{trunk}) = 0$, where the final equality is a consequence of the fact that trunk has Hausdorff dimension strictly less than two [as recalled in Lemma 5.6(d)].

To establish part (b)(ii), it will suffice to check that, given $R > 0$, there exist constants $c_1, c_2, c_3, c_4, c_5, c_6 \in (0, \infty)$ such that, for every $r \in (0, 1)$,

$$\tilde{\mathbf{P}}\Big(\inf_{x \in B_{\mathcal{T}}(\rho_{\mathcal{T}}, R)} \mu_{\mathcal{T}}\big(B_{\mathcal{T}}(x, r)\big) \leq c_1 r^{d_f} \big(\log r^{-1}\big)^{-80}\Big) \leq c_2 r^{c_3},$$

$$\tilde{\mathbf{P}}\Big(\sup_{x \in B_{\mathcal{T}}(\rho_{\mathcal{T}}, R)} \mu_{\mathcal{T}}\big(B_{\mathcal{T}}(x, r)\big) \geq c_4 r^{d_f} \big(\log r^{-1}\big)^{80}\Big) \leq c_5 r^{c_6}.$$

Indeed, one can then apply a simple Borel–Cantelli argument along the subsequence $r_n = 2^{-n}$, $n \in \mathbb{N}$, to deduce the result. We observe that the above inequalities can be deduced from the definition of $\tilde{\mathbf{P}}$ and Corollary 2.11 by applying the coupling of Lemma 5.1. Furthermore, note that part (a)(i) is an elementary consequence of (b)(ii) (see [24], Proposition 1.5.15, e.g.).

Part (b)(iii) can also be obtained using a Borel–Cantelli argument in conjunction with the following: there exist constants $c_1, c_2 \in (0, \infty)$ such that

$$\tilde{\mathbf{P}}\big(\mu_{\mathcal{T}}\big(B_{\mathcal{T}}(\rho_{\mathcal{T}}, r)\big) \geq \lambda r^{d_f}\big) \leq c_1 e^{-c_2 \lambda^{1/3}}, \tag{5.16}$$

$$\tilde{\mathbf{P}}\big(\mu_{\mathcal{T}}\big(B_{\mathcal{T}}(\rho_{\mathcal{T}}, r)\big) \leq \lambda^{-1} r^{d_f}\big) \leq c_1 e^{-c_2 \lambda^{1/9}}, \tag{5.17}$$

for all $r > 0$, $\lambda \geq 1$. Again applying the definition of $\tilde{\mathbf{P}}$ and the coupling of Lemma 5.1, it is possible to deduce the bound at (5.16) from Theorem 2.1(b). The proof for the bound at (5.17) is similar.

The first statement of part (c)(i) depends on Lemmas 5.4 and 5.7. In particular, these two results imply that $\phi_{\mathcal{T}}$ is a homeomorphism from $(\mathcal{T}^o, d_{\mathcal{T}})$ to $(\text{trunk}, d^S_{\text{trunk}})$. To replace the topology generated by d^S_{trunk} with the Euclidean one, we will show that the identity map from $(\text{trunk}, d^S_{\text{trunk}})$ to (trunk, d_E) is also a homeomorphism. Clearly, it is a continuous bijection, and so we need to show its inverse is continuous. To do this, suppose that $x_n, x \in \text{trunk}$ are such that $d_E(x_n, x) \to 0$. Now, in light of $\phi_{\mathcal{T}} : (\mathcal{T}^o, d_{\mathcal{T}}) \to (\text{trunk}, d^S_{\text{trunk}})$ being a homeomorphism, the map $\phi_{\mathcal{T}} : \mathcal{T} \to \mathbb{R}^2$ can be viewed as the extension of the identity map $\text{trunk} \to \mathbb{R}^2$ to a continuous map on the completion of $(\text{trunk}, d^S_{\text{trunk}})$, and it therefore follows from the discussion in [48], Remark 10.15, that $|\phi^{-1}_{\mathcal{T}}(x)|$ is equal to one if x is not contained in $\text{trunk}^\dagger$. In particular, since $\text{trunk} \cap \text{trunk}^\dagger = \emptyset$, there exist unique $y_n, y \in \mathcal{T}$ such that $\phi_{\mathcal{T}}(y_n) = x_n$ and $\phi_{\mathcal{T}}(y) = x$. Moreover, since $x_n, x \in B_E(0, r)$ for some $r < \infty$, there must exist an $R < \infty$ such that $y_n, y \in B_{\mathcal{T}}(\rho_{\mathcal{T}}, R)$—this is an easy consequence of (5.9). Hence, by compactness, for any subsequence n_i, there exists a sub-subsequence $y_{n_{i_j}}$ such that $d_{\mathcal{T}}(y_{n_{i_j}}, y') \to 0$ for some $y' \in \mathcal{T}$. By the continuity of $\phi_{\mathcal{T}}$, it follows that $d_E(\phi_{\mathcal{T}}(y'), \phi_{\mathcal{T}}(y)) = \lim_{j\to\infty} d_E(\phi_{\mathcal{T}}(y_{n_{i_j}}), \phi_{\mathcal{T}}(y)) = \lim_{j\to\infty} d_E(x_{n_{i_j}}, x) = 0$, and so $y' = y$. Noting that y_n, y are necessarily in $\mathcal{T}^o$, Lemmas 5.4 and 5.7 thus yield $d^S_{\text{trunk}}(x_{n_{i_j}}, x) = d^S_{\mathcal{T}}(y_{n_{i_j}}, y) \to 0$. Since the initial subsequence (n_i) was arbitrary, this implies $d^S_{\text{trunk}}(x_n, x) \to 0$, as desired. The denseness of $\phi_{\mathcal{T}}(\mathcal{T}^o)$ in $\mathbb{R}^2$ follows from Lemmas 5.6(b) and 5.7. Furthermore, applying Lemma 5.6(c) together with the homeomorphism between $\mathcal{T}^o$ and trunk yields that $\max_{x\in\mathcal{T}} \deg_{\mathcal{T}}(x) = 3$, which is the first claim of part (c)(ii). To check the remaining claim of part (c)(ii), we note that if x is contained in $\text{trunk}^\dagger$, then $|\phi^{-1}_{\mathcal{T}}(x)|$ is equal to the degree of x in $\text{trunk}^\dagger$ (again, see the discussion in [48], Remark 10.15). Since $\text{trunk}^\dagger$ also has maximum degree 3 [by Lemma 5.6(c) again], this establishes the desired result. [Recall that $|\phi^{-1}_{\mathcal{T}}(x)| = 1$ for $x \notin \text{trunk}^\dagger$.] Finally, point (c)(iii) will be a simple consequence of Lemma 5.2, at least if we can show that $\mathcal{L}(\{x : |\phi^{-1}_{\mathcal{T}}(x)| \geq 2\}) = 0$. However, from our previous observations, we know that the set $\{x : |\phi^{-1}_{\mathcal{T}}(x)| \geq 2\}$ is contained in $\text{trunk}^\dagger$, which has Hausdorff dimension strictly less than two [by Lemma 5.6(d)]. □

6. Simple random walk scaling limit. In this section, we prove Theorem 1.4. The general convergence result for simple random walks on graph trees (see Theorem 6.1 below) that we apply extends [16, 20, 21] from the setting of ordered graph trees (in particular, in those articles graph trees and real trees were encoded by functions). Work is also needed to extend to the noncompact setting of this article.

Let us start by introducing some notation. Let $(T_n)_{n\geq 1}$ be a sequence of finite graph trees. Write d_{T_n} for the shortest path graph distance on T_n, and μ_{T_n} for the counting measure on the vertices of T_n. Suppose that ϕ_n is a map from the vertices of T_n into M—until otherwise noted, we assume that M is a separable normed vector space, and write the metric induced by its norm as d_M. Fix a distinguished vertex ρ_{T_n} of T_n. We extend $(T_n, d_{T_n}, \mu_{T_n}, \phi_{T_n}, \rho_{T_n})$ to an element of $\mathbb{T}_c$ by adding line segments of unit length along edges of the graph tree, and (isometrically) interpolating ϕ_{T_n} between these along the relevant geodesics. The process $(X_t^{T_n})_{t\geq 0}$ is the discrete time simple random walk on T_n, and $P_x^{T_n}$ its law started from x. We extend $(\phi_{T_n}(X_t^{T_n}))_{t\geq 0}$ to an element of $C(\mathbb{R}_+, M)$ by interpolation along geodesics.

The limit space we consider in Theorem 6.1 is the natural generalisation of that of [21]. In particular, let $\mathbb{T}_c^*$ be the collection of those elements $\underline{\mathcal{T}}$ of $\mathbb{T}_c$ such that $\mu_{\mathcal{T}}$ is nonatomic, supported on the leaves of $\mathcal{T}$ [recall that the leaves of a real tree $\mathcal{T}$ are those points $x \in \mathcal{T}$ such that $\mathcal{T} \setminus \{x\}$ is connected, i.e., which have $\deg_{\mathcal{T}}(x) = 1$], and also there exists a constant c such that

$$\liminf_{r\to 0} \inf_{x\in\mathcal{T}} r^{-c} \mu_{\mathcal{T}}\big(B_{\mathcal{T}}(x, r)\big) > 0. \tag{6.1}$$

For a locally compact real tree $(\mathcal{T}, d_{\mathcal{T}})$ equipped with a locally finite Borel measure $\mu_{\mathcal{T}}$ of full support, it is shown in [6] how to construct an associated "Brownian motion" (cf. [32], which deals with the case when $(\mathcal{T}, d_{\mathcal{T}})$ is complete). For readers' convenience, let us briefly summarise this construction. In particular, first define the length measure $\lambda_{\mathcal{T}}$ on $\mathcal{T}$ to be the restriction of one-dimensional Hausdorff measure to $\mathcal{T}^o := \mathcal{T} \setminus \{x \in \mathcal{T} : \deg_{\mathcal{T}}(x) = 1\}$. Moreover, let $\mathcal{A}$ be the collection of locally absolutely continuous functions on $\mathcal{T}$, where we say a function $f : \mathcal{T} \to \mathbb{R}$ is locally absolutely continuous if and only if for every $\varepsilon > 0$ and subset $A \subseteq \mathcal{T}$ with $\lambda_{\mathcal{T}}(A) < \infty$, there exists a $\delta > 0$ such that: if $\gamma_{\mathcal{T}}(x_1, y_1), \dots, \gamma_{\mathcal{T}}(x_k, y_k) \subseteq A$, $k \in \mathbb{N}$ are disjoint arcs with $\sum_{i=1}^k d_{\mathcal{T}}(x_i, y_i) < \delta$, then $\sum_{i=1}^k |f(x_i) - f(y_i)| < \varepsilon$. Given a function $f \in \mathcal{A}$, there exists a unique (up to $\lambda_{\mathcal{T}}$-null sets) function g that is locally in $L^1(\mathcal{T}, \lambda_{\mathcal{T}})$ such that

$$f(y) - f(x) = -\int_{\gamma_{\mathcal{T}}(b_{\mathcal{T}}(\rho_{\mathcal{T}},x,y),x)} g(z)\lambda_{\mathcal{T}}(dz) + \int_{\gamma_{\mathcal{T}}(b_{\mathcal{T}}(\rho_{\mathcal{T}},x,y),y)} g(z)\lambda_{\mathcal{T}}(dz),$$

for all $x, y \in \mathcal{T}$, where $b_{\mathcal{T}}(\rho_{\mathcal{T}}, x, y)$ is the unique branch-point of $\rho_{\mathcal{T}}$, x and y in $\mathcal{T}$ [6], Proposition 1.1. (Note that the above difference can be interpreted as an oriented integral from x to y, and is independent of the choice of $\rho_{\mathcal{T}}$.) The function g in the previous sentence is called the gradient of f, and is denoted ∇f. Next, define

$$\mathcal{E}_{\mathcal{T}}(f, g) := \frac{1}{2} \int \nabla f(x) \nabla g(x) \lambda_{\mathcal{T}}(dx),$$

for all $f, g \in \mathcal{F}_{\mathcal{T}}$, where $\mathcal{F} := \{f \in \mathcal{A} : \nabla f \in L^2(\mathcal{T}, \lambda_{\mathcal{T}})\} \cap L^2(\mathcal{T}, \mu_{\mathcal{T}}) \cap C_\infty(\mathcal{T})$. Here, $C_\infty(\mathcal{T})$ is the space of continuous functions on $\mathcal{T}$ that vanish at infinity.

By [6], Propositions 2.4 and 4.1, and the proof of Theorem 1, $(\mathcal{E}_{\mathcal{T}}, \mathcal{F}_{\mathcal{T}})$ is a local, regular Dirichlet form on $L^2(\mathcal{T}, \mu_{\mathcal{T}})$. The Brownian motion on $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}})$ is the continuous, $\mu_{\mathcal{T}}$-symmetric, strong Markov process $((X_t^{\mathcal{T}})_{t\geq 0}, (P_x^{\mathcal{T}})_{x\in\mathcal{T}})$ associated with this Dirichlet form (see [27]). Clearly, this construction applies to elements of $\mathbb{T}_c^*$, and the additional restriction (6.1) allows one to deduce that this Brownian motion has local times $(L_t^{\mathcal{T}}(x))_{x\in\mathcal{T},t\geq 0}$ which are jointly continuous in t and x (see [21], Lemma 2.2).

THEOREM 6.1. *Let $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$, $(c_n)_{n\geq 1}$ be null sequences with $b_n = o(a_n)$ such that*

$$(T_n, a_n d_{T_n}, b_n \mu_{T_n}, c_n \phi_{T_n}, \rho_{T_n}) \to \underline{\mathcal{T}} \tag{6.2}$$

in $(\mathbb{T}_c, \Delta_c)$, where $\underline{\mathcal{T}}$ is an element of $\mathbb{T}_c^*$. Then

$$\big(c_n \phi_{T_n}(X^{T_n}_{t/a_n b_n})\big)_{t\geq 0} \to \big(\phi_{\mathcal{T}}(X_t^{\mathcal{T}})\big)_{t\geq 0} \tag{6.3}$$

in distribution in $C(\mathbb{R}_+, M)$, where we assume $X_0^{T_n} = \rho_{T_n}$ for each n, and also $X_0^{\mathcal{T}} = \rho_{\mathcal{T}}$.

Since the proof of this result is close to the arguments of [16, 20, 21], we will not give all of the details. For clarity, though, we will break it into three lemmas. The basic idea is to approximate the processes of interest by processes on trees which have finite total length, for which convergence is more straightforward to prove. So, let $\underline{\mathcal{T}}$ be an element of $\mathbb{T}_c^*$, and $(x_i)_{i\geq 1}$ be a dense sequence of vertices in $\mathcal{T}$—these will be fixed throughout the current discussion. (To avoid trivialities, we assume that $\mathcal{T}$ consists of more than one point.) We suppose that $(x_i)_{i\geq 1}$ are distinct, and none is equal to the root $\rho_{\mathcal{T}}$, which we will sometimes also denote by x_0. For each $k \geq 1$, define $\mathcal{T}(k) := \bigcup_{i=1}^k \gamma_{\mathcal{T}}(\rho_{\mathcal{T}}, x_i)$, and let $\pi_{\mathcal{T},\mathcal{T}(k)}$ be the natural projection from $\mathcal{T}$ to $\mathcal{T}(k)$, that is, for $x \in \mathcal{T}$, $\pi_{\mathcal{T},\mathcal{T}(k)}(x)$ is the closest point in $\mathcal{T}(k)$ to x. Taking $\mu^{(k)} := \mu_{\mathcal{T}} \circ \pi^{-1}_{\mathcal{T},\mathcal{T}(k)}$, we define $X^{\mathcal{T}(k),\mu^{(k)}}$ to be Brownian motion on $(\mathcal{T}(k), d_{\mathcal{T}}|_{\mathcal{T}(k)}, \mu^{(k)})$. By [21], Proposition 2.1, if we assume that $X^{\mathcal{T}(k),\mu^{(k)}}$ and $X^{\mathcal{T}}$ are both started from $\rho_{\mathcal{T}}$, then $(X_t^{\mathcal{T}(k),\mu^{(k)}})_{t\geq 0} \to (X_t^{\mathcal{T}})_{t\geq 0}$ in distribution in $C(\mathbb{R}_+, \mathcal{T})$. (This step is one of the places in the proof that the existence of jointly continuous local times for the Brownian motion $X^{\mathcal{T}}$ is used.) Hence, the continuous mapping theorem implies that

$$\big(\phi_{\mathcal{T}}(X_t^{\mathcal{T}(k),\mu^{(k)}})\big)_{t\geq 0} \to \big(\phi_{\mathcal{T}}(X_t^{\mathcal{T}})\big)_{t\geq 0} \tag{6.4}$$

in distribution in $C(\mathbb{R}_+, M)$. Moreover, if we define $\mathcal{B}(k) := \{b_{\mathcal{T}}(\rho_{\mathcal{T}}, x_i, x_j) : i, j \in \{0, \dots, k\}\}$ and $\phi_{\mathcal{T}}^{(k)} : \mathcal{T}(k) \to M$ by setting $\phi_{\mathcal{T}}^{(k)} = \phi_{\mathcal{T}}$ on $\mathcal{B}(k)$ and interpolating along geodesics between these vertices, then one can deduce from the denseness of $(x_i)_{i\geq 1}$ and continuity of $\phi_{\mathcal{T}}$ that $\lim_{k\to\infty} \sup_{x\in\mathcal{T}(k)} d_M(\phi_{\mathcal{T}}^{(k)}(x), \phi_{\mathcal{T}}(x)) = 0$ (cf. [20], Theorem 8.2, and the following discussion). Consequently, (6.4) yields the following lemma.

LEMMA 6.2. *As $k \to \infty$,*

$$(\phi_{\mathcal{T}}^{(k)}(X_t^{\mathcal{T}(k),\mu^{(k)}}))_{t\geq 0} \to (\phi_{\mathcal{T}}(X_t^{\mathcal{T}}))_{t\geq 0}$$

in distribution in $C(\mathbb{R}_+, M)$.

Before describing the connection with discrete objects, let us note that $X^{\mathcal{T}(k),\mu^{(k)}}$ can also be represented as a time change of another Brownian motion on $\mathcal{T}(k)$. In particular, let $\lambda^{(k)}$ be the one-dimensional Hausdorff measure on $(\mathcal{T}(k), d_{\mathcal{T}}|_{\mathcal{T}(k)})$, and $X^{\mathcal{T}(k),\lambda^{(k)}}$ be the associated Brownian motion. Since $\lambda^{(k)}$ satisfies (6.1), this process admits jointly continuous local times $(L_t^{(k)}(x))_{x\in\mathcal{T}(k),t\geq 0}$, from which we define an additive functional $\hat{A}_t^{(k)} := \int_{\mathcal{T}(k)} L_t^{(k)}(x)\mu^{(k)}(dx)$, and its inverse $\hat{\tau}^{(k)}(t) := \inf\{s : \hat{A}_s^{(k)} > t\}$. (We use hatted notation for consistency with [21].) From [21], Lemma 2.4, we then obtain that if $X^{\mathcal{T}(k),\lambda^{(k)}}$ is started from $\rho_{\mathcal{T}}$, then

$$(X_{\hat{\tau}^{(k)}(t)}^{\mathcal{T}(k),\lambda^{(k)}})_{t\geq 0} \tag{6.5}$$

is distributed identically to $X^{\mathcal{T}(k),\mu^{(k)}}$ started from $\rho_{\mathcal{T}}$.

For the next part of the proof of Theorem 6.1, we fix a sequence of metric spaces Z_n, isometric embeddings $\psi_n : \mathcal{T} \to Z_n$, $\psi_n' : (T_n, a_n d_{T_n}) \to Z_n$ and correspondences $\mathcal{C}_n$ between $\mathcal{T}$ and T_n containing $(\rho_{\mathcal{T}}, \rho_{T_n})$ and such that (3.3) holds with $\underline{\mathcal{T}}_n$ replaced by $(T_n, a_n d_{T_n}, b_n\mu_{T_n}, c_n\phi_{T_n}, \rho_{T_n})$ for some sequence $\varepsilon_n \to 0$. [This is possible if we suppose that (6.2) holds.] Moreover, let $x_i^n \in T_n$ be such that $(x_i, x_i^n) \in \mathcal{C}_n$, and define the subtree $T_n(k) \subseteq T_n$ and projection $\pi_{n,k} : T_n \to T_n(k)$ similarly to the continuous case. Using elementary arguments, one can check that

$$\begin{aligned}\lim_{k\to\infty} \limsup_{n\to\infty} a_n \max_{x\in T_n} d_{T_n}(x, \pi_{n,k}(x)) \\ \leq \lim_{k\to\infty} \limsup_{n\to\infty}\Big(\sup_{x\in\mathcal{T}} d_{\mathcal{T}}(x, \pi_{\mathcal{T},\mathcal{T}(k)}(x)) + 5\varepsilon_n\Big) = 0\end{aligned} \tag{6.6}$$

(cf. [16], Lemma 2.7), which says that the subtrees $T_n(k)$ are uniformly good approximations of the full trees T_n. As for processes, we define $X^{n,k} := \pi_{n,k}(X^{T_n})$ and $J^{n,k}$ to be the corresponding jump chain, that is, if $A_0^{n,k} := 0$ and $A_t^{n,k} := \min\{s \geq A_{t-1}^{n,k} : X_s^{T_n} \in T_n(k) \setminus \{X_{A_{t-1}^{n,k}}^{T_n}\}\}$, then $J_t^{n,k} = X_{A_t^{n,k}}^{T_n}$. Conversely, if $\tau^{n,k}(t) := \max\{s : A_s^{n,k} \leq t\}$, then we can write

$$X_t^{n,k} = J_{\tau^{n,k}(t)}^{n,k}. \tag{6.7}$$

Define the local times of $J^{n,k}$ by setting $L_t^{n,k}(x) := \frac{2}{\deg_{n,k}(x)} \sum_{s=0}^{t-1} \mathbf{1}_x(J_s^{n,k})$ for $x \in T_n(k)$ and $t \geq 0$, where $\deg_{n,k}(x)$ is the usual graph degree of x in $T_n(k)$. We use these to define an associated additive functional $\hat{A}^{n,k}$ by setting $\hat{A}_0^{n,k} := 0$ and $\hat{A}_t^{n,k} := \int_{T_n(k)} L_t^{n,k}(x)\mu_{n,k}(dx)$, where $\mu_{n,k} := \mu_{T_n} \circ \pi_{n,k}^{-1}$. Finally, from the inverse $\hat{\tau}^{n,k}(t) := \max\{s : \hat{A}_s^{n,k} \leq t\}$, we define an alternative time-change of $J^{n,k}$ by setting

$$\hat{X}_t^{n,k} = J^{n,k}_{\hat{\tau}^{n,k}(t)}. \tag{6.8}$$

In the next lemma, we describe how methods of [21], Section 3, can be applied to deduce a scaling limit for these processes. The map $\phi_{T_n}^{(k)} : T_n(k) \to M$ is defined analogously to the continuous case.

LEMMA 6.3. *Suppose that* (6.2) *holds, and fix* $k \geq 1$. *Then*

$$\big(c_n \phi_{T_n}^{(k)}\big(\hat{X}^{n,k}_{t/a_n b_n}\big)\big)_{t\geq 0} \to \big(\phi_{\mathcal{T}}^{(k)}\big(X_t^{\mathcal{T}(k),\mu^{(k)}}\big)\big)_{t\geq 0},$$

in distribution in $C(\mathbb{R}_+, M)$.

PROOF. In this proof, we will use an embedding of trees into ℓ^1, 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 $x \in \ell^1$ (the procedure was originally described in [4]). In particular, given a sequence $(\mathcal{T}(k))_{k\geq 1}$ as above it is possible to construct a distance-preserving map $\tilde{\psi} : (\mathcal{T}, d_{\mathcal{T}}) \to (\ell^1, d_{\ell^1})$ that satisfies $\tilde{\psi}(\rho) = 0$ and

$$\pi_k\big(\tilde{\psi}(\sigma)\big) = \tilde{\psi}\big(\pi_{\mathcal{T},\mathcal{T}(k)}(\sigma)\big) \tag{6.9}$$

for every $\sigma \in \mathcal{T}$ and $k \geq 1$, where π_k is the projection map on ℓ^1, that is, $\pi_k(x(1), x(2), \dots) = (x(1), \dots, x(k), 0, 0, \dots)$. Roughly speaking, we first map $\mathcal{T}(1)$ to a line segment of length $d_{\mathcal{T}}(\rho_{\mathcal{T}}, x_1)$ in the first coordinate direction of ℓ^1. Then, given the map $\tilde{\psi}$ on $\mathcal{T}(k)$, map the additional line segment in $\mathcal{T}(k+1)$ to a line segment in the $(k+1)$st coordinate direction of ℓ^1 [i.e., orthogonally to $\tilde{\psi}(\mathcal{T}(k))$], attached at the image in ℓ^1 of the appropriate branch-point. Such a map is determined uniquely by insisting that $\tilde{\psi}(\mathcal{T}) \subseteq \{(x(1), x(2), \dots) \in \ell^1 : x(i) \geq 0, i = 1, 2, \dots\}$. We can of course embed the discrete trees similarly, and we denote the corresponding embeddings by $\tilde{\psi}_n$. It is not difficult to check from our construction that, for every $i \geq 0$,

$$a_n \tilde{\psi}_n\big(x_i^n\big) \to \tilde{\psi}(x_i). \tag{6.10}$$

As a consequence of this and the fact that the maps $\tilde{\psi}$ and $\tilde{\psi}_n$ are isometries, we find that $(T_n, a_n d_{T_n}, b_n \mu_{T_n}, a_n \tilde{\psi}_n, \rho_{T_n}) \to (\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \tilde{\psi}, \rho_{\mathcal{T}})$ in the version of $(\mathbb{T}_c, \Delta_c)$ where maps are into $(M, d_M) = (\ell^1, d_\ell^1)$. Moreover, taking projections π_k yields $(T_n, a_n d_{T_n}, b_n \mu_{T_n}, \pi_k \circ a_n \tilde{\psi}_n, \rho_{T_n}) \to (\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \pi_k \circ \tilde{\psi}, \rho_{\mathcal{T}})$ in the version of $(\mathbb{T}_c, \Delta_c)$ where maps are into $\mathbb{R}^k$. Hence, by applying Lemma 3.7, and noting the characterisation of $\tilde{\psi}$ at (6.9), it follows that

$$b_n \mu_{n,k} \circ (a_n \tilde{\psi}_n)^{-1} \to \mu^{(k)} \circ \tilde{\psi}^{-1} \tag{6.11}$$

weakly as measures on $\mathbb{R}^k$, and the same conclusion also holds in terms of weak convergence of measures on ℓ_1. The two conditions (6.10) and (6.11) enable us to obtain from [21], Proposition 3.1, (see also the proof of [21], Lemma 4.2) that

$$(6.12)\qquad \big(a_n\tilde{\psi}_n(J^{n,k}_{t/a_n^2}), a_nb_n\hat{A}^{n,k}_{t/a_n^2}\big)_{t\geq 0} \to \big(\tilde{\psi}(X_t^{\mathcal{T}(k),\lambda^{(k)}}), \hat{A}_t^{(k)}\big)_{t\geq 0}$$

in distribution in $C(\mathbb{R}_+, \ell^1\times\mathbb{R}_+)$. Since the functions $t\mapsto \hat{A}_t^{(k)}$ are almost-surely continuous and strictly increasing (by [21], Lemma 2.5), one can take an inverse in the second coordinate and compose with the first to obtain

$$(6.13)\qquad \big(a_n\tilde{\psi}_n(\hat{X}^{n,k}_{t/a_nb_n})\big)_{t\geq 0} \to \big(\tilde{\psi}(X_t^{\mathcal{T}(k),\mu^{(k)}})\big)_{t\geq 0}$$

in distribution in $C(\mathbb{R}_+, \ell^1)$, for which it is helpful to recall the expressions at (6.5) and (6.8). Now, from our choice of x_i^n, one can check that, for every $i, j\geq 0$, $c_n\phi_{T_n}(b_{T_n}(\rho_{T_n}, x_i^n, x_j^n)) \to \phi_{\mathcal{T}}(b_{\mathcal{T}}(\rho_{\mathcal{T}}, x_i, x_j))$, where the function b_{T_n} returns the branch-point of three vertices of T_n. This allows one to transfer the convergence of (6.13) into M, and so obtain the result. □

In light of Lemmas 6.2 and 6.3, the proof of Theorem 6.1 is completed by the following lemma (see [13], Theorem 3.2, e.g.).

LEMMA 6.4. *Suppose that* (6.2) *holds. For every* $\varepsilon > 0$,

$$\lim_{k\to\infty}\limsup_{n\to\infty} P_{\rho_{T_n}}^{T_n}\Big(\sup_{s\in[0,t]} d_M\big(c_n\phi_{T_n}(X^{T_n}_{s/a_nb_n}), c_n\phi_{T_n}^{(k)}(\hat{X}^{n,k}_{s/a_nb_n})\big) > \varepsilon\Big) = 0.$$

PROOF. From the definition of ε_n, one can obtain

$$\begin{aligned}
&\lim_{\delta\to 0}\lim_{k\to\infty}\limsup_{n\to\infty} \max_{\substack{x\in T_n, y\in T_n(k):\\ a_nd_{T_n}(x,y)\leq\delta}} c_nd_M\big(\phi_{T_n}(x), \phi_{T_n}^{(k)}(y)\big)\\
&\quad\leq \lim_{\delta\to 0}\lim_{k\to\infty}\limsup_{n\to\infty}\Big(\sup_{\substack{x,y\in\mathcal{T}:\\ d_{\mathcal{T}}(x,y)\leq\delta+\delta_k+18\varepsilon_n}} d_M\big(\phi_{\mathcal{T}}(x), \phi_{\mathcal{T}}(y)\big) + 2\varepsilon_n\Big).
\end{aligned}$$

Here, δ_k is the maximum $d_{\mathcal{T}}$-distance between two adjacent vertices of $\mathcal{B}(k)$, where by saying $x, y\in\mathcal{B}(k)$ are adjacent, we mean that $\gamma_{\mathcal{T}}(x,y)$ contains no element of $\mathcal{B}(k)$ other than x and y. By the continuity of $\phi_{\mathcal{T}}$ and denseness of $(x_i)_{i\geq 1}$, the upper bound above is equal to 0. Hence, the lemma will follow from

$$(6.14)\qquad \lim_{k\to\infty}\limsup_{n\to\infty} P_{\rho_{T_n}}^{T_n}\Big(a_n\sup_{s\in[0,t]} d_{T_n}(X^{n,k}_{s/a_nb_n}, \hat{X}^{n,k}_{s/a_nb_n}) > \varepsilon\Big) = 0,$$

where we have applied (6.6) to replace X^{T_n} by $X^{n,k}$ in this requirement. Now, by making the change from α_n^{-1} to a_n and from n^{-1} to b_n, one can follow the argument of [21], Lemma 4.3, exactly to deduce that

$$\lim_{k\to\infty}\limsup_{n\to\infty} P_{\rho_{T_n}}^{T_n}\Big(a_nb_n\sup_{s\leq t}\big|A^{n,k}_{s/a_n^2} - \hat{A}^{n,k}_{s/a_n^2}\big| > \varepsilon\Big) = 0.$$

Note that we needed the fact that $b_n = o(a_n)$, and also (6.6), (6.12) and $b_n \mu_{\mathcal{T}_n}(T_n) \to \mu_{\mathcal{T}}(\mathcal{T})$ [which follows from our assumption at (6.2)]. Recalling the characterisations of $X^{n,k}$ and $\hat{X}^{n,k}$ given in (6.7) and (6.8), respectively, we can complete the proof of (6.14) by combining the limit above with the convergence statements of (6.13) and Lemma 6.2 (cf. [21], Proposition 4.1). □

The following measurability result will be useful when we come to look at random walks on random trees. Its proof is similar to that of [16], Lemma 8.1(b).

LEMMA 6.5. *The map $\underline{\mathcal{T}} \mapsto P^{\mathcal{T}}_{\rho_{\mathcal{T}}} \circ \phi_{\mathcal{T}}^{-1}$ defines a measurable function from $\mathbb{T}_c^*$ (equipped with the subspace σ-algebra) to the space of probability measures on $C(\mathbb{R}_+, M)$.*

PROOF. Suppose that $\underline{\mathcal{T}_n} \to \underline{\mathcal{T}}$ in $\mathbb{T}_c^*$. A straightforward adaptation of [46], Proposition 10, then yields that if $(x_i^n)_{i\geq 1}$ is a sequence of $\mu_{\mathcal{T}_n}$-random vertices of $\mathcal{T}_n$ and $(x_i)_{i\geq 1}$ is a sequence of $\mu_{\mathcal{T}}$-random vertices of $\mathcal{T}$, then for each fixed $k \geq 1$,

$$(\mathcal{T}_n, d_{\mathcal{T}_n}, \mu_{\mathcal{T}_n}, \phi_{\mathcal{T}_n}, \rho_{\mathcal{T}_n}, (x_i^n)_{i=1}^k) \to (\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \phi_{\mathcal{T}}, \rho_{\mathcal{T}}, (x_i)_{i=1}^k),$$

in distribution in a version of $(\mathbb{T}_c, \Delta_c)$ where metric spaces are marked by k points [so that the supremum in (3.1) is taken over correspondences that include not only the root pairs $(\rho_{\mathcal{T}}, \rho'_{\mathcal{T}})$, but also the marked pairs (x_i, x'_i), $i = 1, \dots, k$, say]. Since the latter space can be checked to be separable in the same way as was discussed for $(\mathbb{T}_c, \Delta_c)$ in the proof of Proposition 3.1, one can apply a Skorohod representation argument to deduce that there exist realisations of the relevant random variables such that the above convergence occurs almost surely. As a consequence, the proof of Lemma 6.3 can be applied to deduce that for fixed $k \geq 1$, as $n \to \infty$,

$$P^{\mathcal{T}_n(k), \mu^{(k)}_{\mathcal{T}_n}}_{\rho_{\mathcal{T}_n}} \circ \phi_{\mathcal{T}_n}^{-1} \to P^{\mathcal{T}(k), \mu^{(k)}_{\mathcal{T}}}_{\rho_{\mathcal{T}}} \circ \phi_{\mathcal{T}}^{-1}$$

in distribution as probability measures on $C(\mathbb{R}_+, M)$, where $P^{\mathcal{T}(k), \mu^{(k)}_{\mathcal{T}}}_{\rho_{\mathcal{T}}}$ is the law of $X^{\mathcal{T}(k), \mu^{(k)}}$ started from $\rho_{\mathcal{T}}$, and the objects indexed by n are defined analogously to the limiting ones. In particular, this establishes that the map from $\underline{\mathcal{T}}$ to the law of $P^{\mathcal{T}(k), \mu^{(k)}_{\mathcal{T}}}_{\rho_{\mathcal{T}}} \circ \phi_{\mathcal{T}}^{-1}$ is continuous, and therefore measurable, on $\mathbb{T}_c^*$. Moreover, since $\mu_{\mathcal{T}}$ is nonatomic and has full support, then we may assume that $(x_i)_{i\geq 1}$ is almost-surely dense in $\mathcal{T}$, and that all the vertices are distinct (and not equal to the root $\rho_{\mathcal{T}}$). Hence, by Lemma 6.3, it holds that

$$P^{\mathcal{T}(k), \mu^{(k)}_{\mathcal{T}}}_{\rho_{\mathcal{T}}} \circ \phi_{\mathcal{T}}^{-1} \to P^{\mathcal{T}}_{\rho_{\mathcal{T}}} \circ \phi_{\mathcal{T}}^{-1}$$

almost-surely as probability measures on $C(\mathbb{R}_+, M)$. Thus, the map from $\underline{\mathcal{T}}$ to the law of $P^{\mathcal{T}}_{\rho_{\mathcal{T}}} \circ \phi_{\mathcal{T}}^{-1}$ is a limit of measurable functions, and so is also measurable on $\mathbb{T}_c^*$. Since $P_{\rho_{\mathcal{T}}}^{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1}$ is a function of only $\underline{\mathcal{T}}$ [and not the particular sequence $(x_i)_{i\geq 1}$], the result follows from a standard argument. □

Suppose that $\underline{\mathcal{T}}$ is a random element of $\mathbb{T}_c$, built on a probability space with probability measure $\mathbf{P}$, and $\mathbf{P}$-a.s. takes values in $\mathbb{T}_c^*$. The previous lemma tells us that the annealed law of the process $\phi_{\mathcal{T}}(X^{\mathcal{T}})$, where the Brownian motion $X^{\mathcal{T}}$ is started from the root, that is, $\int_{\mathbb{T}_c} P_{\rho_{\mathcal{T}}}^{\mathcal{T}} \circ \phi_{\mathcal{T}}^{-1}(\cdot) d\mathbf{P}$ [cf. (1.6)], is a well-defined probability measure on $C(\mathbb{R}_+, M)$. By a Skorohod representation argument, we also obtain the following as an immediate corollary of Theorem 6.1.

COROLLARY 6.6. *Let $(T_n, d_{T_n}, \mu_{T_n}, \phi_{T_n}, \rho_{T_n})$, $n \geq 1$ be a random sequence, and $(a_n)_{n\geq 1}$, $(b_n)_{n\geq 1}$, $(c_n)_{n\geq 1}$ be null sequences with $b_n = o(a_n)$, such that (6.2) holds in distribution, and the limit $\underline{\mathcal{T}}$ almost-surely takes values in $\mathbb{T}_c^*$. Then the annealed laws of the processes*

$$\left(c_n \phi_{T_n}\left(X_{t/a_n b_n}^{T_n}\right)\right)_{t\geq 0}$$

[*cf.* (1.5)] *converge to the annealed law of $\phi_{\mathcal{T}}(X^{\mathcal{T}})$, where we assume that $X_0^{T_n} = \rho_{T_n}$ for each n, and also $X_0^{\mathcal{T}} = \rho_{\mathcal{T}}$.*

As with Lemmas 3.7 and 3.9, Theorem 6.1 and Corollary 6.6 can be extended to the noncompact case with an additional assumption. To begin with the deterministic case, suppose $(T_n)_{n\geq 1}$ is a deterministic sequence of locally finite graph trees for which (6.2) holds in $(\mathbb{T}, \Delta)$, where $\underline{\mathcal{T}}$ is such that $\mu_{\mathcal{T}}$ is nonatomic, supported on the leaves of $\mathcal{T}$, and satisfies (6.1) when the infimum is taken over $B_{\mathcal{T}}(\rho_{\mathcal{T}}, R)$ for any R; we denote the subset of $\mathbb{T}$ whose elements satisfy these properties by $\mathbb{T}^*$. (Note that for an element of $\mathbb{T}^*$, it is possible to define Brownian motion $X^{\mathcal{T}}$ on $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}})$ by the procedure of [6], as described above the statement of Theorem 6.1.) Moreover, assume that, for $t > 0$,

$$\lim_{R\to\infty} \limsup_{n\to\infty} P_{\rho_{T_n}}^{T_n}\left(\tau\left(X^{T_n}, B_{T_n}\left(\rho_{T_n}, a_n^{-1}R\right)\right) \leq t/a_n b_n\right) = 0,$$

where, here and in the following, $\tau(X, A) := \inf\{t \geq 0 : X_t \notin A\}$ is the exit time of a process X from a set A. It is then the case that (6.3) holds. Similarly to Remarks 3.8 and 3.10, we do not include the full details of this argument, but instead restrict our presentation to describing the probabilistic version of the argument needed to handle the simple random walk on $\mathcal{U}$. In particular, given Theorems 1.1 and 1.3, the key additional ingredient for this is the following, where we recall $\mathbb{P}$ is the annealed law of the simple random walk on $\mathcal{U}$, as introduced at (1.5).

PROPOSITION 6.7. *For $t > 0$,*

$$\lim_{R\to\infty} \limsup_{\delta\to 0} \mathbb{P}\left(\tau\left(X^{\mathcal{U}}, B_{\mathcal{U}}(0, \delta^{-\kappa} R)\right) \leq t\delta^{-\kappa d_w}\right) = 0.$$

PROOF. Given the volume growth and resistance estimates of [11], Proposition 4.2, we can apply an identical argument to that used to prove the corresponding limit in [17], Theorem 1.1. □

We are now ready to prove the main simple random walk convergence result for the UST. In the proof, we let $(\mathbf{P}_{\delta_n})_{n\geq 1}$ be a convergent sequence with limit $\tilde{\mathbf{P}}$, and suppose that $\underline{\mathcal{T}}$ is a random variable with law $\tilde{\mathbf{P}}$. Unless otherwise noted, we take (M, d_M) to be $\mathbb{R}^2$ equipped with the Euclidean distance.

PROOF OF THEOREM 1.4(a), (b). As in the proof of Lemma 5.2, the separability of $(\mathbb{T}, \Delta)$ allows us to find realisations of $\underline{\mathcal{U}}_{\delta_n}$, $n \geq 1$, and $\underline{\mathcal{T}}$ built on a common probability space with probability measure $\mathbf{P}^*$ such that $\underline{\mathcal{U}}_{\delta_n} \to \underline{\mathcal{T}}$ holds in $(\mathbb{T}, \Delta)$, $\mathbf{P}^*$-a.s. This yields the existence of a subsequence $(n_i)_{i\geq 1}$ and divergent sequence $(r_j)_{j\geq 1}$ such that, for every r_j, $\mathbf{P}^*$-a.s., we have convergence in $(\mathbb{T}_c, \Delta_c)$ of the radius r_j restrictions along the subsequence $(n_i)_{i\geq 1}$, that is, $\underline{\mathcal{U}}_{\delta_{n_i}}^{(r_j)}$ converges to $\underline{\mathcal{T}}^{(r_j)}$. Moreover, since the Lebesgue measure of those $r \geq 0$ for which $\mu_{\mathcal{T}}(\partial \mathcal{T}^{(r)}) > 0$ is zero, $\tilde{\mathbf{P}}$-a.s., we may further assume that $\mu_{\mathcal{T}}(\partial \mathcal{T}^{(r_j)}) = 0$ for every r_j, $\tilde{\mathbf{P}}$-a.s. Noting that the map $\underline{\mathcal{T}} \mapsto \underline{\mathcal{T}}^{(r)}$ is continuous at those elements of $\mathbb{T}$ satisfying $\mu_{\mathcal{T}}(\partial \mathcal{T}^{(r)}) = 0$ [cf. the proof of Proposition 3.4, and equation (3.5) in particular], this final condition ensures that, for every r_j, $\underline{\mathcal{T}}^{(r_j)}$ is $\underline{\mathcal{T}}$-measurable. Now, from Theorem 1.3(b)(ii), it is the case that $\underline{\mathcal{T}}^{(r)}$ takes values in $\mathbb{T}_c^*$, $\tilde{\mathbf{P}}$-a.s. [Observe that if $r' \leq r/2$ and $x \in \mathcal{T}^{(r)}$, then there exists an $x' \in \mathcal{T}^{(r)}$ such that $B_{\mathcal{T}^{(r)}}(x, r') \supseteq B_{\mathcal{T}}(x', r'/2)$, and so there is no problem in deducing the lower volume estimate of (6.1) for the ball $\mathcal{T}^{(r)}$ from the lower volume estimate from $\mathcal{T}$. The further two properties—that $\mu_{\mathcal{T}}^{(r)}$ is nonatomic and supported on the leaves of $\mathcal{T}^{(r)}$—are immediate from the full $\mathcal{T}$ statements.] Hence, by Corollary 6.6 (with $a_{n_i} = \delta_{n_i}^{\kappa}$, $b_{n_i} = \delta_{n_i}^2$, $c_{n_i} = \delta_{n_i}$), it follows that, for every r_j, the annealed law of

$$\left(\delta_{n_i} X^{B_{\mathcal{U}}(0, \delta_{n_i}^{-\kappa} r_j)}_{\delta_{n_i}^{-\kappa d_w} t}\right)_{t\geq 0},$$

where $X^{B_{\mathcal{U}}(0, \delta_{n_i}^{-\kappa} r_j)}$ is the simple random walk on $B_{\mathcal{U}}(0, \delta_{n_i}^{-\kappa} r_j)$, converges as $i \to \infty$ to the annealed law of $(\phi_{\mathcal{T}}^{(r_j)}(X_t^{\mathcal{T}^{(r_j)}}))_{t\geq 0}$, where $X^{\mathcal{T}^{(r_j)}}$ is Brownian motion on the measured real tree $(\mathcal{T}^{(r_j)}, d_{\mathcal{T}}^{(r_j)}, \mu_{\mathcal{T}}^{(r_j)})$, and we assume processes are started from the roots of the relevant trees. Given the measurability of $\underline{\mathcal{T}}^{(r_j)}$ described above, this limit law can be expressed as

$$\int P_{\rho_{\mathcal{T}}}^{\mathcal{T}^{(r_j)}} \circ \phi_{\mathcal{T}}^{-1}(\cdot)\, d\tilde{\mathbf{P}}. \tag{6.15}$$

Furthermore, since the limit does not depend on the subsequence, we obtain annealed distributional convergence of the rescaled discrete processes along the full sequence $(\delta_n)_{n\geq 1}$.

Next, suppose $X^{\mathcal{U}}$ and $X^{B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j)}$ are coupled so that their sample paths agree up to $\tau(X^{\mathcal{U}}, B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j))$ [e.g., by taking $X^{B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j)}$ to be $X^{\mathcal{U}}$ observed on $B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j)$]. It then holds that, for $t<\infty$ and $\varepsilon>0$,

$$\mathbb{P}\Big(\sup_{s\in[0,t]}\delta_n\big|X^{\mathcal{U}}_{\delta_n^{-\kappa d_w}s}-X^{B_{\mathcal{U}}(0,\delta_n^{-\kappa}r_j)}_{\delta_n^{-\kappa d_w}s}\big|>\varepsilon\Big)\leq\mathbb{P}\big(\tau\big(X^{\mathcal{U}},B_{\mathcal{U}}\big(0,\delta_n^{-\kappa}r_j\big)\big)\leq t\delta^{-\kappa d_w}\big).$$

Hence, we deduce from Proposition 6.7 that the left-hand side converges to 0 as $n\to\infty$ and then $j\to\infty$.

Finally, we need to take care of the situation when $r_j\to\infty$ for the continuous trees. We have already established that $P^{\mathcal{T}^{(r_j)}}_{\rho_{\mathcal{T}}}\circ\phi_{\mathcal{T}}^{-1}$ is $\underline{\mathcal{T}}$-measurable for every r_j. To show that this is also the case for $P^{\mathcal{T}}_{\rho_{\mathcal{T}}}\circ\phi_{\mathcal{T}}^{-1}$, it will suffice to check that $P^{\mathcal{T}^{(r)}}_{\rho_{\mathcal{T}}}\to P^{\mathcal{T}}_{\rho_{\mathcal{T}}}$ as $r\to\infty$, $\tilde{\mathbf{P}}$-a.s. This will follow if we can check that

$$P^{\mathcal{T}}_{\rho_{\mathcal{T}}}\Big(\lim_{r\to\infty}\tau\big(X^{\mathcal{T}},B_{\mathcal{T}}(\rho_{\mathcal{T}},r)\big)=\infty\Big)=1, \tag{6.16}$$

for $\tilde{\mathbf{P}}$-a.e. realisation of $\mathcal{T}$. To this end, first note that

$$\begin{aligned}&P^{\mathcal{T}}_{\rho_{\mathcal{T}}}\Big(\lim_{r\to\infty}\tau\big(X^{\mathcal{T}},B_{\mathcal{T}}(\rho_{\mathcal{T}},r)\big)<\infty\Big)\\&\quad=\lim_{t\to\infty}\lim_{r\to\infty}\lim_{j\to\infty}P^{\mathcal{T}^{(r_j)}}_{\rho_{\mathcal{T}}}\big(\tau\big(X^{\mathcal{T}^{(r_j)}},B_{\mathcal{T}}(\rho_{\mathcal{T}},r)\big)\leq t\big),\end{aligned} \tag{6.17}$$

where we note that the laws of $X^{\mathcal{T}}$ and $X^{\mathcal{T}^{(r_j)}}$ agree up to the exit time of $B_{\mathcal{T}}(\rho_{\mathcal{T}},r)$ whenever $r_j>r$. Now, suppose that $\underline{\mathcal{U}}_{\delta_n}$, $n\geq1$, and $\underline{\mathcal{T}}$ are coupled as in the first part of the proof. It is not difficult to check from the definition of Δ_c that the convergence in $(\mathbb{T}_c,\Delta_c)$ of the radius r_j restrictions still holds $\mathbf{P}^*$-a.s. if $\delta_n\phi_{\mathcal{U}}$ is replaced by $\delta_n^{\kappa}\tilde{\phi}_{\mathcal{U}}$ and $\phi_{\mathcal{T}}$ is replaced by $\tilde{\phi}_{\mathcal{T}}$, where $\tilde{\phi}_{\mathcal{U}}(x):=(d_{\mathcal{U}}(0,x),0)$ and $\tilde{\phi}_{\mathcal{T}}(x):=(d_{\mathcal{T}}(0,x),0)$, respectively. Thus, an application of Theorem 6.1 (with $M=\mathbb{R}$) and a subsequence argument yields that, for every r_j, $\mathbf{P}^*$-a.s., $(\delta_n^{\kappa}d_{\mathcal{U}}(0,X^{B_{\mathcal{U}}(0,\delta_n^{-\kappa}r_j)}_{\delta_n^{-\kappa d_w}t}))_{t\geq0}$ converges to $(d_{\mathcal{T}}(\rho_{\mathcal{T}},X_t^{\mathcal{T}^{(r_j)}}))_{t\geq0}$ in distribution in $C(\mathbb{R}_+,\mathbb{R})$ as $n\to\infty$. Consequently,

$$\begin{aligned}&\int P^{\mathcal{T}^{(r_j)}}_{\rho_{\mathcal{T}}}\big(\tau\big(X^{\mathcal{T}^{(r_j)}},B_{\mathcal{T}}(\rho_{\mathcal{T}},r)\big)\leq t\big)\,d\tilde{\mathbf{P}}\\&\quad=\int P^{\mathcal{T}^{(r_j)}}_{\rho_{\mathcal{T}}}\Big(\sup_{s\in[0,t]}d_{\mathcal{T}}\big(\rho_{\mathcal{T}},X_s^{\mathcal{T}^{(r_j)}}\big)\geq r\Big)\,d\tilde{\mathbf{P}}\\&\quad\leq\int\liminf_{n\to\infty}P_0^{B_{\mathcal{U}}(0,\delta_{n_i}^{-\kappa}r_j)}\Big(\sup_{s\in[0,t]}d_{\mathcal{U}}\big(0,X^{B_{\mathcal{U}}(0,\delta_n^{-\kappa}r_j)}_{\delta_n^{-\kappa d_w}s}\big)\geq\delta_n^{-\kappa}r/2\Big)\,d\mathbf{P}\\&\quad\leq\liminf_{n\to\infty}\mathbb{P}\big(\tau\big(X^{B_{\mathcal{U}}(0,\delta_n^{-\kappa}r_j)},B_{\mathcal{U}}\big(0,\delta_n^{-\kappa}r\big)\big)\leq t\delta_n^{-\kappa d_w}\big),\end{aligned}$$

where we note that the necessary measurability of the law of $d_{\mathcal{T}}(\rho_{\mathcal{T}}, X^{\mathcal{T}^{(r_j)}})$ can be checked similarly to the measurability of $P^{\mathcal{T}^{(r_j)}}_{\rho_{\mathcal{T}}} \circ \phi_{\mathcal{T}}^{-1}$. Taking limits as $j \to \infty$, $r \to \infty$ and then $t \to \infty$, we obtain from another application of Proposition 6.7 that the upper bound above converges to zero. Thus, the dominated convergence theorem yields that the $\tilde{\mathbf{P}}$-expectation of the left-hand side of (6.17) is equal to zero, and so we have established (6.16), as desired. In summary, we have now shown that $P^{\mathcal{T}}_{\rho_{\mathcal{T}}} \circ \phi_{\mathcal{T}}^{-1}$ is $\underline{\mathcal{T}}$-measurable, and so $\tilde{\mathbb{P}}$—the annealed law of $\phi_{\mathcal{T}}(X^{\mathcal{T}})$, as introduced at (1.6)—is well-defined. This establishes part (a) of the theorem. Moreover, since $P^{\mathcal{T}^{(r)}}_{\rho_{\mathcal{T}}} \to P^{\mathcal{T}}_{\rho_{\mathcal{T}}}$, $\tilde{\mathbf{P}}$-a.s., the continuous mapping theorem and the dominated convergence theorem yield that the annealed law at (6.15) converges as $j \to \infty$ to $\tilde{\mathbb{P}}$. Together with the conclusions of the previous two paragraphs, this completes the proof of part (b) (see [13], Theorem 3.2, e.g.). □

REMARK 6.8. In [20], Theorem 8.1, it was shown that the convergence of rescaled graph tree "tours" (i.e., functions encoding trees and embeddings into $\mathbb{R}^d$) implies convergence of spatial trees. It only requires a simple extension of the proof of that result to add the measure, and thereby deduce that convergence of tours also implies convergence in the topology of $(\mathbb{T}_c, \Delta_c)$ (with $\mathbb{R}^2$ replaced by $\mathbb{R}^d$). Consequently, (the $\mathbb{R}^d$ version of) Theorem 6.1 provides an extension of the random walk convergence result of [20], Theorem 8.1, and can also be used to deduce a scaling limit for the simple random walks on critical branching random walks satisfying the various assumptions on the offspring and step distribution detailed in [20], Section 10.

REMARK 6.9. In [21], Theorem 1.1 (see also [16], Theorem 1.1), a scaling limit was established for simple random walks on ordered graph trees. The proofs of these results used the convenience of encoding ordered graph trees by functions to simplify various technical details. Theorem 6.1 provides the additional framework needed to handle unordered graph trees. Indeed, suppose that we have a sequence of finite rooted graph trees such that

$$(T_n, a_n d_{T_n}, b_n \mu_{T_n}, \rho_{T_n}) \to (\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \rho_{\mathcal{T}}), \tag{6.18}$$

for some null sequences $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ with $b_n = o(a_n)$ in the Gromov–Hausdorff–Prohorov topology of [1] for rooted, compact metric spaces equipped with Borel measures. Moreover, assume that the limit satisfies the additional properties on the measure that hold for elements of $\mathbb{T}^*_c$. It is then the case that, by applying the procedure described in the proof of Lemma 6.3, one can find isometric embeddings $\tilde{\psi} : \mathcal{T} \to \ell^1$ and $\tilde{\psi}_n : T_n \to \ell^1$, $n \geq 1$, such that

$$(a_n \tilde{\psi}_n(X^{T_n}_{t/a_n b_n}))_{t\geq 0} \to (\tilde{\psi}(X^{\mathcal{T}}_t))_{t\geq 0},$$

in distribution in $C(\mathbb{R}_+, \ell^1)$, where X^{T_n} is the simple random walk on T_n started from ρ_{T_n}, and $X^{\mathcal{T}}$ is the Brownian motion on $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}})$ started from $\rho_{\mathcal{T}}$. We note that a similar, but slightly stronger result was recently established independently in [7]—the most important difference being that the argument of the latter work required a weaker lower volume growth condition.

The above result can also be extended to the random case by embedding trees into ℓ^1 in a canonical random way—specifically, by choosing the vertices $(x_i)_{i\geq 1} \in \mathcal{T}$ to be a $\mu_{\mathcal{T}}$-random sequence. The only additional complications come from some measurability issues, but these can be resolved using similar ideas to those applied in the proof of Lemma 6.5. To summarise the conclusion, suppose that $(T_n)_{n\geq 1}$ is a sequence of finite rooted (unordered) graph trees for which (6.18) holds in distribution for some random measured compact real tree $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \rho_{\mathcal{T}})$. Moreover, assume that $\mu_{\mathcal{T}}$ is a nonatomic probability measure, supported on the leaves of $\mathcal{T}$ and satisfies (6.1), almost-surely. It can then be checked that the annealed laws of $(a_n \tilde{\psi}_n(X^{T_n}_{t/a_n b_n}))_{t\geq 0}$, where $\tilde{\psi}_n$ is the canonical random isometric embedding of T_n into ℓ^1, converge to the annealed law of Brownian motion on $\mathcal{T}$ randomly isometrically embedded into ℓ^1. For example, taking $a_n = n^{-1/2}$, $b_n = n^{-1}$, this result applies to the model of uniformly random unordered trees with n vertices, in which each vertex has at most $m \in \{2, 3, \dots, \infty\}$ children, as studied in [28]. Moreover, applying the argument of [22], Section 5.2, (where critical Galton–Watson trees were considered), a corollary of this is that the mixing times of the simple random walks on the random discrete trees, when rescaled by $n^{3/2}$, converge in distribution to (a constant multiple of) the mixing time of the limiting diffusion.

7. Heat kernel bounds for the scaling limit. As in Section 5 and the proof of Theorem 1.4(a), (b), let $(\mathbf{P}_{\delta_n})_{n\geq 1}$ be a convergent sequence with limit $\tilde{\mathbf{P}}$, and suppose that $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}}, \phi_{\mathcal{T}}, \rho_{\mathcal{T}})$ is a random variable with law $\tilde{\mathbf{P}}$. It follows from [6], Remark 3.1 and [27], Theorem 1.5.2, that the Dirichlet form $(\mathcal{E}_{\mathcal{T}}, \mathcal{F}_{\mathcal{T}})$ given in Section 6 is the same as that of [32], Section 5. In particular, this is the form associated with the natural "resistance form" on $(\mathcal{T}, d_{\mathcal{T}})$, and so we can apply [33], Theorem 10.4, to deduce the existence of a jointly continuous transition density $(p_t^{\mathcal{T}}(x, y))_{x,y\in\mathcal{T},t>0}$ for the process $X^{\mathcal{T}}$.

Let $R_{\mathcal{T}}$ be the resistance associated with $(\mathcal{E}_{\mathcal{T}}, \mathcal{F}_{\mathcal{T}})$, defined by setting, for two disjoint subsets $A, B \subseteq \mathcal{T}$,

$$R_{\mathcal{T}}(A, B)^{-1} := \inf\{\mathcal{E}_{\mathcal{T}}(f, f) : f \in \mathcal{F}_{\mathcal{T}}, f|_A = 0, f|_B = 1\}.$$

Since $\mathcal{T}$ is complete and, by Theorem 1.3(a)(ii), has a single end at infinity, we deduce from [6], Theorem 4, that $X^{\mathcal{T}}$ is recurrent. As a consequence, combining [27], Theorem 1.6.3 and [6], Proposition 3.5, yields that the resistance between two points corresponds to (a multiple of) the distance between them, that is, $R_{\mathcal{T}}(\{x\}, \{y\}) = 2d_{\mathcal{T}}(x, y)$ for all $x, y \in \mathcal{T}$, $x \neq y$ (see [6], e.g.). Hence, we can use [18], Theorem 3, to obtain estimates for $p_t^{\mathcal{T}}(x, y)$ from the volume estimates of Theorem 1.3(b)(ii).

PROOF OF THEOREM 1.4(c). We have already discussed the claims about the existence and joint continuity of the heat kernel, and the recurrence of $X^{\mathcal{T}}$. So, we will simply present here a few key points that are needed to apply the heat kernel estimates of [18], Theorem 3. As already noted, the resistance metric coincides with (a multiple of) the tree metric $d_{\mathcal{T}}$ in our setting, and so we can replace $R_{\mathcal{T}}$ in the conclusion of [18] by $d_{\mathcal{T}}$. Moreover, the fact that $(\mathcal{T}, d_{\mathcal{T}})$ is a real tree automatically means the "chaining condition" of [18] is satisfied, that is, there exists a constant c_1 such that for all $x, y \in \mathcal{T}$ and all $n \in \mathbb{N}$, there exist $x_0 = x, x_1, \dots, x_N = y$ such that $d_{\mathcal{T}}(x_{i-1}, x_i) \le c_1 d_{\mathcal{T}}(x, y)/n, \forall i = 1, \dots, n$. (Clearly, we can take $c_1 = 1$ and equality in the latter statement.) Finally, note that in [18] volume estimates were assumed to hold uniformly over the entire space, but Theorem 1.3(b)(ii) only gives uniformity over balls of finite radius. However, it is straightforward to check that the arguments of [18] are enough to give the stated heat kernel estimates. □

For the remaining heat kernel estimates, we derive the following tail bound for the resistance from the root to the radius of a ball will be useful.

LEMMA 7.1. *There exist constants* $c_1, c_2, \theta \in (0, \infty)$ *such that for all* $r > 0$, $\lambda \ge 1$,

$$\tilde{\mathbf{P}}\big(R_{\mathcal{T}}\big(\rho_{\mathcal{T}}, B_{\mathcal{T}}(\rho_{\mathcal{T}}, r)^c\big) \le \lambda^{-1} r\big) \le c_1 e^{-c_2 \lambda^{\theta}}. \tag{7.1}$$

PROOF. As we have done several times previously, we will apply a coupling argument, and start by supposing that we have a realisation of random variables such that $\underline{\mathcal{U}}_{\delta_n} \to \underline{\mathcal{T}}$ holds almost-surely along the sequence $(\delta_n)_{n\ge1}$. Let $N_{\mathcal{T}}(r, r/\lambda)$ be the minimum number of $d_{\mathcal{T}}$-balls of radius r/λ required to cover $B_{\mathcal{T}}(\rho_{\mathcal{T}}, r)$. From the definition of $(\mathbb{T}, \Delta)$, it is elementary to check that if $N_{\mathcal{T}}(r, r/\lambda) \ge N_0$, then so is $N_{\mathcal{U}}(4\delta_n^{-\kappa} r, \delta_n^{-\kappa} r/8\lambda)$ for large n. It follows that

$$\tilde{\mathbf{P}}\big(N_{\mathcal{T}}(r, r/\lambda) \ge c\lambda^5\big) \le \limsup_{n\to\infty} \mathbf{P}\big(N_{\mathcal{U}}\big(4\delta_n^{-\kappa} r, \delta_n^{-\kappa} r/8\lambda\big) \ge c\lambda^5\big), \tag{7.2}$$

which by Remark 2.13 is, for a suitable choice of c, bounded above by $c_1 e^{-c_2 \lambda^{1/80}}$. Now, the proof of [35], Lemma 4.1, gives that $R_{\mathcal{T}}(\rho_{\mathcal{T}}, B_{\mathcal{T}}(\rho_{\mathcal{T}}, r)^c) \ge r/8N_{\mathcal{T}}(r, r/4)$, and so, for any $\lambda > 4^5 c$,

$$\tilde{\mathbf{P}}\big(R_{\mathcal{T}}\big(\rho_{\mathcal{T}}, B_{\mathcal{T}}(\rho_{\mathcal{T}}, r)^c\big) \le r/\lambda\big) \le \tilde{\mathbf{P}}\big(N_{\mathcal{T}}\big(r, c^{1/5} r/\lambda^{1/5}\big) \ge 8\lambda\big).$$

Combining this bound with (7.2) yields the result. □

Given (5.16), (5.17) and (7.1), the next two results can be proved in exactly the same way as the corresponding parts of [19], Theorem 1.6 and [19], Proposition 1.7, modulo a different volume growth exponent.

THEOREM 7.2. (a) $\tilde{\mathbf{P}}$-*a.s., there exists a random* $t_0(\mathcal{T}) \in (0, \infty)$ *and deterministic constants* $c_1, c_2, \theta_1, \theta_2 \in (0, \infty)$ *such that*

$$c_1 t^{-8/13}(\log\log t^{-1})^{-\theta_1} \le p_t^{\mathcal{T}}(\rho_{\mathcal{T}}, \rho_{\mathcal{T}}) \le c_2 t^{-8/13}(\log\log t^{-1})^{\theta_2},$$

for all $t \in (0, t_0)$.

(b) *There exist constants* $c_1, c_2 \in (0, \infty)$ *such that*

$$c_1 t^{-8/13} \le \tilde{\mathbf{E}} p_t^{\mathcal{T}}(\rho_{\mathcal{T}}, \rho_{\mathcal{T}}) \le c_2 t^{-8/13},$$

for all $t \in (0, 1)$.

Acknowledgement. The authors thank Sunil Chhita, who wrote the code used to produce Figure 1.

REFERENCES

[1] ABRAHAM, R., DELMAS, J.-F. and HOSCHEIT, P. (2013). A note on the Gromov–Hausdorff–Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.* **18** 1–21. MR3035742

[2] AIZENMAN, M., BURCHARD, A., NEWMAN, C. M. and WILSON, D. B. (1999). Scaling limits for minimal and random spanning trees in two dimensions. *Random Structures Algorithms* **15** 319–367. MR1716768

[3] ALBERTS, T., KOZDRON, M. J. and MASSON, R. (2013). Some partial results on the convergence of loop-erased random walk to SLE(2) in the natural parametrization. *J. Stat. Phys.* **153** 119–141. MR3100817

[4] ALDOUS, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226

[5] ALDOUS, D. J. (1990). The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.* **3** 450–465. MR1069105

[6] ATHREYA, S., ECKHOFF, M. and WINTER, A. (2013). Brownian motion on $\mathbb{R}$-trees. *Trans. Amer. Math. Soc.* **365** 3115–3150. MR3034461

[7] ATHREYA, S., LÖHR, W. and WINTER, A. (2014). Invariance principle for variable speed random walks on trees. Preprint. Available at arXiv:1404.6290.

[8] BARLOW, M. T. (2016). Loop erased walks and uniform spanning trees. *MSJ Memoirs* **34** 1–32. MR3525847

[9] BARLOW, M. T., JÁRAI, A. A., KUMAGAI, T. and SLADE, G. (2008). Random walk on the incipient infinite cluster for oriented percolation in high dimensions. *Comm. Math. Phys.* **278** 385–431. MR2372764

[10] BARLOW, M. T. and MASSON, R. (2010). Exponential tail bounds for loop-erased random walk in two dimensions. *Ann. Probab.* **38** 2379–2417. MR2683633

[11] BARLOW, M. T. and MASSON, R. (2011). Spectral dimension and random walks on the two dimensional uniform spanning tree. *Comm. Math. Phys.* **305** 23–57. MR2802298

[12] BENJAMINI, I., LYONS, R., PERES, Y. and SCHRAMM, O. (2001). Uniform spanning forests. *Ann. Probab.* **29** 1–65. MR1825141

[13] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York. MR1700749

[14] BRODER, A. (1989). Generating random spanning trees. In *Proceedings of the* 30*th Annual Symposium on Foundations of Computer Science* 442–447. IEEE Computer Society, Washington, DC.

[15] BURAGO, D., BURAGO, Y. and IVANOV, S. (2001). *A Course in Metric Geometry*. *Graduate Studies in Mathematics* **33**. Amer. Math. Soc., Providence, RI. MR1835418

[16] CROYDON, D. (2008). Convergence of simple random walks on random discrete trees to Brownian motion on the continuum random tree. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 987–1019. MR2469332

[17] CROYDON, D. and KUMAGAI, T. (2008). Random walks on Galton–Watson trees with infinite variance offspring distribution conditioned to survive. *Electron. J. Probab.* **13** 1419–1441. MR2438812

[18] CROYDON, D. A. (2007). Heat kernel fluctuations for a resistance form with non-uniform volume growth. *Proc. Lond. Math. Soc.* (3) **94** 672–694. MR2325316

[19] CROYDON, D. A. (2008). Volume growth and heat kernel estimates for the continuum random tree. *Probab. Theory Related Fields* **140** 207–238. MR2357676

[20] CROYDON, D. A. (2009). Hausdorff measure of arcs and Brownian motion on Brownian spatial trees. *Ann. Probab.* **37** 946–978. MR2537546

[21] CROYDON, D. A. (2010). Scaling limits for simple random walks on random ordered graph trees. *Adv. in Appl. Probab.* **42** 528–558. MR2675115

[22] CROYDON, D. A., HAMBLY, B. M. and KUMAGAI, T. (2012). Convergence of mixing times for sequences of random walks on finite graphs. *Electron. J. Probab.* **17** 1–32. MR2869250

[23] DUQUESNE, T. and LE GALL, J.-F. (2005). Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* **131** 553–603. MR2147221

[24] EDGAR, G. A. (1998). *Integral, Probability, and Fractal Measures*. Springer, New York. MR1484412

[25] EVANS, S. N. (2008). *Probability and Real Trees*. *Lecture Notes in Math.* **1920**. Springer, Berlin. MR2351587

[26] EVANS, S. N., PITMAN, J. and WINTER, A. (2006). Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields* **134** 81–126. MR2221786

[27] FUKUSHIMA, M., OSHIMA, Y. and TAKEDA, M. (1994). *Dirichlet Forms and Symmetric Markov Processes*. *De Gruyter Studies in Mathematics* **19**. de Gruyter, Berlin. MR1303354

[28] HAAS, B. and MIERMONT, G. (2012). Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. *Ann. Probab.* **40** 2589–2666. MR3050512

[29] HÄGGSTRÖM, O. (1995). Random-cluster measures and uniform spanning trees. *Stochastic Process. Appl.* **59** 267–275. MR1357655

[30] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York. MR1876169

[31] KENYON, R. (2000). The asymptotic determinant of the discrete Laplacian. *Acta Math.* **185** 239–286. MR1819995

[32] KIGAMI, J. (1995). Harmonic calculus on limits of networks and its application to dendrites. *J. Funct. Anal.* **128** 48–86. MR1317710

[33] KIGAMI, J. (2012). Resistance forms, quasisymmetric maps and heat kernel estimates. *Mem. Amer. Math. Soc.* **216** vi+132. MR2919892

[34] KIRCHHOFF, G. (1847). Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Ann. Phys.* **148** 497–508.

[35] KUMAGAI, T. (2004). Heat kernel estimates and parabolic Harnack inequalities on graphs and resistance forms. *Publ. Res. Inst. Math. Sci.* **40** 793–818. MR2074701

[36] KUMAGAI, T. (2014). *Random Walks on Disordered Media and Their Scaling Limits*. *Lecture Notes in Math.* **2101**. Springer, Cham. MR3156983

[37] KUMAGAI, T. and MISUMI, J. (2008). Heat kernel estimates for strongly recurrent random walk on random media. *J. Theoret. Probab.* **21** 910–935. MR2443641

[38] LAWLER, G. F. (1999). Loop-erased random walk. In *Perplexing Problems in Probability* (M. Bramson and R. Durrett, eds.). *Progress in Probability* **44** 197–217. Birkhäuser, Boston, MA. MR1703133

[39] LAWLER, G. F. (2013). *Intersections of Random Walks*. Birkhäuser/Springer, New York. MR2985195

[40] LAWLER, G. F. (2014). The probability that planar loop-erased random walk uses a given edge. *Electron. Commun. Probab.* **19** 1–13. MR3246970

[41] LAWLER, G. F., SCHRAMM, O. and WERNER, W. (2004). Conformal invariance of planar loop-erased random walks and uniform spanning trees. *Ann. Probab.* **32** 939–995. MR2044671

[42] LAWLER, G. F. and SHEFFIELD, S. (2011). A natural parametrization for the Schramm–Loewner evolution. *Ann. Probab.* **39** 1896–1937. MR2884877

[43] LAWLER, G. F. and ZHOU, W. (2013). *SLE* curves and natural parametrization. *Ann. Probab.* **41** 1556–1584. MR3098684

[44] LE GALL, J.-F. (2006). Random real trees. *Ann. Fac. Sci. Toulouse Math.* (6) **15** 35–62. MR2225746

[45] MASSON, R. (2009). The growth exponent for planar loop-erased random walk. *Electron. J. Probab.* **14** 1012–1073. MR2506124

[46] MIERMONT, G. (2009). Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér.* (4) **42** 725–781. MR2571957

[47] PEMANTLE, R. (1991). Choosing a spanning tree for the integer lattice uniformly. *Ann. Probab.* **19** 1559–1574. MR1127715

[48] SCHRAMM, O. (2000). Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.* **118** 221–288. MR1776084

[49] WILSON, D. B. (1996). Generating random spanning trees more quickly than the cover time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing* (*Philadelphia*, *PA*, 1996) 296–303. ACM, New York. MR1427525

M. T. BARLOW
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, B.C., V6T 1Z2
CANADA
E-MAIL: barlow@math.ubc.ca

D. A. CROYDON
DEPARTMENT OF STATISTICS
UNIVERSITY OF WARWICK
COVENTRY, CV4 7AL
UNITED KINGDOM
E-MAIL: d.a.croydon@warwick.ac.uk

T. KUMAGAI
RESEARCH INSTITUTE FOR MATHEMATICAL SCIENCES
KYOTO UNIVERSITY
KYOTO 606-8502
JAPAN
E-MAIL: kumagai@kurims.kyoto-u.ac.jp