

Scaling limit for Brownian motions on the l -level Sierpinski gaskets: The fractal to Euclidean crossover*

David A. Croydon[†] Ben Hambly[‡] Takashi Kumagai[§]

Abstract

In two dimensions, the l -level Sierpinski gasket $SG(l)$ is obtained by splitting an equilateral triangle into a collection of l^2 equilateral triangles of equal size and with the same total area, retaining only the $l(l+1)/2$ triangles with the same orientation as the original triangle, and then iterating this procedure indefinitely. We show that the canonical diffusions on the spaces $SG(l)$, $l \geq 2$, can be rescaled to yield Brownian motion on the initial triangle. Our argument also applies to the analogous higher-dimensional Sierpinski gaskets. Moreover, we prove a local central limit theorem for the associated transition densities. Key to this is the derivation of a Poincaré inequality, in the proof of which we exploit the Euclidean-type mixing that occurs between the bottlenecks present at each scale of the fractal.

Keywords: Sierpinski gasket; fractal; Brownian motion; scaling limit; transition density.

MSC2020 subject classifications: Primary 28A80, Secondary 60J25; 60J35; 60K50.

Submitted to EJP on September 25, 2025, final version accepted on February 27, 2026.

1 Introduction

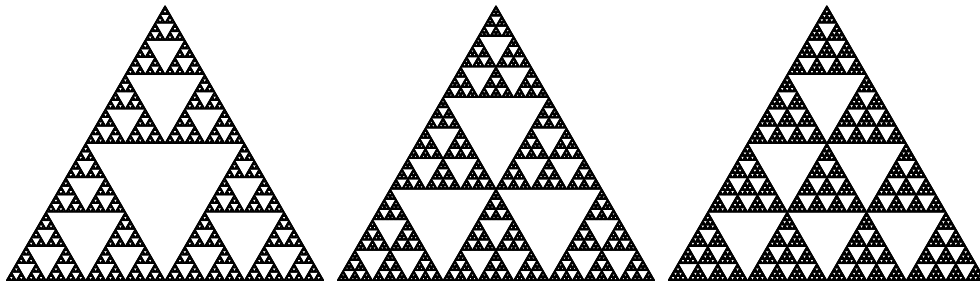
It is now well known that heat flow on fractal spaces, such as Sierpinski gaskets, has clear differences to that on classical Euclidean space. Indeed, unlike the diffusive behaviour of Brownian motion on the latter spaces, the canonical diffusions on many fractals have been shown to be sub-diffusive, with walk dimension strictly larger than 2. (We recall that the walk dimension is the polynomial exponent governing the space-time scaling of the process in question, see (1.1) below for a precise definition in our setting.) In the physics literature, see [30] in particular, the question was raised as to how the

*This research was supported by JSPS Grant-in-Aid for Scientific Research (A) 22H00099, 23KK0050, (C) 19K03540, 24K06758, and the Research Institute for Mathematical Sciences, an International Joint Usage/Research Center located in Kyoto University.

[†]Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606–8502, Japan. E-mail: croydon@kurims.kyoto-u.ac.jp

[‡]Mathematical Institute, University of Oxford, United Kingdom. E-mail: hambly@maths.ox.ac.uk

[§]Department of Mathematics, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan. E-mail: t-kumagai@waseda.jp

Figure 1: The sets $SG(2)$, $SG(3)$ and $SG(4)$ in two dimensions.

crossover from ‘fractal’ to ‘Euclidean’ behaviour occurs when one considers a sequence of fractals increasingly well-approximating Euclidean space. Establishing rigorously a result conjectured in [30], in [19], this question was studied at the level of certain exponents describing the laws of the related stochastic processes. It was established that these exponents, which relate to the short-time, microscopic behaviour of the relevant diffusions, retain the fractal aspects of the model, even asymptotically. (We expand on this discussion in Theorem 1.1 and Remarks 1.4 and 1.6 below.) In this article, we consider the same issue, but at the level of a functional scaling limit. In contrast to the story told in [19] and [30], we demonstrate that for a certain sequence of fractal spaces and a certain scaling regime, the fractality of the spaces is no longer preserved in the limit. Rather, we see that the associated sub-diffusive processes have as a scaling limit a standard Brownian motion on Euclidean space (restricted to lie within a suitable domain). Additionally, we present a corresponding result for the transition densities of the processes in question.

To describe the results in [19] and our main results, let us start by introducing some key notation; full definitions are postponed to Section 2. We will work in d -dimensional Euclidean space, $d \geq 2$, but usually suppress d from the notation as this parameter does not vary. First, let V_0 be the vertices of a d -dimensional regular tetrahedron Δ of side length 1, one of the vertices of which is 0. For each integer $l \geq 2$, we set $SG(l)$ to be the d -dimensional l -level Sierpinski gasket with boundary V_0 (see Figure 1), and suppose that $X^l = (X_t^l)_{t \geq 0}$ is the canonical diffusion, i.e. Brownian motion, on this space, started from 0; the time-scaling of X^l will be fixed so that the expected commute time between 0 and each other vertex of Δ is equal to 1. For each $l, d \geq 2$, the following limit, called the walk dimension, is known to exist:

$$d_w(l) := \left(\lim_{t \rightarrow 0} \frac{\log \mathbf{E} \left(|X_t^l|^2 \right)}{2 \log t} \right)^{-1}; \quad (1.1)$$

moreover, it has been established that $d_w(l) \in (2, \infty)$, see [20, Proposition 5.3], for example. Additionally, each of the processes X^l admits a transition density (or heat kernel) with respect to the canonical flat probability measure on the underlying space, $(p_t^l(x, y))_{x, y \in SG(l), t > 0}$ say. The spectral dimension is defined to be the exponent governing the asymptotic scaling in the eigenvalues and it typically determines the behaviour of the on-diagonal part of the heat kernel at short times. In particular, for our purposes we can define

$$d_s(l) := \lim_{t \rightarrow 0} \frac{2 \log p_t^l(0, 0)}{\log t}.$$

It is known that the spectral dimension exists for Brownian motion on $SG(l)$. The following is the main result of [19].

Theorem 1.1. ([19, Corollary 2.3]) *For each $d \geq 2$, the following hold as $l \rightarrow \infty$.*

$$d_w(l) = \begin{cases} 2 + \frac{\log \log l}{\log l} + \frac{O(1)}{\log l}, & \text{if } d = 2; \\ d + \frac{O(1)}{\log l}, & \text{if } d \geq 3. \end{cases} \quad (1.2)$$

$$d_s(l) = \begin{cases} 2 - \frac{\log \log l}{\log l} + \frac{O(1)}{\log l}, & \text{if } d = 2; \\ 2 - \frac{O(1)}{\log l}, & \text{if } d \geq 3. \end{cases} \quad (1.3)$$

As this theorem shows, $d_w(l) \rightarrow d$ and $d_s(l) \rightarrow 2$ as $l \rightarrow \infty$, meaning the fractal aspects of the model are retained asymptotically for $d \geq 3$. In our first conclusion, we show that this is not the case for the sample paths of the processes when they are observed on an appropriate time scale. To present this result, we suppose that $X^\Delta = (X_t^\Delta)_{t \geq 0}$ is standard Brownian motion on Δ , reflected at the boundary. Unless otherwise stated, we assume that all stochastic processes start from 0. (This is not an essential requirement, but eases the presentation.)

Theorem 1.2. *There exists a sequence of scaling parameters $(\tau^l)_{l \geq 2}$ taking values in $(0, \infty)$ such that, as $l \rightarrow \infty$,*

$$(X_{\tau^l t}^l)_{t \geq 0} \rightarrow (X_t^\Delta)_{t \geq 0}$$

in distribution with respect to the local uniform topology on $C([0, \infty), \Delta)$, the space of continuous functions on Δ (equipped with the Euclidean distance).

To place this result into context, we note there have been substantial developments in the literature for understanding the scaling limits of processes on fractal-like spaces as described by so-called ‘resistance forms’. See [8, 10, 22], for example. However, we can not immediately apply the results in those works to deduce Theorem 1.2 since, whilst the processes X^l , $l \geq 2$, are associated with resistance forms, the limiting process X^Δ is not. Indeed, the resistance metric (see Subsection 2.2 for a definition) degenerates in the limit, and so fails to capture the behaviour of the process X^Δ .

The scaling factors $(\tau^l)_{l \geq 2}$ of Theorem 1.2 will initially be defined in terms of a certain hitting time, see (2.7) below. Combining our definition of τ^l with (1.2), we have the following result. (In the following, we write $f(l) \asymp g(l)$ if there exist $c_1, c_2 > 0$ such that $c_1 g(l) \leq f(l) \leq c_2 g(l)$ for all l , and $f(l) \sim g(l)$ if $\lim_{l \rightarrow \infty} f(l)/g(l)$ exists.)

Theorem 1.3. *It holds that*

$$\tau^l = \frac{1}{2} l^{2-d_w(l)} \asymp \begin{cases} \frac{1}{\log l}, & \text{if } d = 2; \\ l^{2-d}, & \text{if } d \geq 3. \end{cases}$$

In particular, Theorem 1.3 shows that $\tau^l \rightarrow 0$ as $l \rightarrow \infty$, and so we are slowing down the diffusions X^l by an increasing amount as $l \rightarrow \infty$. Intuitively this makes sense as the diffusions X^l hit the vertices of Δ in finite time, whereas the process X^Δ does not.

Remark 1.4. (a) The \asymp in the statement of Theorem 1.3 is a direct consequence of (1.2). It would be interesting to determine whether \asymp could be replaced by \sim ; this would require showing that the $O(1)$ terms in (1.2) are of the form $C + o(1)$ for some constant $C \in \mathbb{R}$. (See Remark 3.3 below for further discussion of this point.)
 (b) We highlight that the walk dimension of the limiting process X^Δ is 2, independent of the dimension. Hence, apart from in dimension two, the walk dimensions $d_w(l)$ do not converge to the walk dimension of X^Δ as $l \rightarrow \infty$. In particular, the fractality of the spaces persists in the walk dimension, but not at the level of processes on the time scale of Theorem 1.2. This difference can be explained by noting that the walk dimension is given by the arbitrarily short-time asymptotics of the diffusions X^l , whereas the scaling limit of Theorem 1.2 depends of the behaviour of X^l at a specified (albeit still short) time scale, as is required to enable the process to suitably explore the space.

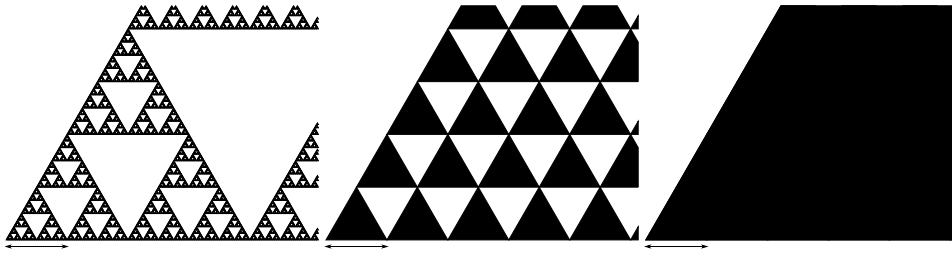


Figure 2: The set $SG_\infty(2)$, the (Hausdorff) limit of the sets $SG_\infty(l)$ (as $l \rightarrow \infty$) and the convex hull of $SG_\infty(l)$ (for any l). The marked line segment shows the distance between vertices in V_0 .

- (c) One might also consider the behaviour of X^l on the infinite versions of the fractals, as given by $SG_\infty(l) := \cup_{n=1}^\infty l^n SG(l)$. In this case, the limiting space is naturally given by an infinite collection of tetrahedra only connected at vertices, see Figure 2 for the two-dimensional case. On the time-scales considered in Theorem 1.2, the natural diffusions on $SG_\infty(l)$, $X^{l,\infty}$ say, will take an increasingly long time to leave the initial tetrahedron. Indeed, the expected time it takes $X_{\tau^l}^{l,\infty}$ to hit a vertex in V_0 other than its starting location is $1/2\tau^l$. Since $1/\tau^l$ diverges, it is easy to see that the limit process must be trapped in the initial tetrahedron forever and, in fact, one readily sees that $X_{\tau^l}^{l,\infty} \rightarrow X^\Delta$ by the same argument as Theorem 1.2. This situation is to be expected given that, as already noted, the limiting Brownian motion X^Δ never hits points.

On the other hand, by considering the process $X^{l,\infty}$ at hitting times of tetrahedral vertices in the limiting space, it is an elementary exercise to check that, for any sequence $(\varepsilon_l)_{l \geq 2}$ such that $l^{-1} \ll \varepsilon_l \ll 1$,

$$\left(\varepsilon_l X_{\varepsilon_l^{-2}t}^{l,\infty}\right)_{t \geq 0} \rightarrow \left(X_t^\Delta\right)_{t \geq 0}$$

in distribution, where X^Δ is the Brownian motion on the cone given by the convex hull of $SG_\infty(l)$ (for any l), reflected at the boundary. (Suitably avoiding the lattice cases when the constants ε_l takes values in $\{l^{-n} : n \in \mathbb{N}\}$, a similar claim can be made for other sequences $\varepsilon_l \rightarrow 0$.)

Finally, if one considers the processes $X^{l,\infty}$ without time-rescaling, we expect that one will see the processes mix within each tetrahedron in asymptotically zero time, and then, after an asymptotically exponential amount of time, jump to an adjacent tetrahedron, uniformly amongst the possible neighbours. (One could regard the process as a continuous-time Markov chain on the graph representing the part of the tetrahedral lattice obtained when one considers each tetrahedron as a point.) In particular, we conjecture that a functional scaling limit with respect to the uniform topology will not hold. However, it should be possible to describe the finite-dimensional distributions of the limiting process in terms of a continuous time random walk on tetrahedra in such a way that, conditional on where this walk is, the limiting process is uniformly distributed on the relevant tetrahedron. We leave checking the details of such a result as an open problem.

As already advertised, in addition to a functional scaling limit, we can also give a scaling limit for the transition densities of X^l , $l \geq 2$. Recall that $(p_t^l(x, y))_{x, y \in SG(l), t > 0}$ is the transition density of Brownian motion on $SG(l)$. Similarly, we write $(p_t^\Delta(x, y))_{x, y \in \Delta, t > 0}$ for the transition density of X^Δ with respect the d -dimensional Hausdorff measure on Δ

normalised to be a probability measure. We are then able to show the following uniform convergence.

Theorem 1.5. *For every compact interval $I \subseteq (0, \infty)$, it holds that*

$$\lim_{l \rightarrow \infty} \sup_{x, y \in \text{SG}(l)} \sup_{t \in I} |p_{\tau^l t}^l(x, y) - p_t^\Delta(x, y)| = 0.$$

Remark 1.6. (a) The proof of Theorem 1.5 highlights the point at which the transition density evolves from the fractal scaling to the Euclidean. Indeed, in Lemma 4.6 below, we show that, for $t \leq l^{-d_w(l)}$, the on-diagonal supremum $\sup_{x \in \text{SG}(l)} p_t^l(x, x)$ is bounded above by an expression of the form $t^{-d_f(l)/d_w(l)}$ (multiplied by a suitable l -dependent constant), whereas for $t \geq l^{-d_w(l)}$, the bound is of the form $t^{-d/2}$. Whilst we only need the upper bound, we expect that the nature of the t dependence is precise (though possibly there will be a different l -dependent constant needed in a corresponding lower bound). This is natural to expect, since $l^{-d_w(l)}$ is precisely the time-scale on which the diffusion X^l leaves a level 1 cell, and thus moves beyond the fractal scale to the more Euclidean structure in which these cells are arranged.

- (b) A similar remark to Remark 1.4(b) could be made concerning the spectral dimensions $d_s(l)$ of the diffusions X^l . As noted in (1.3), $d_s(l) \rightarrow 2$ as $l \rightarrow \infty$. Since the limiting process X^Δ has spectral dimension d , there is only continuity in these exponents in dimension 2. Again, comparing the results of [19] and this paper, we see that the fractality of the spaces is preserved on arbitrarily short time scales, but not on the time scales given by τ^l (which is much longer than the time-scale $l^{-d_w(l)}$ identified in the previous comment). In fact, in the arguments used to prove Theorem 1.5, we are required to consider the interplay between the fractal and Euclidean nature of the spaces not just at the first level, but between each scale level along the sequence l^{-m} , $m \geq 0$.
- (c) Whilst we do not pursue it here, we believe that by appealing to the general framework of [27], it should also be possible to show convergence of the rescaled eigenvalues and eigenfunctions of the generators of the processes X^l , $l \geq 2$, to those associated with the limiting process X^Δ .

The remainder of the article is organised as follows. In Section 2, we present some basic background regarding the construction of the sets $\text{SG}(l)$ and the associated processes X^l , as well as the limiting process X^Δ . Section 3 contains the proof of Theorem 1.2, a key input to which is an estimate on the diameter of $\text{SG}(l)$ with respect to an effective resistance metric. Using some symmetries of the underlying objects, the proof of Theorem 1.3 is also set out in this section. Next, in Section 4, we show the equicontinuity of the rescaled transition densities of the processes X^l , and as a consequence obtain Theorem 1.5. Finally, in Section 5, we discuss some further examples to which the approach of this article should be applicable, and some for which additional work seems to be required. Unless specifically noted, constants such as c, C, \dots may change from line to line.

2 Preliminaries

In this section, we introduce some of the key notation that will be used throughout the rest of the article. As already noted, we will typically suppress the dimensional dependence from the notation.

2.1 The l -level Sierpinski gaskets

To construct an l -level Sierpinski gasket in d dimensions, we start by considering a d -dimensional regular tetrahedron Δ of side length 1, with one vertex equal to 0.

(For each dimension d , the set Δ will be fixed for all l .) For each $l \in \mathbb{N}$, let $N_{d,l}$ be the maximal number of translations of $l^{-1}\Delta$ that can be placed inside Δ and only intersect at vertices. There is only one arrangement that achieves this; we denote the corresponding translations of $l^{-1}\Delta$ by $(\Delta_j)_{j=1}^{N_{d,l}}$. For example, when $d = 2$, these are the collection of upward pointing triangles that appear when Δ is tiled by triangles a factor l smaller. (A similar description can also be given in higher dimensions.) As noted in [19], it holds that $N_{2,l} = l(l+1)/2$, and $N_{d,l}$ can be computed for $d \geq 3$ via the formula

$$N_{d,l} = \sum_{k=1}^l N_{d-1,k}.$$

It is a simple exercise in induction (on d) to check that this implies

$$N_{d,l} = \binom{d+l-1}{d}. \quad (2.1)$$

Now, for each $j = 1, \dots, N_{d,l}$, let ψ_j be the orientation preserving similitude from Δ to Δ_j . For $l \geq 2$, the l -level Sierpinski gasket that we will consider, $SG(l)$, is then the unique compact set K satisfying

$$K = \bigcup_{j=1}^{N_{d,l}} \psi_j(K).$$

See Figure 1 for illustrations of $SG(2)$, $SG(3)$ and $SG(4)$ in two dimensions.

The set $SG(l)$ is readily computed to have Hausdorff dimension given by

$$d_H(SG(l)) = \frac{\log N_{d,l}}{\log l}.$$

Moreover, the corresponding $d_H(SG(l))$ -dimensional Hausdorff measure on $SG(l)$ is a non-trivial, finite measure of full support. We will denote by μ^l the version of this measure that has been normalized to have total mass equal to one. In particular, it holds that $\mu^l(SG(l) \cap \Delta_j) = 1/N_{d,l}$ for each $j = 1, \dots, N_{d,l}$.

In our subsequent arguments, it will be convenient to label the ‘vertices’ of each $SG(l)$. As noted in the introduction, we will denote by V_0 the vertices of the original regular tetrahedron Δ . We will further define the i -th level vertices inductively by setting

$$V_{i+1}^l := \bigcup_{j=1}^{N_{d,l}} \psi_j(V_i^l).$$

Note in particular that if we consider V_1^l to be the vertices of a graph, with edges between nearest-neighbour vertices, then V_1^l is simply the restriction of a tetrahedral lattice with edge length equal to l^{-1} to those vertices within Δ .

2.2 Resistance and Brownian motion

To construct our Brownian motion $X^l = (X_t^l)_{t \geq 0}$ on $SG(l)$, we will apply the general machinery of [22] (see also [23]), which covers a much wider collection of self-similar fractals. (We note that Brownian motion was constructed and detailed heat kernel estimates were obtained earlier on nested fractals, which is a class of fractals including $SG(l)$. See [12, 13, 14, 18, 24, 29], for instance.) We start by considering the bilinear form on V_0 given by setting:

$$\mathcal{E}_0(f, f) := \frac{1}{d+1} \sum_{x,y \in V_0} (f(x) - f(y))^2.$$

In particular, the above expression represents the energy dissipated in the network created when wires of resistance $(d + 1)/2$ are placed between each pair of vertices in V_0 and vertices are held at voltages according to f . The choice of normalisation is made so that the corresponding effective resistance, as defined by setting $R_0(x, x) := 0$ and, for $x \neq y$,

$$R_0(x, y) := (\inf \{ \mathcal{E}_0(f, f) : f(x) = 0, f(y) = 1 \})^{-1}, \tag{2.2}$$

is simply the discrete metric on V_0 , i.e. $R_0(x, y) = \mathbf{1}_{\{x \neq y\}}$. For each $d, l \geq 2$, it is then possible to deduce the existence of a constant $\rho_l = \rho_{d,l} > 1$ such that if a sequence of bilinear forms $(\mathcal{E}_i^l)_{i \geq 1}$ is defined on the sets $(V_i^l)_{i \geq 1}$, respectively, by setting

$$\mathcal{E}_{i+1}^l(f, f) := \rho_l \sum_{j=1}^{N_{d,l}} \mathcal{E}_i^l(f \circ \psi_j, f \circ \psi_j), \tag{2.3}$$

where $\mathcal{E}_0^l := \mathcal{E}_0$, then the corresponding effective resistances R_i^l , defined analogously to (2.2), satisfy

$$R_i^l(x, y) = R_{i'}^l(x, y) \tag{2.4}$$

for every $i > i'$ and $x, y \in V_{i'}^l$. This compatibility means it is possible to introduce a metric R^l on $\cup_{i=1}^\infty V_i^l$ such that $R^l = R_i^l$ on each V_i^l (and $R^l = R_0$ on V_0). By [22, Theorem 3.3.4], one can use continuity to extend R^l to a metric on the whole $SG(l)$ that is equivalent to the Euclidean metric (see [13, Remark 3.7]). Moreover, as per [22, Definition 2.3.2], R^l is a resistance metric on $SG(l)$ and, by [22, Theorem 2.3.6], there is an associated resistance form $(\mathcal{E}^l, \mathcal{F}^l)$, which is a bilinear form on $SG(l)$ satisfying

$$R^l(x, y) := (\inf \{ \mathcal{E}^l(f, f) : f \in \mathcal{F}^l, f(x) = 0, f(y) = 1 \})^{-1}, \tag{2.5}$$

for each $x, y \in SG(l)$ with $x \neq y$. Finally, we note that the results of [22, Section 2.4] show that $(\mathcal{E}^l, \mathcal{F}^l)$ can also be considered as a regular Dirichlet form on $L^2(SG(l), \mu^l)$. Thus, by the classical correspondence (e.g. [15, Theorem 7.2.1]), there exists a related Hunt process X^l , which can be checked to be a diffusion process; this is our Brownian motion. Since points have positive capacity in our setting (by [23, Theorem 9.9]), the law of this process is uniquely defined from any starting point. We recall from the introduction that, for simplicity of presentation, we will henceforth suppose that X^l is started from 0 unless otherwise stated.

2.3 Vertex hitting times and an associated random walk

At the heart of our arguments will be the discrete-time random walk obtained by observing X^l at the hitting times of vertices in V_1^l . To introduce this, we first recall that processes associated with resistance forms (on compact spaces) satisfy the commute time identity (see [10, equation (2.17)], for example). In particular, if $\sigma_x^l := \inf\{t \geq 0 : X_t^l = x\}$, then

$$\mathbf{E}(\sigma_y^l | X_0^l = x) + \mathbf{E}(\sigma_x^l | X_0^l = y) = R^l(x, y), \quad \forall x, y \in SG(l). \tag{2.6}$$

Applying this in conjunction with the strong Markov property for X^l , we obtain that the stopping times $(T_n^l)_{n \geq 0}$ are all well-defined and finite random variables, where $T_0^l = 0$ and

$$T_n^l := \inf \left\{ t > T_{n-1}^l : X_t^l \in V_1^l \setminus \{X_{T_{n-1}^l}^l\} \right\}.$$

Moreover, from the obvious symmetries of the fractal $SG(l)$, we have that the discrete-time process $(Y_n^l)_{n \geq 0}$ given by

$$Y_n^l := X_{T_n^l}^l, \quad \forall n \geq 0,$$

is the simple random walk on V_1^l (considered as a graph with nearest-neighbour edges) started from 0.

As a final piece of notation for this preliminary section, we define

$$\tau^l := dl^2 \mathbf{E}(T_1^l). \tag{2.7}$$

In Section 4, we will later check that this is equivalent to the description of τ^l given in Theorem 1.3. As for Theorem 1.2, we will use a law of large numbers for the sequence $(T_n^l)_{n \geq 0}$ to deduce that

$$T_n^l \approx n \mathbf{E}(T_1^l) = n \tau^l l^{-2} d^{-1}$$

(see Proposition 3.6), which then allows us to show

$$X_{\tau^l t}^l \approx X_{T_{[dl^2 t]}^l}^l = Y_{[dl^2 t]}^l \rightarrow X_t^\Delta, \tag{2.8}$$

where the convergence result is a simple consequence of the fact that Y^l is merely a random walk on a section of a tetrahedral lattice.

3 Proof of Theorems 1.2 and 1.3

In this section, we will prove the functional scaling limit of Theorem 1.2 and the description of the scaling factors given by Theorem 1.3. In order to make rigorous the sketch proof of Theorem 1.2 presented at the end of previous section, we start by giving the scaling limit of the discrete-time random walk Y^l . The proof involves quite standard techniques, but we include it for completeness. (In the case of two dimensions, a much easier argument is possible using the symmetries of the triangle.)

Proposition 3.1. As $l \rightarrow \infty$,

$$\left(Y_{[dl^2 t]}^l \right)_{t \geq 0} \rightarrow \left(X_t^\Delta \right)_{t \geq 0}$$

in distribution with respect to the usual Skorohod J_1 -topology on $D([0, \infty), \Delta)$, the space of cadlag functions on Δ (equipped with the Euclidean distance).

Proof. We will apply the general result for convergence to a reflected diffusion of [7, Theorem 2.1], checking the conditions discussed at the end of [7, Section 2]. (Technically, we should apply this to the continuous time version of Y^l with unit mean exponential holding times, but for the various estimates we require, there is no problem to switch between discrete and continuous time.) We will first show tightness of the sequence $(Y_{[dl^2 t]}^l)_{t \geq 0}$, $l \geq 1$, in the following sense. For every $\varepsilon, T > 0$,

$$\lim_{\delta \rightarrow 0} \limsup_{l \rightarrow \infty} \sup_{x \in V_1^l} \mathbf{P} \left(\sup_{\substack{0 \leq s, t \leq T: \\ |s-t| \leq \delta}} \left| Y_{[dl^2 t]}^l - Y_{[dl^2 s]}^l \right| > \varepsilon \mid Y_0^l = x \right) = 0. \tag{3.1}$$

To this end, it will be convenient to consider the random walk on the graph with vertex set $V = \cup_{l' \geq 0} l' V_1^{l'}$, with edges between nearest-neighbour vertices. Note that V is simply a section of the tetrahedral lattice contained within a cone. Write $Y = (Y_n)_{n \geq 0}$ for the discrete-time simple random walk on this graph, and μ^V for the measure on V that assigns each vertex $x \in V$ a mass equal to its graph degree. It is clear that the graph metric on V is bounded above and below by constant multiples of the Euclidean metric. Moreover, writing $\bar{B}(x, r) := \{y : |y - x| \leq r\}$ for the closed Euclidean ball, then it holds that $\mu^V(\bar{B}(x, r)) \asymp 1 + r^d$ for $x \in V$ and $r \geq 0$. In particular, μ^V is a doubling measure, i.e. there exists a constant C such that

$$\mu^V(\bar{B}(x, 2r)) \leq C \mu^V(\bar{B}(x, r)), \quad \forall x \in V, r \geq 0. \tag{3.2}$$

Applying [2, Lemma 3.29 and Lemma 3.31], we further deduce that, for every $x \in V$ and $r \geq 0$,

$$\frac{1}{\mu^V(\bar{B}(x, r))} \sum_{y, z \in V \cap \bar{B}(x, r)} (f(y) - f(z))^2 \leq CR_{\bar{B}(x, r)}^{-2} \sum_{\substack{y, z \in V \cap \bar{B}(x, r): \\ |y-z|=1}} (f(y) - f(z))^2,$$

where $R_{\bar{B}(x, r)}$ is the relative isoperimetric constant for $\bar{B}(x, r)$ of [2, Definition 3.21]. From [2, Theorem 3.24], we have that

$$R_{\bar{B}(x, r)} \geq \frac{|V \cap \bar{B}(x, r)|}{\kappa_{\bar{B}(x, r)}},$$

where $\kappa_{\bar{B}(x, r)}$ is given by

$$\kappa_{\bar{B}(x, r)} := \min_{\Gamma} \max_{\substack{y, z \in V \cap \bar{B}(x, r): \\ |y-z|=1}} |\{u, v \in V \cap \bar{B}(x, r) : \{y, z\} \in \gamma_{u, v}\}|,$$

with the first infimum being taken over collections Γ of nearest neighbour paths with vertices in $V \cap \bar{B}(x, r)$ that contain, for each pair $u, v \in V \cap \bar{B}(x, r)$ an element $\gamma_{u, v}$ connecting u to v . Adapting the argument of [2, Theorem 3.25] from the square to the tetrahedral lattice, it is an elementary exercise to check that

$$\kappa_{\bar{B}(x, r)} \leq C |V \cap \bar{B}(x, r)| r.$$

Hence, we see that $R_{\bar{B}(x, r)} \geq Cr^{-1}$, which implies in turn that, for every $x \in V$ and $r \geq 0$,

$$\frac{1}{\mu^V(\bar{B}(x, r))} \sum_{y, z \in V \cap \bar{B}(x, r)} (f(y) - f(z))^2 \leq Cr^2 \sum_{\substack{y, z \in V \cap \bar{B}(x, r): \\ |y-z|=1}} (f(y) - f(z))^2. \quad (3.3)$$

Now, (3.3) is the Poincaré inequality for the graph V . It is well-known that in conjunction with the volume doubling property of (3.2) and the fact the graph V has bounded degree, this implies that the random walk Y admits two-sided Gaussian transition density estimates (see [11, Theorem 1.7]). Furthermore, two-sided Gaussian estimates imply the following exit time estimates: for every $x \in V$, $r \geq 1$,

$$\sup_{y \in V \cap \bar{B}(x, r)} \mathbf{E}(\bar{\sigma}^Y(x, r) | Y_0 = y) \leq c_1 r^2, \quad (3.4)$$

$$\mathbf{E}(\bar{\sigma}^Y(x, r) | Y_0 = x) \geq c_2 r^2,$$

where $\bar{\sigma}^Y(x, r) := \inf\{n \geq 0 : Y_n \notin \bar{B}(x, r)\}$ (see [17, Theorem 3.1] and [16, Proposition 6.1]). In turn, these yield the following estimate: for every $x \in V$, $r \geq 1$ and $t > 0$,

$$\mathbf{P}(\bar{\sigma}^Y(x, r) \leq t | Y_0 = x) \leq Ce^{-cr^2/t},$$

see the argument of [4, Theorem 5.5], for example. Clearly, the latter bound implies: for every $x \in V$, $r \geq 1$ and $t > 0$,

$$\mathbf{P}\left(\sup_{0 \leq n \leq t} |Y_n - x| > r \mid Y_0 = x\right) \leq Ce^{-cr^2/t}.$$

By appealing to the symmetry of the graphs V_1^l , it readily follows that: for $\varepsilon \in (0, \frac{1}{2})$, $T > 0$,

$$\begin{aligned} & \limsup_{l \rightarrow \infty} \sup_{x \in V_1^l} \mathbf{P}\left(\sup_{\substack{0 \leq s, t \leq T: \\ |s-t| \leq \delta}} |Y_{[dl^2 t]}^l - Y_{[dl^2 s]}^l| > \varepsilon \mid Y_0^l = x\right) \\ & \leq \limsup_{l \rightarrow \infty} \sup_{x \in V} T\delta^{-1} \mathbf{P}\left(\sup_{0 \leq n \leq 3dl^2 \delta} |Y_n - x| > \varepsilon l \mid Y_0 = x\right) \leq CT\delta^{-1} e^{-c\varepsilon^2/\delta}. \end{aligned}$$

Since the final expression here converges to 0 as $\delta \rightarrow 0$, the claim in (3.1) follows. Appealing to standard criteria for tightness of measures on $D([0, \infty), \Delta)$ (see [21, Chapter 16], for example), we deduce that [7, Assumption 2.6] holds.

For [7, Assumption 2.3], we need to check an exit time bound and a Hölder continuity result for harmonic functions. The first of these is implied by (3.4). As for the second, we again appeal to [17, Theorem 1] to observe that the two-sided Gaussian estimates for the transition density of Y imply an elliptic Harnack inequality, i.e. there exists a constant $C < \infty$ such that, for every $x \in V$, $r \geq 1$ and $u : V \rightarrow \mathbb{R}$ that is harmonic in $V \cap \bar{B}(x, 2r)$ (with respect to Y),

$$\sup_{x \in V \cap \bar{B}(x,r)} u(x) \leq C \inf_{x \in V \cap \bar{B}(x,r)} u(x). \tag{3.5}$$

From this, standard arguments imply the existence of a constant C and $\gamma > 0$ such that

$$|u(x) - u(y)| \leq C \left(\frac{|x - y|}{r} \right)^\gamma \|u\|_\infty, \quad \forall x, y \in V \cap \bar{B}(x, r),$$

for every $u : V \rightarrow \mathbb{R}$ that is harmonic in $V \cap \bar{B}(x, 2r)$. (We give an example of the argument in the proof of Theorem 4.7 below.) By scaling, the corresponding estimate holds for functions $u : V_1^l \rightarrow \mathbb{R}$ that are harmonic with respect to Y^l , with the constant being uniform over different values of l . This confirms [7, Assumption 2.3] in the present setting.

Next, suppose \tilde{V} are the vertices of the tetrahedral lattice obtained by extending V in the obvious way, and let $\tilde{Y} = (\tilde{Y}_n)_{n \geq 0}$ be the associated nearest-neighbour discrete-time random walk, started from 0. It is elementary to check that

$$\left(\varepsilon \tilde{Y}_{\lfloor d\varepsilon^{-2}t \rfloor} \right)_{t \geq 0} \rightarrow (X_t)_{t \geq 0},$$

where the limit process is standard Brownian motion on \mathbb{R}^d , started from 0. From this, it is straightforward to check that, if \bar{Y}^l is Y^l killed on hitting $\partial\Delta$, \bar{X}^Δ is X^Δ killed on hitting $\partial\Delta$, and $x_l \in V_1^l$ satisfies $x_l \rightarrow x \in \Delta$, then

$$\mathbf{P} \left(\left(\bar{Y}_{\lfloor dl^2t \rfloor}^l \right)_{t \geq 0} \in \cdot \mid \bar{Y}_0^l = x_l \right) \rightarrow \mathbf{P} \left((X_t^\Delta)_{t \geq 0} \in \cdot \mid \bar{X}_0^\Delta = x \right),$$

weakly as probability measures on $D([0, \infty), \Delta)$. This confirms condition (ii) of [7, Theorem 2.1].

It remains to check that the invariant measures of Y^l , suitably normalised, converge to the invariant measure of X^Δ , and also that the Dirichlet form of any potential limiting process is bounded above by that of X^Δ . The first of these claims is elementary. As for the second, this is also straightforward to check using the fact that: there exists a constant $c \in (0, \infty)$ such that, for any $f \in C_c^2(\Delta)$,

$$l^{2-d} \sum_{\substack{x, y \in V_1^l: \\ |x-y|=l^{-1}}} (f(x) - f(y))^2 \rightarrow c \int_\Delta |\nabla f(x)|^2 dx,$$

cf. the argument of [7, Proposition 3.16], in which the more challenging example of a percolation cluster was considered. As a consequence, we can conclude that all of the conditions of [7, Theorem 2.1] hold, and the result follows. \square

To check the remaining approximation in (2.8), it will be useful to suitably understand the effective resistance metrics R^l and the scaling constants ρ_l . In this direction, we first recall a result from [19]. (The notation ρ_l was introduced at (2.3).)

Lemma 3.2 ([19, Theorem 2.2]). *If $d = 2$, then $\rho_l \asymp \log l$. If $d \geq 3$, then $\rho_l \asymp 1$.*

Remark 3.3. It will follow from observations made in the proof of Theorem 1.3 that

$$\tau^l = \frac{l^2}{2\rho_l N_{d,l}}.$$

From the expression for $N_{d,l}$ in (2.1), it follows that

$$\tau^l \sim \frac{d!}{2} l^{2-d} \rho_l^{-1}.$$

Hence to replace \asymp by \sim in Theorem 1.3, as suggested might be an interesting problem in Remark 1.4(b), it would be enough to show that $\rho_l \sim c \log l$ when $d = 2$ and that ρ_l has a limit when $d \geq 3$.

Lemma 3.2 will be sufficient for our purposes when $d = 2$. However, for $d \geq 3$, we need slightly more control on the ρ_l , and in particular need that they are uniformly bounded away from 1.

Lemma 3.4. *For all $d \geq 2$,*

$$\inf_{l \geq 2} \rho_l > 1.$$

Proof. Since we know that $\rho_l > 1$ for each fixed l , the result for $d = 2$ immediately follows from Lemma 3.2. Thus we restrict to the case when $d \geq 3$. For each l , suppose V_1^l is the vertex set of a graph, equipped with nearest-neighbour edges. See Figure 3. Moreover, suppose $(R_{V_1^l}(x, y))_{x, y \in V_1^l}$ is the effective resistance on V_1^l when edges are equipped with unit resistors. By (2.3) and (2.4), we have that, for any $x \in V_0 \setminus \{0\}$,

$$R_{V_1^l}(0, x) = \frac{2\rho_l}{d+1} R_1^l(0, x) = \frac{2\rho_l}{d+1} R_0(0, x) = \frac{2\rho_l}{d+1},$$

and so to prove the result, we are required to find a suitable lower bound for $R_{V_1^l}(0, x)$. To this end, fix $x \in V_0 \setminus \{0\}$ and suppose Δ_0^l and Δ_1^l are the vertices in V_1^l at graph distance $\lfloor l/2 \rfloor$ from 0 and x , respectively. (Again, see Figure 3.) Now, by Rayleigh’s monotonicity principle and the parallel law (see [28, Chapter 9], for example),

$$R_{V_1^l}(0, x) \geq R_{V_1^l}(0, \Delta_0^l) + R_{V_1^l}(\Delta_1^l, x) = 2R_{V_1^l}(0, \Delta_0^l).$$

By shorting vertices at each fixed distance from 0, it further holds that

$$R_{V_1^l}(0, \Delta_0^l) \geq \sum_{i=1}^{\lfloor l/2 \rfloor} \frac{1}{dN_{d-1,i}}.$$

Applying (2.1), this yields that

$$R_{V_1^l}(0, \Delta_0^l) \geq \frac{(d-1)!}{d} \sum_{i=1}^{\lfloor l/2 \rfloor} \frac{1}{i(i+1) \dots (i+d-2)},$$

and clearly the right-hand side here converges as $l \rightarrow \infty$ to

$$\frac{(d-1)!}{d} \sum_{i=1}^{\infty} \frac{1}{i(i+1) \dots (i+d-2)} = \frac{d-1}{d(d-2)}.$$

Combining the above observations, we deduce that

$$\liminf_{l \rightarrow \infty} \rho_l \geq \frac{(d+1)(d-1)}{d(d-2)} > 1.$$

Since $\rho_l > 1$ for each fixed l , this completes the proof. □

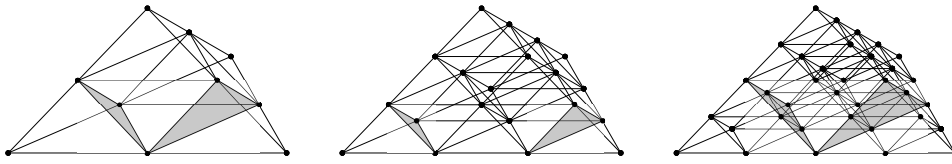


Figure 3: The sets V_1^2 , V_1^3 and V_1^4 in three dimensions. In each case, the set V_0 consists of the four extremal vertices of the tetrahedron. The shaded regions show Δ_0^l and Δ_1^l when the resistance being estimated is that between the vertices in the bottom left and right of the figure.

From the above estimates on ρ_l , we are able to deduce the following uniform bound on the diameters of the spaces $SG(l)$ in terms of their effective resistance metrics. This result is fundamental for the arguments that follow.

Lemma 3.5. For all $d \geq 2$,

$$\sup_{l \geq 2} \sup_{x, y \in SG(l)} R^l(x, y) < \infty.$$

Proof. For a given $i \in \mathbb{N}$, we say that the sets of the form

$$\psi_{j_1} \circ \dots \circ \psi_{j_i}(\Delta), \quad j_1, \dots, j_i \in \{1, \dots, N_{d,l}\},$$

are i -cells; we also say that Δ is a 0-cell. Clearly, for each $x \in SG(l)$, we can find a sequence $(x_i)_{i \geq 0}$ such that, for each $i \geq 0$, $x_i \in V_i^l$ (where $V_0^l := V_0$), and x, x_i and x_{i+1} are in the same i -cell. It thus obviously holds that $x_i \rightarrow x$ with respect to the Euclidean metric, and, since this is topologically equivalent to the effective resistance metric, then $R^l(x_i, x) \rightarrow 0$. Supposing $(y_i)_{i \geq 0}$ is a similarly constructed sequence for $y \in SG(l)$, it follows that

$$\begin{aligned} R^l(x, y) &= \lim_{i \rightarrow \infty} R^l(x_i, y_i) \\ &\leq \lim_{i \rightarrow \infty} \left(R^l(x_0, y_0) + \sum_{m=0}^{i-1} R^l(x_m, x_{m+1}) + \sum_{m=0}^{i-1} R^l(y_m, y_{m+1}) \right) \\ &\leq 1 + 2 \sum_{m=0}^{\infty} \sup_{x', y' \in V_1^l} \sup_{j_1, \dots, j_m \in \{1, \dots, N_{d,l}\}} R^l(\psi_{j_1} \circ \dots \circ \psi_{j_m}(x'), \psi_{j_1} \circ \dots \circ \psi_{j_m}(y')). \end{aligned}$$

Now, by the self-similarity of the construction of the resistance form, i.e. (2.3), it holds that

$$R^l(\psi_{j_1} \circ \dots \circ \psi_{j_m}(x'), \psi_{j_1} \circ \dots \circ \psi_{j_m}(y')) \leq \rho_l^{-m} R^l(x', y').$$

Hence, we have shown that

$$\sup_{x, y \in SG(l)} R^l(x, y) \leq 1 + 2 \sum_{m=0}^{\infty} \rho_l^{-m} \sup_{x', y' \in V_1^l} R^l(x', y').$$

Consequently, since Lemma 3.4 implies that $\sum_{m=0}^{\infty} \rho_l^{-m} = (1 - \rho_l^{-1})^{-1}$ is uniformly bounded over $l \geq 2$, to complete the proof it will suffice to show that the same is true of the supremum on the right-hand side above. Using the same notation as in the proof of Lemma 3.4, we have that

$$\sup_{x', y' \in V_1^l} R^l(x', y') = \frac{d+1}{2\rho_l} \sup_{x', y' \in V_1^l} R_{V_1^l}(x', y').$$

And, by elementary arguments (cf. [28, Proposition 9.16 and Exercise 9.1]), one can check that

$$\sup_{x', y' \in V_1^l} R_{V_1^l}(x', y') \leq \begin{cases} C_d, & \text{for } d \geq 3; \\ C_2 \log l, & \text{for } d = 2, \end{cases}$$

where C_d is a dimension dependent constant. Applying these estimates in conjunction with Lemma 3.2, we obtain the desired result. \square

We are now able to show that the stopping times T_n^l , which were introduced in Subsection 2.3, concentrate on their means in the following manner. The constant τ^l was defined in (2.7).

Proposition 3.6. For any $T > 0$, as $l \rightarrow \infty$,

$$\frac{\sup_{n \leq l^2 T} |T_n^l - n\mathbf{E}(T_1^l)|}{\tau^l} \rightarrow 0$$

in probability.

Proof. Since $T_n^l - n\mathbf{E}(T_1^l)$ is a martingale (in n), we have from Doob's L^2 martingale inequality that

$$\mathbf{P} \left(\frac{\sup_{n \leq l^2 T} |T_n^l - n\mathbf{E}(T_1^l)|}{\tau^l} > \varepsilon \right) \leq \frac{\mathbf{E} \left(\left(T_{\lfloor l^2 T \rfloor}^l - \lfloor l^2 T \rfloor \mathbf{E}(T_1^l) \right)^2 \right)}{(\varepsilon \tau^l)^2} \leq C_{d, \varepsilon, T} l^{-2} \frac{\mathbf{Var}(T_1^l)}{\mathbf{E}(T_1^l)^2},$$

and so it will be sufficient to show that the right-hand side converges to 0 as $l \rightarrow \infty$. For $A \subseteq \text{SG}(l)$, we write

$$\sigma^l(A) := \inf\{t \geq 0 : X_t^l \in A\}.$$

By the scaling of the resistance form $(\mathcal{E}^l, \mathcal{F}^l)$ that follows from (2.3) and the scaling of the measure μ^l , it is possible to check that, when X^l is started from 0, T_1^l is equal in distribution to $\rho_l^{-1} N_{d,l}^{-1} \sigma^l(V_0 \setminus \{0\})$. Moreover, in the proof of Theorem 1.3 below (see (3.6) in particular), we will check that $\mathbf{E}(\sigma^l(V_0 \setminus \{0\})) = 1/2d$. Hence it will be enough to show that

$$l^{-2} \mathbf{E}(\sigma^l(V_0 \setminus \{0\})^2) \rightarrow 0$$

as $l \rightarrow \infty$. To do this, we will check that the expectation on the left-hand side above can be bounded uniformly in l . To this end, we first introduce a parameter

$$t_l := \sup_{x \in \text{SG}(l)} \mathbf{E}(\sigma^l(V_0 \setminus \{0\}) | X_0^l = x).$$

By the commute time identity (see (2.6)), we have that

$$t_l \leq \sup_{x \in \text{SG}(l)} \sup_{y \in V_0 \setminus \{0\}} \mathbf{E}(\sigma_y^l | X_0^l = x) \leq \sup_{x, y \in \text{SG}(l)} R^l(x, y) \leq C,$$

where we have applied Lemma 3.5 to deduce the second inequality with a constant C that is independent of l . Consequently, applying the Markov property and the Markov inequality, we have that

$$\mathbf{P}(\sigma^l(V_0 \setminus \{0\}) \geq 2nt_l) \leq \sup_{x \in \text{SG}(l)} \mathbf{P}(\sigma^l(V_0 \setminus \{0\}) \geq 2t_l | X_0^l = x)^n \leq 2^{-n}, \quad \forall n \in \mathbb{N}.$$

In particular, under \mathbf{P} , $\sigma^l(V_0 \setminus \{0\})/2t_l$ is stochastically dominated by a geometric, parameter $1/2$, random variable, and so it has a second moment that is uniformly bounded in l . Since t_l is also uniformly bounded in l , this completes the proof. \square

Next, define $(H^l(t))_{t \geq 0}$, the right-continuous inverse of T^l , by setting

$$H^l(t) := \inf \{n \geq 0 : T_n^l > t\}.$$

It is a straightforward consequence of the previous proposition (and the definition of τ^l) that we have a similar concentration result for H^l .

Corollary 3.7. For any $T > 0$, as $l \rightarrow \infty$,

$$\sup_{t \leq T} |d^{-1}l^{-2}H^l(\tau^l t) - t| \rightarrow 0$$

in probability.

Putting together Proposition 3.1 and Corollary 3.7, we arrive at the proof of Theorem 1.2.

Proof of Theorem 1.2. Fix $\delta, T > 0$, and suppose that

$$\sup_{t \leq T} |d^{-1}l^{-2}H^l(\tau^l t) - t| \leq \delta.$$

It then holds that

$$\begin{aligned} \sup_{t \leq T} |X_{\tau^l t}^l - Y_{[dl^2 t]}^l| &\leq l^{-1} + \sup_{t \leq T} |Y_{H^l(\tau^l t)}^l - Y_{[dl^2 t]}^l| \\ &\leq l^{-1} + \sup_{\substack{s, t \leq T + \delta: \\ |s - t| \leq \delta}} |Y_{[dl^2 s]}^l - Y_{[dl^2 t]}^l|. \end{aligned}$$

Consequently, for any $\varepsilon, \delta > 0$,

$$\begin{aligned} &\mathbf{P} \left(\sup_{t \leq T} |X_{\tau^l t}^l - Y_{[dl^2 t]}^l| > \varepsilon \right) \\ &\leq \mathbf{P} \left(\sup_{t \leq T} |d^{-1}l^{-2}H^l(\tau^l t) - t| > \delta \right) + \mathbf{P} \left(l^{-1} + \sup_{\substack{s, t \leq T + \delta: \\ |s - t| \leq \delta}} |Y_{[dl^2 s]}^l - Y_{[dl^2 t]}^l| > \varepsilon \right). \end{aligned}$$

By Corollary 3.7, the first of these probabilities converges to 0 as $l \rightarrow \infty$. Moreover, Proposition 3.1 gives that the second probability converges to

$$\mathbf{P} \left(\sup_{\substack{s, t \leq T + \delta: \\ |s - t| \leq \delta}} |X_s^\Delta - X_t^\Delta| > \varepsilon \right),$$

which, by the continuity of the process X^l , converges in turn to 0 as $\delta \rightarrow 0$. In particular, we have checked that, as $l \rightarrow \infty$,

$$\mathbf{P} \left(\sup_{t \leq T} |X_{\tau^l t}^l - Y_{[dl^2 t]}^l| > \varepsilon \right) \rightarrow 0,$$

which, in conjunction with Proposition 3.1, is enough to complete the proof. □

To complete this section, we check the claims of Theorem 1.3, which reexpresses τ^l , as originally defined in (2.7), in terms of the walk dimension, and describes the $l \rightarrow \infty$ asymptotics.

Proof of Theorem 1.3. We start by recalling from the proof of Proposition 3.6 that T_1^l is equal in distribution to $\rho_l^{-1} N_{d,l}^{-1} \sigma^l(V_0 \setminus \{0\})$, where $\sigma^l(V_0 \setminus \{0\})$ is defined in the same proof. Hence, since $d_w(l) = \log(\rho_l N_{d,l}) / \log l$ (by [19, Proposition 2.1], see also [1, Theorem 8.18]),

$$\tau^l = dl^{2-d_w(l)} \mathbf{E}(\sigma^l(V_0 \setminus \{0\})).$$

In particular, to complete the proof of the equality in the statement of the theorem, it suffices to show that

$$\mathbf{E}(\sigma^l(V_0 \setminus \{0\})) = \frac{1}{2d}. \tag{3.6}$$

In checking this, we will apply the commute time identity of (2.6), which implies that, for any $x \in V_0 \setminus \{0\}$,

$$\mathbf{E}(\sigma_x^l) = \frac{1}{2} R^l(0, x) = \frac{1}{2}.$$

Now, we may alternatively write

$$\begin{aligned} \mathbf{E}(\sigma_x^l) &= \mathbf{E}(\sigma_x^l \mathbf{1}_{\{\sigma_x^l = \sigma^l(V_0 \setminus \{0\})\}}) + \mathbf{E}(\sigma_x^l \mathbf{1}_{\{\sigma_x^l > \sigma^l(V_0 \setminus \{0\})\}}) \\ &= \mathbf{E}(\sigma^l(V_0 \setminus \{0\}) \mathbf{1}_{\{\sigma_x^l = \sigma^l(V_0 \setminus \{0\})\}}) + \\ &\quad \mathbf{E}(\sigma^l(V_0 \setminus \{0\}) \mathbf{1}_{\{\sigma_x^l > \sigma^l(V_0 \setminus \{0\})\}}) + \mathbf{E}(\sigma_x^l) \mathbf{P}(\sigma_x^l > \sigma^l(V_0 \setminus \{0\})) \\ &= \mathbf{E}(\sigma^l(V_0 \setminus \{0\})) + \mathbf{E}(\sigma_x^l) \mathbf{P}(\sigma_x^l > \sigma^l(V_0 \setminus \{0\})), \end{aligned}$$

where we have applied the strong Markov property at $\sigma^l(V_0 \setminus \{0\})$ to deduce the second equality. (We also use symmetry to replace the expected time taken by the process X^l to traverse from a vertex of $V_0 \setminus \{0, x\}$ to x with the expected time taken to traverse from 0 to x .) Again appealing to symmetry, we have that

$$\mathbf{P}(\sigma_x^l > \sigma^l(V_0 \setminus \{0\})) = \frac{|V_0| - 2}{|V_0| - 1} = \frac{d - 1}{d}.$$

It follows that

$$\mathbf{E}(\sigma^l(V_0 \setminus \{0\})) = \mathbf{E}(\sigma_x^l) \mathbf{P}(\sigma_x^l = \sigma^l(V_0 \setminus \{0\})) = \frac{1}{2d},$$

which establishes (3.6), as required.

Finally, the \asymp part of the theorem statement is a ready consequence of (1.2). □

4 Proof of Theorem 1.5

In this section, we establish the convergence of transition densities stated as Theorem 1.5. We first note that, by standard results on processes associated with resistance forms, for each $l \geq 2$, the process X^l admits a jointly continuous transition density with respect to μ^l (see [23, Theorem 10.4], for example). As already noted in the introduction, we will denote this function by $(p_t^l(x, y))_{x, y \in \text{SG}(l), t > 0}$. Given the functional scaling limit of Theorem 1.2, to establish the convergence of these objects, it will suffice to check they satisfy the following equicontinuity result. Indeed, the proof of Theorem 1.5 can then be completed by a simple adaptation of the argument of [9, Theorem 1]. (In [9], the focus was on the transition densities of random walks on graphs, but the basic argument applies to more general processes.)

Theorem 4.1. *For every compact interval $I \subseteq (0, \infty)$, it holds that*

$$\lim_{\delta \rightarrow 0} \limsup_{l \rightarrow \infty} \sup_{\substack{x, x', y, y' \in \text{SG}(l): \\ |x-x'|, |y-y'| \leq \delta}} \sup_{t \in I} |p_{\tau^l t}^l(x, y) - p_{\tau^l t}^l(x', y')| = 0.$$

In fact, we will check something slightly stronger than the above statement, namely a uniform Hölder continuity property for the transition densities. To state this, we introduce the following scale function:

$$\Psi_l(r) := r^{d_w(l)} \mathbf{1}_{\{r \leq l^{-1}\}} + 2\tau^l r^2 \mathbf{1}_{\{r > l^{-1}\}},$$

where, as per the characterization of Theorem 1.3, $\tau^l = \frac{1}{2}l^{2-d_w(l)}$. We note that this function is a continuous, strictly increasing function, and so has a well-defined inverse Ψ_l^{-1} , which is given by

$$\Psi_l^{-1}(t) := t^{1/d_w(l)} \mathbf{1}_{\{t \leq l^{-d_w(l)}\}} + (2\tau^l)^{-1/2} t^{1/2} \mathbf{1}_{\{t > l^{-d_w(l)}\}}.$$

Roughly speaking, Ψ_l describes the time-scaling of X^l on different scales: for $r \leq l^{-1}$, the appropriate time-scaling is the fractal one determined by $r^{d_w(l)}$, whereas, for $r > l^{-1}$, the scaling is the usual Euclidean one r^2 , scaled by the l -dependent constant $2\tau^l$. (The 2 here is simply to ensure that Ψ_l and Ψ_l^{-1} are continuous.) It is clear that Theorem 4.1 is a consequence of the following statement. Note that we use the notation $x \vee y := \max\{x, y\}$.

Theorem 4.2. *For any $t_0 > 0$, there exist constants C and $\gamma > 0$ such that*

$$|p_{\tau^l t}^l(x, y) - p_{\tau^l t}^l(x', y')| \leq C t^{-(1+3d/4)} (|x - x'| \vee |y - y'| \vee l^{-1})^\gamma,$$

for all $x, x', y, y' \in \text{SG}(l)$, $t \in [4l^{-2}, t_0]$.

Towards proving Theorem 4.2, we start by deriving an exit time estimate that establishes that Ψ_l indeed gives appropriate control on the time-scaling, and also a version of the elliptic Harnack inequality in the present setting. For the statement, we let $B(x, r)$ be the standard d -dimensional Euclidean ball, centred at x and with radius r , and define

$$\bar{\sigma}^l(x, r) := \sigma^l(B(x, r)^c)$$

to be the exit time of $B(x, r)$ by X^l . Moreover, we say that a function $h_l : \text{SG}(l) \rightarrow \mathbb{R}$ is harmonic for $(\mathcal{E}_l, \mathcal{F}_l)$ in a ball $B(x_0, r)$ if $h_l(X_{t \wedge \bar{\sigma}^l(x_0, r)}^l)$ is a martingale (with respect to the filtration associated with X^l) for every starting point $x \in \text{SG}(l)$. To ease notation in what follows, we write $\mathbf{P}_x^l := \mathbf{P}(\cdot | X_0^l = x)$ and $\mathbf{E}_x^l := \mathbf{E}(\cdot | X_0^l = x)$.

Lemma 4.3. (a) *There exists a constant c_1 , independent of l , such that*

$$\mathbf{E}_x^l \bar{\sigma}^l(x_0, r) \leq c_1 \Psi_l(r), \quad \forall x, x_0 \in \text{SG}(l), r \in (0, \frac{1}{4}).$$

(b) *There exists a constant $\theta > 0$, independent of l , such that the following holds. If $r \in (8l^{-1}, \frac{1}{4})$ and $h_l : \text{SG}(l) \rightarrow \mathbb{R}_+$ is a harmonic function for $(\mathcal{E}_l, \mathcal{F}_l)$ on $B(x_0, 2r)$ for some $x_0 \in \text{SG}(l)$, then*

$$\sup_{y \in B(x_0, r) \cap \text{SG}(l)} h_l(y) \leq \theta \inf_{y \in B(x_0, r) \cap \text{SG}(l)} h_l(y).$$

Proof. We start with the proof of (a). Let $x, x_0 \in \text{SG}(l)$ and $r \in (0, \frac{1}{4})$. Started from x , if X^l has exited $B(x, 2r)$ at a certain time, then it will also have exited $B(x_0, r)$. In particular,

$$\mathbf{E}_x^l \bar{\sigma}^l(x_0, r) \leq \mathbf{E}_x^l \bar{\sigma}^l(x, 2r).$$

Moreover, if $r \in (\frac{1}{4}l^{-1}, \frac{1}{4})$ and the jump chain on V_1^l has exited a ball of radius $l^{-1}(\lceil 2rl \rceil + 1)$ centred at the first hitting location of V_1^l by X^l , then X^l will have exited $B(x, 2r)$. Hence, by the definition of τ^l ,

$$\mathbf{E}_x^l \bar{\sigma}^l(x, 2r) \leq \mathbf{E}_x^l \sigma^l(V_1^l) + \sup_{y \in V} \mathbf{E}(\bar{\sigma}^Y(y, \lceil 2rl \rceil + 1) | Y_0 = y) \mathbf{E}T_1^l,$$

where $\bar{\sigma}^Y$ was defined in the proof of Proposition 3.1, and $\mathbf{E}T_1^l$ was defined in Subsection 2.3. Now, by scaling and the commute time identity, see (2.6), the first term above satisfies

$$\mathbf{E}_x^l \sigma^l(V_1^l) \leq \rho_l^{-1} N_{d,l}^{-1} \sup_{y,z \in \text{SG}(l)} R^l(z,y) \leq C\tau^l l^{-2} \leq C\tau^l r^2,$$

where the second inequality follows from the description of τ^l in Remark 3.3 and the upper bound for the resistance diameter of Lemma 3.5, and the third inequality follows from the choice of r . Furthermore, recalling the definition of the lattice random walk Y from the proof of Proposition 3.1, it holds that

$$\sup_{y \in V} \mathbf{E}(\bar{\sigma}^Y(y, \lceil 2rl \rceil + 1) \mid Y_0 = y) \mathbf{E}T_1^l \leq C\tau^l l^{-2} (\lceil 2rl \rceil + 1)^2 \leq C\tau^l r^2 = \frac{1}{2} C\Psi_l(r),$$

where to deduce the first inequality, we apply the definition of τ^l from (2.7) and the bound on the exit times of Y given in (3.4). Hence we have shown the desired bound for this range of r . Similarly, if $r \in (\frac{1}{4}l^{-(m+1)}, \frac{1}{4}l^{-m}]$, then we have

$$\begin{aligned} \mathbf{E}_x^l \bar{\sigma}^l(x_0, r) &\leq \mathbf{E}_x^l \bar{\sigma}^l(x, 2r) \\ &\leq \mathbf{E}_x^l \sigma^l(V_{m+1}^l) + \rho_l^{-m+1} N_{d,l}^{-m+1} \sup_{y \in V} \mathbf{E}(\bar{\sigma}^Y(y, \lceil 2rl^{m+1} \rceil + 1) \mid Y_0 = y) \\ &\leq \rho_l^{-(m+1)} N_{d,l}^{-(m+1)} \left(\sup_{y,z \in \text{SG}(l)} R^l(z,y) + (\lceil 2rl^{m+1} \rceil + 1)^2 \right) \\ &\leq Cl^{-d_w(l)(m+1)} (1 + r^2 l^{2(m+1)}) \\ &\leq Cr^{d_w(l)}, \end{aligned}$$

where for the final inequality we use the fact that $d_w(l) \geq 2$. This establishes the exit time upper bound.

Let $r \in (8l^{-1}, \frac{1}{4})$. If h_l is harmonic on $B(x_0, r)$ with respect to X^l , then it is harmonic on $\bar{B}(x_0, 2r(1-l^{-1})) \cap V_1^l$ with respect to Y^l . Hence, similarly to (3.5),

$$\sup_{x \in \bar{B}(x_0, \frac{3}{2}r(1-l^{-1})) \cap V_1^l} h_l(x) \leq C \inf_{x \in \bar{B}(x_0, \frac{3}{2}r(1-l^{-1})) \cap V_1^l} h_l(x).$$

Note that this statement follows directly from (3.5) and the remark following [2, Definition 1.44], which explains how there is no problem in incorporating the factor $\frac{3}{2}$ if one restricts the range of r suitably; this is the reason for taking $r \geq 8l^{-1}$ rather than $r \geq l^{-1}$. Now, since harmonic functions take their maximum and minimum values on the boundaries of domains and the 1-cells of $\text{SG}(l)$ only meet at vertices in V_1^l , the above inequality readily implies the desired result. \square

Remark 4.4. It is natural to ask whether the exit time upper bound can be complemented by a lower bound of the form $\mathbf{E}_{x_0}^l \bar{\sigma}^l(x_0, r) \geq c\Psi_l(r)$ for $x_0 \in \text{SG}(l)$, $r > 0$. However, by arguing similarly to the previous proof, we are only able to show that, for $r \in (\frac{1}{4}l^{-(m+1)}, \frac{1}{4}l^{-m}]$,

$$\mathbf{E}_{x_0}^l \bar{\sigma}^l(x_0, r) \geq cr^2 l^{-(d_w(l)-2)(m+1)} \geq cr^{d_w(l)} l^{-(d_w(l)-2)} = cl^{-(d_w(l)-2)} \Psi_l(r).$$

In particular, there is additional l -dependence in the constant. Whilst we believe there should be some l -dependence as one passes along scales of the fractal from l^{-m} to $l^{-(m+1)}$, we do not know if this constant is optimal.

We next derive a version of the Poincaré inequality that is tailored to the different levels of the model with the scaling function $\Psi_l(r)$. Again, the proof follows the standard argument for fractals, but we have to combine this with an estimate for lattices in an appropriate way to handle the situation between the scales l^{-m} and $l^{-(m+1)}$ for each $m \geq 0$. The restriction to the m -cell $\psi_i(\text{SG}(l))$ seems rather awkward, but we believe that it is necessary, since without this, the comparison with the triangular lattice, which allows us to use the bound in (3.3), would not be applicable. Indeed, bottlenecks between m -cells will mean the result will not be true for more general sets of a similar size. Intuitively, mixing within an m -cell of the random walk happens at a much quicker rate than transitions between cells; this suggests functions like the heat kernel will, on the appropriate time-scale, be somewhat step-like – smoothing on each m -cell, before mass passes to a neighbouring one. Indeed, the situation that arises between two neighbouring m -cells is, on the appropriate scale, somewhat analogous to the joining of two copies of \mathbb{Z}^d , $d \geq 2$, at a single point. It is known that the latter graphs violate the Poincaré inequality (see [2, Example 3.32] and [25, Section 3.3]), and we find a similar issue in our setting. Thus we believe it is unavoidable that we have to consider each m -cell $\psi_i(\text{SG}(l))$ separately.

Lemma 4.5. *There exists finite constants $C > 0$ and $C_0 > 1$ such that for $l \geq 2$: if $r \in (\frac{1}{4}l^{-(m+1)}, \frac{1}{4}l^{-m}]$ for some $m \geq 0$, $x \in \psi_i(\Delta)$ for some $i = (i_1, \dots, i_m) \in \{1, \dots, N_{d,l}\}^m$, where we define $\psi_i := \psi_{i_1} \circ \dots \circ \psi_{i_m}$, and $f \in \mathcal{F}^l$, then*

$$\begin{aligned} & \int_{B(x,r) \cap \psi_i(\text{SG}(l))} \left(f(y) - \bar{f}_{B(x,r)}^{l,i} \right)^2 \mu^l(dy) \\ & \leq C \Psi_l(r) \rho_l^{m+1} \sum_{\substack{j \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m), \\ \psi_j(\text{SG}(l)) \cap B(x, 5r) \neq \emptyset}} \mathcal{E}^l(f \circ \psi_j, f \circ \psi_j), \end{aligned} \tag{4.1}$$

where

$$\bar{f}_{B(x,r)}^{l,i} := \frac{1}{\mu^l(B(x,r) \cap \psi_i(\text{SG}(l)))} \int_{B(x,r) \cap \psi_i(\text{SG}(l))} f(y) \mu^l(dy).$$

Proof. In the proof, we will write $B = B(x,r) \cap \psi_i(\text{SG}(l))$, where $i \in \{1, \dots, N_{d,l}\}^m$ is such that $x \in \psi_i(\Delta)$, with m satisfying $r \in (\frac{1}{4}l^{-(m+1)}, \frac{1}{4}l^{-m}]$. Note that the left-hand side of (4.1) can be written

$$\begin{aligned} & \int_B \left(f(y) - \bar{f}_{B(x,r)}^{l,i} \right)^2 \mu^l(dy) \\ & = \frac{1}{2\mu^l(B)} \int_B \int_B (f(y) - f(z))^2 \mu^l(dy) \mu^l(dz) \\ & = \frac{1}{2\mu^l(B)} \sum_{j \in \{1, \dots, N_{d,l}\}^{m+1}} \sum_{k \in \{1, \dots, N_{d,l}\}^{m+1}} \int_{B \cap \psi_j(\text{SG}(l))} \int_{B \cap \psi_k(\text{SG}(l))} (f(y) - f(z))^2 \mu^l(dy) \mu^l(dz). \end{aligned}$$

We next replace $f(y) - f(z)$ by

$$f(y) - f(\psi_j(0)) + f(\psi_j(0)) - f(\psi_k(0)) + f(\psi_k(0)) - f(z)$$

and use that $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$ to deduce that

$$\begin{aligned} & \int_{B \cap \psi_j(\text{SG}(l))} \int_{B \cap \psi_k(\text{SG}(l))} (f(y) - f(z))^2 \mu^l(dy) \mu^l(dz) \\ & \leq 3 \int_{B \cap \psi_j(\text{SG}(l))} (f(y) - f(\psi_j(0)))^2 \mu^l(dy) \mu^l(B \cap \psi_k(\text{SG}(l))) \\ & \quad + 3 (f(\psi_j(0)) - f(\psi_k(0)))^2 \mu^l(B \cap \psi_j(\text{SG}(l))) \mu^l(B \cap \psi_k(\text{SG}(l))) \\ & \quad + 3 \int_{B \cap \psi_k(\text{SG}(l))} (f(z) - f(\psi_k(0)))^2 \mu^l(dz) \mu^l(B \cap \psi_j(\text{SG}(l))). \end{aligned} \tag{4.2}$$

Now, to handle the first term here, we apply (2.5) and Lemma 3.5 to obtain

$$\begin{aligned} \int_{B \cap \psi_j(\text{SG}(l))} (f(y) - f(\psi_j(0)))^2 \mu^l(dy) & \leq N_{d,l}^{-(m+1)} \int_{\text{SG}(l)} (f \circ \psi_j(y) - f \circ \psi_j(0))^2 \mu^l(dy) \\ & \leq N_{d,l}^{-(m+1)} \sup_{y \in \text{SG}(l)} R^l(0, y) \mathcal{E}^l(f \circ \psi_j, f \circ \psi_j) \\ & \leq CN_{d,l}^{-(m+1)} \mathcal{E}^l(f \circ \psi_j, f \circ \psi_j). \end{aligned}$$

Obviously, it is also the case that, if $(j_1, \dots, j_m) \neq (i_1, \dots, i_m)$ or $B \cap \psi_j(\text{SG}(l)) = \emptyset$, then the left-hand side above is zero. Hence

$$\begin{aligned} & \frac{1}{\mu^l(B)} \sum_{j \in \{1, \dots, N_{d,l}\}^{m+1}} \sum_{k \in \{1, \dots, N_{d,l}\}^{m+1}} \int_{B \cap \psi_j(\text{SG}(l))} (f(y) - f(\psi_j(0)))^2 \mu^l(dy) \mu^l(B \cap \psi_k(\text{SG}(l))) \\ & \leq CN_{d,l}^{-(m+1)} \sum_{\substack{j \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m), \\ \psi_j(\text{SG}(l)) \cap B(x, r) \neq \emptyset}} \mathcal{E}^l(f \circ \psi_j, f \circ \psi_j). \end{aligned}$$

The third term in (4.2) can be dealt with in the same way. As for the second term, we have that

$$\begin{aligned} & \frac{1}{\mu^l(B)} \sum_{j, k \in \{1, \dots, N_{d,l}\}^{m+1}} (f(\psi_j(0)) - f(\psi_k(0)))^2 \mu^l(B \cap \psi_j(\text{SG}(l))) \mu^l(B \cap \psi_k(\text{SG}(l))) \\ & \leq \frac{1}{N_{d,l}^{2(m+1)} \mu^l(B)} \sum_{\substack{j, k \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m), \\ (k_1, \dots, k_m) = (i_1, \dots, i_m)}} (f(\psi_j(0)) - f(\psi_k(0)))^2 \mathbf{1}_{\{B \cap \psi_j(\text{SG}(l)) \neq \emptyset, B \cap \psi_k(\text{SG}(l)) \neq \emptyset\}} \\ & \leq \frac{1}{N_{d,l}^{2(m+1)} \mu^l(B)} \sum_{\substack{j, k \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m), \\ (k_1, \dots, k_m) = (i_1, \dots, i_m)}} (f(\psi_j(0)) - f(\psi_k(0)))^2 \mathbf{1}_{\{\psi_j(0), \psi_k(0) \in B(x, r+l^{-(m+1)})\}} \\ & \leq \frac{1}{N_{d,l}^{2(m+1)} \mu^l(B)} \sum_{y, z \in V_{m+1}^l \cap B(x, r+l^{-(m+1)}) \cap \psi_i(\text{SG}(l))} (f(y) - f(z))^2. \end{aligned}$$

Since the graph structure of $V_{m+1}^l \cap \psi_i(\text{SG}(l))$ is simply a part of the triangular lattice, we can apply the discrete-space Poincaré inequality of (3.3) to deduce that there exists a constant C such that

$$\begin{aligned} & \frac{1}{\mu^l(B)} \sum_{j, k \in \{1, \dots, N_{d,l}\}^{m+1}} (f(\psi_j(0)) - f(\psi_k(0)))^2 \mu^l(B \cap \psi_j(\text{SG}(l))) \mu^l(B \cap \psi_k(\text{SG}(l))) \\ & \leq \frac{C}{N_{d,l}^{2(m+1)} \mu^l(B)} (rl^{m+1})^2 |B_{i,m}| \sum_{\substack{y, z \in B_{i,m}: \\ |y-z|=l^{-(m+1)}}} (f(y) - f(z))^2, \end{aligned} \tag{4.3}$$

where we write $B_{i,m} := V_{m+1}^l \cap B(x, r + l^{-(m+1)}) \cap \psi_i(\text{SG}(l))$.

Each of the sums can be bounded as follows:

$$\begin{aligned} & \sum_{\substack{y,z \in B_{i,m}: \\ |y-z|=l^{-(m+1)}}} (f(y) - f(z))^2 \\ & \leq (d+1) \sum_{\substack{j \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m)}} \mathcal{E}_0(f \circ \psi_j, f \circ \psi_j) \mathbf{1}_{\{\psi_j(\text{SG}(l)) \cap B_{i,m} \neq \emptyset\}} \\ & \leq (d+1) \sum_{\substack{j \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m)}} \mathcal{E}^l(f \circ \psi_j, f \circ \psi_j) \mathbf{1}_{\{\psi_j(\text{SG}(l)) \cap B(x, 5r) \neq \emptyset\}}, \end{aligned}$$

where we have used that $\mathcal{E}_0(f \circ \psi_j, f \circ \psi_j) \leq \mathcal{E}^l(f \circ \psi_j, f \circ \psi_j)$ and $B(x, r + l^{-1}) \subseteq B(x, 5r)$ to obtain the second inequality. Moreover, it is elementary to check that $|B_{i,m}| \leq CN_{d,l}^{m+1} \mu^l(B)$. Hence, the expression in (4.3) is bounded above by

$$\frac{C(r l^{m+1})^2}{N_{d,l}^{m+1}} \sum_{\substack{j \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m)}} \mathcal{E}^l(f \circ \psi_j, f \circ \psi_j) \mathbf{1}_{\{\psi_j(\text{SG}(l)) \cap B(x, 5r) \neq \emptyset\}} = \frac{C(r l^{m+1})^2}{\rho_l^{m+1} N_{d,l}^{m+1}} \Sigma,$$

where we use the abbreviation

$$\Sigma := \rho_l^{m+1} \sum_{\substack{j \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m), \\ \psi_j(\text{SG}(l)) \cap B(x, 5r) \neq \emptyset}} \mathcal{E}^l(f \circ \psi_j, f \circ \psi_j).$$

Combining this estimate with the bounds on the first and third terms, we conclude that

$$\int_B (f(y) - \bar{f}_{B(x,r)}^{l,i})^2 \mu^l(dy) \leq C \rho_l^{-(m+1)} N_{d,l}^{-(m+1)} (1 + (r l^{m+1})^2) \Sigma.$$

Writing all the scaling factors in terms of l and applying the assumption that $r \geq \frac{1}{4} l^{-(m+1)}$, the right-hand side here is bounded above by

$$C r^2 l^{-(d_w(l)-2)(m+1)} \Sigma.$$

Finally, for $m \geq 1$, we use the fact that $r \geq \frac{1}{4} l^{-(m+1)}$ and $d_w(l) \geq 2$, to deduce this is bounded above by $r^{d_w(l)}$. For $m = 0$, we recall that $\tau^l = \frac{1}{2} l^{2-d_w(l)}$ (see Theorem 1.3) to obtain the bound $C \tau^l r^2$. In particular, in either case, the bound can be written simply as $C \Psi_l(r) \Sigma$, and so the proof is complete. \square

We continue by applying the above Poincaré inequality to give an on-diagonal heat kernel estimate, following closely the argument of [18, 26]. The final claim of the following lemma is the one that we will require in order to establish the main results of this section. Note the naturalness of the appearance of l^{-2} in (4.4); this is the time taken by $X_{\tau^l}^l$ to get beyond the fractal scale.

Lemma 4.6. *There exists a constant c , independent of l , such that*

$$\sup_{x \in \text{SG}(l)} p_t^l(x, x) \leq \begin{cases} c(t/\tau^l)^{-d/2}, & \forall t \in [l^{-d_w(l)}, \tau^l]; \\ c l^{d-d_f(l)} t^{-d_f(l)/d_w(l)}, & \forall t \in (0, l^{-d_w(l)}]. \end{cases}$$

In particular, there exists a constant c , independent of l , such that

$$\sup_{x \in \text{SG}(l)} p_{\tau^l t}^l(x, x) \leq c t^{-d/2}, \quad \forall t \in [l^{-2}, 1]. \tag{4.4}$$

Proof. Let $u_0 \in L^2(\text{SG}(l), \mu^l)$ with $u_0 \geq 0$ and $\|u_0\|_1 = 1$. (In this proof, we suppress the dependence on l in the notation of norms.) Set $u_t(x) = (P_t^l u_0)(x)$, where $(P_t^l)_{t \geq 0}$ is the semigroup associated with X^l , and $g(t) = \|u_t\|_2^2$; note that g is continuous and decreasing. Define

$$V_l(r) = \begin{cases} l^{-d+d_f(l)} r^{d_f(l)}, & \text{if } r \leq l^{-1}; \\ r^d, & \text{if } r \geq l^{-1}. \end{cases}$$

(We highlight that, from (2.1), it is possible to check $l^{-d+d_f(l)} \rightarrow 1$ as $l \rightarrow \infty$, and so this term is bounded above and below by constants, uniformly in l ; we include it to ensure the function V_l is continuous.) Now, suppose $m \geq 0$ and $r \in (\frac{1}{4}l^{-(m+1)}, \frac{1}{4}l^{-m}]$. For each $i \in \{1, \dots, N_{d,l}\}^m$, it is possible to cover $\psi_i(\text{SG}(l))$ by balls of the form $B(x, r)$ where $x \in \psi_i(\Delta)$ and, for each $j \in \{1, \dots, N_{d,l}\}^{m+1}$ with $(j_1, \dots, j_m) = (i_1, \dots, i_m)$, $\psi_j(\text{SG}(l))$ has a non-empty intersection with at most a constant number C_0 of the sets $B(x, 5r)$; the constant C_0 can be chosen independently of l, m and r . Write Γ_i for the centres of the balls in this cover of $\psi_i(\text{SG}(l))$. Applying the self-similarity of the Dirichlet form \mathcal{E}^l , as follows from (2.3), we then deduce that

$$\begin{aligned} \frac{d}{dt}g(t) &= -2\mathcal{E}^l(u_t, u_t) \\ &= -2\rho_l^{m+1} \sum_{j \in \{1, \dots, N_{d,l}\}^{m+1}} \mathcal{E}^l(u_t \circ \psi_j, u_t \circ \psi_j) \\ &\leq -2C_0^{-1} \sum_{i \in \{1, \dots, N_{d,l}\}^m} \sum_{x \in \Gamma_i} \rho_l^{m+1} \sum_{\substack{j \in \{1, \dots, N_{d,l}\}^{m+1}: \\ (j_1, \dots, j_m) = (i_1, \dots, i_m), \\ \psi_j(\text{SG}(l)) \cap B(x, 5r) \neq \emptyset}} \mathcal{E}^l(u_t \circ \psi_j, u_t \circ \psi_j) \\ &\leq -C\Psi_l(r)^{-1} \sum_{i \in \{1, \dots, N_{d,l}\}^m} \sum_{x \in \Gamma_i} \int_{B(x,r) \cap \psi_i(\text{SG}(l))} (u_t(y) - \bar{u}_t^{x,i})^2 \mu^l(dy), \end{aligned}$$

where $\bar{u}_t^{x,i}$ is the average of u_t on the set $B(x, r) \cap \psi_i(\text{SG}(l))$ (defined similarly to the definition of $\bar{f}_{B(x,r)}^{l,i}$ in the statement of Lemma 4.5) and, to deduce the final equality, we have applied Lemma 4.5. Now, we observe that the inner integral can be written

$$\int_{B(x,r) \cap \psi_i(\text{SG}(l))} u_t(y)^2 \mu^l(dy) - \mu^l(B(x, r) \cap \psi_i(\text{SG}(l))) (\bar{u}_t^{x,i})^2.$$

Since the sets $B(x, r), x \in \Gamma_i, i \in \{1, \dots, N_{d,l}\}^m$ form a cover of $\text{SG}(l)$, it holds that

$$\sum_{i \in \{1, \dots, N_{d,l}\}^m} \sum_{x \in \Gamma_i} \int_{B(x,r) \cap \psi_i(\text{SG}(l))} u_t(y)^2 \mu^l(dy) \geq \int_{\text{SG}(l)} u_t(y)^2 \mu^l(dy) = g(t).$$

Moreover, as the semigroup is conservative, $\|u_t\|_1 = 1$, which readily implies

$$\bar{u}_t^{x,i} \leq \frac{1}{\mu^l(B(x, r) \cap \psi_i(\text{SG}(l)))} \leq \frac{C}{V_l(r)}.$$

Putting these estimates together yields

$$\frac{d}{dt}g(t) \leq -c_1\Psi_l(r)^{-1} (g(t) - c_2V_l(r)^{-1}),$$

for any $r \leq \frac{1}{4}$. Therefore, if $g(t) > c_2V_l(r)^{-1}$,

$$-\frac{d}{dt} \log(g(t) - c_2V_l(r)^{-1}) \geq 2c_1\Psi_l(r)^{-1}. \tag{4.5}$$

Next, fix $\alpha \in (0, 1)$, set $r_n := \alpha^n$ and define $s_n = \inf\{t \geq 0 : g(t) \leq c_2 V_l(r_n)^{-1}\}$ for $n \in \mathbb{N}$. Then (4.5) holds with $r = r_n$ for $0 < t < s_n$. With $r = r_n$, integrating (4.5) over t from s_{n+2} to s_{n+1} , we thus obtain

$$\begin{aligned} 2c_1 \Psi_l(r_n)^{-1} (s_{n+1} - s_{n+2}) &\leq -\log(g(s_{n+1}) - c_2 V_l(r_n)^{-1}) + \log(g(s_{n+2}) - c_2 V_l(r_n)^{-1}) \\ &= \log\left(\frac{c_2 V_l(r_{n+2})^{-1} - c_2 V_l(r_n)^{-1}}{c_2 V_l(r_{n+1})^{-1} - c_2 V_l(r_n)^{-1}}\right) \\ &= \log\left(\frac{(V_l(r_n)/V_l(r_{n+2})) - 1}{(V_l(r_n)/V_l(r_{n+1})) - 1}\right) \\ &\leq c_3, \end{aligned}$$

for some $c_3 > 0$ independent of l , where the last inequality is due to the fact that

$$\alpha^{-1} \leq \alpha^{-d_f(l)} \leq V_l(r_n)/V_l(r_{n+1}) \leq \alpha^{-d},$$

which can easily be obtained from the definition of $V_l(r)$ and the observation that $d_f(l) \geq 1$. Hence we have shown that $s_{n+1} - s_{n+2} \leq c_4 \Psi_l(r_n)$, and iterating this gives

$$s_n \leq c_4 \sum_{k=n-1}^{\infty} \Psi_l(r_k) \leq c_5 \Psi_l(r_n).$$

This implies that $g(c_5 \Psi_l(r_n)) \leq g(s_n) = c_2 V_l(r_n)^{-1}$. It follows in turn that if $c_5 \Psi_l(r_{n+1}) \leq t < c_5 \Psi_l(r_n)$, then

$$g(t) \leq c_2 V_l(r_n)^{-1} \leq \begin{cases} c_6 (t/\tau^l)^{-d/2}, & \forall t \in [l^{-d_w(l)}, \tau^l]; \\ c_6 t^{d-d_f(l)} t^{-d_f(l)/d_w(l)}, & \forall t \in (0, l^{-d_w(l)}]. \end{cases}$$

(Here, we note that by Remark 1.4(a), it holds that $c_7 \Psi_l(r_n) \leq \Psi_l(r_{n+1})$ for some $c_7 > 0$ which is independent of l .) Using the fact that $\|P_t^l\|_{1 \rightarrow \infty} = \|P_t^l\|_{1 \rightarrow 2}^2$, we deduce the desired heat kernel upper bound. \square

Now the basic preparations are in place, we are ready to start developing continuity results. We proceed by applying the elliptic Harnack inequality of Lemma 4.3 to give a uniform Hölder continuity property for harmonic functions.

Theorem 4.7. *There exist constants c_1 and $\gamma' > 0$, independent of l , such that if h_l is bounded and harmonic for $(\mathcal{E}_l, \mathcal{F}_l)$ in a ball $B(x_0, 2r)$ for some $x_0 \in \text{SG}(l)$ and $r \in (8l^{-1}, \frac{1}{4})$, then*

$$|h_l(x) - h_l(y)| \leq c_1 \left(\frac{|x - y| \vee l^{-1}}{r}\right)^{\gamma'} \|h_l\|_{\infty}, \quad \forall x, y \in B(x_0, r) \cap \text{SG}(l). \quad (4.6)$$

Proof. We follow the arguments in [3, Lemma 3.8, Theorem 3.9]. Let $h := h_l$ be a bounded harmonic function for $(\mathcal{E}_l, \mathcal{F}_l)$ in $B(x_0, 2r)$ for some $x_0 \in \text{SG}(l)$ and $r \in (8l^{-1}, \frac{1}{4})$. For $m \in \mathbb{N} \cup \{0\}$ and $x \in B(x_0, r) \cap \text{SG}(l)$, set

$$O_m(x, h) := \sup_{y \in B(x, 2^{-m}r) \cap \text{SG}(l)} h(y) - \inf_{y \in B(x, 2^{-m}r) \cap \text{SG}(l)} h(y).$$

We claim that there exists a constant $p \in (0, 1)$, independent of l, x, x_0, r, m and h , such that

$$O_{m+1}(x, h) \leq p O_m(x, h), \quad (4.7)$$

at least whenever $2^{-(m+1)}r \geq 8l^{-1}$. When proving (4.7), by rescaling, we may assume $-1 \leq h \leq 1$ and $O_m(x, h) = 2$. Define

$$\begin{aligned} T &:= \inf\{t \geq 0 : X_t^l \in \partial B(x, 2^{-m}r) \cap \text{SG}(l)\}, \\ A &:= \{y \in \partial B(x, 2^{-m}r) \cap \text{SG}(l) : h(y) \leq 0\}. \end{aligned}$$

Clearly either $\mathbf{P}_x^l(X_T^l \in A) \geq 1/2$ or $\mathbf{P}_x^l(X_T^l \in A^c) \geq 1/2$ holds; without loss of generality we may assume (by multiplying h by -1 if needed) the former option holds. Then, for $y \in B(x, 2^{-m-1}r) \cap \text{SG}(l)$,

$$\begin{aligned} h(y) &= \mathbf{E}_y^l(h(X_T^l)) \\ &= \mathbf{E}_y^l(h(X_T^l) : X_T^l \in A^c) + \mathbf{E}_y^l(h(X_T^l) : X_T^l \in A) \\ &\leq \mathbf{P}_y^l(X_T^l \in A^c) \\ &= 1 - \mathbf{P}_y^l(X_T^l \in A). \end{aligned}$$

Now, since $\mathbf{P}^l(X_T^l \in A)$ is harmonic on $B(x, 2^{-m}r) \cap \text{SG}(l)$, when $2^{-(m+1)}r \geq 8l^{-1}$, we can apply Lemma 4.3(ii) to obtain that

$$\mathbf{P}_y^l(X_T^l \in A) \geq \theta^{-1} \mathbf{P}_x^l(X_T^l \in A) \geq (2\theta)^{-1}$$

for some $\theta \geq 1$. This yields that $h(y) \leq 1 - (2\theta)^{-1}$, and hence

$$O_{m+1}(x, h) \leq 1 - \frac{1}{2\theta} - (-1) = 2 - \frac{1}{2\theta} = pO_m(x, h),$$

where $p := 1 - (4\theta)^{-1} \in (0, 1)$. This proves (4.7).

We now prove (4.6). Let $x, y \in B(x_0, r) \cap \text{SG}(l)$. Suppose that there exists an $m \in \mathbb{N} \cup \{0\}$ so that $8l^{-1} \leq 2^{-(m+1)}r \leq |x - y| < 2^{-m}r$ (which is possible if $y \in B(x, r) \setminus B(x, 16l^{-1})$). Then, using (4.7),

$$|h(x) - h(y)| \leq O_m(x, h) \leq p^m O_0(x, h) \leq 2p^m \|h\|_\infty \leq c_1 \left(\frac{|x - y|}{r}\right)^\gamma \|h\|_\infty,$$

where $\gamma' = \log p^{-1} / \log 2$ and $c_1 = 2^{1+\gamma}$. Moreover, if $y \in B(x, 16l^{-1}) \cap \text{SG}(l)$, then one can simply note that h_l takes its maximum and minimum values in $B(x, 16l^{-1}) \cap \text{SG}(l)$ on the boundary of this set, and thus the difference between $h(x)$ and $h(y)$ is also obtained there. Together with the above bound, this observation implies

$$|h(x) - h(y)| \leq c_1 \left(\frac{|x - y| \vee l^{-1}}{r}\right)^\gamma \|h\|_\infty.$$

For $y \notin B(x, r)$, we apply the triangle inequality to deduce that $|h(x) - h(y)| \leq |h(x) - h(x_0)| + |h(y) - h(x_0)|$, and then apply the above estimate to each part. \square

Following the arguments in [5, Section 3], our subsequent goal is to show the λ -potentials associated with X^l , $l \geq 2$, are uniformly Hölder continuous. Specifically, we define

$$U_\lambda^l f(x) = \mathbf{E}_x^l \int_0^\infty e^{-\lambda t} f(X_t^l) dt.$$

Proposition 4.8. There exist constants c_1 and γ , independent of l and λ , such that if $f_l : \text{SG}(l) \rightarrow \mathbb{R}$ is a bounded, measurable function, then

$$|U_\lambda^l f_l(x) - U_\lambda^l f_l(y)| \leq c_1 (\tau^l + \lambda^{-1}) (|x - y| \vee l^{-1})^\gamma \|f_l\|_\infty, \quad \forall x, y \in \text{SG}(l).$$

Proof. Fix $x_0 \in \text{SG}(l)$ and $r \in (8l^{-1}, \frac{1}{8})$. Suppose $x, y \in B(x_0, r) \cap \text{SG}(l)$. Define $\bar{\sigma}_r := \bar{\sigma}^l(x_0, 2r)$ and $f := f_l$. By applying the strong Markov property at time $\bar{\sigma}_r$,

$$\begin{aligned} U_\lambda^l f(x) &= \mathbf{E}_x^l \int_0^{\bar{\sigma}_r} e^{-\lambda t} f(X_t^l) dt + \mathbf{E}_x^l ((e^{-\lambda \bar{\sigma}_r} - 1) U_\lambda^l f(X_{\bar{\sigma}_r}^l)) + \mathbf{E}_x^l U_\lambda^l f(X_{\bar{\sigma}_r}^l) \\ &= I_1 + I_2 + I_3, \end{aligned}$$

and similarly when x is replaced by y . By Lemma 4.3, we have

$$|I_1| \leq \|f\|_\infty \mathbf{E}_x^l \bar{\sigma}_r \leq c\Psi_l(r)\|f\|_\infty.$$

Moreover, by the observation that $|1 - e^{-u}| \leq u$ for all $u \geq 0$ and Lemma 4.3,

$$|I_2| \leq \lambda \mathbf{E}_x^l \bar{\sigma}_r \|U_\lambda^l f\|_\infty \leq c\Psi_l(r)\|f\|_\infty.$$

Similar bounds hold when x is replaced by y , and so we deduce that

$$|U_\lambda^l f(x) - U_\lambda^l f(y)| \leq c\Psi_l(r)\|f\|_\infty + |\mathbf{E}_x^l U_\lambda^l f(X_{\bar{\sigma}_r}^l) - \mathbf{E}_y^l U_\lambda^l f(X_{\bar{\sigma}_r}^l)|. \tag{4.8}$$

Now, the function $z \rightarrow \mathbf{E}_z^l U_\lambda^l f(X_{\bar{\sigma}_r}^l)$ is bounded and harmonic for $(\mathcal{E}_l, \mathcal{F}_l)$ in $B(x_0, 2r) \cap SG(l)$, and thus, by Theorem 4.7, the second term in (4.8) is bounded by

$$c \left(\frac{|x - y| \vee l^{-1}}{r} \right)^\gamma \|U_\lambda^l f\|_\infty \leq c\lambda^{-1} \left(\frac{|x - y| \vee l^{-1}}{r} \right)^\gamma \|f\|_\infty.$$

Summarising, we have shown that

$$|U_\lambda^l f(x) - U_\lambda^l f(y)| \leq c \left(\Psi_l(r) + \lambda^{-1} \left(\frac{|x - y| \vee l^{-1}}{r} \right)^\gamma \right)$$

for any $x, y \in B(x_0, r) \cap SG(l)$, where $x_0 \in SG(l)$ and $r \in (8l^{-1}, \frac{1}{8})$. Moreover, by applying the triangle inequality, it is straightforward to extend the range of r from $(8l^{-1}, \frac{1}{8})$ to $(8l^{-1}, r_0]$ for any $r_0 > 0$; in particular, we take $r_0 = 8$. Hence, for $x, y \in SG(l)$, we can take $x_0 = x$ and $r = (|x - y| \vee 8^{(2+\gamma')/\gamma'} l^{-1})^{\gamma'/(2+\gamma')} \in (8l^{-1}, r_0)$ to conclude

$$\begin{aligned} & |U_\lambda^l f(x) - U_\lambda^l f(y)| \\ & \leq c \left(\Psi_l \left(\left(|x - y| \vee 8^{(2+\gamma')/\gamma'} l^{-1} \right)^{\gamma'/(2+\gamma')} \right) + \lambda^{-1} \left(|x - y| \vee l^{-1} \right)^{2\gamma'/(2+\gamma')} \right) \|f\|_\infty \\ & \leq c(\tau^l + \lambda^{-1}) \left(|x - y| \vee l^{-1} \right)^{2\gamma'/(2+\gamma')} \|f\|_\infty, \end{aligned}$$

which completes the proof. □

We transfer Proposition 4.8 to the semigroup by appealing to spectral theory. In particular, for each $l \geq 2$, let $(P_t^l)_{t \geq 0}$ be the semigroup of X^l . By the standard spectral theorem, there exist projection operators $(E_\nu^l)_{\nu \geq 0}$ on the space $L^2(SG(l), \mu^l)$ such that: for $f \in L^2(SG(l), \mu^l)$,

$$\begin{aligned} f &= \int_0^\infty dE_\nu^l(f), \\ P_t^l f &= \int_0^\infty e^{-\nu t} dE_\nu^l(f), \\ U_\lambda^l f &= \int_0^\infty \frac{1}{\lambda + \nu} dE_\nu^l(f). \end{aligned} \tag{4.9}$$

(Note that the integration above is with respect to ν .)

Proposition 4.9. If f_l is in $L^2(SG(l), \mu^l)$, then $P_{\tau^l t}^l f_l$ is almost-surely equal to a function that is Hölder continuous. Specifically, for this version of $P_{\tau^l t}^l f_l$, it holds that

$$|P_{\tau^l t}^l f_l(x) - P_{\tau^l t}^l f_l(y)| \leq Ct^{-(1+d/2)} (|x - y| \vee l^{-1})^\gamma \|f_l\|_2, \quad \forall x, y \in SG(l), t \in [2l^{-2}, 1],$$

where γ is the constant of Proposition 4.8 and C is a constant independent of all the other variables.

Proof. Write $\langle f, g \rangle$ for the inner product in $L^2(\text{SG}(l), \mu^l)$. Note that in what follows t is fixed. As before, we write $f = f_l$. Define

$$h = \int_0^\infty ((\tau^l)^{-1} + \nu) e^{-\nu\tau^l t} dE_\nu^l(f).$$

Since $\sup_\nu ((\tau^l)^{-1} + \nu)^2 e^{-2\nu\tau^l t} \leq 1 + (\tau^l t)^{-2} =: c_{1,t}(l)$, then

$$\int_0^\infty ((\tau^l)^{-1} + \nu)^2 e^{-2\nu\tau^l t} d\langle E_\nu^l(f), E_\nu^l(f) \rangle \leq c_{1,t}(l) \int_0^\infty d\langle E_\nu^l(f), E_\nu^l(f) \rangle = c_{1,t}(l) \|f\|_2^2,$$

and so we see that h is an element of $L^2(\text{SG}(l), \mu^l)$.

Next, suppose $g \in L^1(\text{SG}(l), \mu^l)$. Then $\|P_{\tau^l t}^l g\|_1 \leq \|g\|_1$, and

$$|P_{\tau^l t}^l g(x)| = \left| \int p_{\tau^l t}^l(x, y) g(y) \mu_l(dy) \right| \leq c_{2,t}(l) \|g\|_1,$$

where

$$c_{2,t}(l) := \sup_{x, y \in \text{SG}(l)} p_{\tau^l t}^l(x, y).$$

Note that this constant is finite by the continuity of the heat kernel. It follows that

$$\|P_{\tau^l t}^l g\|_2 \leq \|P_{\tau^l t}^l g\|_\infty \leq c_{2,t}(l) \|g\|_1.$$

Using Cauchy-Schwarz and the fact that $\sup_\nu ((\tau^l)^{-1} + \nu) e^{-\nu\tau^l t} = c_{1,t}(l)^{1/2} < \infty$, we have

$$\begin{aligned} \langle h, g \rangle &= \int_0^\infty ((\tau^l)^{-1} + \nu) e^{-\nu\tau^l t} d\langle E_\nu^l(f), E_\nu^l(g) \rangle \\ &\leq \left(\int_0^\infty ((\tau^l)^{-1} + \nu)^2 e^{-2\nu\tau^l t} d\langle E_\nu^l(f), E_\nu^l(f) \rangle \right)^{1/2} \times \left(\int_0^\infty e^{-2\nu\tau^l t} d\langle E_\nu^l(g), E_\nu^l(g) \rangle \right)^{1/2} \\ &\leq \left(c_{1,t/2}(l) \int_0^\infty d\langle E_\nu^l(f), E_\nu^l(f) \rangle \right)^{1/2} \times \left(\int_0^\infty e^{-2\nu\tau^l t} d\langle E_\nu^l(g), E_\nu^l(g) \rangle \right)^{1/2} \\ &= c_{1,t/2}(l)^{1/2} \|f\|_2 \|P_{\tau^l t/2}^l g\|_2 \\ &\leq c'_t(l) \|f\|_2 \|g\|_1, \end{aligned}$$

where $c'_t(l) := c_{1,t/2}^{1/2} c_{2,t/2}(l)$. By considering the supremum over $g \in L^1(\text{SG}(l), \mu^l)$ with L^1 -norm less than 1, we obtain that $\|h\|_\infty \leq c'_t(l) \|f\|_2$. Additionally, by (4.9),

$$U_{(\tau^l)^{-1}}^l h = \int_0^\infty e^{-\nu\tau^l t} dE_\nu^l(f) = P_{\tau^l t}^l f,$$

and so the Hölder continuity of $P_{\tau^l t}^l f$ follows from Proposition 4.8. In particular, we have that the constant of the result arises from the following bound: for $t \in [2l^{-2}, 1]$,

$$2c_1 \tau^l c'_t(l) \leq 2c_1 (1 + t^{-1}) \sup_{x, y \in \text{SG}(l)} p_{\tau^l t/2}^l(x, y) \leq Ct^{-(1+d/2)},$$

where c_1 is constant of Proposition 4.8 and we have applied Lemma 4.6 to deduce the final inequality. \square

We are now in a position to prove Theorems 4.1 and 4.2.

Proof of Theorems 4.1 and 4.2. Fix $y \in \text{SG}(l)$ and let $f_l(z) = p_{\tau^{l_t}/2}^l(z, y)$. Since

$$P_{\tau^{l_t}/2}^l f_l(x) = \int p_{\tau^{l_t}/2}^l(x, z) f_l(z) \mu^l(dz) = \int p_{\tau^{l_t}/2}^l(x, z) p_{\tau^{l_t}/2}^l(z, y) \mu^l(dz) = p_{\tau^{l_t}}^l(x, y),$$

we obtain from Proposition 4.9 that

$$\begin{aligned} |p_{\tau^{l_t}}^l(x, y) - p_{\tau^{l_t}}^l(x', y)| &= \left| P_{\tau^{l_t}/2}^l f_l(x) - P_{\tau^{l_t}/2}^l f_l(x') \right| \\ &\leq C t^{-(1+d/2)} (|x - x'| \vee l^{-1})^\gamma \|f_l\|_2, \\ &= C t^{-(1+d/2)} (|x - x'| \vee l^{-1})^\gamma p_{\tau^{l_t}}^l(y, y)^{1/2}, \end{aligned}$$

for all $x, y \in \text{SG}(l)$, $t \in [4l^{-2}, 1]$. By Lemma 4.6, the term $p_{\tau^{l_t}}^l(y, y)^{1/2}$ is bounded above by $Ct^{-d/4}$, and thus the result of Theorem 4.2 follows for the range $t \in [4l^{-2}, 1]$. By simple modifications of the previous arguments, the constant 1 is readily replaced by an arbitrary constant t_0 .

Finally, let $I = [t_1, t_0]$, where $0 < t_1 < t_0$. Assume that l is large enough so that $4l^{-2} \leq t_1$. Then, from Theorem 4.2,

$$\sup_{\substack{x, x', y, y' \in \text{SG}(l): \\ |x-x'|, |y-y'| \leq \delta}} \sup_{t \in I} |p_{\tau^{l_t}}^l(x, y) - p_{\tau^{l_t}}^l(x', y')| \leq C \sup_{t \in I} t^{-(1+3d/4)} (\delta \vee l^{-1})^\gamma.$$

Taking the limit as $l \rightarrow \infty$ and then $\delta \rightarrow 0$ yields Theorem 4.1, as desired. □

5 Other examples

5.1 Triangles

We have worked with the family $\text{SG}(l)$, which has a part of the triangular lattice as the initial structure in each fractal. This could be modified to, for example, a sequence of triangles attached to the outer edge of the unit triangle, see Figure 4. This would have a mass of $3(n - 1)$ triangles of side length $1/n$ at stage n . In this case, the fractal dimension of the limit would be 1 and we should see that the associated diffusions converge to a one-dimensional Brownian motion on the boundary of a unit triangle. (Actually, for this example, the limiting process is described by a resistance form, and so we could alternatively deduce this conclusion by applying the convergence result of [10, Theorem 1.3]. Moreover, the corresponding local limit theorem would follow by checking the equicontinuity of the heat kernel using the resistance-based approach of [9, Section 4].) Other higher-dimensional analogues of the construction would also be possible. For instance, in three-dimensions, taking those first-stage tetrahedra touching the outer boundary of the initial one would yield a two-dimensional Brownian motion on the boundary of the tetrahedron. Or, taking only those first-stage tetrahedra intersecting the edges of the initial one would yield a one-dimensional Brownian motion on the edges of the tetrahedron. We believe that straightforward modifications to the arguments of this paper would yield the corresponding results in these cases.

5.2 Vicsek sets

We discuss here how our analysis can also be transferred to a family of finitely ramified fractals based on the square and the cube. These we call Vicsek sets and label them $\text{VS}(l)$, where $l = 2n + 1$ is the number of divisions of the side, see Figure 5. There is much that is the same as in the case of the Sierpinski gasket, but we need to be a little more careful in certain places.

We start by recalling the resistance estimates from [19].

Lemma 5.1 ([19, Section 4.1]). *If $d = 2$, then $\rho_l \asymp \log l$. If $d \geq 3$, then $\rho_l \asymp 1$.*

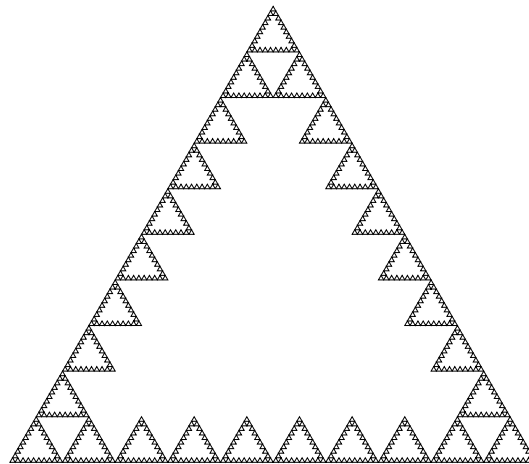


Figure 4: A variation on the $SG(l)$ containing only the triangles touching the outer boundary of Δ , shown at the second stage of construction for $l = 10$.

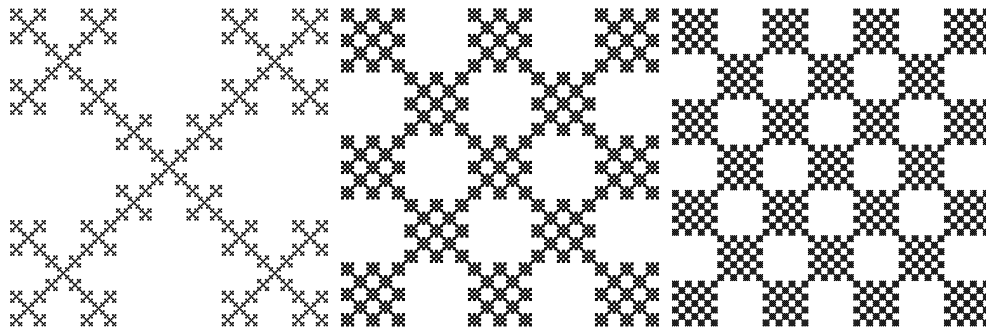


Figure 5: The sets $VS(2)$, $VS(3)$ and $VS(4)$ in two dimensions, shown at the 5th, 4th and 3rd stages of construction, respectively.

5.2.1 Two-dimensional case

The construction is to take a unit square and divide it into a $l \times l$ grid. We then construct a checkerboard of squares of side $1/l$. Retain all the squares that are connected diagonally, taking steps through corner vertices, to one of the outer corners, and discard the remaining ones. (See [6] for a more detailed definition.) Repeat this process within each remaining square to build a fractal with fractal dimension $\log((l^2 + 1)/2)/\log(l)$. We see that as $l \rightarrow \infty$, the fractal dimension converges to 2 as expected.

The first step is to show the existence of diffusions on the fractals in the sequence. For the case of the square, this is a consequence of the fact that it is a family of nested fractals. By the argument of Lindstrom [29], for each n , there exists a fixed point w_l such that the network with conductance 1 on the horizontal and vertical edges and w_l on the diagonal edges, will satisfy the fixed point problem for the Dirichlet form

$$\mathcal{E}_l^{(0)}(f, f) = \sum_{i=1}^{(l^2+1)/2} \rho_l \mathcal{E}_{l,i}^{(0)}(f \circ \phi_i, f \circ \phi_i).$$

(The graph upon which the Dirichlet form $\mathcal{E}_l^{(0)}$ is based is shown in Figure 6.) From this fixed point, we can construct a sequence of Dirichlet forms $(\mathcal{E}_l, \mathcal{F}_l)$ each defined on $VS(l)$. For each l , the corresponding Hunt process $(X_t^l)_{t \geq 0}$ is the Brownian motion on the $VS(l)$.

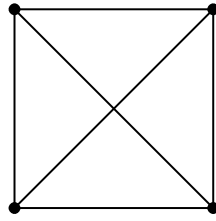


Figure 6: The graph upon which $\mathcal{E}_l^{(0)}$ is built for the two-dimensional Vicsek set.

We now let $l \rightarrow \infty$. It is shown in [19] that

$$\rho_l \asymp \log l.$$

This makes it straightforward to follow the same steps as for the Sierpinski gasket case to show that there will be a sequence of scaling parameters τ^l such that as $l \rightarrow \infty$

$$(X_{\tau^l t}^l)_{t \geq 0} \rightarrow (B_t)_{t \geq 0},$$

where B is reflected Brownian motion on the unit square.

5.2.2 Three-dimensional case

We consider the case of cubes where we break an initial cube up into a collection of $(2n + 1)^3$ cubes of side length $1/(2n + 1)$. Similarly to the two-dimensional case, retain all cubes that are connected diagonally, taking steps through corner vertices, to one of the outer corners, and discard the remaining ones. (Again, see [6] for a more detailed definition.) The Hausdorff dimension of the resulting fractal is easily seen to be $\log((n + 1)^3 + n^3)/\log(2n + 1)$, which converges to 3 as expected.

The first issue is the existence of a Brownian motion on these sets as they are not nested fractals in dimension 3 (or indeed higher dimensions), see the lecture notes [1] for a discussion of the class of nested fractals. However, the structure of these sets is such that we can apply the same arguments as used by [29] to show existence.

Let $\mathcal{P} = \{(p_1, p_2, p_3) : p_1 \geq p_2 \geq p_3 \geq 0, 3p_1 + 3p_2 + p_3 = 1\}$ be the set of probabilities for a random walk on the unit cube where p_1 is the probability of crossing an edge, p_2 the probability of crossing a diagonal on a face and p_3 is the probability of crossing the main diagonal across the cube. Note that these probabilities are proportional to the conductivities (c_1, c_2, c_3) between the vertices. That is, if we fix the edge conductivity $c_1 = 1$, then for the diagonal across a face $c_2 = p_2/p_1$ and, for the main diagonal across the cube, $c_3 = p_3/p_1$.

Proposition 5.2. For each l , there is a fixed point for the renormalization map in the set \mathcal{P} . (In particular, one can construct a sequence of consistent resistance metrics on the different levels of the fractal, analogous to (2.4).)

Proof. This follows the same argument as [29]. We show that the map on transition probabilities induced by considering the trace of the random walk on V_1^l onto the vertex set V_0 is a continuous map from \mathcal{P} to \mathcal{P} and hence, by Brouwer's fixed point theorem, the fixed point exists. (The map on the probabilities is clearly continuous in p .)

Let $X^{l,1}$ denote the Markov chain on V_1^l moving according to the transition probabilities given by an element of \mathcal{P} . The renormalization map is equivalent to the map on transition probabilities given by $f_i(p) = \mathbb{P}^0(X_{T_1}^{l,1} = A_i)$, where $T_1 = \inf\{n \geq 0 : X^{l,1} \in V_0 \setminus \{0\}\}$ and the vertices A_i are as in Figure 7. Now, by considering the paths the random walk can take between the vertices and using the reflection symmetry, it is possible to check

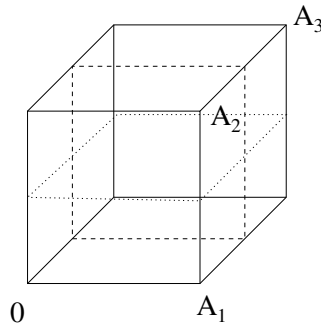


Figure 7: The cube with corners labelled as in the proof of Proposition 5.2.

that, if the vertex is further away from 0, the probability of reaching it first will be smaller. To see this for the edge and the diagonal across a face, we consider the bisector of the face. To move from 0 to A_2 , we must cross the bisector, shown as the dotted line in the figure, so taking a reflection in this plane we can map any path from 0 to A_2 to a path from 0 to A_1 . Hence, not only are there are more paths in V_1^l to A_1 , those that hit the bisector have a higher probability than the corresponding reflected path that goes from 0 to A_2 . Thus the probability $f_1(p)$ of arriving at A_1 must be greater than the probability $f_2(p)$ of arriving at A_2 . Similarly, for the long diagonal, we can take the plane bisecting A_2 and A_3 , as indicated on the figure by the dashed line, and any path on V_1^l from 0 to A_3 can be reflected into a path from 0 to A_2 , which therefore has a higher probability of being reached, i.e. $f_3(p) < f_2(p)$. \square

Given Proposition 5.2, we have the existence of the processes exactly as for nested fractals. The next task is to prove that the resistance scale factor ρ_l is strictly larger than one.

Lemma 5.3. For $d = 3$,

$$\inf_{l \geq 2} \rho_l > 1.$$

Proof. We follow an approach similar to the Sierpinski gasket case. For each l , suppose V_1^l is the vertex set of a graph, equipped with all edges within each 0-cell. Moreover, suppose $(R_{V_1^l}(x, y))_{x, y \in V_1^l}$ is the effective resistance on V_1^l when edges are equipped with the fixed point resistors, and define R_{V_0} similarly on V_0 . To fix the normalisation, we consider the case when the fixed point resistances on the ‘edges’ (as opposed to the ‘diagonals’) are given by 1. As in the proof of Lemma 3.4, we will short all the vertices at a fixed graph distance from 0.

As the resistance scale factor will be independent of which edge is considered, we focus on the main diagonal. Using the labelling from Figure 7, we note that

$$R_{V_1^l}(0, A_3) = \rho_l R_{V_0}(0, A_3).$$

Firstly, write a for the resistance along the diagonal across a face and b for the resistance along the long diagonal across the cube. (Recall that the resistance along an edge is equal to 1.) By a standard calculation we get

$$R_{V_0}(0, A_3) = \frac{1}{2} \frac{b(5ab + 2a + b)}{(a + 2b + ab)(3b + 1)}.$$

To prove the result, we need a good lower bound on $R_{V_1^l}(0, A_3)$. Suppose B_0^l and B_1^l are the vertices in V_1^l at graph distance $\lfloor l/2 \rfloor$ from 0 and A_3 , respectively. (Similar to

Figure 3 for the Sierpinski gasket.) By Rayleigh’s monotonicity principle and the parallel law,

$$R_{V_1^l}(0, A_3) \geq R_{V_1^l}(0, B_0^l) + R_{V_1^l}(B_1^l, A_3) = 2R_{V_1^l}(0, B_0^l).$$

By shorting vertices at each fixed graph distance from 0, we have

$$\begin{aligned} R_{V_1^l}(0, B_0^l) &\geq \sum_{i=1}^{\lfloor (l-3)/4 \rfloor} \frac{2}{((i+1)^3 - i^3 - 1)(4 + 8/a + 4/b) + 1 + 3/a + 3/b} \\ &\geq \frac{1}{2(1 + 2/a + 1/b)} \sum_{i=1}^{\lfloor (l-3)/4 \rfloor} \frac{1}{(i+1)^3 - i^3} \\ &\rightarrow \frac{1}{2(1 + 2/a + 1/b)} \frac{\pi}{\sqrt{3}} \tanh\left(\frac{\pi}{2\sqrt{3}}\right), \quad l \rightarrow \infty. \end{aligned}$$

Writing

$$\eta(a, b) = \frac{1}{2R_{V_0}(0, A_3)(1 + 2/a + 1/b)},$$

we have

$$\liminf_{l \rightarrow \infty} \rho_l \geq \frac{2\pi}{\sqrt{3}} \tanh\left(\frac{\pi}{2\sqrt{3}}\right) \eta(a, b).$$

Note that, for any $a, b \geq 1$, it holds that $1/2 = \eta(1, 1) \leq \eta(a, b) \leq \lim_{a', b' \rightarrow \infty} \eta(a', b') = 3/5$ and it is also the case that $\frac{\pi}{\sqrt{3}} \tanh\left(\frac{\pi}{2\sqrt{3}}\right) \approx 1.305$. Combining the above observations, as $b \geq a \geq 1$ from the fixed point argument, we deduce that

$$\liminf_{l \rightarrow \infty} \rho_l > 1.$$

Since $\rho_l > 1$ for each fixed l , this completes the proof. □

Applying Lemma 5.3, the following is established in exactly the same way as for Lemma 3.5.

Lemma 5.4. For $d = 2, 3$,

$$\sup_{l \geq 2} \sup_{x, y \in VS(l)} R^l(x, y) < \infty.$$

Once we have this result, we have the convergence of the processes in the same way as for the Sierpinski gasket.

Remark 5.5. We can approximate the Sierpinski carpet by a sequence of Vicsek sets by taking the side lengths as 3^l and removing the appropriate pieces of checkerboard to approximate the carpet. We would hope to be able to recover Brownian motion on the carpet as a limit of these processes. In two-dimensions the existence of fixed points for the sequence holds as the approximating sequence consists of nested fractals. Establishing convergence would need some proof of the limit $\rho_{3^l} \rightarrow \rho_{SC}$, which may be more challenging. (Again, an alternative approach for this example, in which the limiting process is described by a resistance form, would be to apply [10, Theorem 1.3].)

References

- [1] M. T. Barlow, *Diffusions on fractals*, Lectures on probability theory and statistics (Saint-Flour, 1995), Lecture Notes in Math., vol. 1690, Springer, Berlin, 1998, pp. 1–121. MR1668115
- [2] M. T. Barlow, *Random walks and heat kernels on graphs*, London Mathematical Society Lecture Note Series, vol. 438, Cambridge University Press, Cambridge, 2017. MR3616731
- [3] M. T. Barlow and R. F. Bass, *The construction of Brownian motion on the Sierpiński carpet*, Ann. Inst. H. Poincaré Probab. Statist. **25** (1989), no. 3, 225–257. MR1023950

- [4] M. T. Barlow and R. F. Bass, *Random walks on graphical Sierpinski carpets*, Random walks and discrete potential theory (Cortona, 1997), Sympos. Math., vol. XXXIX, Cambridge Univ. Press, Cambridge, 1999, pp. 26–55. MR1802425
- [5] R. F. Bass, M. Kassmann, and T. Kumagai, *Symmetric jump processes: localization, heat kernels and convergence*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 1, 59–71. MR2641770
- [6] S. Cao, R. S. Strichartz, and M. Wei, *Eigenvalues of Laplacians on higher dimensional Vicsek set graphs*, Fractals **30** (2021), no. 1, Paper No. 2250019.
- [7] Z.-Q. Chen, D. A. Croydon, and T. Kumagai, *Quenched invariance principles for random walks and elliptic diffusions in random media with boundary*, Ann. Probab. **43** (2015), no. 4, 1594–1642. MR3353810
- [8] D. A. Croydon, *Scaling limits of stochastic processes associated with resistance forms*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018), no. 4, 1939–1968. MR3865663
- [9] D. A. Croydon and B. M. Hambly, *Local limit theorems for sequences of simple random walks on graphs*, Potential Anal. **29** (2008), no. 4, 351–389. MR2453564
- [10] D. A. Croydon, B. M. Hambly, and T. Kumagai, *Time-changes of stochastic processes associated with resistance forms*, Electron. J. Probab. **22** (2017), Paper No. 82. MR3718710
- [11] T. Delmotte, *Parabolic Harnack inequality and estimates of Markov chains on graphs*, Rev. Mat. Iberoamericana **15** (1999), no. 1, 181–232. MR1681641
- [12] R. L. Dobrushin and S. Kusuoka, *Statistical mechanics and fractals*, Lecture Notes in Mathematics, vol. 1567, Springer-Verlag, Berlin, 1993, With a foreword by Ze Pei Jiang, Shi Jian Yan, Ping Cheng and Rong Wu. MR1295841
- [13] P. J. Fitzsimmons, B. M. Hambly, and T. Kumagai, *Transition density estimates for Brownian motion on affine nested fractals*, Comm. Math. Phys. **165** (1994), no. 3, 595–620. MR1301625
- [14] M. Fukushima, *Dirichlet forms, diffusion processes and spectral dimensions for nested fractals*, Ideas and methods in mathematical analysis, stochastics, and applications (Oslo, 1988), Cambridge Univ. Press, Cambridge, 1992, pp. 151–161. MR1190496
- [15] M. Fukushima, Y. Oshima, and M. Takeda, *Dirichlet forms and symmetric Markov processes*, extended ed., De Gruyter Studies in Mathematics, vol. 19, Walter de Gruyter & Co., Berlin, 2011. MR2778606
- [16] A. Grigor'yan and A. Telcs, *Sub-Gaussian estimates of heat kernels on infinite graphs*, Duke Math. J. **109** (2001), no. 3, 451–510. MR1853353
- [17] A. Grigor'yan and A. Telcs, *Harnack inequalities and sub-Gaussian estimates for random walks*, Math. Ann. **324** (2002), no. 3, 521–556. MR1938457
- [18] B. M. Hambly and T. Kumagai, *Transition density estimates for diffusion processes on post critically finite self-similar fractals*, Proc. London Math. Soc. (3) **78** (1999), no. 2, 431–458. MR1665249
- [19] B. M. Hambly and T. Kumagai, *Asymptotics for the spectral and walk dimension as fractals approach Euclidean space*, Fractals **10** (2002), no. 4, 403–412. MR1951472
- [20] N. Kajino and M. Murugan, *On singularity of energy measures for symmetric diffusions with full off-diagonal heat kernel estimates*, Ann. Probab. **48** (2020), no. 6, 2920–2951. MR4164457
- [21] O. Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR1876169
- [22] J. Kigami, *Analysis on fractals*, Cambridge Tracts in Mathematics, vol. 143, Cambridge University Press, Cambridge, 2001. MR1840042
- [23] J. Kigami, *Resistance forms, quasisymmetric maps and heat kernel estimates*, Mem. Amer. Math. Soc. **216** (2012), no. 1015, vi+132. MR2919892
- [24] T. Kumagai, *Estimates of transition densities for Brownian motion on nested fractals*, Probab. Theory Related Fields **96** (1993), no. 2, 205–224. MR1227032

- [25] T. Kumagai, *Random walks on disordered media and their scaling limits*, Lecture Notes in Mathematics, vol. 2101, Springer, Cham, 2014, Lecture notes from the 40th Probability Summer School held in Saint-Flour, 2010, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. MR3156983
- [26] S. Kusuoka and X.Y. Zhou, *Dirichlet forms on fractals: Poincaré constant and resistance*, Probab. Theory Related Fields **93** (1992), no. 2, 169–196. MR1176724
- [27] K. Kuwae and T. Shioya, *Convergence of spectral structures: a functional analytic theory and its applications to spectral geometry*, Comm. Anal. Geom. **11** (2003), no. 4, 599–673. MR2015170
- [28] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009, With a chapter by James G. Propp and David B. Wilson. MR2466937
- [29] T. Lindstrøm, *Brownian motion on nested fractals*, Mem. Amer. Math. Soc. **83** (1990), no. 420, iv+128. MR0988082
- [30] S. Milošević, D. Stassinopoulos, and H. E. Stanley, *Asymptotic form of the spectral dimension at the fractal to lattice crossover*, J. Phys. A **21** (1988), no. 6, 1477–1482. MR0939752