Convergence of mixing times for sequences of random walks on finite graphs

D.A. Croydon*, B.M. Hambly† and T. Kumagai‡

March 4, 2011

Abstract

We establish conditions on sequences of graphs which ensure that the mixing times of the random walks on the graphs in the sequence converge. The main assumption is that the graphs, associated measures and heat kernels converge in a suitable Gromov-Hausdorff sense. With this result we are able to establish the convergence of the mixing times on the largest component of the Erdős-Rényi random graph in the critical window, sharpening previous results for this random graph model. Our results also enable us to establish convergence in a number of other examples, such as finitely ramified fractal graphs, Galton-Watson trees and the range of a high-dimensional random walk.

1 Introduction

The geometric and analytic properties of random graphs have been the subject of much recent research. One strand of this development has been to examine sequences of random subgraphs of vertex transitive graphs that are, in some sense, at or near criticality. A key example is the percolation model and, for bond percolation above the upper critical dimension, we expect to see mean field behaviour in the sequence of finite graphs in the critical window. That is the natural scaling exponents for the volume and diameter of the graph and for the mixing time are of the same order as those for the Erdős-Rényi random graph in the critical window, as given in [35]. This mean field behaviour is seen in other natural models of sequences of critical random graphs. For example [6] obtained general conditions for this behaviour and showed they hold for examples such as the n-cube, while the high dimensional torus is treated in [23]. Motivated by these results we will focus on the asymptotic behaviour of mixing times for random walks on sequences of finite graphs. We consider general sequences of graphs but under some strong conditions which will enable us to establish the convergence of the mixing time.

In order to demonstrate our main result we consider the Erdős-Rényi random graph. Let $G(N,p)$ be the random subgraph of the complete graph on N labelled vertices $\{1,\dots,N\}$ in which each edge is present with probability p independently of the other edges. It is a classical result that if we set $p=c/N$, then as $N\to\infty$, if $c>1$ there is a giant component containing a positive fraction of the vertices, while for $c<1$ the largest component is of size $\log N$. However, if $p=N^{-1}+\lambda N^{-4/3}$ for some $\lambda\in\mathbb{R}$, we are in the so-called critical window, and it is known that the largest connected component $\mathcal{C}^N$, is of order $N^{2/3}$. The recent work of [1] has shown that the scaling limit of the graph, $\mathcal{M}$, exists and can be constructed from the continuum random tree.

*Dept of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom; d.a.croydon@warwick.ac.uk.
†Mathematical Institute, 24-29 St Giles', Oxford, OX1 3LB, United Kingdom; hambly@maths.ox.ac.uk.
‡RIMS, Kyoto University, Kyoto 606-8502, Japan; kumagai@kurims.kyoto-u.ac.jp.

For the Erdős-Rényi random graph above criticality, [19] and [4] established mixing time bounds for the simple random walk on the giant component. For the random graph in the critical window, the following result on the mixing time $t^1_{\mathrm{mix}}(\mathcal{C}^N)$ (a precise definition will be given later in (1.8), see also Remark 1.3) was obtained by Nachmias and Peres ([35, Theorem 1.1]).

Theorem 1.1. *Let $\mathcal{C}^N$ be the largest connected component of $G(N,(1+\lambda N^{-1/3})/N)$ for some $\lambda \in \mathbb{R}$. Then, for any $\epsilon > 0$, there exists $A = A(\epsilon, \lambda) < \infty$ such that for all large N,*

$$P(t^1_{\mathrm{mix}}(\mathcal{C}^N) \notin [A^{-1}N, AN]) < \epsilon.$$

It is natural to ask for more refined results on the behaviour of the family of mixing times. The purpose of this paper is to give a general criteria for the convergence of mixing times for a sequence of random walks on finite graphs in the setting where the graphs can be embedded nicely in a compact metric space. Due to the recent work of [1] and [9] we can apply our main result to the case of the Erdős-Rényi random graph, to obtain the following result.

Theorem 1.2. *Fix $p \in [1,\infty]$. If $t^p_{\mathrm{mix}}(\rho^N)$ is the L^p-mixing time of the random walk on $\mathcal{C}^N$ started from its root ρ^N, then*

$$N^{-1}t^p_{\mathrm{mix}}(\rho^N) \to t^p_{\mathrm{mix}}(\rho),$$

in distribution, where $t^p_{\mathrm{mix}}(\rho) \in (0,\infty)$ is the L^p-mixing time of the Brownian motion on $\mathcal{M}$ started from ρ.

We will later illustrate our main result with a number of other examples of random walks on sequences of finite graphs. In order to state it, though, we start by describing the general framework in which we work. Firstly, let (F, d_F) be a compact metric space and let π be a non-atomic Borel probability measure on F with full support. We will assume that balls $B_F(x,r) := \{y \in F : d_F(x,y) < r\}$ are π-continuity sets (i.e. $\pi(\partial B_F(x,r)) = 0$ for every $x \in F$, $r > 0$). Secondly, take $X^F = (X^F_t)_{t\geq 0}$ to be a π-symmetric Hunt process on F. We suppose the following:

- X^F is conservative, (1.1)
- there exists a jointly continuous transition density $(q_t(x,y))_{x,y\in F,t>0}$ of X^F, (1.2)
- for every $x,y \in F$ and $t > 0$, $q_t(x,y) > 0$, (1.3)
- for every $x \in F$ and $t > 0$, $q_t(x,\cdot)$ is not identically equal to 1, (1.4)

where conditions (1.3) and (1.4) are assumed to exclude various trivial cases, and by transition density we mean the kernel $q_t(x,y)$ such that

$$\mathbf{E}_x[f(X^F_t)] = \int_F q_t(x,y)f(y)\pi(dy),$$

for all bounded continuous function f on F. Furthermore, we will say that the transition density $(q_t(x,y))_{x,y\in F,t>0}$ converges to stationarity in an L^p sense for some $p \in [1,\infty]$ if it holds that

$$\lim_{t\to\infty} D_p(x,t) = 0, \tag{1.5}$$

for every $x \in F$, where $D_p(x,t) := \|q_t(x,\cdot) - 1\|_{L^p(\pi)}$. If this previous condition is satisfied, then it is possible to check that the L^p-mixing time of F,

$$t^p_{\mathrm{mix}}(F) := \inf\left\{t > 0 : \sup_{x\in F} D_p(x,t) \leq 1/4\right\}, \tag{1.6}$$

is a finite quantity (see Section 3). Finally, note that $t^p_{\rm mix}(F) \le t^{p'}_{\rm mix}(F)$ for $p \le p'$, which can easily be shown using the Hölder inequality.

We continue by introducing some general notation for graphs and their associated random walks. First, fix $G = (V(G), E(G))$ to be a finite connected graph with at least two vertices, where $V(G)$ denotes the vertex set and $E(G)$ the edge set of G, and suppose d_G is a metric on $V(G)$. In some examples, d_G will be a rescaled version of the usual shortest path graph distance, by which we mean that $d_G(x, y)$ is some multiple of the number of edges in the shortest path from x to y in G, but this is not always the most convenient choice. Define a symmetric weight function $\mu^G : V(G)^2 \to \mathbb{R}_+$ that satisfies $\mu^G_{xy} > 0$ if and only if $\{x, y\} \in E(G)$. The discrete time random walk on the weighted graph G is then the Markov chain $((X^G_m)_{m\ge0}, \mathbf{P}^G_x, x \in V(G))$ with transition probabilities $(P_G(x, y))_{x,y\in V(G)}$ defined by $P_G(x, y) := \mu^G_{xy}/\mu^G_x$, where $\mu^G_x := \sum_{y\in V(G)} \mu^G_{xy}$. If we define a measure π^G on $V(G)$ by setting, for $A \subseteq V(G)$, $\pi^G(A) := \sum_{x\in A} \mu^G_x / \sum_{x\in V(G)} \mu^G_x$, then π^G is the invariant probability measure for X^G. The transition density of X^G, with respect to π^G, is given by $(p^G_m(x, y))_{x,y\in V(G),m\ge0}$, where

$$p^G_m(x, y) := \frac{\mathbf{P}^G_x(X_m = y)}{\pi^G(\{y\})}.$$

Due to parity concerns for bipartite graphs, we will consider a smoothed version of this function $(q^G_m(x, y))_{x,y\in V(G),m\ge0}$ obtained by setting

$$q^G_m(x, y) := \frac{p^G_m(x, y) + p^G_{m+1}(x, y)}{2}, \tag{1.7}$$

and define the L^p-mixing time of G by

$$t^p_{\rm mix}(G) := \inf\left\{ m > 0 : \sup_{x\in V(G)} D^G_p(x, m) \le 1/4 \right\}, \tag{1.8}$$

where $D^G_p(x, m) := \|q^G_m(x, \cdot) - 1\|_{L^p(\pi^G)}$. Finally, in the case that we are considering a sequence of graphs $(G^N)_{N\ge1}$, we will usually abbreviate π^{G^N} to π^N and q^{G^N} to q^N, etc.

Remark 1.3. In [35], the mixing time of $\mathcal{C}^N$ is defined in terms of the total variation distance, that is

$$T_{\rm mix}(\mathcal{C}^N) = \min\{t : \|P_t(x, \cdot) - \pi(\cdot)\|_{\rm TV} \le 1/8, \quad \forall x \in V(\mathcal{C}^N)\}, \tag{1.9}$$

where $P_t(x, B) = \sum_{y\in B} q^N_t(x, y)\pi(y)$ for $B \subset V(\mathcal{C}^N)$ and $\|\mu - \nu\|_{\rm TV} = \max_{B\subset V(\mathcal{C}^N)} |\mu(B) - \nu(B)|$ for probability measures μ, ν on $V(\mathcal{C}^N)$. (To be precise, 1/8 in (1.9) is 1/4 in [35], but this only affects the constants in the results.) However, noting that

$$\|\mu - \nu\|_{\rm TV} = \frac{1}{2} \sum_{x\in V(\mathcal{C}^N)} |\mu(\{x\}) - \nu(\{x\})|,$$

(see, for example [33, Proposition 4.2]), one sees that $T_{\rm mix}(\mathcal{C}^N) = t^1_{\rm mix}(\mathcal{C}^N)$. Also note that [35] considers the lazy walk on the graph to avoid parity issues, but the same techniques will apply to the mixing time defined in terms of the smoothed heat kernel introduced at (1.7).

We are now ready to state the assumption under which we are able to prove the convergence of mixing times for the random walks on a sequence of graphs. This explains that, when suitably rescaled, the discrete state spaces, invariant measures and transition densities of a sequence of graphs converge to (F, d_F), π and $(q_t(x, y))_{x,y\in F,t>0}$, respectively. Its formulation involves a spectral Gromov-Hausdorff topology, the definition of which is postponed until Section 2, and a useful sufficient condition for it will be given in Proposition 2.4. Note that we extend the definition of the discrete transition densities on graphs to all positive times by linear interpolation of $(q^G_m(x,y))_{m\geq 0}$ for each pair of vertices $x, y \in V(G)$. Note also that the extended transition densities are different from those of continuous time Markov chains.

Assumption 1. $(G^N)_{N\geq 1}$ *is a sequence of finite connected graphs with at least two vertices for which there exists a sequence* $(\gamma(N))_{N\geq 1}$ *such that, for any compact interval* $I \subset (0,\infty)$*,*

$$\left(\left(V(G^N), d_{G^N}\right), \pi^N, \left(q^N_{\gamma(N)t}(x,y)\right)_{x,y\in V(G^N),t\in I}\right) \to \left((F,d_F), \pi, (q_t(x,y))_{x,y\in F,t\in I}\right)$$

in a spectral Gromov-Hausdorff sense.

Our main conclusion is then the following.

Theorem 1.4. *Suppose that Assumption 1 is satisfied. If* $p \in [1,\infty]$ *is such that the transition density* $(q_t(x,y))_{x,y\in F,t>0}$ *converges to stationarity in an* L^p *sense, then* $t^p_{\mathrm{mix}}(F) \in (0,\infty)$ *and*

$$\gamma(N)^{-1} t^p_{\mathrm{mix}}(G^N) \to t^p_{\mathrm{mix}}(F). \tag{1.10}$$

In Section 3.2, we will explain how to derive a variation of Theorem 1.4 that concerns the convergence of mixing times of processes started at a distinguished point in the state space.

We emphasize that a key part of our paper is to verify Assumption 1 and apply Theorem 1.4 in various interesting examples (including the Erdős-Rényi random graphs in the critical window as mentioned above). Therefore, we devote considerable space to applying our results to such.

The organization of the paper is as follows. In Section 2, we give a precise definition of the spectral Gromov-Hausdorff convergence and give some of its basic properties. In Section 3, we prove Theorem 1.4 and derive a variation of the theorem for distinguished starting points. Some sufficient conditions for (1.1)-(1.5) are given in Section 4. A selection of examples where the assumptions of Theorem 1.4 can be verified, and hence we have convergence of the mixing time sequence, are given in Section 5. In the Appendix, we introduce some geometric conditions on graphs for upper and lower bounds on the mixing times for the corresponding symmetric Markov chains. Some useful conditions to derive tail estimates of mixing times on random graphs are also given there.

2 Spectral Gromov-Hausdorff convergence

The aim of this section is to define a spectral Gromov-Hausdorff distance on triples consisting of a metric space, a measure and a heat kernel-type function that will allow us to make precise Assumption 1. We will also derive an equivalent characterisation of this assumption that will be applied in the subsequent section when proving our mixing time convergence result, and present a sufficient condition for Assumption 1 that will be useful when it comes to checking it in examples. Note that we do not need to assume (1.3), (1.4) in this section, and only use (1.1) to deduce Proposition 2.4 from a result of [14].

First, for a compact interval $I \subset (0,\infty)$, let $\tilde{\mathcal{M}}_I$ be the collection of triples of the form (F,π,q), where $F = (F,d_F)$ is a non-empty compact metric space, π is a Borel probability measure on F and $q = (q_t(x,y))_{x,y\in F,t\in I}$ is a jointly continuous real-valued function of (t,x,y). We say two elements, (F,π,q) and (F',π',q'), of $\tilde{\mathcal{M}}_I$ are equivalent if there exists an isometry

$f : F \to F'$ such that $\pi \circ f^{-1} = \pi'$ and $q'_t \circ f = q_t$ for every $t \in I$, by which we mean $q'_t(f(x), f(y)) = q_t(x, y)$ for every $x, y \in F$, $t \in I$. Define $\mathcal{M}_I$ to be the set of equivalence classes of $\tilde{\mathcal{M}}_I$ under this relation. We will often abuse notation and identify an equivalence class in $\mathcal{M}_I$ with a particular element of it. Now, set

$$\begin{aligned} &\Delta_I\left((F,\pi,q),(F',\pi',q')\right)\\ &:= \inf_{Z,\phi,\phi',\mathcal{C}} \Bigg\{ d_H^Z(\phi(F),\phi'(F')) + d_P^Z(\pi\circ\phi^{-1},\pi'\circ\phi'^{-1}) \\ &\quad + \sup_{(x,x'),(y,y')\in\mathcal{C}} \left(d_Z(\phi(x),\phi'(x')) + d_Z(\phi(y),\phi'(y')) + \sup_{t\in I}\left|q_t(x,y)-q'_t(x',y')\right| \right) \Bigg\}, \end{aligned}$$

where the infimum is taken over all metric spaces $Z = (Z, d_Z)$, isometric embeddings $\phi : F \to Z$, $\phi' : F' \to Z$, and correspondences $\mathcal{C}$ between F and F', d_H^Z is the Hausdorff distance between compact subsets of Z, and d_P^Z is the Prohorov distance between Borel probability measures on Z. Note that, by a correspondence $\mathcal{C}$ between F and F', we mean a subset of $F \times F'$ such that for every $x \in F$ there exists at least one $x' \in F'$ such that $(x, x') \in \mathcal{C}$ and conversely for every $x' \in F'$ there exists at least one $x \in F$ such that $(x, x') \in \mathcal{C}$.

Before proceeding to check the above definition gives us a metric and that the corresponding space is separable, let us make a few remarks about the inspiration for it. In the infimum characterising Δ_I, the first term is simply that used in the standard Gromov-Hausdorff distance (see [7, Definition 7.3.10], for example). The second term is that considered by the authors of [22] in defining their 'Gromov-Prohorov' distance between metric measure spaces. The final term is closely related to one used in [16, Section 6], when defining a distance between spatial trees – real trees equipped with a continuous function. Indeed, the notion of a correspondence is quite standard in the Gromov-Hausdorff setting as a way to relate two compact metric spaces. One can, for example, alternatively define the Gromov-Hausdorff distance between compact metric spaces as half the infimum of the distortion of the correspondences between them (see [7, Theorem 7.3.25]).

Lemma 2.1. *For any compact interval $I \subset (0, \infty)$, $(\mathcal{M}_I, \Delta_I)$ is a separable metric space.*

Proof. Fix a compact interval $I \subset (0, \infty)$. That Δ_I is a non-negative function and is symmetric is obvious. To prove that it is also the case that $\Delta_I ((F, \pi, q), (F', \pi', q')) < \infty$ for any choice of $(F, \pi, q), (F', \pi', q') \in \mathcal{M}_I$, simply consider Z be the disjoint union of F and F', setting $d_Z(x, x') := \mathrm{diam}(F, d_F) + \mathrm{diam}(F', d'_F)$ for any $x \in F, x' \in F'$, and suppose $\mathcal{C} = F \times F'$.

We next show that Δ_I is positive definite. Suppose $(F, \pi, q), (F', \pi', q') \in \mathcal{M}_I$ are such that $\Delta_I ((F, \pi, q), (F', \pi', q')) = 0$. For every $\varepsilon > 0$, we can thus choose $Z, \phi, \phi', \mathcal{C}$ such that the sum of quantities in the defining infimum of Δ_I is bounded above by ε. Moreover, there exists a $\delta \in (0, \varepsilon]$ such that

$$\sup_{\substack{x_1,x_2,y_1,y_2\in F:\\ d_F(x_1,x_2),d_F(y_1,y_2)\leq\delta}} \sup_{t\in I} |q_t(x_1, y_1) - q_t(x_2, y_2)| \leq \varepsilon. \tag{2.1}$$

Now, let $(x_i)_{i=1}^\infty$ be a dense sequence of disjoint elements of F (in the case F is finite, we suppose that the sequence terminates after having listed the $\#F$ elements). By the compactness of F, there exists an integer N_ε such that $(B_F(x_i, \delta))_{i=1}^{N_\varepsilon}$ is a cover for F. Define $A_1 := B_F(x_1, \delta)$, and $A_i := B_F(x_i, \delta)\backslash \cup_{j=1}^{i-1} B_F(x_i, \delta)$ for $i = 2, \dots, N_\varepsilon$, so that $(A_i)_{i=1}^{N_\varepsilon}$ is a disjoint cover of F, and then consider a function $f_\varepsilon : F \to F'$ obtained by setting

$$f_\varepsilon(x) := x'_i$$

on A_i, where x_i' is chosen such that $(x_i, x_i') \in \mathcal{C}$ for each $i = 1, \dots, N_\varepsilon$. Clearly, by definition, f_ε is a measurable function. It is further the case that it satisfies, for any $x \in F$,

$$d_Z(\phi(x), \phi'(f_\varepsilon(x))) \leq d_Z(\phi(x), \phi(x_i)) + d_Z(\phi(x_i), \phi'(x_i')) \leq 2\varepsilon,$$

where we assume above that $i \in \{1, \dots, N_\varepsilon\}$ is such that $x \in A_i$. From this, it readily follows that:

$$\sup_{x,y \in F} |d_F(x,y) - d_{F'}(f_\varepsilon(x), f_\varepsilon(y))| \leq 4\varepsilon \tag{2.2}$$

and

$$d_P^{F'}(\pi \circ f_\varepsilon^{-1}, \pi') \leq 3\varepsilon, \tag{2.3}$$

where $d_P^{F'}$ is the Prohorov distance on F'. We also have, by applying (2.1), that

$$\sup_{x,y \in F, t \in I} \left| q_t(x,y) - q_t'(f_\varepsilon(x), f_\varepsilon(y)) \right| \leq 2\varepsilon. \tag{2.4}$$

To continue, we use a diagonalisation argument to deduce the existence of a sequence $(\varepsilon_n)_{n \geq 1}$ such that $f_{\varepsilon_n}(x_i)$ converges to some limit $f(x_i) \in F'$ for every $i \geq 1$. From (2.2), we obtain that $d_{F'}(f(x_i), f(x_j)) = d_F(x_i, x_j)$ for every $i, j \geq 1$, and so we can extend the map f continuously to the whole of F ([7, Proposition 1.5.9]). This construction immediately implies that f is distance preserving. Moreover, reversing the roles of F and F', we are able to find a distance preserving map from F' to F. Hence f must be an isometry. To check that (F, π, q) and (F', π', q') are equivalent, it therefore remains to check that $\pi \circ f^{-1} = \pi'$ and $q_t' \circ f = q_t$ for every $t \in I$. Fix $\varepsilon > 0$ and recall that the definition of $(x_i)_{i=1}^{N_\varepsilon}$ means that it is an ε-net for F. Let $\varepsilon' \in (0, \varepsilon]$ be such that $d_{F'}(f_{\varepsilon'}(x_i), f(x_i)) \leq \varepsilon$ for every $i = 1, \dots, N_\varepsilon$. Then,

$$d_{F'}(f_{\varepsilon'}(x), f(x)) \leq d_{F'}(f_{\varepsilon'}(x), f_{\varepsilon'}(x_i)) + d_{F'}(f_{\varepsilon'}(x_i), f(x_i)) + d_{F'}(f(x_i), f(x)) \leq 7\varepsilon, \tag{2.5}$$

where we are again assuming that $i \in \{1, \dots, N_\varepsilon\}$ is such that $x \in A_i$, and have applied (2.2) and the distance-preserving property of f. In particular, this implies that

$$d_P^{F'}(\pi \circ f^{-1}, \pi') \leq d_P^{F'}(\pi \circ f^{-1}, \pi \circ f_{\varepsilon'}^{-1}) + d_P^{F'}(\pi \circ f_{\varepsilon'}^{-1}, \pi') \leq 10\varepsilon,$$

where we use (2.3) to deduce the second inequality. Since $\varepsilon > 0$ was arbitrary, this yields that $\pi \circ f^{-1} = \pi'$. Finally, (2.4) and (2.5) imply that

$$\sup_{x,y \in F, t \in I} \left| q_t(x,y) - q_t'(f(x), f(y)) \right| \leq 2\varepsilon + \sup_{\substack{x_1', x_2', y_1', y_2' \in F': \\ d_{F'}(x_1', x_2'), d_{F'}(y_1', y_2') \leq 7\varepsilon}} \sup_{t \in I} \left| q_t'(x_1', y_1') - q_t(x_2', y_2') \right|,$$

and so $q_t' \circ f = q_t$ for every $t \in I$ follows from the continuity properties of q'.

For the triangle inequality, we closely follow the proof of [22, Lemma 5.2]. Let $(F^{(i)}, \pi^{(i)}, q^{(i)})$ be an element of $\mathcal{M}_I$, $i = 1, 2, 3$. Suppose that $\Delta_I((F^{(1)}, \pi^{(1)}, q^{(1)}), (F^{(2)}, \pi^{(2)}, q^{(2)})) < \delta_1$, so that we can find a metric space Z_1, isometric embeddings $\phi_{1,1} : F^{(1)} \to Z_1$ and $\phi_{2,1} : F^{(2)} \to Z_1$ and correspondence $\mathcal{C}_1$ between $F^{(1)}$ and $F^{(2)}$ such that the sum of quantities in the defining infimum of Δ_I is bounded above by δ_1. If $\Delta_I((F^{(2)}, \pi^{(2)}, q^{(2)}), (F^{(3)}, \pi^{(3)}, q^{(3)})) < \delta_2$, we define Z_2 ,$\phi_{2,2}$, $\phi_{3,2}$, $\mathcal{C}_2$ in an analogous way. Now, set Z to be the disjoint union of Z_1 and Z_2, and define a distance on it by setting $d_Z|_{Z_i \times Z_i} = d_{Z_i}$ for $i = 1, 2$, and for $x \in Z_1$, $y \in Z_2$,

$$d_Z(x,y) := \inf_{z \in F^{(2)}} \left(d_{Z_1}(x, \phi_{2,1}(z)) + d_{Z_2}(\phi_{2,2}(z), y) \right).$$

Abusing notation slightly, it is then the case that, after points separated by a 0 distance have been identified, (Z, d_Z) is a metric space into which there is a natural isometric embedding ϕ_i of Z_i, $i = 1, 2$. In this space, we have that

$$\begin{aligned} &d_H^Z(\phi_1(\phi_{1,1}(F^{(1)})), \phi_2(\phi_{3,2}(F^{(3)}))) \\ &\leq \; d_H^{Z_1}(\phi_{1,1}(F^{(1)}), \phi_{2,1}(F^{(2)})) + d_H^{Z_2}(\phi_{2,2}(F^{(2)}), \phi_{3,2}(F^{(3)})), \end{aligned}$$

where we have applied the fact that $\phi_1(\phi_{2,1}(y)) = \phi_2(\phi_{2,2}(y))$ for every $y \in F^{(2)}$, and so $\phi_1(\phi_{2,1}(F^{(2)})) = \phi_2(\phi_{2,2}(F^{(2)}))$ as subsets of Z. A similar bound applies to the embedded measures. Now, let

$$\mathcal{C} := \{(x, z) \in F^{(1)} \times F^{(3)} : \exists y \in F^{(2)} \text{ such that } (x, y) \in \mathcal{C}_1, (y, z) \in \mathcal{C}_2\},$$

then if $(x, z) \in \mathcal{C}$,

$$d_Z(\phi_1(\phi_{1,1}(x)), \phi_2(\phi_{3,2}(z))) \leq d_{Z_1}(\phi_{1,1}(x), \phi_{2,1}(y)) + d_{Z_2}(\phi_{2,2}(y), \phi_{3,2}(z)),$$

where $y \in F^{(2)}$ is chosen such that $(x, y) \in \mathcal{C}_1$ and $(y, z) \in \mathcal{C}_2$, and we again note $\phi_1(\phi_{2,1}(y)) = \phi_2(\phi_{2,2}(y))$. Proceeding in the same fashion, one can deduce a corresponding bound involving $q^{(i)}$, $i = 1, 2, 3$. Putting these pieces together, it is elementary to deduce that

$$\Delta_I((F^{(1)}, \pi^{(1)}, q^{(1)}), (F^{(3)}, \pi^{(3)}, q^{(3)})) \leq \delta_1 + \delta_2,$$

and the triangle inequality follows.

To complete the proof, we only need to show separability. This is straightforward, however, as for any element of $\mathcal{M}_I$, one can construct an approximating sequence that incorporates only: metric spaces with a finite number of points and rational distances between them, probability measures on these with a rational mass at each point, and functions that are defined (at each coordinate pair) to be equal to rational values at a finite collection of rational time points and are linear between these. □

We will say that if a sequence in $\mathcal{M}_I$ converges to a limit in this space with respect to the metric Δ_I, then the convergence is in a spectral Gromov-Hausdorff sense. We note that in the framework of compact Riemannian manifolds, different but related notions of spectral distances were introduced by Bérard, Besson and Gallot ([5]) and by Kasue and Kumura ([24]). Moreover, by applying our characterisation of spectral Gromov-Hausdorff convergence, we are able to deduce that if Assumption 1 holds, then we can isometrically embed all the rescaled graphs, measures and transition densities upon them in a common metric space so that they converge to the relevant limit objects in a more standard way, as follows.

Lemma 2.2. *Suppose that Assumption 1 is satisfied. For any compact interval $I \subset (0, \infty)$, there exist isometric embeddings of $(V(G^N), d_{G^N})$, $N \geq 1$, and (F, d_F) into a common metric space (E, d_E) such that*

$$\lim_{N \to \infty} d_H^E(V(G^N), F) = 0, \tag{2.6}$$

$$\lim_{N \to \infty} d_P^E(\pi^N, \pi) = 0, \tag{2.7}$$

and also,

$$\lim_{N \to \infty} \sup_{x,y \in F} \sup_{t \in I} \left| q^N_{\gamma(N)t}(g_N(x), g_N(y)) - q_t(x, y) \right| = 0, \tag{2.8}$$

where, for brevity, we have identified the spaces $(V(G^N), d_{G^N})$, $N \geq 1$, and (F, d_F), and the measures upon them with their isometric embeddings in (E, d_E). For each $x \in F$, we define $g_N(x)$ to be a vertex in $V(G^N)$ minimising $d_E(x, y)$ over $y \in V(G^N)$.

Proof. Fix a compact interval $I \subset (0,\infty)$. By Assumption 1, for each $N \geq 1$ it is possible to find metric spaces (E_N, d_N), isometric embeddings $\phi_N : (V(G^N), d_{G^N}) \to (E_N, d_N)$, $\phi'_N : (F, d_F) \to (E_N, d_N)$ and correspondences $\mathcal{C}_N$ between $V(G^N)$ and F such that, identifying the original objects and their embeddings,

$$d_H^{E^N}(V(G^N), F) + d_P^{E^N}(\pi^N, \pi) + \sup_{(x,x'),(y,y') \in \mathcal{C}_N} \left(d_N(x,x') + d_N(y,y') + \sup_{t \in I} \left| q^N_{\gamma(N)t}(x,y) - q_t(x',y') \right| \right) \leq \varepsilon_N, \quad (2.9)$$

where $\varepsilon_N \to 0$. Now, proceeding similarly to the proof of the triangle inequality in Lemma 2.1, set E to be the disjoint union of E^N, $N \geq 1$, and define a distance on it by setting $d_E|_{E^N \times E^N} = d_N$ for $N \geq 1$, and for $x \in E^N$, $x' \in E^{N'}$, $N \neq N'$, set

$$d_E(x,x') := \inf_{y \in F} \left(d_N(x,y) + d_{N'}(y,x') \right).$$

Quotienting out points that are separated by distance 0 results in a metric space (E, d_E) (again, this is a slight abuse of notation), into which we have natural isometric embeddings of the metric spaces $(V(G^N), d_{G^N})$, $N \geq 1$, and (F, d_F). Moreover, in the metric space (E, d_E), it readily follows from (2.9) that the relevant isometrically embedded objects satisfy (2.6) and (2.7). To prove (2.8), first note that for every $x \in V(G^N)$, $N \geq 1$, there exists an $x' \in F$ such that $(x,x') \in \mathcal{C}_N$. This implies that $d_E(x,x') \leq \varepsilon_N$, and so, for any $\delta > 0$,

$$\sup_{\substack{x,y,z \in V(G^N): \\ d_{G^N}(y,z) \leq \delta}} \sup_{t \in I} \left| q^N_{\gamma(N)t}(x,y) - q^N_{\gamma(N)t}(x,z) \right| \leq 2\varepsilon_N + \sup_{\substack{x,y,z \in F: \\ d_F(y,z) \leq \delta + 2\varepsilon_N}} \sup_{t \in I} \left| q_t(x,y) - q_t(x,z) \right|. \quad (2.10)$$

Now, for every $x \in F$ and $N \geq 1$, there exists an $x' \in V(G^N)$ such that $(x', x) \in \mathcal{C}_N$, and so $d_E(x', x) \leq \varepsilon_N$. Therefore, since $g_N(x)$ is the closest vertex of $V(G^N)$ to x,

$$g_N(x) \in B_E(x, 2\varepsilon_N) \cap V(G^N) \subseteq B_E(x', 3\varepsilon_N) \cap V(G^N) = B_{V(G^N)}(x', 3\varepsilon_N).$$

Consequently,

$$\begin{aligned} &\sup_{x,y \in F} \sup_{t \in I} \left| q^N_{\gamma(N)t}(g_N(x), g_N(y)) - q_t(x,y) \right| \\ &\quad \leq \varepsilon_N + 2 \sup_{\substack{x,y,z \in V(G^N): \\ d_{G^N}(y,z) \leq 3\varepsilon_N}} \sup_{t \in I} \left| q^N_{\gamma(N)t}(x,y) - q^N_{\gamma(N)t}(x,z) \right| \\ &\quad \leq 5\varepsilon_N + 2 \sup_{\substack{x,y,z \in F: \\ d_F(y,z) \leq 5\varepsilon_N}} \sup_{t \in I} \left| q_t(x,y) - q_t(x,z) \right|, \end{aligned}$$

where the second inequality is an application of (2.10). Letting $N \to \infty$ and applying the joint continuity of $(q_t(x,y))_{x,y \in F, t > 0}$, we obtain the desired result. □

For our later convenience, let us note a useful tightness condition for the rescaled transition densities that was essentially established in the proof of the previous result. Since the result readily follows from the bound at (2.10), we will not explain its proof further.

Lemma 2.3. *Suppose that Assumption 1 holds. For any compact interval $I \subset (0,\infty)$,*

$$\lim_{\delta\to 0}\limsup_{N\to\infty}\sup_{\substack{x,y,z\in V(G^N):\\ d_{G^N}(y,z)\leq\delta}}\sup_{t\in I}\left|q^N_{\gamma(N)t}(x,y)-q^N_{\gamma(N)t}(x,z)\right|=0. \tag{2.11}$$

It is straightforward to reverse the conclusions of the previous two lemmas to check that if (2.6), (2.7), (2.8) and (2.11) hold, then so does Assumption 1. Thus, in examples, it will suffice to check these equivalent conditions when seeking to verify Assumption 1. In fact, it is further possible to weaken these assumptions slightly by appealing to a local limit theorem from [14]. To be precise, because we are assuming that the transition densities of the graph satisfy the tightness condition of (2.11), we can apply [14, Theorem 15], to replace the local convergence statement of (2.8) with a central limit-type convergence statement. Note that, although in [14] it was assumed that the metric on G^N was a rescaled graph distance, exactly the same argument yields the corresponding conclusion in our setting, and so we simply state the result.

Proposition 2.4 (cf. [14, Theorem 15])**.** *Suppose that $(V(G^N),d_{G^N})$, $N\geq 1$, and (F,d_F) can be isometrically embedded into a common metric space (E,d_E) in such a way that (2.6) and (2.7) are both satisfied. Moreover, assume that there exists a dense subset F^* of F such that, for any compact interval $I\subset(0,\infty)$, $x\in F^*$, $y\in F$, $r>0$,*

$$\lim_{N\to\infty}\mathbf{P}^{G^N}_{g_N(x)}\left(X^{G^N}_{\lfloor\gamma(N)t\rfloor}\in B_E(y,r)\right)=\mathbf{P}^F_x\left(X^F_t\in B_E(y,r)\right) \tag{2.12}$$

uniformly for $t\in I$, and also (2.11) holds. Then Assumption 1 holds.

To complete this section, let us observe that [14] also provides two ways to check (2.11): one involving a resistance estimate on the graphs in the sequence ([14, Proposition 17]), and one involving the parabolic Harnack inequality ([14, Proposition 16]). Since the first of these two methods will be applied in several of our examples later, let us recall the result here. To allow us to state the result, we define $R_{G^N}(x,y)$ to be the resistance between x and y in $V(G^N)$ (see (A.1)), when we suppose that G^N is an electrical network with conductances of edges being given by the weight function μ^{G^N}. This defines a metric on $V(G^N)$, for which we may check the following.

Lemma 2.5 (cf. [14, Proposition 17])**.** *Suppose that there exists a sequence $(\alpha(N))_{N\geq 1}$ and constants $\kappa,c_1,c_2,c_3\in(0,\infty)$ such that*

$$R_{G^N}(x,y)\leq c_1\left(\alpha(N)d_{G^N}(x,y)\right)^\kappa,\qquad \forall x,y\in V(G^N),$$

and also

$$c_2\gamma(N)\leq\alpha(N)^\kappa\beta(N)\leq c_3\gamma(N),$$

where $\beta(N):=\sum_{x,y\in V(G^N)}\mu^{G^N}_{xy}$, then (2.11) holds.

3 Convergence of L^p-mixing times

3.1 Proof of Theorem 1.4

In this subsection we prove the mixing time convergence result of Theorem 1.4. Throughout, we will suppose that Assumption 1 holds and that the graphs G^N and limiting metric space F have been embedded into a common metric space (E,d_E) in the way described by Lemma 2.2.

Recall from the introduction the definition of $D_p(x,t) = \|q_t(x,\cdot) - 1\|_{L^p(\pi)}$, the L^p-distance from stationarity of the process X^F started from x at time t. By applying the continuity of $(q_t(x,y))_{x,y\in F,t>0}$, compactness of F and finiteness of π, it is easy to check that this quantity is finite for every $x \in F$ and $t > 0$. The next lemma collects together a number of other basic properties of $D_p(x,t)$ that we will apply later (the first part is a minor extension of [8, Proposition 3.1], in our setting).

Lemma 3.1. *Let $p \in [1,\infty]$. For every $x \in F$, the function $t \mapsto D_p(x,t)$ is continuous and strictly decreasing. Furthermore, we have*

$$\lim_{t\to 0} D_p(x,t) \geq 2. \tag{3.1}$$

Proof. That the function $t \mapsto D_p(x,t)$ is continuous is clear from (1.2). We now check that it is strictly decreasing. First, a standard argument involving an application of Jensen's inequality and the invariance of π allows one to deduce that $\|P_t f\|_{L^p(\pi)} \leq \|f\|_{L^p(\pi)}$ for any $f \in L^p(F,\pi)$, where $(P_t)_{t\geq 0}$ is the semigroup naturally associated with the transition density $(q_t(x,y))_{x,y\in F,t>0}$. Now, suppose $f \in L^p(F,\pi)$ is such that $\|P_t f\|_{L^p(\pi)} = \|f\|_{L^p(\pi)}$, and define $f_1(y) := |P_t f(y)|^p$ and $f_2(y) := P_t(|f|^p)(y)$. By our assumption on f, we have that $\int_F f_1 d\pi = \int_F f_2 d\pi$. Furthermore, Jensen's inequality implies $f_1(y) \leq f_2(y)$. Thus, it must be the case that $f_1(y) = f_2(y)$, π-a.e. In particular, because π is a probability measure, there exists a $y \in F$ such that $f_1(y) = f_2(y)$. In the case $p > 1$, this equality readily implies that f is constant $q_t(y,z)\pi(dz)$-a.e. Recalling the assumption that $q_t(y,z) > 0$ everywhere, namely (1.3), it must therefore hold that f is constant π-a.e. Observing that for $s,t > 0$ we can write $D_p(x,s+t) = \|P_s(q_t(x,\cdot) - 1)\|_{L^p(\pi)}$, and noting the condition (1.4), it follows that $D_p(x,s+t) < D_p(x,t)$, as desired. For $p = 1$, the result $f_1(y) = f_2(y)$ implies that f is either non-negative or non-positive, π-a.e. Consequently, for $D_p(x,s+t) = \|P_s(q_t(x,\cdot) - 1)\|_{L^p(\pi)} = D_p(x,t)$, we would require that $q_t(x,\cdot) - 1$ is either non-negative or non-positive. Since $\int_F (q_t(x,y) - 1)\pi(dy) = 0$ (due to (1.1)) and (1.4) holds, this can not occur, which completes the proof of strict monotonicity.

To establish the limit at (3.1), it will suffice to prove the result in the case $p = 1$ (obtaining the result for other values of p is then simply Jensen's inequality). Let $x \in F$ and $r > 0$, then

$$\begin{aligned} D_1(x,t) &\geq \int_{B_E(x,r)} (q_t(x,y) - 1)\pi(dy) + \int_{B_E(x,r)^c} (1 - q_t(x,y))\pi(dy) \\ &= 2\mathbf{P}_x\left(X_t^F \in B_E(x,r)\right) - 2\pi(B_E(x,r)), \end{aligned}$$

where (1.1) is used in the last equality. Since X^F is a Hunt process, the first term here converges to 2 as $t \to 0$. Furthermore, because π is non-atomic, the second term can be made arbitrarily small by suitable choice of r. The result follows. □

We continue by defining the L^p-mixing time at $x \in F$ by setting

$$t^p_{\mathrm{mix}}(x) := \inf\{t > 0 : D_p(x,t) \leq 1/4\}.$$

In fact, the previous lemma yields that $t^p_{\mathrm{mix}}(x)$ is the unique value of $t \in (0,\infty)$ such that $D_p(x,t) = 1/4$ (when (1.5) holds at x). Similarly, define the L^p-mixing time of $x \in V(G^N)$ by setting

$$t^{N,p}_{\mathrm{mix}}(x) := \inf\{t > 0 : D^N_p(x,t) \leq 1/4\},$$

where $D^N_p(x,m) = \|q^N_m(x,\cdot) - 1\|_{L^p(\pi^N)}$. That the discrete mixing times at a point converge when suitably rescaled to the continuous mixing time there is the conclusion of the following proposition.

Proposition 3.2. *Suppose that Assumption 1 is satisfied. If $p \in [1, \infty]$ is such that (1.5) holds for $x \in F$, then*

$$\lim_{N\to\infty} \gamma(N)^{-1} t_{\mathrm{mix}}^{N,p}(g_N(x)) = t_{\mathrm{mix}}^{p}(x),$$

where, as in the statement of Lemma 2.2, $g_N(x)$ is a vertex in $V(G^N)$ that minimises the distance $d_E(x, y)$ over $V(G^N)$.

Proof. Suppose $p \in [1, \infty]$ is such that (1.5) holds for $x \in F$, set $t_0 := t_{\mathrm{mix}}^{p}(x) \in (0, \infty)$, and fix $\varepsilon > 0$. By (1.2) and the tightness of Lemma 2.3, there exists a $\delta > 0$ such that

$$\sup_{t\in I} \sup_{\substack{y,z\in F:\\ d_E(y,z)<2\delta}} \left||q_t(x, y) - 1|^p - |q_t(x, z) - 1|^p\right| < \varepsilon, \tag{3.2}$$

$$\limsup_{N\to\infty} \sup_{t\in I} \sup_{\substack{y,z\in V(G^N):\\ d_{G^N}(y,z)<3\delta}} \left| |q^N_{\gamma(N)t}(g_N(x), y) - 1|^p - |q^N_{\gamma(N)t}(g_N(x), z) - 1|^p \right| < \varepsilon, \tag{3.3}$$

where $I := [t_0/2, 2t_0]$. Moreover, by the compactness of F, there exists a finite collection of balls $(B_E(x_i, \delta))_{i=1}^k$ covering F. Define $A_1 := B(x_1, 2\delta)$, and $A_i := B_E(x_i, 2\delta) \setminus \cup_{j=1}^{i-1} B_E(x_i, 2\delta)$ for $i = 2, \dots, k$, so that $(A_i)_{i=1}^k$ is a disjoint cover of the δ-enlargement of F.

We observe

$$|D_p(x, t)^p - D_p^N(g_N(x), \gamma(N)t)^p| \le T_1 + T_2 + T_3 + T_4,$$

where

$$\begin{aligned}
T_1 &:= \left| \int_F |q_t(x, y) - 1|^p \pi(dy) - \sum_{i=1}^k |q_t(x, x_i) - 1|^p \pi(A_i) \right|, \\
T_2 &:= \left| \sum_{i=1}^k |q_t(x, x_i) - 1|^p \pi(A_i) - \sum_{i=1}^k |q_t(x, x_i) - 1|^p \pi^N(A_i) \right|, \\
T_3 &:= \left| \sum_{i=1}^k |q_t(x, x_i) - 1|^p \pi^N(A_i) - \sum_{i=1}^k |q^N_{\gamma(N)t}(g_N(x), g_N(x_i)) - 1|^p \pi^N(A_i) \right|, \\
T_4 &:= \left| \sum_{i=1}^k |q^N_{\gamma(N)t}(g_N(x), g_N(x_i)) - 1|^p \pi^N(A_i) - \int_{V(G^N)} |q^N_{\gamma(N)t}(g_N(x), y) - 1|^p \pi^N(dy) \right|.
\end{aligned}$$

Now, suppose $t \in I$. From (3.2), we immediately deduce that $T_1 \le \varepsilon$. For T_2, we first observe that the fact balls are π-continuity sets implies that so are the sets A_i, $i = 1, \dots, k$. Hence $\pi^N(A_i) \to \pi(A_i)$ for each $i = 1, \dots, k$, and so $T_2 \le \varepsilon$ for large N. That $T_3 \le \varepsilon$ for large N is a straightforward consequence of Lemma 2.2. Finally, applying the fact that $d_H^E(F, V(G^N)) \to 0$, we deduce that, for large N, $(A_i)_{i=1}^k$ is a disjoint cover for $V(G^N)$. Since $g_N(x_i) \in B_E(x_i, \delta)$ for large N, we also have that $d_{G^N}(y, g_N(x_i)) \le 3\delta$, uniformly over $y \in A_i$, $i = 1, \dots, k$. Thus we can appeal to (3.3) to deduce that it is also the case that $T_4 \le \varepsilon$ for large N. In fact, each of these bounds can be assumed to hold uniformly over $t \in I$, thereby demonstrating that

$$\lim_{N\to\infty} \sup_{t\in I} \left| D_p(x, t) - D_p^N(g_N(x), \gamma(N)t) \right| = 0. \tag{3.4}$$

Since $t \mapsto D_p^N(g_N(x), \gamma(N)t)$ is a decreasing function in t for every N (cf. [8, Proposition 3.1]) and $t \mapsto D_p(x, t)$ is strictly decreasing, the proposition follows. □

Remark 3.3. In the case $p = 2$, the proof of the previous result greatly simplifies. In particular, we note that

$$D_2(x,t)^2 = \|q_t(x,\cdot) - 1\|_2^2 = q_{2t}(x,x) - 1, \tag{3.5}$$

and similarly

$$D_2^N(x,\gamma(N)t)^2 = \|q^N_{\gamma(N)t}(x,\cdot) - 1\|_2^2 = q^N_{2\gamma(N)t}(x,x) - 1.$$

Hence the limit at (3.4) is an immediate consequence of the local limit result of (2.8), and we do not have to concern ourselves with estimating the relevant integrals directly.

To extend the above proposition to the corresponding result for the mixing times of the entire spaces, we will appeal to the following lemma, which establishes a continuity property for the L^p-mixing times from fixed starting points in the limiting space, and a related tightness property for the discrete approximations.

Lemma 3.4. *Suppose $p \in [1,\infty]$ is such that (1.5) holds for $x \in F$, then the following statements are true.*
(a) The function $y \mapsto t^p_{\text{mix}}(y)$ is continuous at x.
(b) Under Assumption 1, it is the case that

$$\lim_{\delta\to 0}\limsup_{N\to\infty} \sup_{\substack{y\in V(G^N):\\ d_{G^N}(g_N(x),y)<\delta}} \gamma(N)^{-1}\left|t^{N,p}_{\text{mix}}(y) - t^{N,p}_{\text{mix}}(g_N(x))\right| = 0.$$

Proof. Consider $p \in [1,\infty]$ such that (1.5) holds for $x \in F$, so that $t_0 := t^p_{\text{mix}}(x)$ is finite, and let $\varepsilon \in (0, t_0/2)$. Since the function $t \mapsto D_p(x,t)$ is strictly decreasing (by Lemma 3.1), there exists an $\eta > 0$ such that $D_p(x,t_0-\varepsilon) > D_p(x,t_0) + \eta = 1/4 + \eta$ and also $D_p(x,t_0+\varepsilon) < 1/4 - \eta$. By the continuity of $(q_t(x,y))_{x,y\in F,t>0}$, there also exists a $\delta > 0$ such that

$$\sup_{t\in[t_0-\varepsilon,t_0+\varepsilon]} \sup_{\substack{y\in F:\\ d_F(x,y)<\delta}} |D_p(x,t) - D_p(y,t)| < \eta.$$

Hence if $y \in B_F(x,\delta)$, then

$$D_p(y,t_0-\varepsilon) > D_p(x,t_0-\varepsilon) - \eta > \frac{1}{4},$$

$$D_p(y,t_0+\varepsilon) < D_p(x,t_0+\varepsilon) + \eta < \frac{1}{4}.$$

This implies that $t^p_{\text{mix}}(y) \in [t_0-\varepsilon, t_0+\varepsilon]$, and (a) follows.

The proof of part (b) is similar. In particular, choose η as above and note that (3.4) implies that $D_p^N(g_N(x),\gamma(N)(t_0-\varepsilon)) > 1/4 + \eta/2$ and $D_p^N(g_N(x),\gamma(N)(t_0+\varepsilon)) < 1/4 - \eta/2$ for large N. Furthermore, by the transition density tightness of Lemma 2.3, there exists a $\delta > 0$ such that

$$\sup_{t\in[t_0-\varepsilon,t_0+\varepsilon]} \sup_{\substack{y\in V(G^N):\\ d_{G^N}(g_N(x),y)<\delta}} \left|D_p^N(g_N(x),\gamma(N)t) - D_p^N(y,\gamma(N)t)\right| < \frac{\eta}{2},$$

for large N. Hence if N is large and $y \in V(G^N)$ is such that $d_{G^N}(g_N(x),y) < \delta$, then $D_p^N(y,\gamma(N)(t_0-\varepsilon)) > 1/4$, and $D_p^N(y,\gamma(N)(t_0+\varepsilon)) < 1/4$. This implies that $\gamma(N)^{-1}t^{N,p}_{\text{mix}}(y) \in [t_0-\varepsilon, t_0+\varepsilon]$. Since it is trivially true that, once N is large enough, this result can be applied with $y = g_N(x)$, the result follows. □

We are now ready to give the proof of our main result.

Proof of Theorem 1.4. Observe that, under the assumptions of the theorem, Lemma 3.4(a) implies that the function $(t^p_{\mathrm{mix}}(x))_{x\in F}$ is continuous. Since F is compact, the supremum of $(t^p_{\mathrm{mix}}(x))_{x\in F}$ is therefore finite. Now, it is an elementary exercise to check that we can write the L^p-mixing time of F, as defined at (1.6), in the following way:

$$t^p_{\mathrm{mix}}(F)=\sup_{x\in F}t^p_{\mathrm{mix}}(x). \tag{3.6}$$

Consequently $t^p_{\mathrm{mix}}(F)\in(0,\infty)$, as desired.

To complete the proof, we are required to demonstrate the convergence statement of (1.10). Fix $\varepsilon>0$. For every $x\in F$, Proposition 3.2 and Lemma 3.4(b) allow us to choose $\delta(x)>0$ and $N(x)<\infty$ such that

$$\sup_{N\geq N(x)}\left|\gamma(N)^{-1}t^{N,p}_{\mathrm{mix}}(g_N(x))-t^p_{\mathrm{mix}}(x)\right|\leq\varepsilon,$$

$$\sup_{N\geq N(x)}\ \sup_{\substack{y\in V(G^N):\\ d_{G^N}(g_N(x),y)<4\delta(x)}}\gamma(N)^{-1}\left|t^{N,p}_{\mathrm{mix}}(g_N(x))-t^{N,p}_{\mathrm{mix}}(y)\right|\leq\varepsilon.$$

Since $(B_E(x,\delta(x)))_{x\in F}$ is an open cover for F, by compactness it admits a finite subcover $(B_E(x,\delta(x)))_{x\in\mathcal{X}}$. Moreover, because $d_H^E(F,V(G^N))\to 0$, there exists an $N_0>0$ such that if $N\geq N_0$, then $(B_E(x,2\delta(x)))_{x\in\mathcal{X}}$ is a cover for $V(G^N)$. Applying this choice of $\mathcal{X}$, we have for $N\geq N_0\vee\max_{x\in\mathcal{X}}N(x)$ that

$$\gamma(N)^{-1}t^p_{\mathrm{mix}}(G^N)\leq\sup_{x\in\mathcal{X}}\gamma(N)^{-1}t^{N,p}_{\mathrm{mix}}(g_N(x))+\varepsilon\leq\sup_{x\in\mathcal{X}}t^p_{\mathrm{mix}}(x)+2\varepsilon\leq t^p_{\mathrm{mix}}(F)+2\varepsilon,$$

where we note that, similarly to (3.6), the L^p-mixing time of the graph G^N can be written as

$$t^p_{\mathrm{mix}}(G^N)=\sup_{x\in V(G^N)}t^{N,p}_{\mathrm{mix}}(x).$$

Furthermore, if $x_0\in F$ is chosen such that $t^p_{\mathrm{mix}}(x_0)\geq t^p_{\mathrm{mix}}(F)-\varepsilon$, then, for large N,

$$\gamma(N)^{-1}t^{N,p}_{\mathrm{mix}}(G^N)\geq\gamma(N)^{-1}t^{N,p}_{\mathrm{mix}}(g_N(x_0))\geq t^p_{\mathrm{mix}}(x_0)-\varepsilon\geq t^p_{\mathrm{mix}}(F)-2\varepsilon,$$

where we have again made use of Proposition 3.2. Since $\varepsilon>0$ was arbitrary, we are done. □

3.2 Distinguished starting points

In the case when convergence of transition densities is only known with respect to a single distinguished starting point, it is only possible to prove a convergence result for the mixing time from that point. Our goal in this subsection is to describe the topology in which we can do this.

Consider, for a compact interval $I\subset(0,\infty)$, the space of triples of the form (F,π,q), where $F=(F,d_F,\rho)$ is a non-empty compact metric space with distinguished vertex ρ, π is a Borel probability measure on F and $q=(q_t(x,y))_{x,y\in F,t\in I}$ is a jointly continuous real-valued function of (t,x,y); this is the same as the collection $\tilde{\mathcal{M}}_I$ defined in Section 2, though we have added the supposition that the metric spaces are pointed. We say two such elements, (F,π,q) and (F',π',q'), are equivalent if there exists an isometry $f:F\to F'$ such that $f(\rho)=\rho'$, $\pi\circ f^{-1}=\pi'$ and $q'_t\circ f=q_t$ for every $t\in I$. By following the proof of Lemma 2.1, one can check that it is possible to define a metric on the equivalence classes of this relation by simply including in the definition of Δ_I the condition that the correspondence $\mathcal{C}$ must contain (ρ, ρ'). We define convergence in a spectral pointed Gromov-Hausdorff sense to be with respect to this metric. The distinguished starting point version of Assumption 1 is then as follows.

Assumption 2. *Let $(G^N)_{N\geq 1}$ be a sequence of finite connected graphs with at least two vertices and one, ρ^N say, distinguished, for which there exists a sequence $(\gamma(N))_{N\geq 1}$ such that, for any compact interval $I \subset (0, \infty)$,*

$$\left(\left(V(G^N), d_{G^N}, \rho^N\right), \pi^N, \left(q^N_{\gamma(N)t}(\rho^N, x)\right)_{x\in V(G^N), t\in I}\right)$$

converges to $((F, d_F, \rho), \pi, (q_t(\rho, x))_{x\in F, t\in I})$ in a spectral pointed Gromov-Hausdorff sense, where ρ is a distinguished point in F.

The following result can then be proved in an almost identical fashion to Proposition 3.2. In doing this it is useful to note that if Assumption 1 is replaced by Assumption 2, then we are able to include in the conclusions of Lemma 2.2 that ρ^N converges to ρ in E.

Theorem 3.5. *Suppose that Assumption 2 is satisfied. If $p \in [1, \infty]$ is such that (1.5) holds for $x = \rho$, then*

$$\gamma(N)^{-1} t^{N,p}_{\text{mix}}(\rho^N) \to t^p_{\text{mix}}(\rho).$$

4 Convergence to stationarity of the transition density

Before continuing to present example applications of the mixing time convergence results proved so far, we describe how to check the L^p convergence to stationarity of the transition density of X^F in the case when we have a spectral decomposition for it and a spectral gap. In the same setting, we will also explain how to check the non-triviality conditions on the transition density that were made in the introduction.

Write the generator of the conservative Hunt process X^F as $-\Delta$, and suppose that Δ has a compact resolvent. Then there exists a complete orthonormal basis of $L^2(F, \pi)$, $(\varphi_k)_{k\geq 1}$ say, such that $\Delta\varphi_k = \lambda_k\varphi_k$ for all $k \geq 0$, $0 \leq \lambda_0 \leq \lambda_1 \leq \dots$ and $\lim_{k\to\infty} \lambda_k = \infty$. By expanding as a Fourier series, we can consequently write the transition density of X^F as

$$\begin{aligned} q_t(x, y) &= \sum_{k\geq 0} \left(\int_F q_t(x, z)\varphi_k(z)\pi(dz)\right) \varphi_k(y) \\ &= \sum_{k\geq 0} P^F_t \varphi_k(x)\varphi_k(y) \\ &= \sum_{k\geq 0} e^{-\lambda_k t}\varphi_k(x)\varphi_k(y), \end{aligned}$$

where $(P^F_t)_{t\geq 0}$ is the associated semigroup, and the final equality holds as a simple consequence of the fact that $\frac{d}{dt}(P^F_t \varphi_k) = -P^F_t \Delta\varphi_k = -\lambda_k P^F_t \varphi_k$. Now by (1.1), it holds that $1 = P^F_t 1$ is in the domain of Δ. A standard argument thus yields $\Delta 1 = \Delta P^F_t 1 = -\frac{d}{dt}(P^F_t 1) = 0$, and so there is no loss of generality in presupposing that $\lambda_0 = 0$ and $\varphi_0 \equiv 1$ in this setting. The only additional assumption we make on the transition density $(q_t(x, y))_{x,y\in F, t>0}$ is that it is jointly continuous in (t, x, y) (i.e. (1.2) holds).

Lemma 4.1. *Suppose that the operator Δ has a compact resolvent, so that the above spectral decomposition holds. If there is a spectral gap, i.e. $\lambda_1 > 0$, then $(q_t(x,y))_{x,y\in F,t>0}$ converges to stationarity in an L^p sense (namely* (1.5) *holds) for any $p \in [1,\infty]$.*

Proof. Recall from (3.5) that $D_2(x,t)^2 = q_{2t}(x,x) - 1$. Under the assumptions of the lemma, it follows that

$$D_2(x,t)^2 = \sum_{k\geq 1} e^{-2\lambda_k t}\varphi_k(x)^2 \to 0, \tag{4.1}$$

as $t \to \infty$, which completes the proof of the result for $p = 2$. To extend this to any p, we first use Cauchy-Schwarz to deduce

$$\begin{aligned}(q_t(x,y)-1)^2 &= \left(\sum_{k\geq 1} e^{-\lambda_k t}\varphi_k(x)\varphi_k(y)\right)^2 \\ &\leq \sum_{k\geq 1} e^{-\lambda_k t}\varphi_k(x)^2 \sum_{k\geq 1} e^{-\lambda_k t}\varphi_k(y)^2 \\ &= (q_t(x,x)-1)(q_t(y,y)-1).\end{aligned}$$

Consequently, we have that

$$\begin{aligned}D_\infty(x,t)^2 &= \sup_{y\in F}(q_t(x,y)-1)^2 \\ &\leq (q_t(x,x)-1)\sup_{y\in F}(q_t(y,y)-1) \\ &\leq D_2(x,t/2)^2 \sup_{y\in F} D_\infty(y,1)\end{aligned}$$

for any $t \geq 1$, where the second inequality involves an application of the monotonicity property proved as part of Lemma 3.1. Now, by (1.2), the term $\sup_{y\in F} D_\infty(y,1)$ is a finite constant, and so combining the above bound with (4.1) implies that $D_\infty(x,t) \leq CD_2(x,t/2) \to 0$ as $t \to \infty$. The result for general $p \in [1,\infty]$ is an immediate consequence of this. □

We now give a lemma that explains how to check conditions (1.3) and (1.4).

Lemma 4.2. *Suppose that the operator Δ has a compact resolvent and there is a spectral gap, then the conditions* (1.3) *and* (1.4) *are automatically satisfied.*

Proof. Firstly, assume that $q_t(x,y) = 0$ for some $x,y \in F$, $t > 0$. If $s \in (0,t)$, then the Chapman-Kolmogorov equations yield $0 = q_t(x,y) = \int_F q_s(x,z)q_{t-s}(z,y)\pi(dz)$. Since π has full support, using (1.2), it follows that $q_s(x,z)q_{t-s}(z,y) = 0$ for every $z \in F$. In particular, $q_s(x,y)q_{t-s}(y,y) = 0$. Noting that $q_{t-s}(y,y) = D_2^2(y,t/2) + 1 \geq 1$, we deduce that $q_s(x,y) = 0$. Now, define a function $f : (0,\infty) \to \mathbb{R}_+$ by setting $f(s) := q_s(x,y)$. Letting $(\lambda'_i)_{i\geq 0}$ represent the distinct eigenvalues of Δ, we can write

$$f(s) = \sum_{i\geq 0} a_i e^{-\lambda'_i s},$$

where $a_i := \sum_{j:\lambda_j=\lambda'_i}\varphi_j(x)\varphi_j(y)$. In fact, since Cauchy-Schwarz implies $\sum_{i\geq 0}|a_i e^{-\lambda'_i s}| \leq (q_s(x,x)q_s(y,y))^{1/2} < \infty$, this series converges absolutely whenever $s \in (0,\infty)$. Thus $f(z) := \sum_{i\geq 0} a_i e^{-\lambda'_i z}$ defines an analytic function on the whole half-plane $\Re(z) > 0$. By our previous observation regarding $q_s(x,y)$, this analytic function is equal to 0 on the set $(0,t]$, and therefore it must be 0 everywhere on $\Re(z)>0$. However, this contradicts the fact that $f(t)=q_t(x,y)\to 1$ as $t\to\infty$, which was proved in Lemma 4.1. Hence, $q_t(x,y)>0$ for every $x,y\in F$, $t>0$.

Secondly, suppose that $q_t(x,\cdot)\equiv 1$ for some $x\in F$ and $t>0$. Then $1=q_t(x,x)=1+\sum_{i\geq 1}\varphi_i(x)^2e^{-\lambda_i t}$, and so $\varphi_i(x)=0$ for every $i\geq 1$. This implies that $q_t(x,x)=1$ for every $t>0$. However, by following the proof of (3.1), one can deduce that

$$\lim_{t\to 0}(q_t(x,x)-1)=\lim_{t\to 0}D_2^2(x,t/2)\geq \lim_{t\to 0}D_1^2(x,t/2)\geq 2,$$

and so the previous conclusion can not hold. Consequently, we have shown that $q_t(x,\cdot)\not\equiv 1$ for any $x\in F$, $t>0$, as desired. □

To summarise, the above results demonstrate that to verify all the conditions on the transition density that are required to apply our mixing time convergence results, it will suffice to check that the conservative Hunt process X^F has a jointly continuous transition density and the corresponding non-negative self-adjoint operator, Δ, has a compact resolvent and exhibits a spectral gap. As the following corollary explains, this is a particularly useful observation in the case that the Dirichlet form $(\mathcal{E},\mathcal{F})$ associated with X^F is a resistance form (see [26, Definition 3.1]).

Corollary 4.3. *Suppose that X^F is a π-symmetric Hunt process on F such that the associated Dirichlet form $(\mathcal{E},\mathcal{F})$ is a resistance form, then* (1.1)-(1.5) *are automatically satisfied.*

Proof. The fact that X^F is conservative is clear since for a resistance form $1\in\mathcal{F}$ and $\mathcal{E}(1,1)=0$. That (1.2) holds is proved in [26, Lemma 10.7]. Moreover, we can check that the non-negative operator corresponding to $(\mathcal{E},\mathcal{F})$ has a compact resolvent (see [26, Lemma 9.7] and [28, Theorem B.1.13]) and exhibits a spectral gap (this is an easy consequence of the fact that, for a resistance form, $\mathcal{E}(f,f)=0$ if and only if f is constant). Thus, by Lemma 4.1 and Lemma 4.2, the transition density of X^F also satisfies (1.3)-(1.5). □

5 Examples

We now proceed to apply our mixing time convergence results to a number of examples: lattice models in a box, self-similar graphs with fractal weights, critical Galton-Watson trees, the critical Erdős-Rényi random graph, and the range of the random walk in high dimensions. In the third and fourth of these, we will also describe how these can be applied to relate tail asymptotics for mixing time distributions of the discrete and continuous models. Note that the general techniques we apply for estimating the relevant mixing times are postponed until the appendix.

5.1 Lattice models in a box

As a simple application of our main mixing time convergence result, we consider G^N to be a discrete box of side-length N, by which we mean that $V(G^N)=\{1,\dots,N\}^d$ and vertices $x,y\in V(G^N)$ are connected by an edge if and only if $\sum_{i=1}^d|x_i-y_i|=1$. We write d_{G^N} to represent the Euclidean metric on $V(G^N)$, denote the stationary probability measure of the simple random walk on G^N by π^N, and the corresponding smoothed transition density, as defined at (1.7), by $(q_m^N(x,y))_{x,y\in V(G^N),m\geq 0}$. In this setting, standard convergence results imply that, for any compact interval $I\subset(0,\infty)$,

$$\left(\left(G^N,N^{-1}d_{G^N}\right),\pi^N,(q^N_{N^2t}(x,y))_{x,y\in V(G^N),t\in I}\right)$$

converges in $(\mathcal{M}_I,\Delta_I)$ to the triple consisting of: $[0,1]^d$ equipped with the Euclidean metric, d-dimensional Lebesgue measure on this set and the transition density of Brownian motion on $[0,1]^d$ reflected at the boundary. In particular, Assumption 1 holds in our setting, and because (1.1)-(1.5) are satisfied, we are immediately able to apply Theorem 1.4 to deduce the following.

Theorem 5.1. *Fix $p\in[1,\infty]$. If $t^p_{\mathrm{mix}}(\{1,\dots,N\}^d)$ is the L^p-mixing time of the simple random walk on $\{1,\dots,N\}^d$, then*

$$N^{-2}t^p_{\mathrm{mix}}(\{1,\dots,N\}^d)\rightarrow t^p_{\mathrm{mix}}([0,1]^d),$$

where $t^p_{\mathrm{mix}}([0,1]^d)$ is the L^p-mixing time of the Brownian motion on $[0,1]^d$ reflected at the boundary.

We note that the same result holds if $\{1,\dots,N\}^d$ is replaced by the discrete d-dimensional torus $(\mathbb{Z}/N\mathbb{Z})^d$, and $[0,1]^d$ is replaced by the continuous d-dimensional torus $(\mathbb{R}/\mathbb{Z})^d$.

5.2 Self-similar fractal graphs with random weights

Although we will also briefly comment upon Sierpinski carpet-type graphs at the end of this section, the examples that we consider here are primarily those based on nested fractals, the definition of which we now recall. Suppose $(\psi_i)_{i=1}^K$ is a family of L^{-1}-similitudes on $\mathbb{R}^d$ for some $L>1$, by which we mean that, for each i, ψ_i is a map from $\mathbb{R}^d$ to $\mathbb{R}^d$ that satisfies $|\psi_i(x)-\psi_i(y)|=L^{-1}|x-y|$, for every $x,y\in\mathbb{R}^d$, where $|\cdot-\cdot|$ is the usual Euclidean distance on $\mathbb{R}^d$. We assume that the collection $(\psi_i)_{i=1}^K$ satisfies the open set condition; this means that there exists a non-empty bounded set $O\subseteq\mathbb{R}^d$ such that $(\psi_i(O))_{i=1}^K$ are disjoint and $\cup_{i=1}^K\psi_i(O)\subseteq O$. Since $(\psi_i)_{i=1}^K$ is a family of contraction maps, there exists a unique non-empty compact set F such that $F=\cup_{i=1}^K\psi_i(F)$. Write the set of fixed points of $(\psi_i)_{i=1}^K$ as Ξ, and define the collection of essential fixed points of $(\psi_i)_{i=1}^K$ by

$$V_0:=\{x\in\Xi:\ \exists i,j\in\{1,\dots,K\},\ i\neq j\text{ and }y\in\Xi\text{ such that }\psi_i(x)=\psi_j(y)\}.$$

Throughout, we assume that $\#V_0\geq 2$. The compact set F is then said to be a nested fractal if it satisfies the following connectivity, finite ramification and symmetry properties.

- For any $i,j\in\{1,\dots,K\}$, there exists a sequence $i=i_0,i_1,\dots,i_m=j$ such that
  $$\psi_{i_{l-1}}(V_0)\cap\psi_{i_l}(V_0)\neq\emptyset,$$
  for every $l=1,\dots,m$.
- If $i_1\dots i_n$ and $j_1\dots j_n$ are distinct sequences in $\{1,\dots,K\}$, then
  $$\psi_{i_1\dots i_n}(F)\cap\psi_{j_1\dots j_n}(F)=\psi_{i_1\dots i_n}(V_0)\cap\psi_{j_1\dots j_n}(V_0),$$
  where $\psi_{i_1\dots i_n}:=\psi_{i_1}\circ\dots\circ\psi_{i_n}$.
- If $x,y\in V_0$, then the reflection in the hyperplane $H_{xy}:=\{z\in\mathbb{R}^d:\ |z-x|=|z-y|\}$ maps V_n to itself, where
  $$V_n:=\bigcup_{i_1,\dots,i_n=1}^K\psi_{i_1\dots i_n}(V_0).$$

We will suppose d_F is the intrinsic shortest path metric on F defined in [18, Section 3] (we assume that the size vector introduced there is simply $\tilde{\mathbf{r}} = (1, \dots, 1)$), and note that this induces the same topology as the Euclidean metric. Moreover, we suppose π is the $(\ln K / \ln L)$-Hausdorff measure on F with respect to the Euclidean metric, normalised to be a probability measure. This measure is non-atomic, has full support and satisfies $\pi(\partial B(x,r)) = 0$ for every $x \in F$, $r > 0$ (see [14, Lemma 25]).

We now define a sequence of graphs $(G^N)_{N \geq 0}$ by setting

$$V(G^N) := \bigcup_{i_1, \dots, i_N = 1}^{K} \psi_{i_1 \dots i_N}(V_0)$$

and

$$E(G^N) := \left\{ \{ \psi_{i_1 \dots i_N}(x), \psi_{i_1 \dots i_N}(y) \} : \, x, y \in V_0, \, x \neq y, \, i_1, \dots, i_N \in \{1, \dots, K\} \right\}.$$

We set $d_{G^N} := d_F|_{V(G^N) \times V(G^N)}$, so that $(V(G^N), d_{G^N})$ clearly converges to (F, d_F) with respect to the Hausdorff distance between compact subsets of F. Weights μ^N on the edges of G^N will be selected randomly from a distribution that satisfies uniform boundedness and cell independence. By uniform boundedness, we mean that there exist deterministic constants $c_1, c_2 \in (0, \infty)$ such that, almost-surely,

$$c_1 \leq \mu^N_{xy} \leq c_2, \qquad \forall \{x, y\} \in E(G^N),$$

and we define cell independence to be the property that, for each $N \geq 0$, the collections

$$\left\{ \left(\mu^N_{\psi_{i_1 \dots i_N}(x) \psi_{i_1 \dots i_N}(y)} \right)_{x, y \in V_0, x \neq y} \right\}_{i_1, \dots, i_N \in \{1, \dots, K\}}$$

are independent and have the same distribution as $\left(\mu^0_{xy} \right)_{x, y \in V_0, x \neq y}$. Note that we still require $\mu^N_{xy} = \mu^N_{yx}$ for every $x, y \in V(G^N)$, and $\mu^N_{xy} = 0$ if $\{x, y\} \notin E(G^N)$. By the procedure described in the introduction, we define from these weights a sequence of random measures $(\pi^N)_{N \geq 0}$ on the vertex sets of our graphs in the sequence $(G^N)_{N \geq 0}$. That π^N weakly converges to π as Borel probability measures on F, almost-surely, can be checked by applying the argument of [14, Lemma 26].

To describe the scaling limit of the random walks associated with the random weights μ^N, we appeal to the homogenisation result of [29]. To describe this, we first introduce the Dirichlet form associated with the walk on the level N graph by setting, for $f \in \mathbb{R}^{V(G^N)}$,

$$\mathcal{E}^N(f, f) := \sum_{i_1, \dots, i_N = 1}^{K} \sum_{x, y \in V_0, x \neq y} \mu^N_{\psi_{i_1 \dots i_N}(x) \psi_{i_1 \dots i_N}(y)} \left(f(\psi_{i_1 \dots i_N}(x)) - f(\psi_{i_1 \dots i_N}(y)) \right)^2. \tag{5.1}$$

Let $\Lambda^N = (\Lambda^N_{xy})_{x, y \in V_0, x \neq y}$ be the collection of weights such that the associated random walk on G^0 is the trace of X^{G^N} onto V_0. It is then proved in [29, Theorem 3.4], that there exists a deterministic $C = (C_{xy})_{x, y \in V_0, x \neq y}$ and resistance scaling factor $\lambda \in (0, \infty)$ such that

$$\lim_{n \to \infty} \lambda^N \Lambda^N = C,$$

where the limit is an L^1-limit in the space of non-negative weights on the complete graph with vertex set V_0. Moreover, C satisfies $C_{xy} > 0$ for every $x, y \in V_0$, $x \neq y$, and is self-similar under the natural renormalisation operation. Now, suppose $\mathcal{E}_C^N$ is a quadratic form on $\mathbb{R}^{V(G^N)}$ which satisfies (5.1) with $\mu^N_{\psi_{i_1\dots i_N}(x)\psi_{i_1\dots i_N}(y)}$ replaced by C_{xy} in each summand, then define

$$\mathcal{E}(f,f)=\lim_{N\to\infty}\lambda^N\mathcal{E}_C^N(f|_{V(G^N)},f|_{V(G^N)})$$

for $f\in\mathcal{F}$, where $\mathcal{F}$ is the subset of $C(F,\mathbb{R})$ such that the right-hand side above exists finitely. It is known that $(\mathcal{E},\mathcal{F})$ is a local, regular Dirichlet form on $L^2(F,\pi)$, which is also a resistance form (see [28], for example). Thus, by Corollary 4.3, the associated π-symmetric diffusion X^F satisfies (1.1)-(1.5).

We now explain how to verify (2.12) in this setting. First, note that if we can prove a version of [30, Theorem 3.6] (or [29, Theorem 3.5]) with respect to X^{G^N} on our compact fractal F, then a probabilistic version of (2.12) can be proved similarly to [14, Proposition 30(i)]. (Here, our X^{G^N} is a discrete time Markov chain with π^N as the invariant measure, whereas in [29] and [30] the continuous-time Markov chains with normalised counting measure as the invariant measure were considered. However, since both measures are comparable and they converge to π almost-surely, the difference can be easily resolved.) So, we will explain how to prove a version of [30, Theorem 3.6]. Take two distinct points a_1,a_2 from V_0. Let $\sigma^0_{a_1}$ be the first hitting time of a_1, and for each $i\in\mathbb{N}$, define inductively

$$\begin{aligned}\sigma^i_{a_2}(X^{G^N}) &:= \inf\{m\ge\sigma^{i-1}_{a_1}(X^{G^N}):X^{G^N}_m=a_2\},\\ \sigma^i_{a_1}(X^{G^N}) &:= \inf\{m\ge\sigma^{i}_{a_2}(X^{G^N}):X^{G^N}_m=a_1\}.\end{aligned}$$

Then, we can write, for continuous $f:F\to\mathbb{R}$,

$$\begin{aligned}&\mathbf{E}^{G^N}_{x_N}[f(X^{G^N}_{t(K\lambda)^N})]\\ &= \mathbf{E}^{G^N}_{x_N}[f(X^{G^N}_{t(K\lambda)^N}):t(K\lambda)^N<\sigma^{(0)}_{a_1}]+\mathbf{E}^{G^N}_{x_N}[f(X^{G^N}_{t(K\lambda)^N}):\sigma^{(0)}_{a_1}\le t(K\lambda)^N<\sigma^{(1)}_{a_2}]\\ &\quad+\sum_{i=1}^\infty\mathbf{E}^{G^N}_{x_N}[f(X^{G^N}_{t(K\lambda)^N}):\sigma^{(i)}_{a_2}\le t(K\lambda)^N<\sigma^{(i)}_{a_1}] \qquad (5.2)\\ &\quad+\sum_{i=1}^\infty\mathbf{E}^{G^N}_{x_N}[f(X^{G^N}_{t(K\lambda)^N}):\sigma^{(i)}_{a_1}\le t(K\lambda)^N<\sigma^{(i+1)}_{a_2}], \qquad (5.3)\end{aligned}$$

where $x_N\in V(G^N)$ converges to $x\in F$, say. The first summand in the right hand side of (5.2) can be written in terms of the process X^{G^N} killed at a_1, and so by tracing the proof of [14, Proposition 30(i)] line by line, we can check it converges to the corresponding expectation involving X^F killed on hitting a_1. Similarly, the first summand in (5.3) can be written as

$$\begin{aligned}&\mathbf{E}^{G^N}_{x_N}[f(X^{G^N}_{t(K\lambda)^N}):\sigma^{(0)}_{a_1}\le t(K\lambda)^N<\sigma^{(1)}_{a_2}]\\ &= \mathbf{E}^{G^N}_{x_N}[1_{\{\sigma^{(0)}_{a_1}\le t(K\lambda)^N\}}\mathbf{E}^{G^N}_{a_1}[f(X^{G^N}_{t(K\lambda)^N-\sigma^{(0)}_{a_1}})1_{\{t(K\lambda)^N-\sigma^{(0)}_{a_1}<\sigma^{(1)}_{a_2}\circ\theta_{\sigma^{(0)}_{a_1}}\}}|\mathcal{F}_{\sigma^{(0)}_{a_1}}]],\end{aligned}$$

where θ is the shift map. Given $\sigma^{(0)}_{a_1}=s$, $\mathbf{E}^{G^N}_{a_1}[f(X^{G^N}_{t(K\lambda)^N-s})1_{\{t(K\lambda)^N-s<\sigma^{(1)}_{a_2}\}}]$ can be written in terms of the process started at a_1 and killed on hitting a_2, independently of the distribution of $\sigma^{(0)}_{a_1}$. Thus the second term in the right hand side of (5.2) converges to $\mathbf{E}^F_x[f(X^F_t):\sigma^{(0)}_{a_1}(X^F)\le t<\sigma^{(1)}_{a_2}(X^F)]$. We can prove convergence of the rest of the terms similarly. Moreover, by applying the estimate for the exit time of the random walks from balls stated as part of [14, Lemma 27], for example, it is straightforward to check that there exists a $t_0>0$ such that $\mathbf{P}^{G^N}_{a_1}((K\lambda)^{-N}\sigma^{(1)}_{a_2}\le t_0)$ and $\mathbf{P}^{G^N}_{a_2}((K\lambda)^{-N}\sigma^{(0)}_{a_1}\le t_0)$ are both bounded above by $1/2$, uniformly in N. As a consequence of this, one can show that the terms in the sums at (5.2) and (5.3) decay exponentially, uniformly in N, and hence that the right hand side of (5.2) converges to $\mathbf{E}_x^F[f(X_t^F)]$ as $N \to \infty$. Convergence of the finite dimensional distributions can be shown similarly and we obtain the desired version of [30, Theorem 3.6].

Finally, a probabilistic version of the tightness condition of (2.11) is easily checked by applying (a probabilistic version of) Lemma 2.5, using known resistance estimates for nested fractals (cf. [14, Proposition 30(ii)]), and so Assumption 1 holds in probability due to Proposition 2.4. Thus we are able to apply Theorem 1.4 to deduce the following.

Theorem 5.2. *If $t_{\mathrm{mix}}(G^N)$ is the mixing time of the random walk on the level N approximation to the nested fractal F equipped with random weights satisfying uniform boundedness and cell independence, then*

$$(K\lambda)^{-N} t_{\mathrm{mix}}(G^N) \to t_{\mathrm{mix}}(F)$$

in probability, where $t_{\mathrm{mix}}(F)$ is the mixing time of the diffusion X^F.

Let us remark that, if the weights $C = (C_{xy})_{x,y\in V_0, x\neq y}$ are the unique collection of weights that is invariant under the symmetries of the nested fractal (i.e. for every map h which is a reflection in a hyperplane of the form H_{xy}, $x, y \in V_0$, the collection $(C_{h(x)h(y)})_{x,y\in V_0,x\neq y}$ is identical to $(C_{xy})_{x,y\in V_0,x\neq y}$, see [34, 37] relevant uniqueness result), then the resulting diffusion X^F is the so-called Brownian motion on the nested fractal F. This is the case if we assume that $(\mu^0_{xy})_{x,y\in V_0,x\neq y}$ is invariant in distribution (so that $(\mu^0_{h(x)h(y)})_{x,y\in V_0,x\neq y}$ is equal in distribution to $(\mu^0_{xy})_{x,y\in V_0,x\neq y}$ for reflections h of the form described).

Finally, variations on the above mixing time convergence result can also be established for examples along the lines of those appearing in [14, Sections 7.4 and 7.5]. These include: an almost-sure statement for Vicsek set-type graphs (which complements the mixing time bounds for deterministic versions of these graphs proved in [21]); a convergence of mixing times for deterministic Sierpinski carpet graphs; and a subsequential limit for Sierpinski carpets with random weights. Since many of the ideas needed for these applications are similar to those discussed above, we omit the details.

5.3 Critical Galton-Watson trees

The connection between critical Galton-Watson processes and α-stable trees is now well-known, and so we will be brief in introducing it. Let ξ be a mean 1 random variable whose distribution is aperiodic (not supported on a sub-lattice of $\mathbb{Z}$). Furthermore, suppose that ξ is in the domain of attraction of a stable law with index $\alpha \in (1, 2)$, by which we mean that there exists a sequence $a_N \to \infty$ such that

$$\frac{\xi[N] - N}{a_N} \to \Xi, \tag{5.4}$$

in distribution, where $\xi[N]$ is the sum of N independent copies of ξ and the limit random variable satisfies $\mathbf{E}(e^{-\lambda\Xi}) = e^{-\lambda^\alpha}$. If $\mathcal{T}_N$ is a Galton-Watson tree with offspring distribution ξ conditioned to have total progeny N, then it is the case that

$$N^{-1} a_N \mathcal{T}_N \to \mathcal{T}^{(\alpha)}, \tag{5.5}$$

in distribution with respect to the Gromov-Hausdorff distance between compact metric spaces, where $\mathcal{T}^{(\alpha)}$ is an α-stable tree normalised to have total mass equal to 1 (see [32, Theorem 4.3], which is a corollary of a result originally proved in [15]). Note that the left-hand side here is shorthand for the metric space $(V(\mathcal{T}_N), N^{-1}a_N d_{\mathcal{T}_N})$, where $V(\mathcal{T}_N)$ is the vertex set of $\mathcal{T}_N$ and $d_{\mathcal{T}_N}$ is the shortest path graph distance on this set.

The α-stable tree $\mathcal{T}^{(\alpha)}$ is almost-surely a compact metric space. Moreover, there is a natural non-atomic probability measure upon it, $\pi^{(\alpha)}$ say, which has full support, and appears as the limit of the uniform measure on the approximating graph trees. Usefully, we can decompose this measure in terms of a collection of measures of level sets of the tree. More specifically, in the construction of the α-stable tree from an excursion we can naturally choose a root $\rho \in \mathcal{T}^{(\alpha)}$. We define $\mathcal{T}^{(\alpha)}(r) := \{x \in \mathcal{T}^{(\alpha)} : d_{\mathcal{T}^{(\alpha)}}(\rho, x) = r\}$ to be the collection of vertices at height r above this vertex. For almost-every realisation of $\mathcal{T}^{(\alpha)}$, there then exists a cadlag sequence of finite measures on $\mathcal{T}^{(\alpha)}$, $(\ell^r)_{r>0}$, such that ℓ^r is supported on $\mathcal{T}^{(\alpha)}(r)$ for each r and

$$\pi^{(\alpha)} = \int_0^\infty \ell^r dr$$

(see [16, Section 4.2]). Clearly this implies that $\pi^{(\alpha)}(\partial B_{\mathcal{T}^{(\alpha)}}(\rho, r)) = 0$ for every $r > 0$, for almost-every realisation of $\mathcal{T}^{(\alpha)}$. Since α-stable trees satisfy a root-invariance property (see [16, Theorem 4.8]), one can easily extend this result to hold for $\pi^{(\alpha)}$-a.e. $x \in \mathcal{T}^{(\alpha)}$. Although this is not quite the assumption of the introduction that $\pi^{(\alpha)}(\partial B_{\mathcal{T}^{(\alpha)}}(x, r)) = 0$ for every $x \in \mathcal{T}^{(\alpha)}$, $r > 0$, by a minor tweak of the proof of Proposition 3.2, we are still able to apply our mixing time convergence results in the same way.

Upon almost-every realisation of the metric measure space $(\mathcal{T}^{(\alpha)}, \pi^{(\alpha)})$, it is possible to define a corresponding Brownian motion $X^{(\alpha)}$ (to do this, apply [27, Theorem 5.4], in the way described in [10, Section 2.2]). This is a conservative $\pi^{(\alpha)}$-symmetric Hunt process, and the associated Dirichlet form $(\mathcal{E}^{(\alpha)}, \mathcal{F}^{(\alpha)})$ is actually a resistance form. Thus we can again apply Corollary 4.3 to confirm that (1.1)-(1.5) hold for some corresponding transition density, $q^{(\alpha)}$ say. Now, in [13], it was demonstrated that if $\mathbf{P}^{\mathcal{T}_N}_{\rho^N}$ is the law of the random walk on $\mathcal{T}_N$ started from its root (original ancestor) ρ_N and π^N is its stationary probability measure, then, after embedding all the objects into an underlying Banach space in a suitably nice way, the conclusion of (5.5) can be extended to the distributional convergence of

$$\left(N^{-1}a_N\mathcal{T}_N, \pi^N(Na_N^{-1}\cdot), \mathbf{P}^{\mathcal{T}_N}_{\rho^N}\left(\left(N^{-1}a_N X^{\mathcal{T}_N}_{\lfloor N^2 a_N^{-1} t\rfloor}\right)_{t\in[0,1]} \in \cdot\right)\right)$$

to $(\mathcal{T}^{(\alpha)}, \pi^{(\alpha)}, \mathbf{P}^{(\alpha)}_\rho)$, where $\mathbf{P}^{(\alpha)}_\rho$ is the law of $X^{(\alpha)}$ started from ρ. By applying the fixed starting point version of the local limit result of Proposition 2.4 (cf. [14, Theorem 1]), similarly to the argument of [14, Section 7.2], for the Brownian continuum random tree, which corresponds to the case $\alpha = 2$, one can obtain from this a distributional version of Assumption 2. (The tightness condition of (2.11) is easily checked by applying Lemma 2.5.)

Lemma 5.3. *For any compact interval $I \subset (0, \infty)$,*

$$\left(\left(V(\mathcal{T}_N), N^{-1}a_N d_{\mathcal{T}^N}, \rho^N\right), \pi^N, \left(q^N_{N^2 a_N^{-1} t}(\rho^N, x)\right)_{x\in V(\mathcal{T}_N), t\in I}\right)$$

converges in distribution to $((\mathcal{T}^{(\alpha)}, d_{\mathcal{T}^{(\alpha)}}, \rho), \pi^{(\alpha)}, (q^{(\alpha)}_t(\rho, x))_{x\in\mathcal{T}^{(\alpha)}, t\in I})$ in a spectral pointed Gromov-Hausdorff sense.

Consequently, since the space in which the above convergence in distribution occurs is separable, it is straightforward to apply Theorem 3.5 to deduce from this the following mixing time convergence result. We remark that the $\sqrt{2}$ that appears in the finite variance result is simply an artefact of the particular scaling we have described here, and could alternatively have been absorbed in the scaling of metrics.

Theorem 5.4. *Fix $p \in [1,\infty]$. If $t^p_{\mathrm{mix}}(\rho^N)$ is the L^p-mixing time of the random walk on $\mathcal{T}_N$ started from its root ρ^N, then*

$$N^{-2}a_N t^p_{\mathrm{mix}}(\rho^N) \to t^p_{\mathrm{mix}}(\rho),$$

in distribution, where $t^p_{\mathrm{mix}}(\rho) \in (0,\infty)$ is the L^p-mixing time of the Brownian motion on $\mathcal{T}^{(\alpha)}$ started from ρ. In particular, in the case when the offspring distribution has finite variance σ, it is the case that

$$\frac{\sigma}{\sqrt{2}} N^{-3/2} t^p_{\mathrm{mix}}(\rho^N) \to t^p_{\mathrm{mix}}(\rho),$$

in distribution.

Remark 5.5. We note that it was only for convenience that the convergence of the random walks on the trees $\mathcal{T}_N$, $N \geq 1$, to the Brownian motion on $\mathcal{T}^{(\alpha)}$ was proved from a single starting point in [13]. We do not anticipate any significant problems in extending this result to hold for arbitrary starting points. Indeed, the first step would be to make the obvious adaptations to the proof of [13, Lemma 4.2] to extend the result, which demonstrates convergence of simple random walks (and related additive functionals) on subtrees of $\mathcal{T}_N$ consisting of a finite number of branch segments to the corresponding continuous objects, from the case when all the random walks start from the root to an arbitrary starting point version. An argument identical to the remainder of [13, Section 4] could then be used to obtain the convergence of simple random walks on the whole trees, at least in the case when the starting point of the diffusion is in one of the finite subtrees considered. However, since the union of the finite subtrees is dense in the limiting space, we could subsequently use the heat kernel continuity properties to obtain the non-pointed spectral Gromov-Hausdorff version of Lemma 5.3. However we do not pursue this approach here as it would require a substantial amount of space and new notation that is not relevant to the main ideas of this article. Were it to be checked, though, Theorem 1.4 would imply, for any $p \in [1,\infty]$, the distributional convergence of $t^p_{\mathrm{mix}}(\mathcal{T}_N)$, the L^p-mixing time of the random walk on $\mathcal{T}_N$, when rescaled appropriately, to $t^p_{\mathrm{mix}}(\mathcal{T}^{(\alpha)}) \in (0,\infty)$, the L^p-mixing time of the Brownian motion on $\mathcal{T}^{(\alpha)}$.

Finally, we use our mixing time convergence result to establish asymptotic bounds for the distributions of mixing times of graphs in the sequence $(\mathcal{T}_N)_{N\geq 1}$ in the case when we have a finite offspring distribution.

Corollary 5.6. *In the case when the offspring distribution has finite variance, there exist constants $c_1, c_2, c_3, c_4 \in (0,\infty)$ such that*

$$\limsup_{N\to\infty} \mathbf{P}\left(N^{-3/2} t^\infty_{\mathrm{mix}}(\mathcal{T}_N) \geq \lambda\right) \leq c_1 e^{-c_2\lambda^2}, \qquad \forall \lambda \geq 0, \tag{5.6}$$

and also

$$\limsup_{N\to\infty} \mathbf{P}\left(N^{-3/2} t^1_{\mathrm{mix}}(\rho_N) \leq \lambda^{-1}\right) \leq c_3 e^{-c_4\lambda^{1/25}}, \qquad \forall \lambda \geq 0. \tag{5.7}$$

Proof. To prove (5.6), we apply the general mixing time upper bound of Lemma A.1 to deduce that

$$\mathbf{P}\left(N^{-3/2} t^\infty_{\mathrm{mix}}(\mathcal{T}_N) \geq \lambda\right) \leq \mathbf{P}\left(8N^{-1/2}\mathrm{diam}_{d_{\mathcal{T}_N}}(\mathcal{T}_N) \geq \lambda\right),$$

where $\mathrm{diam}_{d_{\mathcal{T}_N}}(\mathcal{T}_N)$ is the diameter of $\mathcal{T}_N$ with respect to $d_{\mathcal{T}_N}$, and we note that $\#E(\mathcal{T}_N)$ is equal to $2(N-1)$. By (5.5), the right-hand side here converges to $\mathbf{P}(8\,\mathrm{diam}_{d_{\mathcal{T}^{(2)}}}(\mathcal{T}^{(2)}) \geq \lambda)$. By construction, the diameter of the continuum random tree $\mathcal{T}^{(2)}$ is bounded above by twice the supremum of the Brownian excursion of length 1. We can thus use the known distribution of the latter random variable (see [25], for example) to deduce the relevant bound.

For (5.7), we first apply Theorem 5.4 to deduce that

$$\limsup_{N\to\infty} \mathbf{P}\left(N^{-3/2}t^1_{\mathrm{mix}}(\rho_N) \le \lambda^{-1}\right) \le \mathbf{P}\left(t^1_{\mathrm{mix}}(\rho) \le \lambda^{-1}\right).$$

Now, for the continuum random tree, define

$$J(\lambda) = \{r > 0 : \lambda^{-1}r^2 \le \pi^{(2)}(B_{\mathcal{T}^{(2)}}(\rho, r)) \le \lambda r^2,\ R^{(2)}_{\mathcal{T}}(\rho, B_{\mathcal{T}^{(2)}}(\rho, r)^c) \ge \lambda^{-1}r\},$$

where $R_{\mathcal{T}^{(2)}}$ is the resistance on the continuum random tree (see [11, (20)]). Then

$$\mathbf{P}(r \in J(\lambda)) \ge 1 - e^{-c\lambda}, \qquad \forall r \in (0, \tfrac{1}{2}], \lambda \ge 1,$$

(see [11, Lemmas 4.1 and 7.1]). As a consequence of this, we can apply the continuous version of the mixing time lower bound discussed in Remark A.6 (with $H_0 = 0$, $H_1 = H_2 = H_3 = 1$, $H'_2 = 3$, $\alpha_i = 1$ and $d_i = 2$) to deduce the desired result. □

Remark 5.7. The above proof already gives an estimate for the lower tail of $t^1_{\mathrm{mix}}(\rho)$. That the bound corresponding to (5.6) holds for the limiting tree, i.e.

$$\mathbf{P}\left(t^{\infty}_{\mathrm{mix}}(\mathcal{T}^{(2)}) \ge \lambda\right) \le c_1 e^{-c_2\lambda^2},$$

can be proved similarly to the discrete case (more details are given in Remark A.2).

5.4 Critical Erdős-Rényi random graph

Closely related to the random trees of the previous section is the Erdős-Rényi random graph at criticality. In particular, let $G(N, p)$ be the random graph in which every edge of the complete graph on N labelled vertices $\{1, \dots, N\}$ is present with probability p independently of the other edges. Supposing $p = N^{-1} + \lambda N^{-4/3}$ for some $\lambda \in \mathbb{R}$, so that we are in the so-called critical window, it is known that the largest connected component $\mathcal{C}^N$, equipped with its shortest path graph metric $d_{\mathcal{C}^N}$, satisfies

$$\left(V(\mathcal{C}^N), N^{-1/3}d_{\mathcal{C}^N}\right) \to (\mathcal{M}, d_{\mathcal{M}})$$

in distribution, again with respect to the Gromov-Hausdorff distance between compact metric spaces, where $(\mathcal{M}, d_{\mathcal{M}})$ is a random compact metric space [1]. (In fact, this and all the results given in this subsection hold for a family of i-th largest connected components for all $i \in \mathbb{N}$. For notational simplicity, we only discuss the largest connected component $\mathcal{C}^N$.) Moreover, in [9], it was shown that the associated random walks started from a root vertex ρ^N satisfy a distributional convergence result of the form

$$\left(N^{-1/3}X^{\mathcal{C}^N}_{\lfloor Nt\rfloor}\right)_{t\ge0} \to \left(X^{\mathcal{M}}_t\right)_{t\ge0},$$

where $X^{\mathcal{M}}$ is a diffusion on the space $\mathcal{M}$ started from a distinguished vertex $\rho \in \mathcal{M}$. Although the invariant probability measures of the random walks, π^N say, were not considered in [9], it is not difficult to extend this result to include them since the hard work regarding their convergence has already been completed (see [9, Lemma 6.3], in particular). Hence, by again applying the fixed starting point version of the local limit result of Proposition 2.4 (using Lemma 2.5 again to deduce the relevant tightness condition), we are able to obtain the analogue of Lemma 5.3 in this setting.

Lemma 5.8. *For any compact interval $I \subset (0,\infty)$,*

$$\left(\left(V(\mathcal{C}^N), N^{-1/3}d_{\mathcal{C}^N}, \rho^N\right), \pi^N, \left(q^N_{Nt}(\rho^N,x)\right)_{x\in V(\mathcal{T}_N),t\in I}\right),$$

converges in distribution to $((\mathcal{M}, d_\mathcal{M}, \rho), \pi^\mathcal{M}, (q^\mathcal{M}_t(\rho,x))_{x\in\mathcal{M},t\in I})$, where $\pi^\mathcal{M}$ is the invariant probability measure of $X^\mathcal{M}$ and $(q^\mathcal{M}_t(x,y))_{x,y\in\mathcal{M},t>0}$ is its transition density with respect to this measure, in a spectral pointed Gromov-Hausdorff sense.

In order to proceed as above, we must of course check that $\pi^\mathcal{M}$ and $q^\mathcal{M}$ satisfy a number of technical conditions. To do this, first observe that a typical realisation of $\mathcal{M}$ looks like a (rescaled) typical realisation of the Brownian continuum random tree $\mathcal{T}^{(2)}$ glued together at a finite number of pairs of points [1]. Since $\pi^\mathcal{M}$ can be considered as the image of the canonical measure $\pi^{(2)}$ on $\mathcal{T}^{(2)}$ under this gluing map, it is elementary to obtain from the statements of the previous section regarding $\pi^{(2)}$ that $\pi^\mathcal{M}$ is almost-surely non-atomic, has full support and satisfies $\pi^\mathcal{M}(\partial B_\mathcal{M}(x,r)) = 0$ for $\pi^\mathcal{M}$-a.e. $x \in \mathcal{M}$ and every $r > 0$, as desired. For $q^\mathcal{M}$, we simply observe that because the Dirichlet form corresponding to $X^\mathcal{M}$ is a resistance form ([9, Proposition 2.1]), we can once again apply Corollary 4.3 to establish conditions (1.1)-(1.5).

Given these results, pointwise mixing time convergence follows from Theorem 3.5.

Theorem 5.9. *Fix $p \in [1,\infty]$. If $t^p_{\mathrm{mix}}(\rho^N)$ is the L^p-mixing time of the random walk on $\mathcal{C}^N$ started from its root ρ^N, then*

$$N^{-1}t^p_{\mathrm{mix}}(\rho^N)\to t^p_{\mathrm{mix}}(\rho),$$

in distribution, where $t^p_{\mathrm{mix}}(\rho) \in (0,\infty)$ is the L^p-mixing time of the Brownian motion on $\mathcal{M}$ started from ρ.

Remark 5.10. As discussed in Remark 5.5, we do not expect any major barriers in extending the above result to arbitrary starting points. The first task in doing this would be to adapt the convergence result proved in [9] regarding the convergence of simple random walks on subgraphs of $\mathcal{C}^n_1$ formed of a finite number of line segments ([9, Lemma 6.4]) to arbitrary starting points. One could then extend this to obtain the desired convergence result for simple random walks on the entire space using ideas from [9, Section 7] and heat kernel continuity. It would also be necessary to introduce a new Gromov-Hausdorff-type topology to state the result, as the one used in [9] is only suitable for the pointed case. Again, we suspect taking these steps will simply be a lengthy technical exercise, and choose not to follow them through here. We do though reasonably expect that $t^p_{\mathrm{mix}}(\mathcal{C}^N)$, the L^p-mixing time of the random walk on $\mathcal{C}^N$, when rescaled appropriately, converges in distribution to $t^p_{\mathrm{mix}}(\mathcal{M}) \in (0,\infty)$, the L^p-mixing time of the Brownian motion on $\mathcal{M}$, for any $p \in [1,\infty]$.

Finally, for the largest component of the Erdős-Rényi random graph in the critical window, one can prove that there exists constants $c_1, c_2 \in (0,\infty)$ such that

$$\sup_{N\geq 1} \mathbf{P}\left(N^{-1}t^\infty_{\mathrm{mix}}(\mathcal{C}^N) \geq \lambda\right) \leq c_1 e^{-c_2\lambda}, \qquad \forall \lambda \geq 0,$$

(indeed, by [35, Proposition 1.4] and [36, Theorem 1], this result is an application of Proposition A.7 with $p_1(\lambda) = c_3 e^{-c_4\lambda^{3/2}}$ and $p_2(\lambda) = c_5 e^{-c_6\lambda^3}$), and also

$$\sup_{N\geq N_0} \mathbf{P}\left(N^{-1}t^1_{\mathrm{mix}}(\mathcal{C}^N) \leq \lambda^{-1}\right) \leq c_7\lambda^{-\theta}, \qquad \forall \lambda \geq 0, \tag{5.8}$$

for suitable constants $c_7, N_0, \theta \in (0,\infty)$ (see Proposition A.8). It does not, however, seem possible to apply current estimates for the graphs $(\mathcal{C}^N)_{N\geq 1}$ and techniques for bounding mixing times to replace $t^1_{\mathrm{mix}}(\mathcal{C}^N)$ by $t^1_{\mathrm{mix}}(\rho^N)$ in the latter estimate (see Remark A.9), or even prove that the sequence $(N/t^1_{\mathrm{mix}}(\rho^N))_{N\geq 1}$ is tight, i.e.

$$\lim_{\lambda\to\infty}\limsup_{N\to\infty}\mathbf{P}\left(N^{-1}t^1_{\mathrm{mix}}(\rho^N)\leq\lambda^{-1}\right)=0.$$

That this final statement is nonetheless true is a simple consequence of Theorem 5.9.

5.5 Random walk on range of random walk in high dimensions

Let $S=(S_n)_{n\geq 0}$ be the simple random walk on $\mathbb{Z}^d$ started from 0, built on an underlying probability space with probability measure $\mathbf{P}$, and define the range of S up to time N to be the graph G^N with vertex set

$$V(G^N):=\{S_n:0\leq n\leq N\}\,,\tag{5.9}$$

and edge set

$$E(G^N):=\{\{S_{n-1},S_n\}:1\leq n\leq N\}\,.\tag{5.10}$$

In this section, we will explain how to prove that if $d\geq 5$, which is an assumption henceforth, then the mixing times of the sequence of graphs $(G^N)_{N\geq 1}$ grows asymptotically as cN^2, $\mathbf{P}$-a.s., where c is a deterministic constant. Since doing this primarily depends on making relatively simple adaptations of the high-dimensional scaling limit result of [12] for the random walk on the entire range of S (i.e. the $N=\infty$ case) to the finite length setting, we will be brief with the details.

First, suppose that $S=(S_n)_{n\in\mathbb{Z}}$ is a two-sided extension of $(S_n)_{n\geq 0}$ such that $(S_{-n})_{n\geq 0}$ is an independent copy of $(S_n)_{n\geq 0}$. The set of cut-times for this process,

$$\mathcal{T}:=\left\{n:S_{(-\infty,n]}\cap S_{[n+1,\infty)}=\emptyset\right\},$$

is known to be infinite $\mathbf{P}$-a.s. ([17]). Thus we can write $\mathcal{T}=\{T_n:n\in\mathbb{Z}\}$, where $\dots T_{-1}<T_0\leq 0<T_1<T_2<\dots$. The corresponding set of cut-points is given by $C:=\{C_n:n\in\mathbb{Z}\}$, where $C_n:=S_{T_n}$. For these objects, an ergodicity argument can be applied to obtain that, $\mathbf{P}$-a.s., as $|n|\to\infty$,

$$\frac{T_n}{n}\to\tau(d):=\mathbf{E}(T_1|0\in\mathcal{T})\in[1,\infty),\tag{5.11}$$

$$\frac{d_G(0,C_n)}{|n|}\to\delta(d):=\mathbf{E}(d_G(0,C_1)|0\in\mathcal{T})\in[1,\infty),$$

where d_G is the shortest path graph distance on the range G of the entire two-sided walk S, which is defined analogously to (5.9) and (5.10). In particular, see [12, Lemma 2.2], for a proof of the same convergence statements under the measure $\mathbf{P}(\cdot|0\in\mathcal{T})$, and note that the conditioning can be removed by using the relationship between $\mathbf{P}$ and $\mathbf{P}(\cdot|0\in\mathcal{T})$ described in [12, Lemma 2.1]. Given these results, it is an elementary exercise to check that the metric space $(V(G^N),\tau(d)\delta(d)^{-1}N^{-1}d_{G^N})$, where d_{G^N} is the shortest path graph distance on G^N, converges $\mathbf{P}$-a.s. with respect to the Gromov-Hausdorff distance to the interval $[0,1]$ equipped with the Euclidean metric. Moreover, the same ideas readily yield an extension of this result to a spectral Gromov-Hausdorff one including that π^N, the invariant measure of the associated simple random walk, converges to Lebesgue measure on $[0,1]$.

Now, for a fixed realisation of G, let $X=(X_n)_{n\geq 0}$ be the simple random walk on G started from 0. Define the hitting times by X of the set of cut-points $\mathcal{C}$ by $H_0:=\min\{m\geq 0:X_m\in\mathcal{C}\}$, and, for $n\geq 1$, $H_n:=\min\{m>H_{n-1}:X_m\in\mathcal{C}\}$. We use these times to define a useful indexing process $Z=(Z_n)_{n\geq 0}$ taking values in $\mathbb{Z}$. In particular, if $n<H_0$, define Z_n to be the unique $k\in\mathbb{Z}$ such that $X_{H_0}=C_k$. Similarly, if $n\in[H_{m-1},H_m)$ for some $m\geq 1$, then define Z_n to be

the unique $k \in \mathbb{Z}$ such that $X_{H_m} = C_k$. Noting that this definition precisely coincides with the definition of Z in [12], from Lemma 3.5 of that article we have that: for $\mathbf{P}$-a.e. realisation of G,

$$\left(N^{-1}\tau(d)Z_{\lfloor tN^2 \rfloor}\right)_{t\geq 0} \to (B_{t\kappa_2(d)})_{t\geq 0}, \tag{5.12}$$

in distribution, where $(B_t)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}$ started from 0, and $\kappa_2(d) \in (0,\infty)$ is the deterministic constant defined in [12]. To deduce from (5.12) the following scaling limit for X^N, the simple random walk on G^N, we proceed via a time-change argument that is essentially a reworking of parts of [12, Section 3].

Lemma 5.11. *For $\mathbf{P}$-a.e. realisation of S, if X^N is started from 0, then*

$$\left(\tau(d)\delta(d)^{-1}N^{-1}d_{G^N}\left(0, X^N_{\lfloor \kappa_2(d)^{-1}N^2 t\rfloor}\right)\right)_{t\geq 0} \to \left(B^{[0,1]}_t\right)_{t\geq 0},$$

in distribution, where $B^{[0,1]} = (B^{[0,1]}_t)_{t\geq 0}$ is Brownian motion on $[0,1]$ started at 0 and reflected at the boundary.

Proof. The following proof can be applied to any typical realisation of S. To begin with, define a process $(A^{Z,N}_n)_{n\geq 0}$ by setting

$$A^{Z,N}_n := \sum_{m=0}^{n-1} \mathbf{1}_{\{Z_m \in [0, T^{-1}_N]\}},$$

where $T^{-1}_N := \max\{n : T_n \leq N\}$. From (5.11), we have that $T^{-1}_N \sim \tau(d)^{-1}N$. Combining this observation with (5.12), one can check that, simultaneously with (5.12), $(N^{-2}A^N_{\lfloor tN^2\rfloor})_{t\geq 0}$ converges in distribution to $(\kappa_2(d)^{-1}A^B_{t\kappa_2(d)})_{t\geq 0}$, where

$$A^B_t := \int_0^t \mathbf{1}_{\{B_s \in [0,1]\}} ds$$

(cf. [12, Lemma 3.5]).

We now apply the above result to establish a scaling limit for the process X observed on the vertex set $V(\tilde{G}^N) := \{S_n : T_1 \leq n \leq T^{-1}_N\}$. Specifically, set

$$A^N_n := \sum_{m=0}^{n-1} \mathbf{1}_{\{X_m, X_{m+1} \in V(\tilde{G}^N)\}}.$$

Similarly to the proof of [12, Lemma 3.6], one can check that

$$\sup_{0\leq m\leq n} \left|A^N_m - A^{Z,N}_m\right| \leq \sum_{m=0}^{n} \mathbf{1}_{\{Z_m \in [0,1,2]\cup[T^{-1}_N-2, T^{-1}_N-1, T^{-1}_N]\}}.$$

It is therefore a simple consequence of (5.12) that $N^{-2}\sup_{0\leq m\leq TN^2}\left|A^N_m - A^{Z,N}_m\right|$ converges to 0 in probability as $N\to\infty$ for any $T\in(0,\infty)$. Since we know from equation (16) of [12] that

$$N^{-1}\sup_{0\leq m\leq TN^2} \left|d_G(0, X_m) - \delta(d)Z_m\right|$$

also converges to 0 in probability, we readily obtain

$$\left(\tau(d)\delta(d)^{-1}N^{-1}d_G\left(0, \tilde{X}^N_{\lfloor N^2 t\rfloor}\right)\right)_{t\geq 0} \to \left(B^{[0,1]}_{\kappa_2(d)t}\right)_{t\geq 0}, \tag{5.13}$$

in distribution, where $\tilde{X}^N = (\tilde{X}^N_n)_{n\geq 0}$ is the random walk X observed on $V(\tilde{G}^N)$ – this is defined precisely by setting $\tilde{X}^N_n := X_{\alpha^N(n)}$, where $\alpha^N(n) := \max\{A^N_m \leq n\}$. We remark that the particular limit process $B^{[0,1]}$ arises as a consequence of the fact that $(B_{\alpha^B(t)})_{t\geq 0}$, where α^B is the right-continuous inverse of A^B, has exactly the distribution of $B^{[0,1]}$.

Finally, since the process $\tilde{X}^N$ is identical in law to the simple random walk X^N observed on $V(\tilde{G}^N)$, to replace $\tilde{X}^N$ by X^N in (5.13) it will suffice to check that X^N spends only an asymptotically negligible amount of time in $V(G^N)\backslash V(\tilde{G}^N)$. Since doing this requires only a simple adaptation of the proof of [12, Lemma 3.8], we omit the details. To complete the proof, one then needs to replace d_G by d_{G^N}, but this is straightforward since

$$N^{-1}\sup_{0\leq n\leq N}|d_G(0,S_n)-d_{G^N}(0,S_n)|\leq N^{-1}\left(T_1+T_{T_N^{-1}+1}-T_{T_N^{-1}}\right)\to 0,$$

as $N\to\infty$. □

Although the previous lemma only contains a convergence statement for the random walks started from the particular vertex 0, there is no difficulty in extending this to the case when X^N is started from a point $x^N_0 \in V(G^N)$ such that $d_{G^N}(0,x^N_0)\sim \tau(d)^{-1}\delta(d)Nx_0$, and $B^{[0,1]}$ is started from $x_0\in[0,1]$. Applying the local limit result of Proposition 2.4 (to establish (2.11), we once again appeal to Lemma 2.5), we are able deduce from this that Assumption 1 holds for $\mathbf{P}$-a.e. realisation of the original random walk.

Lemma 5.12. *For $\mathbf{P}$-a.e. realisation of S, if $I\subset(0,\infty)$ is a compact interval, then*

$$\left(\left(V(G^N),\tau(d)\delta(d)^{-1}N^{-1}d_{G^N}\right),\pi^N,\left(q^N_{\kappa_2(d)^{-1}N^2t}(x,y)\right)_{x,y\in V(G^N),t\in I}\right),$$

converges in $(\mathcal{M}_I,\Delta_I)$ to the triple consisting of: $[0,1]$ equipped with the Euclidean metric, Lebesgue measure on this set and the transition density of Brownian motion on $[0,1]$ reflected at the boundary.

Since it is clear that (1.1)-(1.5) hold in this case, we can therefore apply Theorem 1.4 to obtain the desired convergence of mixing times.

Theorem 5.13. *Fix $p\in[1,\infty]$. If $t^p_{\mathrm{mix}}(S_{[0,N]})$ is the L^p-mixing time of the simple random walk on the range of S up to time N, then $\mathbf{P}$-a.s.,*

$$\kappa_2(d)N^{-2}t^p_{\mathrm{mix}}(S_{[0,N]})\to t^p_{\mathrm{mix}}([0,1]),$$

where $t^p_{\mathrm{mix}}([0,1])$ is the L^p-mixing time of the Brownian motion on $[0,1]$ reflected at the boundary.

A Appendix: Mixing time estimates

In this appendix, we give some sufficient condition to derive upper and lower estimates for mixing times of random walks on finite graphs, primarily using techniques adapted from [35]. We will also indicate how these can be transferred to the continuous setting (see Remarks A.2 and A.6). We start by fixing our notation. Let $G=(V(G),E(G))$ be a finite connected graph and μ^G be a weight function, as in the introduction. Suppose here that d_G is the shortest path metric on the graph G, and denote, for a distinguished vertex $\rho\in V(G)$,

$$B(R)=\{y:d_G(\rho,y)<R\},\ V(R):=\sum_{x\in B(R)}\sum_{y:y\sim x}\mu^G_{xy}=\pi^G(B(R))\mu(G),\quad R\in(0,\infty),$$

where we write $x \sim y$ if $\mu^G_{xy} > 0$ and set $\mu(G) := \sum_{x,y \in V(G)} \mu^G_{xy}$. For the Markov chain X^G, let

$$\tau_R = \tau_{B(\rho,R)} = \min\{n \geq 0 : X^G_n \notin B(R)\}.$$

We define a quadratic form $\mathcal{E}$ by

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

and let $H^2 = \{f \in \mathbb{R}^{V(G)} : \mathcal{E}(f,f) < \infty\}$. For disjoint subsets A, B of G, the effective resistance between them is then given by:

$$R_{\text{eff}}(A,B)^{-1} = \inf\{\mathcal{E}(f,f) : f \in H^2, f|_A = 1, f|_B = 0\}. \tag{A.1}$$

If we further define $R_{\text{eff}}(x,y) = R_{\text{eff}}(\{x\},\{y\})$, and $R_{\text{eff}}(x,x) = 0$, then one can check that $R_{\text{eff}}(\cdot,\cdot)$ is a metric on $V(G)$ (see [28, Section 2.3]). We will call this the resistance metric. The resistance metric enjoys the following important (but easy to deduce) estimate,

$$|f(x) - f(y)|^2 \leq R_{\text{eff}}(x,y)\mathcal{E}(f,f), \qquad \forall f \in L^2(G, \mu^G).$$

Moreover, it is easy to verify that if $c_1^{-1} := \inf_{x,y \in G : x \sim y} \mu^G_{xy} > 0$, then

$$R_{\text{eff}}(x,y) \leq c_1 d_G(x,y) \qquad \forall x,y \in G. \tag{A.2}$$

Let $v, r : \{0, 1, \cdots, \operatorname{diam}_{d_G}(G) + 1\} \to [0,\infty)$ be strictly increasing functions with $v(0) = r(0) = 0$, $v(1) = r(1) = 1$, which satisfy

$$C_1^{-1}\Big(\frac{R}{R'}\Big)^{d_1} \leq \frac{v(R)}{v(R')} \leq C_1\Big(\frac{R}{R'}\Big)^{d_2}, \quad C_2^{-1}\Big(\frac{R}{R'}\Big)^{\alpha_1} \leq \frac{r(R)}{r(R')} \leq C_2\Big(\frac{R}{R'}\Big)^{\alpha_2}$$

for all $0 < R' \leq R \leq \operatorname{diam}_{d_G}(G) + 1$, where $C_1, C_2 \geq 1$, $1 \leq d_1 \leq d_2$ and $0 < \alpha_1 \leq \alpha_2 \leq 1$. In what follows, $v(\cdot)$ will give the volume growth order and $r(\cdot)$ the resistance growth order. For convenience, we extend them to functions on $[0, \operatorname{diam}_{d_G}(G) + 1]$ by linear interpolation.

Finally, in the appendix, we adopt the convention that if we cite elsewhere the constant c_1 of Lemma A.4 (for example), we denote it as $c_{A.4.1}$.

A.1 Upper bound

In this subsection, we give an upper bound of the mixing times that is a reworking of [35, Corollary 4.2], in our setting. Note that, since $t^p_{\text{mix}}(\rho) \leq t^p_{\text{mix}}(G)$ and $t^p_{\text{mix}}(G) \leq t^{p'}_{\text{mix}}(G)$ for $p \leq p'$, it will be enough to estimate $t^\infty_{\text{mix}}(G)$.

Lemma A.1. *For any weighted graph (G, μ^G),*

$$t^\infty_{\text{mix}}(G) \leq 4\operatorname{diam}_R(G)\mu(G),$$

where $\operatorname{diam}_R(G)$ is the diameter of G with respect to the resistance metric R_{eff}.

Proof. First, note that by [2, Proposition 3 in Chapter 2], we have that

$$\mathbf{E}^G_x\left(\sum_{m=0}^{\infty} \mathbf{1}_{\{X^G_m = x, m < S\}}\right) = \pi(x)\mathbf{E}^G_x(S), \tag{A.3}$$

for any stopping time S with $X_S^G = x$. Taking S to be the first hitting time of x after time $2m-1$, and writing $\Pi(x, 2m)$ to represent the law of X_{2m}^G when X^G is started from x, we obtain that

$$\mathbf{E}_{\Pi(x,2m)}^G(\sigma_x) = \sum_{l=0}^{2m-1} \left(p_l^G(x,x) - 1\right) = 2\sum_{l=0}^{m-1} \left(q_{2l}^G(x,x) - 1\right) \geq 2m\left(q_{2m}^G(x,x) - 1\right),$$

where σ_x is the first hitting time of x, and the inequality holds because $q_{2l}^G(x,x)$ is decreasing in l (see the proof of [14, Lemma 9], for example). Since by Cauchy-Schwarz, $|q_{2m}^G(x,y) - 1| \leq (q_{2m}^G(x,x)-1)^{1/2}(q_{2m}^G(y,y)-1)^{1/2}$, it follows that

$$\sup_{x\in V(G)} D_\infty^G(x,2m) = \sup_{x,y\in V(G)} \left|q_{2m}^G(x,y) - 1\right| \leq \sup_{x\in V(G)} (q_{2m}^G(x,x) - 1) \leq \sup_{x,y\in V(G)} \frac{\mathbf{E}_x^G(\sigma_y)}{2m}. \quad \text{(A.4)}$$

By applying the commute time identity for random walks on graphs, $\mathbf{E}_x^G(\sigma_y) + \mathbf{E}_y^G(\sigma_x) = R_{\mathrm{eff}}(x,y)\mu(G)$, this implies $\sup_{x\in V(G)} D_\infty^G(x,2m) \leq \mathrm{diam}_R(G)\mu(G)/2m$, and the result follows. □

Remark A.2. (1) The first inequality of (A.4) and (3.5) implies the following known fact for mixing times of symmetric Markov chains; $t_{\mathrm{mix}}^\infty(G) \leq 2\, t_{\mathrm{mix}}^2(G;1/2)$, where $t_{\mathrm{mix}}^2(G;1/2)$ is the L^2-mixing time of G with $1/2$ instead of $1/4$ in the definition (1.8).
(2) Essentially the same argument can be applied to deduce the corresponding mixing time upper bound in the continuous setting when we suppose that we have a process whose Dirichlet form is a resistance form. In particular, suppose that this is the case for X^F, as defined in the introduction. Let S be the first hitting time of $x \in F$ after time t, then, for any $f \in L^1(F,\pi)$,

$$\mathbf{E}_x\left(\int_0^S f(X_s)ds\right) \quad = \quad \|f\|_{L^1(\pi)}\mathbf{E}_x(S),$$

which can be obtained by applying an ergodicity argument similar to that used to prove (A.3). Writing $\Pi(x,t)$ to represent the law of X_t^F when X^F is started from x, the expectation on the right-hand side here satisfies $\mathbf{E}_x(S) = t + \mathbf{E}_{\Pi(x,t)}(\tau_x) \leq t + \sup_{y\in F} R_{\mathrm{eff}}(x,y)$, where to deduce the upper bound, we have applied that the commute time identity $\mathbf{E}_x(\tau_y) + \mathbf{E}_y(\tau_x) = R_{\mathrm{eff}}(x,y)$ also holds for resistance forms (since we are assuming π to be a probability measure, it does not appear explicitly in this version of the identity). Moreover, if f is positive, the left-hand side is bounded below as follows: $\mathbf{E}_x(\int_0^S f(X_s)ds) \geq \int_0^t \int_F q_s(x,y)f(y)\pi(dy)ds$. Combining these bounds, we have proved that, for positive $f \in L^1(F,\pi)$ such that $\|f\|_{L^1(\pi)} \neq 0$,

$$\frac{\int_0^t \int_F q_s(x,y)f(y)\pi(dy)ds}{\|f\|_{L^1(\pi)}} \leq t + \mathrm{diam}_R(F).$$

By choosing a sequence of suitable functions whose support converges to $\{x\}$, the joint continuity of $(q_t(x,y))_{x,y\in F,t>0}$ allows us to deduce from this that

$$tq_t(x,x) \leq \int_0^t q_s(x,x)ds \leq t + \mathrm{diam}_R(F),$$

where the first inequality holds because $q_t(x,x)$ is decreasing in t. The remainder of the proof is identical to the graph case.

A.2 Lower bound

In this subsection, we give the mixing time lower bound. By the same reasoning as in the first paragraph of the previous subsection, it will be enough to estimate $t^1_{\rm mix}(\rho)$. Our argument depends on some estimates for hitting times that are modifications of results in [3, 31].

To begin with, let $B = B(R)$ and define

$$g_B(x,y) = \mu_y^{-1} \sum_{k=0}^{\infty} \mathbf{P}_x^G(X_k = y, k < \tau_B).$$

Then, it is easy to show that

$$\mathbf{E}_z^G \tau_B = \sum_{y\in B} g_B(z,y)\mu_y, \qquad R_{\rm eff}(x, B^c) = g_B(x,x)$$

(see, for example [3, (2.19),(2.20)]). Also, if A and B are disjoint subsets of G and $x \notin A \cup B$, then (see [3, (2.14)])

$$\mathbf{P}_x^G(T_A < T_B) \le \frac{R_{\rm eff}(x,B)}{R_{\rm eff}(x,A)}, \tag{A.5}$$

where T_A is the hitting time of $A \subset G$. If $C_3 := 2^{-2/\alpha_1} C_2^{-1/\alpha_2}$ and $C_4 := 8^{-1} C_1^{-1} C_3^{d_2}$, we can then prove the following.

Lemma A.3. *Let $\lambda \ge 1$ and $H_0, \cdots, H_3 > 0$.*
(a) Suppose that

$$R_{\rm eff}(\rho, y) \le \lambda^{H_0} r(d_G(\rho,y)), \ \forall y \in B(R), \quad \textit{and} \quad V(R) \le \lambda^{H_1} v(R), \tag{A.6}$$

then

$$\mathbf{E}_x^G \tau_R \le 2\lambda^{H_0+H_1} v(R) r(R) \quad \textit{for } x \in B(R). \tag{A.7}$$

(b) Suppose (A.6) *and also*

$$R_{\rm eff}(\rho, B(R)^c) \ge \lambda^{-H_2} r(R) \quad \textit{and} \quad V(C_3\lambda^{-(H_0+H_2)/\alpha_1}R) \ge \lambda^{-H_3} v(C_3\lambda^{-(H_0+H_2)/\alpha_1}R), \tag{A.8}$$

then

$$\mathbf{E}_x^G \tau_R \ge 2C_4 \lambda^{-H_2'-H_3} v(R) r(R) \quad \textit{for } x \in B(C_3\lambda^{-(H_0+H_2)/\alpha_1}R), \tag{A.9}$$

where $H_2' = H_2 + (H_0+H_2)d_2/\alpha_1$.
(c) Suppose (A.6) *and* (A.8)*, and let $x \in B(C_3\lambda^{-(H_0+H_2)/\alpha_1}R)$, then*

$$\mathbf{P}_x^G(\tau_R > n) \ge \frac{2C_4\lambda^{-H_2'-H_3} v(R) r(R) - n}{2\lambda^{H_0+H_1} v(R) r(R)} \quad \textit{for } n \ge 0. \tag{A.10}$$

Proof. Using (A.6), we have $R_{\rm eff}(z, B^c) \le R_{\rm eff}(0,z) + R_{\rm eff}(0, B^c) \le 2\lambda^{H_0} r(R)$ for any $z \in B$. So,

$$\mathbf{E}_z^G \tau_B = \sum_{y\in B} g_B(z,y)\mu_y \le \sum_{y \in B} g_B(z,z)\mu_y = R_{\rm eff}(z, B^c) V(R) \le 2\lambda^{H_0+H_1} v(R) r(R),$$

which gives (A.7). In order to prove (A.9), we first establish the following: for $0 < \varepsilon \le 1/(2C_2\lambda^{H_0+H_2})^{1/\alpha_1} = 2^{1/\alpha_1} C_3 \lambda^{-(H_0+H_2)/\alpha_1}$ and $y \in B(\varepsilon R)$, we have

$$\mathbf{E}_y^G(T_\rho < \tau_R) \ge 1 - \frac{C_2\varepsilon^{\alpha_1}\lambda^{H_0+H_2}}{1 - C_2\varepsilon^{\alpha_1}\lambda^{H_0+H_2}} \ge 1 - 2C_2\varepsilon^{\alpha_1}\lambda^{H_0+H_2}. \tag{A.11}$$

Indeed, by the first inequalities of (A.6) and (A.8), we have

$$R_{\mathrm{eff}}(y,B(R)^c)\geq R_{\mathrm{eff}}(\rho,B(R)^c)-R_{\mathrm{eff}}(\rho,y)\geq\lambda^{-H_2}r(R)-\lambda^{H_0}r(\varepsilon R)\geq\frac{r(\varepsilon R)}{C_2\varepsilon^{\alpha_1}\lambda^{H_2}}-\lambda^{H_0}r(\varepsilon R).$$

So, by (A.5),

$$\mathbf{P}^G_y(\tau_R<T_\rho)\leq\frac{R_{\mathrm{eff}}(y,\rho)}{R_{\mathrm{eff}}(y,B(R)^c)}\leq\frac{\lambda^{H_0}r(\varepsilon R)}{\frac{r(\varepsilon R)}{C_2\varepsilon^{\alpha_1}\lambda^{H_2}}-\lambda^{H_0}r(\varepsilon R)}\leq\frac{C_2\varepsilon^{\alpha_1}\lambda^{H_0+H_2}}{1-C_2\varepsilon^{\alpha_1}\lambda^{H_0+H_2}},$$

and (A.11) is obtained. Now, if $y\in B'=B(C_3\lambda^{-(H_0+H_2)/\alpha_1}R)$, then the bound at (A.11) gives that $\mathbf{P}^G_y(T_\rho<\tau_B)\geq\frac12$, so

$$g_B(\rho,y)=g_B(\rho,\rho)\mathbf{P}^G_y(T_\rho<\tau_B)\geq\tfrac12 g_B(\rho,\rho)=\tfrac12 R_{\mathrm{eff}}(\rho,B^c)\geq\tfrac12\lambda^{-H_2}r(R).$$

By the second inequality of (A.8), we have

$$\mu(B')\geq\lambda^{-H_3}v(C_3\lambda^{-(H_0+H_2)/\alpha_1}R)\geq C_1^{-1}C_3^{d_2}\lambda^{-2(H_0+H_2)d_2/\alpha_1-H_3}v(R),$$

and therefore we obtain,

$$\begin{aligned}\mathbf{E}^G_\rho\tau_B&\geq\sum_{y\in B'}g_B(\rho,y)\mu_y\\&\geq\tfrac12 g_B(\rho,\rho)\mu(B')\\&\geq\tfrac12 C_1^{-1}C_3^{d_2}\lambda^{-H_2-(H_0+H_2)d_2/\alpha_1-H_3}v(R)r(R)\\&=4C_4\lambda^{-H_2'-H_3}v(R)r(R).\end{aligned}$$

Moreover, for $x\in B'$ we have that $\mathbf{E}^G_x\tau_B\geq\mathbf{P}^G_x(T_\rho<\tau_B)\mathbf{E}^G_\rho\tau_B$, which gives (A.9).
Finally, by the Markov property, (A.7) and (A.9),

$$\begin{aligned}2C_4\lambda^{-H_2'-H_3}v(R)r(R)\leq\mathbf{E}^G_x\tau_R&\leq n+\mathbf{E}^G_x[\mathbf{1}_{\{\tau_R>n\}}\mathbf{E}^G_{X_n}(\tau_R)]\\&\leq n+2\lambda^{H_0+H_1}v(R)r(R)\mathbf{P}^G_x(\tau_R>n).\end{aligned}$$

Rearranging this gives (A.10). □

The following estimate is a modification of [31, Proposition 3.5 (a)] (see [3, (2.4)] for the important special case $v(R)=R^2$, $r(R)=R$). Note that for $R>\mathrm{diam}_{d_G}(G)$, it is the case that $\tau_R=\infty$, and so (A.12) trivially holds.

Proposition A.4. *Let* $0<\varepsilon\leq C_3\lambda^{-(H_0+H_2)/\alpha_1}$, *and suppose* (A.6) *and* (A.8) *for* R *and* εR, *then*

$$\mathbf{P}^G_y\big(\tau_R\leq C_4\lambda^{-H_2'-H_3}v(\varepsilon R)r(\varepsilon R)\big)\leq c_1\lambda^{H_0+\sum_{i=0}^3H_i+H_2'}\varepsilon^{\alpha_1},\qquad\text{for } y\in B(\varepsilon R).\tag{A.12}$$

Proof. We take a kind of bootstrap from (A.10) and (A.11). Let $t_0>0$, and set

$$q(y)=\mathbf{P}^G_y(\tau_R\leq T_\rho),\qquad a(y)=\mathbf{P}^G_y(\tau_R\leq t_0).$$

Then

$$a(y) = \mathbf{P}^G_y(\tau_R \le t_0) = \mathbf{P}^G_y(\tau_R \le t_0, \tau_R \le T_\rho) + \mathbf{P}^G_y(\tau_R \le t_0, \tau_R > T_\rho)$$
$$\le \mathbf{P}^G_y(\tau_R \le T_\rho) + \mathbf{P}^G_y(T_\rho < \tau_R, \tau_R - T_\rho \le t_0)$$
$$\le q(y) + (1 - q(y))a(\rho) \le q(y) + a(\rho), \tag{A.13}$$

using the strong Markov property for the second inequality. Starting the Markov chain X at ρ, we have

$$a(\rho) = \mathbf{P}^G_\rho(\tau_R \le t_0) \le \mathbf{E}^G_\rho[1_{\{\tau_{\varepsilon R} \le t_0\}} \mathbf{P}^G_{X_{\tau_{\varepsilon R}}}(\tau_R \le t_0)] \le \mathbf{P}^G_\rho(\tau_{\varepsilon R} \le t_0) \max_{y \in \partial B(\varepsilon R)} a(y). \tag{A.14}$$

Combining (A.13) and (A.14) gives

$$a(\rho) \le \frac{\max_{y \in \partial B(\varepsilon R)} q(y)}{\mathbf{P}^G_\rho(\tau_{\varepsilon R} > t_0)}. \tag{A.15}$$

Further, using (A.11) with 2ε, we have

$$q(y) \le \frac{C_2(2\varepsilon)^{\alpha_1}\lambda^{H_0+H_2}}{1 - C_2(2\varepsilon)^{\alpha_1}\lambda^{H_0+H_2}} \le 2C_2(2\varepsilon)^{\alpha_1}\lambda^{H_0+H_2}. \tag{A.16}$$

Let $t_0 = C_4\lambda^{-H_2'-H_3}v(\varepsilon R)r(\varepsilon R)$; then using (A.10) for the ball $B(\varepsilon R)$ (note that (A.6) and (A.8) for εR are assumed here), we obtain

$$\mathbf{P}^G_\rho(\tau_{\varepsilon R} > t_0) \ge c_0\lambda^{-H_0-H_1-H_2'-H_3}.$$

combining this with (A.16), (A.15) and (A.13) completes the proof of (A.12). □

Note that, we may and will take $c_1 > 1/(2C_3^{\alpha_1})$ in (A.12). Using Proposition A.4, we have the following.

Proposition A.5. *i) For $\lambda, R > 1$, assume that $\mu(G) \ge 4V(R)$, and that* (A.6), (A.8) *hold for R, then*

$$t^1_{\mathrm{mix}}(G) > C_4\lambda^{-H_2'-H_3}v(R)r(R). \tag{A.17}$$

ii) For $\lambda, R > 1$, assume that $\mu(G) \ge 4V(R)$, and (A.6), (A.8) *hold for R and $\varepsilon_0(\lambda)R$, where $\varepsilon_0(\lambda) := (2c_{A.4.1})^{-1/\alpha_1}\lambda^{-(H_0+\sum_{i=0}^3 H_i+H_2')/\alpha_1}$. Then*

$$t^1_{\mathrm{mix}}(\rho) > C_4\lambda^{-H_2'-H_3}v(\varepsilon_0(\lambda)R)r(\varepsilon_0(\lambda)R).$$

Proof. i) We follow the argument in [35, Lemma 5.4]. Let $t \in \mathbb{N}$. If $\mathbf{P}^G_x(\tau_B \le t) \ge 1/2$ for all $x \in B(R-1)$, then τ_R/t is stochastically dominated by a geometric random variable with parameter $1/2$, so that $\mathbf{E}^G_\rho[\tau_R] \le 2t$. By this and (A.9), we see that for $t = C_4\lambda^{-H_2'-H_3}v(R)r(R)$, there exists some $x \in B(R-1)$ such that $\mathbf{P}^G_x(\tau_B \le t) \le 1/2$. Further, since $\mu(G) \ge 4V(R)$, $\pi(B(R)) = V(B(R))/\mu(G) \le 1/4$. Combining these observations, we obtain

$$D_1(x,t) \ge 2\mathbf{P}^G_x(\tau_R \ge t) - 2\pi(B(R)) \ge 1 - \frac{1}{2} > \frac{1}{4}, \tag{A.18}$$

so that (A.17) follows.

ii) Take $\varepsilon = \varepsilon_0(\lambda)$ in Proposition A.4 and let $t = C_4\lambda^{-H_2'-H_3}v(\varepsilon R)r(\varepsilon R)$. Then, since $0 < \varepsilon \le C_3\lambda^{-(H_0+H_2)/\alpha_1}$ (this is because we take $c_{A.4.1} > 1/(2C_3^{\alpha_1})$), by (A.12) we have $\mathbf{P}^G_\rho(\tau_R \le t) \le c_{A.4.1}\lambda^{H_0+\sum_{i=0}^3 H_i+H_2'}\varepsilon^{\alpha_1} = 1/2$. The rest is the same as the proof of i) except that we take $x = \rho$ in (A.18). □

Remark A.6. Similarly to the remark at the end of Section A.1, an analogous argument for proving a mixing time lower bound applies in the continuous case when we have a process X^F whose Dirichlet form is a resistance form on F. In order to avoid repetition, we omit the details of this here.

A.3 Random graph case

We now consider a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ carrying a family of random weighted graphs $G^N(\omega) = (V(G^N(\omega)), E(G^N(\omega)), \mu^{N(\omega)}; \omega \in \Omega)$. We assume that, for each $N \in \mathbb{N}$ and $\omega \in \Omega$, $G^N(\omega)$ is a finite, connected graph containing a marked vertex ρ^N, and $\#V(G^N(\omega)) \le M_N$ for some non-random constant $M_N < \infty$. Let $d_{G^N(\omega)}(\cdot,\cdot)$ be a graph distance, $B(R) := B_\omega(\rho^N, R)$, and $V(R) := V_\omega(\rho^N, R)$. We write $X = (X_n, n \ge 0, P_\omega^x, x \in G^N(\omega))$ for the random walk on $G^N(\omega)$, and denote by $p_n^\omega(x,y)$ its transition density with respect to π^ω. Furthermore, we introduce a strictly increasing function $h : \mathbb{N} \cup \{0\} \to [0,\infty)$ with $h(0) = 0$, which will roughly describe the diameter of G^N with respect to the graph distance. We then set $\gamma(\cdot) = v(h(\cdot)) \cdot r(h(\cdot))$. Finally, for $i = 1, 2$, we suppose $p_i : [1,\infty) \to [0,1]$ are functions such that $\lim_{\lambda\to\infty} p_i(\lambda) = 0$. We then have the following.

Proposition A.7. *(1) Suppose that the following holds:*

$$\mathbf{P}(\mathrm{diam}_R(G^N) \ge \lambda r(h(N))) \le p_1(\lambda), \quad \mathbf{P}(\mu^N(G^N) \ge \lambda v(h(N))) \le p_2(\lambda), \tag{A.19}$$

then

$$\mathbf{P}(t_{\mathrm{mix}}^\infty(G^N) \ge \lambda\gamma(N)) \le \inf_{\theta\in[0,1]} (p_1(\lambda^\theta/8) + p_2(\lambda^{1-\theta})).$$

(2) Suppose there exist $c_1 \le 1$ and $J \ge (1+H_1)/d_2$ such that the following holds:

$$\mathbf{P}((\text{A.6}) \wedge (\text{A.8}) \textit{ for } R = c_1\lambda^{-J}h(N)) \ge 1 - p_1(\lambda), \quad \mathbf{P}(\mu^N(G^N) < \lambda^{-1}v(h(N))) \le p_2(\lambda),$$

then there exist $c_2, p_0 > 0$ such that

$$\mathbf{P}(t_{\mathrm{mix}}^1(G^N) \le c_2\lambda^{-p_0}\gamma(N)) \le 2p_1(\lambda) + p_2(\lambda/(4C_1c_1^{d_2})).$$

(3) Suppose there exist $c_1 \le 1$ and $J \ge (1+H_1)/d_2$ such that the following holds:

$$\begin{aligned}&\mathbf{P}((\text{A.6}) \wedge (\text{A.8}) \textit{ for } R = c_1\lambda^{-J}h(N) \textit{ and for } \varepsilon_0(\lambda)R) \ge 1 - p_1(\lambda),\\ &\mathbf{P}(\mu^N(G^N) < \lambda^{-1}v(h(N))) \le p_2(\lambda),\end{aligned} \tag{A.20}$$

where $\varepsilon_0(\lambda)$ is as in Proposition A.5 ii), then there exist $c_2, p_0 > 0$ such that

$$\mathbf{P}(t_{\mathrm{mix}}^1(\rho^N) \le c_2\lambda^{-p_0}\gamma(N)) \le 2p_1(\lambda) + p_2(\lambda/(4C_1c_1^{d_2})).$$

Proof. By Lemma A.1, we have for any $\theta \in [0,1]$ that

$$\begin{aligned}\mathbf{P}\left(t_{\mathrm{mix}}^\infty(G^N) \ge \lambda\gamma(N)\right) &\le \mathbf{P}\left(8\mathrm{diam}_R(G^N)\mu^N(G^N) \ge \lambda\gamma(N)\right)\\ &\le \mathbf{P}\left(8\mathrm{diam}_R(G^N) \ge \lambda^\theta r(h(N))\right) + \mathbf{P}\left(\mu^N(G^N) \ge \lambda^{1-\theta}v(h(N))\right)\\ &\le p_1(\lambda^\theta/8) + p_2(\lambda^{1-\theta}),\end{aligned}$$

which implies the conclusion of (1).

For (2), let $R = c_1\lambda^{-J}h(N)$ and define

$$\begin{aligned} t &:= C_4\lambda^{-H_2'-H_3}v(R)r(R) = C_4\lambda^{-H_2'-H_3}v(c_1\lambda^{-J}h(N))r(c_1\lambda^{-J}h(N)) \\ &\geq C_4\lambda^{-H_2'-H_3}C_1^{-1}C_2^{-2}(c_1\lambda^{-J})^{d_2+\alpha_2}v(h(N))r(h(N)) =: c_2\lambda^{-p_0}\gamma(N). \end{aligned}$$

Then by Proposition A.5 i),

$$\begin{aligned} &\mathbf{P}(t^1_{\text{mix}}(G^N) \leq c_1\lambda^{-p_0}\gamma(N)) \leq \mathbf{P}(t^1_{\text{mix}}(G^N) \leq t) \\ &\quad\leq \mathbf{P}(\text{either (A.6) or (A.8) do not hold for } R = c_1\lambda^{-J}h(N)) + \mathbf{P}(\mu^N(G^N) < 4V(R)) \\ &\quad\leq p_1(\lambda) + \mathbf{P}(\mu^N(G^N) < 4V(R)). \end{aligned}$$

Note that

$$4\lambda^{H_1}v(R) = 4\lambda^{H_1}v(c_1\lambda^{-J}h(N)) \leq 4\lambda^{H_1}C_1(c_1\lambda^{-J})^{d_2}v(h(N)) \leq 4C_1c_1^{d_2}\lambda^{-1}v(h(N)),$$

where we used $J \geq (1 + H_1)/d_2$ in the last inequality. Using this, we have

$$\begin{aligned} &\mathbf{P}(\mu^N(G^N) < 4V(R)) \\ &\quad\leq \mathbf{P}(\mu^N(G^N) < 4\lambda^{H_1}v(R)) + \mathbf{P}(\lambda^{H_1}v(R) \leq V(R)) \\ &\quad\leq \mathbf{P}(\mu^N(G^N) < 4C_1c_1^{d_2}\lambda^{-1}v(h(N))) + p_1(\lambda) \\ &\quad\leq p_2(\lambda/(4C_1c_1^{d_2})) + p_1(\lambda), \end{aligned}$$

which implies the conclusion of (2). The proof of (3) is almost the same, so we omit it. □

To illustrate this result, we consider the case when the random graphs $G^N(\omega)$ are obtained as components of percolation processes on finite graphs, thereby recovering [35, Theorem 1.2(c)]. (In [35], it was actually the lazy random walk was considered to avoid parity concerns, but the same techniques apply when we consider $q^G_m(\cdot,\cdot)$ as in (1.7) instead.) We note that, this setting includes taking $G^N(\omega)$ to be $\mathcal{C}^N$, the largest component of the Erdős-Rényi random graph in the critical window, as introduced in Section 5.4, and hence the following proposition establishes the estimate at (5.8).

Proposition A.8. *Let G^N be a graph with N vertices and with the maximum degree $d \in [3, N-1]$. Let $\mathcal{C}^N$ be the largest component of the percolation subgraph of G^N for $0 < p < 1$. Let $p \leq \frac{1+\lambda n^{-1/3}}{d-1}$ for some fixed $\lambda \in \mathbb{R}$, and assume that there exist $c_1, \theta_1 \in (0,\infty)$ and $K_1 \in \mathbb{N}$ such that*

$$\mathbf{P}(\#\mathcal{C}^N \leq A^{-1}N^{2/3}) \leq c_1A^{-\theta_1}, \qquad \forall A, N \geq K_1, \tag{A.21}$$

then there exist $c_2, \theta_2 \in (0,\infty)$ and $K_2 \in \mathbb{N}$ such that, for all $p \in [1,\infty]$,

$$\mathbf{P}(A^{-1}N \leq t^p_{\text{mix}}(\mathcal{C}^N) \leq AN) \geq 1 - c_2A^{-\theta_2}, \qquad \forall A, N \geq K_2. \tag{A.22}$$

Proof. We only indicate how to apply previous propositions. First, the upper bound of $t^p_{\text{mix}}(\mathcal{C}^N)$ can be obtained by Proposition A.7 (1) with $v(R) = R^2, r(R) = R, h(N) = N^{1/3}$ and $p_1(A) = c_0A^{-q_0}, p_2(A) = c_0'A^{-q_0'}$ for some $c_0, c_0', q_0, q_0' > 0$. Indeed, (A.19) holds because of [35, Theorem 2.1 (a),(b), Theorem 6.1] and the fact $\operatorname{diam}(\mathcal{C}^N) \geq \operatorname{diam}_R(\mathcal{C}^N)$, which is due to (A.2).

The lower bound is more complicated. Using Proposition 5.5–5.7 and (5.1) in [35] with

$$\beta = \lambda^{-1/4}, L = \lambda^{H_2}, \alpha = \lambda^{H_1}, r = R, h = C_3\lambda^{-H_2}R, m = \lambda^{-H_3}(C_3\lambda^{-H_2}R)^2,$$

and then taking $R = c_1\lambda^{-J}N^{1/3}$, $H_0 = 0$ (due to (A.2)), $H_1 = H_2 = 2, H_3 = 4, J = (1+H_1)/2 = 3/2$, we see that for each $v \in G^N$,

$$\mathbf{P}(\sharp\mathcal{C}(v) > \lambda^{-1/4}N^{2/3} \text{ and } \mathcal{A}) \le c_4\lambda^{-1/2}N^{-1/3},$$

where

$$\mathcal{A} = \{V(v, C_3\lambda^{-2}R) \le \lambda^{-5}(C_3\lambda^{-2}R)^2,\ R_{\text{eff}}(v, B(v,R)^c) \le \frac{R}{8\lambda^2},\ \#E(B(v,R)) \ge \lambda^2R^2\}.$$

This corresponds to [35, (5.3)]. Now using Proposition A.5 i) and arguing similarly to the proof of [35, Theorem 2.1 (c.2)], we have

$$\mathbf{P}(\exists v \in G^N \text{ with } \sharp\mathcal{C}(v) > \lambda^{-1/4}N^{2/3} \text{ and } t^1_{\text{mix}}(\mathcal{C}(v)) \le C_4\lambda^{-29/2}N) \le c_4\lambda^{-1/4}.$$

This together with (A.21) implies the desired lower bound of $t^p_{\text{mix}}(\mathcal{C}^N)$. □

The proofs of this proposition and Corollary 5.6 highlight why it is useful to have a general theory where the exponents $H_0, \cdots, H_3$ can vary.

Remark A.9. As mentioned at the end of Subsection 5.4, it does not seem possible to apply current estimates for the graphs $(\mathcal{C}^N)_{N\ge1}$ and techniques for bounding mixing times to replace $A^{-1}N \le t^p_{\text{mix}}(\mathcal{C}^N)$ by $A^{-1}N \le t^p_{\text{mix}}(\rho^N)$ in (A.22). The major difficulty is to verify the first inequality of (A.20) for $\varepsilon_0(\lambda)R$. Indeed, even if we choose $H_0, \cdots, H_3$ large (which increases the chance that (A.6) and (A.8) hold for R), $\varepsilon_0(\lambda)$ gets small accordingly, so that the probability $\mathbf{P}(\text{(A.6)} \wedge \text{(A.8) for } \varepsilon_0(\lambda)R)$ does not increase.

Finally, below is a list of exponents for each example in Section 5.

Section	$v(R)$	$r(R)$	$h(N)$	$\gamma(N)$
5.2	$R^{\log K/\log L}$	$R^{\log\lambda/\log L}$	L^N	$(K\lambda)^N$
5.3 with $a_N = N^{1/\alpha}, \alpha \in (1,2]$	$R^{\alpha/(\alpha-1)}$	R	$N^{1-1/\alpha}$	$N^{2-1/\alpha}$
5.4	R^2	R	$N^{1/3}$	N
5.5	R	R	N	N^2

Here the Euclidean distance is used instead of the intrinsic shortest path metric for the examples in Section 5.2. Note that when $\alpha = 2$ in Section 5.3 (the finite variance case), the growth of $v(R)$ and $r(R)$ is of the same order as in Section 5.4. The difference of scaling exponents of mixing times (namely $\gamma(N)$) is due to the difference of scaling exponents for graph distances (namely $h(N)$). We also observe that the convergence to a stable law at (5.4) forces the scaling constants to be of the form $a_N = N^{1/\alpha}L(N)$ for some slowly varying function L (see [20, Section 35]), and hence the above table captures all the most important first order behaviour for the examples in Section 5.3.

References

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt, *The continuum limit of critical random graphs*, Probab. Theory Related Fields, to appear.

[2] D. Aldous and J. Fill, *Reversible Markov chains and random walks on graphs*, Preprint `http://www.stat.berkeley.edu/~aldous/RWG/book.html`

[3] M.T. Barlow, A.A. Járai, T. Kumagai and G. Slade, *Random walk on the incipient infinite cluster for oriented percolation in high dimensions*, Comm. Math. Phys. **278** (2008), 385–431.

[4] I. Benjamini, G. Kozma and N. Wormald, *The mixing time of the giant component of a random graph*, preprint.

[5] P. Bérard, G. Besson and S. Gallot, *Embedding Riemannian manifolds by their heat kernel*, Geom. Funct. Anal. **4** (1994), 373–398.

[6] C. Borgs, J.T. Chayes, R. van der Hofstad, G. Slade and J. Spencer, Random subgraphs of finite graphs: I. The scaling window under the triangle condition, *Random Structures Algorithms,* **27** 137–184, 2005.

[7] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001.

[8] G.-Y. Chen and L. Saloff-Coste, *The cutoff phenomenon for ergodic Markov processes*, Electron. J. Probab. **13** (2008), 26–78.

[9] D. A. Croydon, *Scaling limit for the random walk on the largest connected component of the critical random graph*, Preprint.

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

[11] ——, *Volume growth and heat kernel estimates for the continuum random tree*, Probab. Theory Related Fields **140** (2008), 207–238.

[12] ——, *Random walk on the range of random walk*, J. Stat. Phys. **136** (2009), 349–372.

[13] ——, *Scaling limits for simple random walks on random ordered graph trees*, Adv. in Appl. Probab. **42** (2010), 528–558.

[14] D. A. Croydon and B. M. Hambly, *Local limit theorems for sequences of simple random walks on graphs*, Potential Anal. **29** (2008), 351–389.

[15] T. Duquesne, *A limit theorem for the contour process of conditioned Galton-Watson trees*, Ann. Probab. **31** (2003), 996–1027.

[16] T. Duquesne and J.-F. Le Gall, *Probabilistic and fractal aspects of Lévy trees*, Probab. Theory Related Fields **131** (2005), 553–603.

[17] P. Erdős and S. J. Taylor, *Some intersection properties of random walk paths*, Acta Math. Acad. Sci. Hungar. **11** (1960), 231–248.

[18] P. J. Fitzsimmons, B. M. Hambly, and T. Kumagai, *Transition density estimates for Brownian motion on affine nested fractals*, Comm. Math. Phys. **165** (1994), 595–620.

[19] N. Fountoulakis and B.A. Reed, *The evolution of the mixing rate of a simple random walk on the giant component of a random graph*, Random Structures Algorithms, **33** (2008), 6886.

[20] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1954, Translated and annotated by K. L. Chung. With an Appendix by J. L. Doob.

[21] S. Goel, R. Montenegro, and P. Tetali, *Mixing time bounds via the spectral profile*, Electron. J. Probab. **11** (2006), 1–26.

[22] A. Greven, P. Pfaffelhuber, and A. Winter, *Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees)*, Probab. Theory Related Fields **145** (2009), 285–322.

[23] M. Heydenreich and R. van der Hofstad, *Random graph asymptotics on high-dimensional tori II. Volume, diameter and mixing time*, Probab. Theory Related Fields, to appear.

[24] A. Kasue and H. Kumura, *Spectral convergence of Riemannian manifolds*, Tôhoku Math. J. **46** (1994), 147–179.

[25] D. P. Kennedy, *The distribution of the maximum Brownian excursion*, J. Appl. Probab. **13** (1976), 371–376.

[26] J. Kigami, *Resistance forms, quasisymmetric maps and heat kernel estimates*, Memoirs AMS, to appear.

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

[28] ———, *Analysis on fractals*, Cambridge Tracts in Mathematics, vol. 143, Cambridge University Press, Cambridge, 2001.

[29] T. Kumagai, *Homogenization on finitely ramified fractals*, Stochastic analysis and related topics in Kyoto, Adv. Stud. Pure Math., vol. 41, Math. Soc. Japan, Tokyo, 2004, pp. 189–207.

[30] T. Kumagai and S. Kusuoka, *Homogenization on nested fractals*, Probab. Theory Related Fields **104** (1996), 375-398.

[31] T. Kumagai and J. Misumi, *Heat kernel estimates for strongly recurrent random walk on random media*, J. Theoret. Probab. **21** (2008), 910–935.

[32] J.-F. Le Gall, *Random real trees*, Ann. Fac. Sci. Toulouse Math. (6) **15** (2006), 35–62.

[33] D. Levin, Y. Peres and E. Wilmer, *Markov chains and mixing times*, Amer. Math. Soc., Providence, RI, 2009.

[34] V. Metz, *Renormalization contracts on nested fractals*, J. Reine Angew. Math. **480** (1996), 161–175.

[35] A. Nachmias and Y. Peres, *Critical random graphs: diameter and mixing time*, *Ann. Probab.* **36** (2008), 1267–1286.

[36] A. Nachmias and Y. Peres, *The critical random graph, with martingales*, Israel J. Math. 176 (2010), 29–41.

[37] C. Sabot, *Existence and uniqueness of diffusions on finitely ramified self-similar fractals*, Ann. Sci. École Norm. Sup. (4) **30** (1997), 605–673.