Heat Kernel Estimates and Law of the Iterated Logarithm for Symmetric Random Walks on Fractal Graphs

Ben M. Hambly and Takashi Kumagai

Abstract. We study two-sided heat kernel estimates on a class of fractal graphs which arise from a subclass of finitely ramified fractals. These fractal graphs do not have spatial symmetry in general, and we find that there is a dependence on direction in the estimates. We will give a new form of expression for the heat kernel estimates using a family of functions which can be thought of as a "distance for each direction". As an application, we give a law of the iterated logarithm which shows that the directional dependence leads to non-uniform behaviour in the typical paths of the random walk.

1. Introduction

There is a long history of work on Gaussian bounds for the heat kernel on spaces with smooth structure. The development of analysis on fractals showed that there are sets where heat kernels naturally have sub-Gaussian bounds. Graphs provide a setting in which to further investigate the generality of these results. One area of recent research activity is in determining analytic conditions which are equivalent to heat kernel estimates of sub-Gaussian type ([**BB1, BCG, GT1, GT2**]). Our interest here will be in exploring the range of behaviour that can be exhibited by the heat kernel in a class of self-similar graphs (cf. [**BB2, HK2, Jo**]).

The class of graphs we consider consists of infinite graphs generated using a set of contraction maps all with the same scale parameter. Instead of iterating forward the set of maps to obtain a fractal set, we take one of the contractions and invert it, then apply it to a finite set (the natural boundary of the fractal) and iterate to obtain an infinite self-similar set of vertices. These can then be connected in a way compatible with the fractal structure to form a fractal graph. The graphs we consider will be finitely ramified in the sense that any subset can be disconnected by removing a finite number of points and hence we will call them uniformly finitely ramified graphs. Some examples are shown in Figure 1 in Section 3. Note that

---

2000 *Mathematics Subject Classification.* Primary 60J10, 60J35; Secondary 60G17, 31C25.

Key words and phrases. Random walks on graphs, sub-Gaussian heat kernel bounds, law of iterated logarithm, geodesic distance, diffusion on fractals.

The second author was supported in part by Japan Association for Mathematical Sciences and Grant-in-Aid for Scientific Research (C)(2) 14540113.

these graphs are based on triangles but they do not have the full symmetry of the triangle and this leads to a variety of interesting behaviour.

In a previous paper [**HK2**] we considered the class of uniform finitely ramified fractal graphs and obtained an expression for the heat kernel. It was shown that in the case of nested fractal graphs, in which the underlying fractal set is invariant under the full symmetry group of the basic cell, this expression is equivalent to the usual formulation of sub-Gaussian estimates. That is there are constants d_s, d_w and $c_{1.1}, c_{1.2}$ (which differ in the upper and lower bounds), such that

$$p_k(x,y) \asymp c_{1.1} k^{-d_s/2} \exp(-c_{1.2}(\frac{d(x,y)^{d_w}}{k})^{1/(d_w-1)}), \quad \forall x, y, d(x,y) \le k,$$

where $d(x,y)$ is the usual graph distance (see (2.8) for the meaning of $\asymp$ here). The exponent $d_s = 2d_f/d_w$ is called the spectral dimension, d_w is the walk dimension, which describes the scaling in the time for the random walk to exit balls and d_f is the Hausdorff dimension of the underlying fractal. We also showed that, without symmetry, the estimates cannot in general be expressed so simply in terms of the graph metric. Indeed, for fractals of this type, where the random walk on the graph is strongly recurrent, the natural metric is the resistance metric, in the sense that the parabolic Harnack inequality holds with respect to this metric. When we write the estimates in terms of the usual graph distance there is a dependence on direction. In this paper we will give a more complete description of this dependence and show how the estimates can be expressed in a form involving functions, one for each direction in the graph, which are like distance functions.

In particular, for the case where there are two directions (each corresponds to a certain type of edge in the graph, for instance diagonal and horizontal for our Sierpinski gasket based examples in Figure 1 in Section 3), we typically have constants $c_{1.3}, c_{1.4}$ (differing in the upper and lower bounds) such that

$$p_k(x,y) \asymp c_{1.3} k^{-d_s/2} \exp(-c_{1.4} \sum_{i=1}^{2} (\frac{d_i(x,y)^{d_w^i}}{k})^{1/(d_w^i-1)}) \quad \forall x, y, d_i(x,y) \le k,$$

where the d_w^i is the walk dimension in the direction i and d_i is a function which is similar to a metric. Though this is typical, it may also be the case that there is a logarithmic correction as well. The full result for the two type case is given in Theorem 3.5.

Finally, once we have such estimates for the heat kernel, we can determine what information they give concerning the random walk itself. Here we will show that the directional dependence leads to two laws of the iterated logarithm, one for each direction. This means that extreme fluctuations of the path in one direction are larger than those in the other.

We believe that similar expressions for the heat kernel estimates are possible for all the p.c.f. graphs and p.c.f. fractals which have a non-degenerate harmonic structure by extending the arguments in this paper. We discuss the diffusion case in Section 5 and partially answer in the affirmative two conjectures made in [**HK3**].

2. Uniform finitely ramified graphs and their quadratic forms

2.1. Uniform finitely ramified graphs. For $\alpha > 1$ and $I = \{1, 2, \cdots, N\}$, let $\{\Psi_i\}_{i \in I}$ be a family of α-*similitudes* on $\mathbb{R}^D$. An α-similitude is a map $\Psi_i \mathbf{x} = \alpha^{-1} U_i \mathbf{x} + \gamma_i, \ \mathbf{x} \in \mathbb{R}^D$ where U_i is a unitary map and $\gamma_i \in \mathbb{R}^D$. We will impose several assumptions on this family. First, we assume

(H-0) $\{\Psi_i\}_{i\in I}$ satisfies the open set condition,

that is there is a non-empty, bounded open set W such that $\{\Psi_i(W)\}_{i\in I}$ are disjoint and $\cup_{i\in I}\Psi_i(W) \subset W$. As $\{\Psi_i\}_{i\in I}$ is a family of contraction maps, there exists a unique non-void compact set $\hat{K}$ such that $\hat{K} = \cup_{i\in I}\Psi_i(\hat{K})$. We assume

(H-1) $\hat{K}$ is connected.

Let Fix be the set of fixed points of the Ψ_i's, $i \in I$. A point $x \in Fix$ is called an *essential fixed point* if there exist $i, j \in I$, $i \neq j$ and $y \in Fix$ such that $\Psi_i(x) = \Psi_j(y)$. Let I_F be the set of $i \in I$ for which the fixed point of Ψ_i is an essential fixed point. We write $\hat{V}_0$ for the set of essential fixed points. We also let $\Psi_{i_1,\dots,i_n} = \Psi_{i_1} \circ \cdots \circ \Psi_{i_n}$. We make one further important assumption, that the family of contraction maps has a finite ramification property;

(H-2) If $\{i_1, \dots, i_n\}, \{j_1, \dots, j_n\}$ are distinct sequences, then

$$\Psi_{i_1,\dots,i_n}(\hat{K}) \bigcap \Psi_{j_1,\dots,j_n}(\hat{K}) = \Psi_{i_1,\dots,i_n}(\hat{V}_0) \bigcap \Psi_{j_1,\dots,j_n}(\hat{V}_0).$$

DEFINITION 2.1. ([**HK2**]) A (compact) uniform finitely ramified fractal (u.f.r. fractal for short) $\hat{K}$ is a set determined by α-similitudes $\{\Psi_i\}_{i\in I}$ satisfying the assumptions (H-0), (H-1), (H-2) and with $\sharp\hat{V}_0 \geq 2$.

If we also assume the following symmetry condition, then $\hat{K}$ is called a (compact) nested fractal, as introduced and discussed in [**Lind, Kus**].

(SYM) If $x, y \in \hat{V}_0$, then the reflection in the hyperplane $H_{xy} = \{z \in \mathbb{R}^D : |z - x| = |z - y|\}$ maps $\hat{V}_n$ to itself, where $\hat{V}_n = \cup_{i_1,\cdots,i_n\in I}\Psi_{i_1,\dots,i_n}(\hat{V}_0)$.

Thus u.f.r. fractals form a class of fractals which is wider than nested fractals, and is included in the class of p.c.f. self-similar sets ([**Kig**]).

Next we define unbounded u.f.r. fractals. We assume without loss of generality that $\Psi_1(\mathbf{x}) = \alpha^{-1}\mathbf{x}$ and that $\mathbf{0}$ belongs to $\hat{V}_0$. Let $K = \cup_{n=1}^{\infty}\alpha^n\hat{K}$. Then, clearly $\Psi_1(K) = K$. We call K an unbounded uniform finitely ramified fractal. Let $V = V_0 = \cup_{n=0}^{\infty}\alpha^n\hat{V}_n$ and $V_n = \alpha^n V$ for $n \in \mathbb{Z}$. (Note that this labelling is the opposite to the one given in [**HK3**]. As n gets bigger, the graph distance between each vertex of V_n gets bigger and $V_n \subset V_{n-1}$.) Then, $K = Cl(\cup_{n\in\mathbb{Z}}V_n)$. For each $l, n \geq 0$ and $i_1, \cdots, i_l \in I$, we call a set of the form $\alpha^{n+l}\Psi_{i_1,\cdots,i_l}(\hat{V}_0)$ an n-cell and $\alpha^{n+l}\Psi_{i_1,\cdots,i_l}(\hat{K})$ an n-complex.

We now introduce uniform finitely ramified graphs. These will be graphs with vertices V and a collection of edges B. In order to define the edges, we first define e_{xy} for an edge between x and y in V and then set $\hat{B}_0 := \{e_{xy} : x \neq y \in \hat{V}_0\}$. Then inside each 0-cell we place a copy of $\hat{B}_0$ and we denote by B the set of all the edges determined in this way. We call the graph (V, B) a uniform finitely ramified (u.f.r.) graph. If we construct the graph starting from a nested fractal, then it will be called a nested fractal graph.

2.2. Quadratic forms. An *electrical network* (V, C) on an arbitrary infinite connected graph with vertices V is an assignment to each edge e_{xy} of a positive number $C_{xy} = C_{yx}$, the conductance between x and y. For our fractal graphs a *basic electrical network* (V, C) on (V, B) is defined as follows. Firstly we assign a conductance to each edge $e_{xy} \in \hat{B}_0$. The conductance matrix C is constructed on the whole graph by putting on each edge in each 0-cell the same conductance as that of the corresponding edge in $\hat{B}_0$ and setting $C_{xy} = 0$ if $e_{x,y} \notin B$. Note that, for our fractal electrical networks, there exists $c_{2.1}, c_{2.2} > 0$ such that

$$c_{2.1} \le C_{xy} \le c_{2.2} \qquad \text{for all } \ e_{xy} \in B. \tag{2.1}$$

We next define a quadratic form on (V, B) associated with the electrical network. For $x, y \in V$, we write $x \sim y$ if $e_{xy} \in B$. For each $f, g \in l(V) := \{h : h \text{ is a function on } V\}$, we define

$$\mathcal{E}_C(f,g) = \frac{1}{2} \sum_{\substack{x,y \in V \\ e_{xy} \in B}} (f(x) - f(y))(g(x) - g(y)) C_{xy}. \tag{2.2}$$

We sometimes abbreviate $\mathcal{E}_C(f,f)$ as $\mathcal{E}_C(f)$. Now, define $\mu_x = \sum_{y \in V} C_{xy}$ for each $x \in V$. Set $\mu(A) = \sum_{x \in A} \mu_x$ for each $A \subset V$; μ is then a measure on V. For each $e_{xy} \in B$, define $P_{xy} = C_{xy}/\mu_x$, the transition probability matrix of the Markov chain corresponding to $\mathcal{E}_C$. To be precise, the process corresponding to $\mathcal{E}_C$ is a continuous time Markov chain in which jumps occur along edge e_{xy} at rate C_{xy}. In this paper, we will consider instead the induced discrete time random walk, the discrete time Markov chain which moves at unit time intervals to any vertex y connected to x with the probabilities for these jumps given by $\{P_{xy}\}$. We denote this induced random walk by $\{X_k\}_{k \ge 0}$. The random walk is reversible with respect to μ, indeed,

$$P_{xy}\mu_x = C_{xy} = C_{yx} = P_{yx}\mu_y.$$

The discrete Laplace operator corresponding to the random walk can be defined as

$$\mathcal{L}f(x) = \sum_y P_{xy} f(y) - f(x) = \frac{1}{\mu_x} \sum_y (\nabla_{xy} f) C_{xy},$$

where $\nabla_{xy} f = f(y) - f(x)$.

Let $\mathcal{Q}_M = \mathcal{Q}_M(\hat{V}_0)$ be the set of all $Q = \{q_{ij} : i, j \in \hat{V}_0$ such that $q_{ii} = 0, q_{ij} = q_{ji} \ge 0$ for any $i, j \in \hat{V}_0\}$. We denote by $\mathrm{Int}(\mathcal{Q}_M)$ the subset of $\mathcal{Q}_M$ such that every Q has $q_{ii} = 0, q_{ij} = q_{ji} > 0$ for all $i \sim j \in \hat{V}_0$. Also, let $\mathcal{Q}_{irr}$ be the set of $Q \in \mathcal{Q}_M$ such that $\mathcal{E}_Q(\xi, \xi) = 0$ if and only if ξ is constant. As our quadratic form on the graph V is constructed from a basic electrical network, the conductances are determined by placing $C \in \mathcal{Q}_M$ within each 0-cell, and we denote the form by $\mathcal{E}_C(\cdot, \cdot)$. Given the graph on V_1 which has the same structure as (V_0, B) and $C' \in \mathrm{Int}(\mathcal{Q}_M(\alpha \hat{V}_0))$, we can define a form on V_1 in the same way:

$$\mathcal{E}^{(1)}_{C'}(f,g) := \frac{1}{2} \sum_{\substack{x,y \in V_1 \\ e_{xy} \in \alpha B}} (f(x) - f(y))(g(x) - g(y)) C'_{xy} \qquad \text{for all } f, g \in l(V_1),$$

where $\{C'_{xy}\}_{x,y}$ is given by placing C' in each 1-cell. As our graph is finitely ramified, we know that for each $C \in \mathcal{Q}_M(\hat{V}_0)$, there exists $C' \in \mathcal{Q}_M(\alpha \hat{V}_0)$ so that

$$\mathcal{E}^{(1)}_{C'}(v) = \inf\{\mathcal{E}_C(f) : f \in l(V_0), f|_{V_1} = v\} \qquad \text{for all } v \in l(V_1),$$

see [**Bar, Kig**] for the proof. We can define a decimation map F from $\mathcal{Q}_M = \mathcal{Q}_M(\hat{V}_0)$ to itself by setting $F(C) = C'$. Note that F is homogeneous, as $F(\theta C) = \theta F(C)$ for all $\theta > 0$ and $C \in \mathcal{Q}_M$, however F is in general a non-linear map. In order to study the asymptotic properties of the form, it is important to observe the dynamics of the iteration of F (see [**Kum1**] and the references therein). By Schauder's fixed point theorem, we know that there exists $Q \in \mathcal{Q}_M$ (with $q_{ij} > 0$ for some $i \neq j$) and $\rho_Q^{-1} > 0$ such that $F(Q) = \rho_Q^{-1} Q$. Throughout this paper, we assume the following.

ASSUMPTION 2.2. (1) *For each $Q \in \mathcal{Q}_{irr}$, there exists $l = l(Q) \in \mathbb{N}$ such that $F^n(Q) \in Int(\mathcal{Q}_M)$ for all $n \geq l$.*
(2) *There exists $Q_0 \in Int(\mathcal{Q}_M)$ and $\rho_{Q_0} > 0$ such that $F(Q_0) = \rho_{Q_0}^{-1} Q_0$.*

REMARK 2.3. (1) By Corollary 6.20 of [**Bar**], $\rho_{Q_0} > 0$ is uniquely determined, that is if $Q_1, Q_2 \in \mathcal{Q}_{irr}$ satisfies $F(Q_j) = \rho_{Q_j}^{-1} Q_j$ $(j = 1, 2)$ with $\rho_{Q_1}, \rho_{Q_2} > 0$, then $\rho_{Q_1} = \rho_{Q_2} = \rho_{Q_0}$. In the class of fractal graphs we consider, we can prove $\rho_{Q_0} > 1$ (see [**Kig**] etc.). For the rest of the paper we denote ρ_{Q_0} by ρ.
(2) For $x, y \in \hat{K}$, the set of vertices $\{x_0, \cdots, x_m\}$ is called an n-chain from x to y if $x_0 = x, x_m = y$, $x_j \in \hat{V}_n$ for $1 \leq j \leq m-1$ and x_i, x_{i+1} are in the same $(-n)$-complex for $0 \leq i \leq m-1$. A sufficient condition for Assumption 2.2 1) is the following.

> (H-3) There exists $l \in \mathbb{N}$ such that for each $x, y \in \hat{V}_0$, there is an l-chain $\{x_0, \cdots, x_m\}$ from x to y so that for each $1 \leq i \leq m-2$, there is a l-cell C_i with $x_i, x_{i+1} \in C_i$ and $\hat{V}_0 \cap C_i = \emptyset$.

Indeed, if (H-3) holds, it is easy to show $F^n(Q) \in \mathrm{Int}(\mathcal{Q}_M)$ for $n \geq l$, $Q \in \mathcal{Q}_{irr}$ by observing the corresponding Markov chain on $\hat{V}_n$.
(3) Every nested fractal satisfies Assumption 2.2 1) and 2). Indeed, (H-3) can be shown using (SYM) and [**Kus**] Lemma 2.10 (or [**Lind**] Proposition *IV*.11), establishing 1). The existence of the fixed point 2) is proved in [**Kus**] Theorem 3.10 and in [**Lind**] Theorem *V*.5.
(4) Note that any non-degenerate fixed point of ρF is not necessarily unique even for nested fractals. In [**Bar**] Example 6.13, a one parameter family of non-degenerate fixed points on the Vicsek set is given.
(5) Using [**Kum1**] Theorem 3.4, we see that the following, which corresponds to Assumption 2.3 in [**HK2**], always holds under our Assumption 2.2.

For all $Q \in \mathrm{Int}(\mathcal{Q}_M)$, there exist $c_{1,Q}, c_{2,Q} > 0$ such that,

$$c_{1,Q}\rho^{-n}(Q_0)_{ij} \leq (F^n(Q))_{ij} \leq c_{2,Q}\rho^{-n}(Q_0)_{ij} \qquad \text{for all } \ n \in \mathbb{N}, \ i, j \in \hat{V}_0.$$

Note that for the construction of diffusion processes on finitely ramified fractals, the standard approach is to start with quadratic forms (random walks) whose conductances (transition probabilities) are invariant under the decimation map. See [**Bar, Kig**] for details.

The natural metric on the graph obtained by counting the number of steps in the shortest path between points is denoted by $d(x, y)$ for $x, y \in V$.

The effective resistance between $x \neq y \in V$ is defined by

$$R(x,y)^{-1} = \inf\{\mathcal{E}_C(f,f) : f \in l(V), f(x) = 1, f(y) = 0\}. \tag{2.3}$$

We define $R(x, x) = 0$ for each $x \in V$. Note that for any $x \neq y \in V$, $R(x, y)$ is positive and finite. Let $\hat{B}_n = \{e_{xy} : x \neq y \in \alpha^n \hat{V}_0\}$ and define edges B_n by placing a copy of $\hat{B}_n$ in each of the n-cells in V_n. Under Assumption 2.2, there exist constants $c_{2.3}, c_{2.4}$ such that

$$c_{2.3}\rho^n \leq R(x,y) \leq c_{2.4}\rho^n, \ \text{ for all } \ n \in \mathbb{N}, \ e_{xy} \in B_n.$$

Let $B_R(x,r)$ be the ball centred at x and radius r in the resistance metric. Note that $B_R(x,r)$ is not necessarily connected (see Remark 7.19 of [**Bar**]). Let $S = \log N/\log\rho$, the Hausdorff dimension of the fractal in the resistance metric. The mass of balls in the resistance metric is then controlled as there exist $c_{2.5}, c_{2.6}, r_0 > 0$ such that ([**HK2**] Lemma 3.2).

$$c_{2.5} r^S \le \mu(B_R(x,r)) \le c_{2.6} r^S \qquad \text{for all } r \ge r_0. \tag{2.4}$$

2.3. Heat kernel estimates. Let $P_k(x,y)$ be the transition function after k steps, $P_k(x,y) = P(X_k = y | X_0 = x)$, where $\{X_k; k \in \mathbb{N}\}$ is the random walk on V associated with the Dirichlet form $\mathcal{E}_C$. The heat kernel $p_k(x,y)$ we will discuss is defined by

$$p_k(x,y) = P_k(x,y)/\mu_y.$$

Note that by the reversibility of $\{X_k; k \in \mathbb{N}\}$, we have $p_k(x,y) = p_k(y,x)$.

In order to give our expression for the heat kernel estimates we will require a shortest path counting function. Let $N_m(x,y)$ denote the number of edges in the shortest path on V_m from x to y. By definition we have $d(x,y) = N_0(x,y)$. If $x, y \notin V_m$, we define the shortest path counting function to be $N_m(x,y) = \max_{x_1 \in \partial D_m(x), y_1 \in \partial D_m(y)} N_m(x_1, y_1)$, where $D_l(x)$ is an l-complex containing x. Define the time scale factor $\tau = \rho N$ and let

$$l(k,n) = \inf\{j : N_j(x,y)\tau^j \ge k\} \wedge n. \tag{2.5}$$

We can now state our heat kernel estimates.

THEOREM 2.4 ([**HK2**]: Theorem 4.10). *There exist constants $c_{2.7}, \cdots, c_{2.10} > 0$ such that for all $x, y \in V$ and $k \ge d(x,y)$, if $\rho^{n-1} \le R(x,y) \le \rho^n$, then*

$$p_k(x,y) \le c_{2.7} k^{-\frac{S}{S+1}} \exp(-c_{2.8} N_{l(k,n)}(x,y)), \tag{2.6}$$

$$p_k(x,y) + p_{k+1}(x,y) \ge c_{2.9} k^{-\frac{S}{S+1}} \exp(-c_{2.10} N_{l(k,n)}(x,y)). \tag{2.7}$$

From now on we will express the pair of estimates (2.6), (2.7) as

$$p_k(x,y) \asymp c_1 k^{-\frac{S}{S+1}} \exp(-c_2 N_{l(k,n)}(x,y)). \tag{2.8}$$

Note that this theorem contains the diagonal estimate as $0 \le N_{l(k,n)}(x,x) \le 1$ for all $n, k \in \mathbb{N}$ and $x \in V$. For $x, y \in V$ such that $c_1\rho^n \le R(x,y) \le c_2\rho^n$ for some fixed constants, define a chemical exponent with respect to the resistance metric, for $0 \le l < n$ as $d_l^c(x,y) = (\log_\rho N_l(x,y))/(n-l)$. For $l = n$ we can choose the exponent arbitrarily and thus define $d_n^c(x,y) = 1$. Using the definition of $l(k,n)$ we can write Theorem 2.4 in terms of the resistance metric and this chemical exponent as follows.

COROLLARY 2.5. *There exist constants $c_{2.11}, c_{2.12} > 0$ such that for $x, y \in V$ and $k \ge d(x,y)$, with n, l as above, then*

$$p_k(x,y) \asymp c_{2.11} k^{-\frac{S}{S+1}} \exp\left(-c_{2.12} \left(\frac{R(x,y)^{S+1}}{k} \right)^{\frac{d^c_{l(k,n)}(x,y)}{S+1-d^c_{l(k,n)}(x,y)}} \right).$$

For nested fractal graphs, where the fractals have spatial symmetry, there are constants $d_c, c_{2.13}, c_{2.14} > 0$ (d_c is called the chemical exponent) such that for $e_{xy} \in B_n$,

$$c_{2.13}\rho^{(n-m)d_c} \le N_m(x,y) \le c_{2.14}\rho^{(n-m)d_c}.$$

Hence, writing $d_w = (S+1)/d_c$, we can express our estimates in the following form.

$$p_k(x,y) \asymp c_{2.15} k^{-S/(S+1)} \exp\left(-c_{2.16} \left(\frac{d(x,y)^{d_w}}{k} \right)^{1/(d_w-1)} \right). \tag{2.9}$$

We note that when there is no spatial symmetry, we do not necessarily have a single exponent d_w (as in (2.9)) for the off-diagonal estimates. See [**HK2**] for details.

We conclude this section with a result about the parabolic Harnack inequality. Let $\beta > 0$. We say $(\mathcal{E}_C, \mathbb{L}^2(V, d\mu))$ satisfies $(PHI(\beta))$, a parabolic Harnack inequality of order β, if whenever $u(n,x) \geq 0$ is defined on $[0,4N] \times \bar{B}(y,2r)$ and satisfies $u(n+1,x) - u(n,x) = \mathcal{L}u(n,x)$ for $(n,x) \in [0,4N] \times B(y,2r)$, then

$$\max_{\substack{N \leq n \leq 2N \\ x \in B(y,r)}} u(n,x) \leq c_{2.17} \min_{\substack{3N \leq n \leq 4N \\ x \in B(y,r)}} (u(n,x) + u(n+1,x)),$$

where $N \geq 2r$ and $c_{2.18} r^{\beta} \leq N \leq c_{2.19} r^{\beta}$ (cf. [**BB1, GT1, GT2, HK2**]). By Corollary 2.5 and a standard argument, we can deduce the following.

PROPOSITION 2.6.
$(\mathcal{E}_C, \mathbb{L}^2(V, d\mu))$ satisfies $(PHI(S+1))$ with respect to the resistance metric.

3. Heat kernel estimates and chemical exponents

In this section we will give a new form of expression for the heat kernel estimates which incorporates the dependence of the estimate on the direction. Our first step is to explain what we mean by a direction for our u.f.r. fractal graphs.

We begin by giving the elements of $\hat{B}_0$ a type. The classification uses the spatial symmetry of the fractal graph, that is, if there is a reflection which maps $\hat{V}_n$ to itself and maps $e \in \hat{B}_0$ to $e' \in \hat{B}_0$, then e and e' are of the same type, otherwise they are different. Let s_0 be the total number of types; clearly $s_0 \leq \sharp V_0 (\sharp V_0 - 1)/2$. By labelling each copy of $\hat{B}_0$ using the types, we can classify each element of the set of edges B by type. Also, by self-similarity, we can classify each edge which consists of two elements in an n-cell for any $n \in \mathbb{Z}$.

For $1 \leq i \leq s_0$ and $m \geq 1$, let $\pi_m(i)$ be the set of m-chains from x to y which do not contain multiple points, where $e_{xy} \in \hat{B}_0$ is of type i. Set

$$S^i = \{\mathbf{u}(\pi) \mid \pi \in \pi_1(i), \pi \text{ is minimal}\},$$

where we define $\mathbf{u}(\pi) = (\mathbf{u}(\pi)_1, \cdots, \mathbf{u}(\pi)_{s_0})$ to be an s_0-dimensional vector such that $\mathbf{u}(\pi)_j$ is the number of type j steps (edges) in the path π $(1 \leq j \leq s_0)$. Here, we say $\pi \in \pi_m(i)$ is minimal when there is no $\pi' \in \pi_m(i)$, $\pi' \neq \pi$ such that $\mathbf{u}(\pi')_j \leq \mathbf{u}(\pi)_j$ for all $1 \leq j \leq s_0$. Let

$$M_{mat} = \{A : A \text{ is an } s_0 \times s_0 \text{ matrix such that for each } j, (j\text{-th row of } A) \in S^j.\},$$

and define $\lambda_{min} = \min_{A \in M_{mat}} \{\text{maximum eigenvalue of } A\}$. Throughout this section, we call a matrix $A \in M_{mat}$ whose maximum eigenvalue attains λ_{min} the minimal step matrix, and define $M_{min} = \{A \in M_{mat} : A \text{ is a minimal step matrix.}\}$.

In this paper, we will only treat the case where the number of types is 2. This is purely for simplicity, but it does include essentially all the different forms for the heat kernel estimates. Note that the possible structures for the minimal step matrix are the following.

$$\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,b,c,d \in \mathbb{N} \cup \{0\}, a^2 + b^2 \geq 1, c^2 + d^2 \geq 1\}.$$

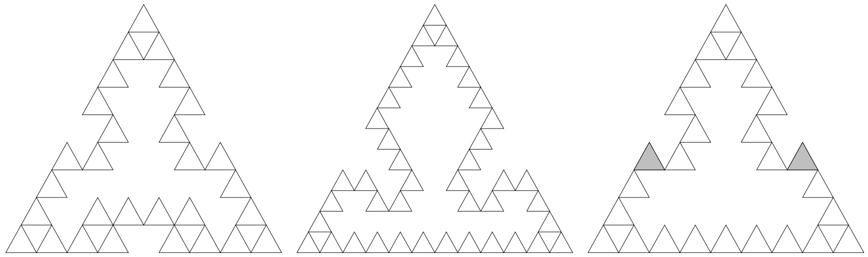

FIGURE 1. Examples

We first classify the possible structures.

Case (0) $a=0$; in this case there are three distinct structures which are possible for the minimal step matrix;

$$\text{(0-1)} \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} \text{(0-2)} \begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \text{(0-3)} \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix},$$

where $b,c,d \in \mathbb{N}$. By simple calculations, we see that the eigenvector corresponding to the maximum eigenvalue is a constant multiple of ${}^t(1,z)$ for some $z>0$ where ${}^t\mathbf{v}$ denotes the transpose of $\mathbf{v}$. The case (0') $d=0$ can be treated similarly.

We next consider the case $ad \neq 0$. There are four (essentially three) cases as indicated below;

$$\text{(1)} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{(2)} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \text{(2')} \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \text{(3)} \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix},$$

where $a,b,c,d \in \mathbb{N}$. Cases (2), (2') and (3) can each be decomposed into 3 subcases ($*$-1) $a>d$, ($*$-2) $a=d$ and ($*$-3) $a<d$. Note that cases (2') and (3-3) can be reduced to the cases (2) and (3-1) respectively by switching the role of type 1 and type 2.

We now describe how we choose and fix one element, denoted by A_{min}, from M_{min} (cf. Remark 3.6 (a)):

Case (a); there is a reducible matrix in M_{min}. (A non-negative matrix A is called irreducible if for each i,j, there exists $n=n(i,j)$ such that $(A^n)_{ij}>0$. A matrix which is not irreducible is called reducible.) By switching the role of type 1 and type 2 if necessary, we can write the matrix as in the case (2).

(a-1); there is a minimal step matrix with $d<\lambda_{min}$ (thus $a=\lambda_{min}$). Then define this matrix as A_{min}. Such a matrix in M_{min}, if it exists, is uniquely determined.

(a-2); (a-1) does not hold (thus $d=\lambda_{min}$ and $a \le \lambda_{min}$). If there is a matrix in M_{min} with $a<\lambda_{min}$, then take the matrix with $b=0$ (if it exists) as A_{min}, otherwise take the matrix where $(\lambda_{min}-a)/b$ attains a maximum (if the original matrix is as in the case (2'), take the matrix where $c/(\lambda_{min}-d)$ attains a maximum) within such matrices in M_{min}. If there is no matrix in M_{min} with $a<\lambda_{min}$, then take the matrix with $a=\lambda_{min}$ as A_{min}. Note that in such a case, one can prove that $M_{min}=\{(\begin{smallmatrix} \lambda_{min} & b \\ 0 & \lambda_{min} \end{smallmatrix})\}$ and $\lambda_{min}=\min_{\mathbf{b}\in S^1}\mathbf{b}_1$.

Case (b); all the matrices in M_{min} are irreducible. Then, take an arbitrary element of M_{min} as A_{min}.

In Figure 1, we indicate several concrete examples of the different cases. These are cases where diagonal moves correspond to type 1 paths and horizontal moves correspond to type 2 paths. The left figure is the case (1), the middle figure (cf. Figure 2 in [**HK2**]) is the case (2-1) and the right figure is the case (2-2). If the contraction maps corresponding to the shaded triangles have a rotation of 60 (or -60) degrees (so that the shaded triangles are rotations of the original one), then the right figure is the case (3). We note that it is easy to check that Assumption 2.2 holds for these cases. The corresponding A_{min}'s are as follows.

$$\begin{pmatrix} 9 & 1 \\ 2 & 8 \end{pmatrix}, \begin{pmatrix} 14 & 2 \\ 0 & 12 \end{pmatrix}, \begin{pmatrix} 9 & 1 \\ 0 & 9 \end{pmatrix}, \begin{pmatrix} 10 & 0 \\ 0 & 9 \end{pmatrix}.$$

For the cases (1), (2-3) and (3-2), the eigenvector corresponding to the maximum eigenvalue is again a constant multiple of ${}^t(1,z)$ for some $z > 0$. For the cases (2-1) and (3-1), the maximum eigenvalue is a, the corresponding eigenvector is a constant multiple of ${}^t(1,0)$ and the Jordan cell for the eigenvalue is 1-dimensional. For the case (2-2), the maximum eigenvalue is a, the corresponding eigenvector is a constant multiple of ${}^t(1,0)$ and the Jordan cell for the eigenvalue is 2-dimensional.

We now determine the relationship between A_{min} and the asymptotic growth of the minimum number of edges in a path for each direction. For $\mathbf{v} \geq 0$, define

$$(G\ \mathbf{v})_i = \min_{\mathbf{b}\in S^i} \mathbf{b}\cdot\mathbf{v} \qquad i = 1,2,$$

and hence $G\mathbf{v} = \min_{A\in M_{mat}} A\mathbf{v}$. Let $\lambda_{min} > 0$, $\mathbf{x} \geq \mathbf{0}$ be the maximum eigenvalue of A_{min} and its corresponding eigenvector, so that

$$G\ \mathbf{x} = A_{min}\mathbf{x} = \lambda_{min}\mathbf{x}. \tag{3.1}$$

When $\mathbf{x} \geq \mathbf{0}$ is not strictly positive, we can prove this as follows. Without loss of generality, we may assume that $\mathbf{x}_2 = 0$ and hence the matrix A_{min} is as in the case (2). Thus, $\lambda_{min} = a$. We can further see that $a = \min_{\mathbf{b}\in S^1} \mathbf{b}_1$. Indeed, if this does not hold for $d < \lambda_{min}$, by substituting the element of S^1 whose first element is minimal into the first row of A_{min}, we can obtain a matrix whose maximum eigenvalue is less than $\lambda_{min} = a$, which is a contradiction. For $d = \lambda_{min}$, this fact is guaranteed by our choice of A_{min}. We thus obtain

$$(G\ \begin{pmatrix} 1 \\ 0 \end{pmatrix})_1 = \min_{\mathbf{b}\in S^1} \mathbf{b}\cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \min_{\mathbf{b}\in S^1} \mathbf{b}_1 = a.$$

Clearly $(G\ {}^t(1,0))_2 = 0$ and we obtain (3.1). When $\mathbf{x} > \mathbf{0}$, we can prove (3.1) using the properties of non-negative matrices, but omit the somewhat lengthy proof here.

Let $x, y \in V_n$ be in the same n-cell and assume the edge e_{xy} is of type i, $(i = 1,2)$. Clearly $N_n(x,y) = (G^{\circ n}\mathbf{1})_i$ where $G^{\circ n}$ is the n-th iteration of G and $\mathbf{1} = {}^t(1,1)$. Thus, when the eigenvector $\mathbf{x}$ of λ_{min} is strictly positive (that is, cases (0), (1), (2-3), (3-2)), using (3.1) we have constants such that

$$c_1\lambda_{min}^n\mathbf{x}_i = c_1(G^{\circ n}\mathbf{x})_i \leq (G^{\circ n}\mathbf{1})_i \leq c_2(G^{\circ n}\mathbf{x})_i = c_2\lambda_{min}^n\mathbf{x}_i.$$

Hence there is no directional dependence for the asymptotic order of the minimum number of edges in the shortest path and the problem is reduced to the case of a single type as discussed in [**Kum2, FHK**]. In particular, the heat kernel estimate has the same form as (2.9).

We thus consider only the cases (2-1), (2-2) and (3-1) in the following. For each case, the n-th power of A_{min} is given by

$$\text{(2-1)} \begin{pmatrix} a^n & \frac{b}{a-d}(a^n - d^n) \\ 0 & d^n \end{pmatrix} \text{(2-2)} \begin{pmatrix} a^n & bna^{n-1} \\ 0 & a^n \end{pmatrix} \text{(3-1)} \begin{pmatrix} a^n & 0 \\ 0 & d^n \end{pmatrix}.$$

Let $\pi(x,y)$ be the shortest path (with respect to the graph distance) from x to y in V. Note that each edge in the path is either of type 1 or type 2. We let $d_1(x,y)$ be the number of type 1 steps in $\pi(x,y)$. We decompose the path $\pi(x,y)$ into connected components C_i that only consist of type 2 steps and define $d_2(x,y) = \max_i |C_i|$. In other words, $d_2(x,y)$ is the maximum length of a connected chain of type 2 steps in the shortest path. Note that $d_i(\cdot,\cdot)$ is not a distance.

In order to define two associated shortest path counting functions, we decompose the shortest path. For $x,y \in V$ such that $\rho^{n-1} \le R(x,y) < \rho^n$, we can decompose the shortest path from x to y as follows. Let x_i be the first element of V_i on the shortest path from x ($x_i = x$ if $x \in V_i$). Define y_i in the same way by looking at the path from y. Then there is an $L = L(x,y) \le n$ such that there in no element in V_{L+1} on the path between x_L and y_L. By the self-similarity of the graph, it is easy to see that either $x_L = y_L$ or the number of elements in V_L on the path between x_L and y_L is uniformly bounded by some positive constant. We define $N^1_m(x,y)$ as the number of type 1 steps in the shortest path on V_m from x_m to y_m if $x_m \ne y_m$. We decompose the shortest path on V_m from x_m to y_m into connected components $C_{m,i}$ that only consist of type 2 steps and define $N^2_m(x,y) = \max_i |C_{m,i}|$ if $x_m \ne y_m$. Finally we set $N^i_m(x,y) = 1$ if $x_m = y_m$ and $m \ne 0$ (this is for consistency with the definition of $N_m(x,y)$). By definition we have $d_i(x,y) = N^i_0(x,y)$ for $i = 1,2$.

We now change notation and let $\alpha_1 = a, \alpha_2 = d$. Note that $\alpha_1 > \alpha_2$. For $x,y \in V$ such that $\rho^{n-1} \le R(x,y) < \rho^n$, define $n_i = n_i(x,y) \le n$ so that $\alpha_i^{n_i} \le d_i(x,y) < \alpha_i^{n_i+1}$ for $i = 1,2$. Note that $N^i_{m'}(x,y) = 1$ for $m' > n_i$. Note also that $n_1 \vee n_2 = n$.

Let $\bar N^1_m(x,y) = N^1_m(x,y) \log N^1_m(x,y)$ for the case (2-2), $\bar N^1_m(x,y) = N^1_m(x,y)$ otherwise. We let $\bar N^2_m(x,y) = N^2_m(x,y)$ for all the cases. Then the following holds.

LEMMA 3.1. *There exists $c_{3.1}, c_{3.2} > 0$ such that for all $m \ge 0$ and $x,y \in V$,*

$$c_{3.1} \sum_{i=1}^{2} \bar N^i_m(x,y) \le N_m(x,y) \le c_{3.2} \sum_{i=1}^{2} \bar N^i_m(x,y).$$

PROOF. Here and in the following, we sometimes abbreviate $N_m(x,y)$ (resp. $\bar N^i_m(x,y)$) as N_m (resp. $\bar N^i_m$). Firstly, when $x,y \in V_n$ are in the same n-cell and e_{xy} is of type i ($i = 1,2$), then $d_i(x,y) = \alpha_i^n$ and $N^i_m = \alpha_i^{n-m} = \alpha_i^{-m} d_i(x,y)$ for all $m \le n$. Next, for $x,y \in V$ such that $\rho^{n-1} \le R(x,y) < \rho^n$, we can decompose the shortest path from x to y into $\{x, x_1, \cdots, x_L, y_L, \cdots, y_1, y\}$ as discussed above. Using the first observation and the self similarity of the graph, we see that

$$N^1_m = \left(\sum_{j=m}^{L-1} \alpha_1^{-m} d_1(x_j, x_{j+1}) + \sum_{j=m}^{L-1} \alpha_1^{-m} d_1(y_j, y_{j+1}) + \alpha_1^{-m} d_1(x_L, y_L) \right) \vee 1,$$

$$\max\{\alpha_2^{-m} d_2(x_j, x_{j+1}), \alpha_2^{-m} d_2(y_j, y_{j+1}), \alpha_2^{-m} d_2(x_L, y_L) : m \le j \le L-1\} \le$$

$$N^2_m \le \left(\sum_{j=m}^{L-1} \alpha_2^{-m} d_2(x_j, x_{j+1}) + \sum_{j=m}^{L-1} \alpha_2^{-m} d_2(y_j, y_{j+1}) + \alpha_2^{-m} d_2(x_L, y_L) \right) \vee 1.$$

For $i = 1,2$, let q_i be the largest $j \ge 0$ such that either $d_i(x_j, x_{j+1})$ or $d_i(y_j, y_{j+1})$ is non-zero. When $d_i(x_L, y_L) \ne 0$, we define $q_i = L$ and when all these values are 0 we define $q_i = m$. Note that there are constants such that $c_1 \alpha_i^j \le d_i(x_j, x_{j+1}) \le c_2 \alpha_i^j$

if it is not zero. The same fact is true for $d_i(y_j, y_{j+1})$ and $d_i(x_L, y_L)$ (in this case $j = L$). Thus, using the estimates of N_m^i above, we have constants such that $c_3\alpha_i^{q_i-m} \le N_m^i \le c_4\alpha_i^{q_i-m}$. Since $\alpha_i^{n_i} \le d_i(x,y) < \alpha_i^{n_i+1}$, we see that $n_i + c_5 \le q_i \le n_i + c_6$ when $n_i \ge m$ and

$$c_7[\alpha_i^{n_i-m} \vee 1] \le N_m^i \le c_8[\alpha_i^{n_i-m} \vee 1] \qquad i = 1,2, \tag{3.2}$$

where $[x]$ is the maximum integer not greater than x.

We next compute N_m. Note that we are in the case $A_{min} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ with $a \ge d$ ($d = a$ for the case (2-2) and $b = 0$ for the case (3-1)). Firstly, when $x, y \in V_n$ are in the same n-cell and e_{xy} is of type i, ($i = 1,2$), then $N_m = (G^{\circ(n-m)}\ {}^t(1,1))_i$. Clearly,

$$G^{\circ(n-m)}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \le A_{min}^{n-m}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \le \begin{cases} c_9\begin{pmatrix} a^{n-m} \\ d^{n-m} \end{pmatrix} & \text{for (2-1), (3-1),} \\ c_9\begin{pmatrix} (n-m)a^{n-m} \\ a^{n-m} \end{pmatrix} & \text{for (2-2).} \end{cases} \tag{3.3}$$

Now, since $a = \min_{\mathbf{b}\in S^1} \mathbf{b}_1$ and S^1, S^2 are finite sets, by taking $\gamma \ge 1$ large enough, we have $\gamma a + b = \min_{\mathbf{b}\in S^1}(\gamma\mathbf{b}_1 + \mathbf{b}_2)$ and $d = \min_{\mathbf{b}\in S^2}(\gamma\mathbf{b}_1 + \mathbf{b}_2)$. In particular,

$$G\begin{pmatrix} \gamma \\ 1 \end{pmatrix} = \begin{pmatrix} \gamma a + b \\ d \end{pmatrix} \ge d\begin{pmatrix} \gamma \\ 1 \end{pmatrix}. \tag{3.4}$$

On the other hand,

$$G^{\circ(n-m)}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \ge \frac{1}{\gamma+1}G^{\circ(n-m)}\begin{pmatrix} \gamma+1 \\ 1 \end{pmatrix} \ge \frac{a^{n-m}}{\gamma+1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\gamma+1}G^{\circ(n-m)}\begin{pmatrix} \gamma \\ 1 \end{pmatrix},$$

where we apply (3.1) to obtain the last inequality. Applying (3.4), we have the same lower bound (with different constant) as (3.3) for (2-1) and (3-1). For the case of (2-2), we have

$$G\begin{pmatrix} (\gamma a + (l+1)b)a^l \\ a^{l+1} \end{pmatrix} \ge a^{l+1}G\begin{pmatrix} \gamma \\ 1 \end{pmatrix} + (l+1)ba^lG\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} (\gamma a + (l+2)b)a^{l+1} \\ a^{l+2} \end{pmatrix},$$

for each $l \ge 0$. Applying this inductively, we have the same lower bound as (3.3) (with different constant).

Finally we consider the general case $x, y \in V$ with $\rho^{n-1} \le R(x,y) < \rho^n$. By the definition of n_i, $\alpha_i^{n_i} \le d_i(x,y) < \alpha_i^{n_i+1}$ for $i = 1,2$. Using the same decomposition of the shortest path, by similar arguments to the case of N_m^i and by applying the result for $x, y \in V_n$ being in the same n-cell, we have the following (since the arguments are similar, we leave the details for the reader); N_m can be estimated from above and below by some (uniform) constant multiple of the expressions given below (note that $a = \alpha_1, b = \alpha_2$),

$$\text{(2-1) and (3-1)} \ \sum_{i=1}^{2}[\alpha_i^{n_i-m} \vee 1], \qquad \text{(2-2)} \ (n_1 - m)[\alpha_1^{n_1-m}] + [\alpha_1^{n_2-m} \vee 1].$$

Combining this with (3.2) and by the definition of $\bar{N}_m^i$, we obtain the result. □

For later convenience, we take $c_{3.1} \le 1$ and $c_{3.2} \ge 1$. For each $c > 0$, let

$$l_c^i = l_c^i(k,n) = \inf\{j : \bar{N}_j^i(x,y)\tau^j \ge ck\} \wedge n.$$

We will omit c when $c = 1$.

LEMMA 3.2. (a) *Assume* $l^i < n_i$*, then* $l_{c_{3.1}} \le l^i$ *where* $i = 1, 2$.
(b) *Let* $j_0 = j_0(l, x, y)$ *be such that* $\max_i \bar{N}_l^i = \bar{N}_l^{j_0}$*. Assume* $l < n$*, then* $l^{j_0}_{1/(2c_{3.2})} \le l$.

PROOF. For (a), using Lemma 3.1 and the fact that $l_i < n_i$, we have

$$k \le \bar{N}^i_{l^i} \tau^{l^i} \le \frac{1}{c_{3.1}} N_{l^i} \tau^{l^i},$$

giving the result. For (b), note that $N_l/2 \le c_{3.2} \bar{N}_l^{j_0}$. Since $l < n$, we have

$$\frac{k}{2} \le \frac{N_l}{2} \tau^l \le c_{3.2} \bar{N}_l^{j_0} \tau^l,$$

giving the result. □

LEMMA 3.3. (a) *For* $c \ge 1$*, there exists* $c_{3.3}(c) \ge 1$ *such that*

$$N_{l_c}(x, y) \le N_l(x, y) \le c_{3.3}(c) N_{l_c}(x, y).$$

(b) *For each* $c < 1$*, there exists* $c_{3.4}(c) \le 1$ *such that*

$$c_{3.4}(c) N_{l_c}(x, y) \le N_l(x, y) \le N_{l_c}(x, y).$$

Both (a) *and* (b) *also hold for* $\bar{N}^i$.

PROOF. When $c \ge 1$, $l_c \ge l$ so that $N_{l_c} \le N_l$ and the first inequality of (a) is clear. To prove the second inequality of (a), we first consider the case (2-1) and (3-1). As discussed in the proof of Lemma 3.1, we have constants c_1, c_2 such that

$$c_1 \sum_{i=1}^2 [\alpha_i^{n_i - m} \vee 1] \le N_m \le c_2 \sum_{i=1}^2 [\alpha_i^{n_i - m} \vee 1].$$

If $l = n$, then $l_c = n$ and the result is clear. We thus consider the case $l < n$. In this case, $N_l \tau^l \ge k$ so that

$$k \le c_2 \sum_{i=1}^2 [\alpha_i^{n_i - l} \vee 1] \tau^l \le 2c_2 [\alpha_{j_0}^{n_{j_0} - l} \vee 1] \tau^l,$$

where $[\alpha_{j_0}^{n_{j_0} - l} \vee 1] := \max_j [\alpha_j^{n_j - l} \vee 1]$. Now, note that $N_{n-m} \le c_3 \tau^{m/2}$ for $e_{xy} \in B_n$ (see the paragraph just after Corollary 4.11 in [**HK2**]). From this, we see $\alpha_j^2 \le \tau$ for $j = 1, 2$, especially $\tau \alpha_j^{-1} > 1$. Take $L \in \mathbb{N}$ as the minimum number that satisfies $(\tau \alpha_j^{-1})^L c_1 > 2c_2 c$. If $n_{j_0} - (l + L) > 0$, then

$$ck \le \frac{(\tau \alpha_j^{-1})^L c_1}{2c_2} k \le c_1 \alpha_{j_0}^{n_{j_0} - (l+L)} \tau^{l+L} \le N_{l+L} \tau^{l+L},$$

so that $l_c \le l + L$ and the result holds. On the other hand, if $n_{j_0} - (l + L) \le 0$, then there exists $c_4 > 0$ such that $n \le l + c_4$. Then, $N_l \le c_5$ for some $c_5 > 0$ and the result holds. The case (2-2) can be proved similarly.

(b) can be proved in the same way as (a). □

The following comparison is important.

PROPOSITION 3.4. *There exists* $c_{3.5}, c_{3.6} > 0$ *such that the following holds.*

$$c_{3.5} \sum_{i=1}^2 \bar{N}^i_{l^i}(x, y) \le N_l(x, y) \le c_{3.6} \sum_{i=1}^2 \bar{N}^i_{l^i}(x, y).$$

PROOF. We first consider the case $l < n$ and $l_i < n_i$ for $i = 1, 2$. We note that $\bar{N}^i_{m_1} \le \bar{N}^i_{m_2}$ $(i = 1, 2)$ and $N_{m_1} \le N_{m_2}$ for $m_2 \le m_1$. By Lemma 3.1, Lemma 3.2 (a) and Lemma 3.3 (b), we have

$$c_{3.1} \sum_{i=1}^{2} \bar{N}^i_{l^i} \le c_{3.1} \sum_{i=1}^{2} \bar{N}^i_{l_{c_{3.1}}} \le N_{l_{c_{3.1}}} \le (c_{3.4}(c_{3.1}))^{-1} N_l.$$

On the other hand, by Lemma 3.1, Lemma 3.2 (b) and Lemma 3.3 (b), we have

$$N_l \le 2c_{3.2} \bar{N}^{j_0}_l \le 2c_{3.2} \bar{N}^{j_0}_{l^{j_0}_{(2c_{3.2})^{-1}}} \le 2c_{3.2} \sum_{i=1}^{2} \bar{N}^i_{l^i_{(2c_{3.2})^{-1}}} \le c_1 \sum_{i=1}^{2} \bar{N}^i_{l^i}.$$

Combining them, we obtain the desired result.

We next consider the case $l = n$ where $N_l = N_n = 1$. Using Lemma 3.1, we have

$$c_{3.1} \bar{N}^i_j \tau^j \le N_j \tau^j < k, \qquad \forall j \le n, \ i = 1, 2.$$

Thus, $l^i_{1/c_{3.1}} = n$ for $i = 1, 2$. Using Lemma 3.3 (a), we obtain the desired result.

We now consider the case $l^1 = n_1$, where $\bar{N}^1_{l^1} = \bar{N}^1_{n_1} = 1$ and

$$\bar{N}^1_j \tau^j < k, \qquad \forall j \le n \tag{3.5}$$

holds. Assume first that $l^2 = n_2$. Then, using Lemma 3.1, we have $N_j \tau^j / c_{3.2} < 2k$ for all $j \le n$ so that $l_{2c_{3.2}} = n$. Thus, with the help of Lemma 3.3, the result can be proved in the same way as in the case $l = n$. Now assume that $l^2 < n_2$. By the definition of l^2, there exists $c_2 \ge 1$ such that

$$c_2 k > \bar{N}^2_{l^2} \tau^{l^2} \ge k.$$

Combining this with (3.5) and Lemma 3.1, we obtain

$$c_{3.2}(c_2 + 1)k \ge c_{3.2}(\bar{N}^1_{l^2} + \bar{N}^2_{l^2})\tau^{l^2} \ge N_{l^2} \tau^{l^2} \ge c_{3.1}(\bar{N}^1_{l^2} + \bar{N}^2_{l^2})\tau^{l^2} \ge c_{3.1} k.$$

Thus we have $l_{c_{3.1}} \le l^2 \le l_{c_{3.2}(c_2+1)}$. Using Lemma 3.1 and Lemma 3.3, we obtain the result.

The case where $l^2 = n_2$ can be treated similarly and the lemma is proved. $\square$

By (2.8) and Proposition 3.4, we have the following.

$$p_k(x, y) \asymp c_1 k^{-\frac{S}{S+1}} \exp(-c_2 \sum_i \bar{N}^i_{l^i(k,n)}(x, y)). \tag{3.6}$$

For $i = 1, 2$, let $d^i_c := \log_\rho \alpha_i$ and define $d^i_w = (S + 1)/d^i_c$. For the case (0), (1), (2-2), (2-3) and (3-2), define $d_c := \log_\rho \lambda_{min}, d_w = (S + 1)/d_c$ where λ_{min} is the maximum eigenvalue of A_{min}. By Theorem 2.4 and Proposition 3.4, we have the following expression for the heat kernel estimates.

THEOREM 3.5. *There exists constants $c_{3.7}, c_{3.8} > 0$ such that for $x, y \in V$ and $k \ge d(x, y)$, the following holds.*
Case A: In the cases (2-1), (3-1) *and* (3-3):

$$p_k(x, y) \asymp c_{3.7} k^{-S/(S+1)} \exp\left\{ -c_{3.8} \left(\left(\frac{d_1(x, y)^{d^1_w}}{k} \right)^{\frac{1}{d^1_w - 1}} + \left(\frac{d_2(x, y)^{d^2_w}}{k} \right)^{\frac{1}{d^2_w - 1}} \right) \right\}.$$

Case B: In the case (2-2):

$$p_k(x,y) \asymp c_{3.7} k^{-S/(S+1)}$$
$$\times \quad \exp\left\{-c_{3.8}\left(\left(\frac{(d_1(x,y)\log(\frac{d_1(x,y)}{k}))^{d_w}}{k}\right)^{\frac{1}{d_w-1}} + \left(\frac{d_2(x,y)^{d_w}}{k}\right)^{\frac{1}{d_w-1}}\right)\right\}.$$

Case C: In the remaining cases (i.e., (0), (1), (2-3) *and* (3-2)):

$$p_k(x,y) \asymp c_{3.7} k^{-S/(S+1)} \exp\left(-c_{3.8}\left(\frac{d(x,y)^{d_w}}{k}\right)^{\frac{1}{d_w-1}}\right).$$

Proof. In this proof, we write $a_n \asymp b_n$ if there exists $c_1, c_2 > 0$ such that $c_1 a_n \le b_n \le c_2 a_n$. We first prove Case A. Using Proposition 3.4, it is enough to prove the following,

$$N^i_{l^i} \asymp \left(\frac{d_i(x,y)^{d^i_w}}{k}\right)^{\frac{1}{d^i_w-1}} \qquad \forall i = 1, 2. \tag{3.7}$$

Assume $l^i < n_i$. Then, by the definition of l_i, we have $N^i_{l^i-1}\tau^{l^i-1} < k \le N^i_{l^i}\tau^{l^i}$ so that $k \le N^i_{l^i}\tau^{l^i} < c_3 k$ for some $c_3 > 1$. Combining this with the fact $N^i_m \asymp [\alpha_i^{n_i-m} \vee 1]$, we have

$$\frac{\log(k\alpha_i^{-n_i})}{\log(\tau/\alpha_i)} + c_4 \le l^i \le \frac{\log(k\alpha_i^{-n_i})}{\log(\tau/\alpha_i)} + c_5.$$

Thus, noting that $d_i(x,y) \asymp \alpha_i^{n_i}$, we have

$$\begin{aligned} N^i_{l^i} &\asymp \alpha_i^{n_i - l^i} \asymp \alpha_i^{n_i}\alpha_i^{-\frac{\log(k\alpha_i^{-n_i})}{\log(\tau/\alpha_i)}} = \alpha_i^{n_i}\left(\frac{\alpha_i^{n_i}}{k}\right)^{\frac{\log\alpha_i}{\log(\tau/\alpha_i)}} \\ &\asymp \alpha_i^{n_i\frac{\log\tau}{\log(\tau/\alpha_i)}} k^{-\frac{\log\alpha_i}{\log(\tau/\alpha_i)}} \asymp \left(\frac{d_i(x,y)^{d^i_w}}{k}\right)^{\frac{1}{d^i_w-1}}. \end{aligned} \tag{3.8}$$

Here we have used the fact that $d^i_w = (S+1)/\log_\rho \alpha_i$ and $\rho^{S+1} = \alpha_i^{d^i_w} = \tau$, so that $d^i_w = \log\tau/\log\alpha_i$, $\log\alpha_i/\log(\tau/\alpha_i) = 1/(d^i_w - 1)$ and $\log\tau/\log(\tau/\alpha_i) = d^i_w/(d^i_w-1)$. Thus (3.7) is proved when $l^i < n_i$.

Next, assume $l^i \ge n_i$ (thus $l^i = n$). Then $N^i_{l^i} \le c_6$ for some $c_6 > 0$. On the other hand, using the last asymptotic equivalence in (3.8), we see that $(d_i(x,y)^{d^i_w}/k)^{\frac{1}{d^i_w-1}}$ is bounded from above by some constant. Thus (3.7) is proved in this case.

We next prove Case B. Using Proposition 3.4, it is enough to prove the following,

$$\bar{N}^1_{l^1} \asymp \left(\frac{(d_1(x,y)\log(\frac{d_1(x,y)}{k}))^{d^1_w}}{k}\right)^{\frac{1}{d^1_w-1}}, \qquad N^2_{l^2} \asymp \left(\frac{d_2(x,y)^{d^2_w}}{k}\right)^{\frac{1}{d^2_w-1}}. \tag{3.9}$$

The second asymptotic equivalence is proved in the same way as Case A, so we consider the first asymptotic equivalence. Assume $l^1 < n_1$. Then, by the definition of l_1 and the fact $\bar{N}^1_m \asymp (n_1 - m)[\alpha_1^{n_1-m}]$, we have $\bar{N}^1_{l^1} \asymp (n_1 - l^1)\alpha_1^{n_1-l^1}$ and $(n_1 - l^1)\alpha_1^{n_1-l^1}\tau^{l^1} \asymp k$. Thus, $(n_1 - l^1)(\alpha_1/\tau)^{n_1-l^1} \asymp k/\tau^{n_1}$. Note that by simple calculations, we see that if for fixed $0 < \lambda < 1$, $x\lambda^x \asymp y$ for all $1 \le x$ and $0 < y < 1$, then there exists $c_7, c_8 > 0$ such that

$$\frac{\log(1/y)}{\log(1/\lambda)} + \frac{\log\log(1/y)}{\log(1/\lambda)} - c_7 \le x \le \frac{\log(1/y)}{\log(1/\lambda)} + \frac{\log\log(1/y)}{\log(1/\lambda)} + c_8.$$

Thus, $x(\lambda\tau)^x \asymp y^{-\log(\lambda\tau)/\log(1/\lambda)}(\log(1/y))^{\log\tau/\log(1/\lambda)}$. Applying this with $x = n_1 - l^1, y = k/\tau^{n_1}$ and $\lambda = \alpha_1/\tau$, we have

$$\bar{N}^1_{l^1} \asymp (n_1 - l^1)\alpha_1^{n_1 - l^1} \asymp (\alpha_1^{n_1 d^1_w}/k)^{\frac{1}{d^1_w - 1}}(\log(\alpha_1^{n_1 d^1_w}/k))^{\frac{d^1_w}{d^1_w - 1}}.$$

Noting that $d_1(x,y) \asymp \alpha_1^{n_1}$, we obtain (3.9). The case $l^1 \geq n_1$ can be shown easily.

We omit the proof of Case C as it is similar and simpler (see [**HK3**] etc.). $\square$

REMARK 3.6. (a) The choice of A_{min} given in this section is not necessarily unique. Even so, if there is more than one choice of A_{min}, then the matrices chosen have the same eigenspace for the maximum eigenvalue. This can be proved by using the properties of non-negative matrices.
(b) In this section we discussed the case where there are two types of steps for simplicity. We note that even if there are more than two types, if all the corresponding minimal step matrices can be decomposed into one or two irreducible components, and if (3.1) holds, then the arguments given in this section can be applied and we have the same type of heat kernel estimates.
(c) We believe that the general case of n edge types could be treated similarly, but there are more possibilities for the form of the heat kernel estimates. For instance, when the Jordan cell for the maximum eigenvalue of some irreducible component of the minimal step matrix is p-dimensional, then the $(1,p)$-th element of the n-th power of the matrix is of order $n^p a^n$. Thus we should define $\bar{N}^1_m := N^1_m(\log N^1_m)^p$ and the corresponding part of the heat kernel estimate (in the exponential) should be as follows

$$-c_{3.9}\left(\frac{\{d_1(x,y)(\log(\frac{d_1(x,y)}{k}))^p\}^{d^1_w}}{k}\right)^{\frac{1}{d^1_w - 1}}.$$

4. Law of the iterated logarithm

In this section, as an application of the heat kernel estimates, we discuss the law of the iterated logarithm for the sample paths of the random walk. We will only treat the case (3) of the last section, as this case already includes interesting aspects of the process we treat.

THEOREM 4.1. *For case* (3), *there exists* $c_{4.1}(i) > 0$ *such that*

$$\limsup_{k\to\infty} \frac{d_i(X_0, X_k)}{k^{1/d^i_w}(\log\log k)^{1-1/d^i_w}} = c_{4.1}(i), \qquad \mathbb{P}^0 - a.s.\ , \ i = 1, 2.$$

The proof will follow from some preliminary results. We begin with a little notation. An n-graph complex is the subgraph of V contained in an n-cell. We write $D^1_n(y)$ for the n-graph complex containing y and the neighbouring n-graph complexes. We will write $\partial_i D^1_n(y)$ for the boundary points $z \in D^1_n(y)$ which have $d_i(y,z) \geq \alpha_i^n$. Let $T_A = \inf\{k : X_k \notin A\}$ be the exit time from a set A and define a sequence of crossing times by $T^m_0 = \inf\{k : X_k \in V_m\}$ and for $i \geq 1$, let $T^m_i = \inf\{k > T^m_{i-1} : X_k \in V_m \backslash \{X_{T^m_{i-1}}\}\}$ and $W^m_i = T^m_i - T^m_{i-1}$.

LEMMA 4.2. *For all* $x \in V$ *there exist constants* $c_{4.2}, c_{4.3}$ *such that for* $i = 1, 2$,

$$P^x(d_i(x, X_k) > \lambda) \leq c_{4.2}\exp\left(-c_{4.3}\left(\frac{\lambda^{d^i_w}}{k}\right)^{1/(d^i_w - 1)}\right), \ \ k > \lambda.$$

PROOF. The proof of this result uses hitting time estimates. Let n be such that $\alpha_i^{n-1} \le \lambda \le \alpha_i^n$. Observe that, by the roughly uniform exit law of $D_n(y)$, there exists a c_1 such that

$$\begin{aligned} P^x(d_i(x,X_k) > \lambda) &\le P^x(T_{D_n(x)} < k, X_{T_{D_n(x)}} \in \partial_i D_n^1(y)) \\ &\le c_1 \sup_{y\in V_{n-1}} P^y(T_{D_n(y)} < k | X_{T_{D_n(y)}} \in \partial_i D_n^1(y)). \end{aligned}$$

By the crossing time estimates of [**HK2**] Lemma 4.4 and Proposition 4.5, as used in the proof of [**HK2**] Theorem 4.6, we have the following for all $0 \le m \le n$,

$$\begin{aligned} P^y(T_{D_n(y)} < k | X_{T_{D_n(y)}} \in \partial_i D_n^1(y)) &\le P^y(\sum_{i=1}^{\alpha_i^{n-m}} W_i^m < k) \\ &\le \exp(-c_2\alpha_i^{n-m} + c_3(\alpha_i^{n-m}\tau^{-m}k)^{1/2}) \end{aligned}$$

If we choose $l := \inf\{j : \alpha_i^{n-j}\tau^j \ge k\} \wedge n$, then

$$P^y(T_{D_n(y)} < k | X_{T_{D_n(y)}} \in \partial_i D_n^1(y)) \le \exp(-c_4\alpha_i^{n-l}).$$

Rewriting this, using the definitions of l and n, we have the estimate. □

As we are in the case (3) we can find vertices $x_i^\lambda \in V$ such that $d_i(0,x_i^\lambda) = [\lambda]$ and $d_j(0,x_i^\lambda) = 0$ for $j \ne i$. We will write

$$B_i(x,r) = \{y : d_i(0,y) \ge d_i(0,x), R(x,y) \le r^{1/(S+1)}\},$$

and

$$B_i'(x,r) = \{y : d_i(0,y) < d_i(0,x), R(x,y) \le r^{1/(S+1)}\}.$$

LEMMA 4.3. *There exist constants $c_{4.4}, c_{4.5}$ such that for $i = 1,2$,*

$$P^0(X_k \in B_i(x_i^\lambda, k^{1/(S+1)})) \ge c_{4.4} \exp\left(-c_{4.5}\left(\frac{\lambda^{d_w^i}}{k}\right)^{1/(d_w^i-1)}\right), \quad k > \lambda.$$

PROOF. For this lower estimate we will use our heat kernel lower bound. Writing $B_i = B_i(x_i^\lambda, k^{1/(S+1)})$, we have

$$\begin{aligned} P^0(X_k \in B_i) &= \sum_{y\in B_i} p_k(0,y)\mu(y) \\ &\ge \mu(B_i) \inf_{y\in B_i} k^{-S/(S+1)} \exp\left(-c\sum_j \left(\frac{d_j(0,y)^{d_w^j}}{k}\right)^{1/(d_w^j-1)}\right) \\ &\ge c_1 \exp\left(-\sup_{y\in B_i}\sum_j \left(\frac{d_j(0,y)^{d_w^j}}{k}\right)^{1/(d_w^j-1)}\right). \qquad (4.1) \end{aligned}$$

If we now observe that within the set B_i we have constants such that

$$\sup_{y\in B_i} d_j(0,y) \le \begin{cases} c_1\lambda + c_2 k^{1/d_w^i} & j = i \\ c_3 k^{1/d_w^j} & j \ne i. \end{cases}$$

Thus, placing this in (4.1), we have

$$\sup_{y\in B_i}\left(\frac{d_j(0,y)^{d_w^j}}{k}\right)^{1/(d_w^j-1)} \leq \begin{cases} c_5\left(\frac{c_1\lambda^{d_w^i}}{k}+c_4\right)^{1/(d_w^i-1)} & j=i \\ c_3^{1/(d_w^j-1)} & j\neq i \end{cases}$$

and hence the result. □

Let $h_i^c(k) = ck^{1/d_w^i}(\log\log k)^{1-1/d_w^i}$ and write $h_i(k)$ for $h_i^1(k)$.

COROLLARY 4.4. *For each $i=1,2$, there exists $N>1$ such that, if $\tilde{x}_i^n$ satisfies $d_i(0,\tilde{x}_i^n) = [h_i^c(N^n)]$, then there are constants $c_{4.6}, c_{4.7}, c_{4.8}$ such that*

$$P^0(X_{N^n}\in B_i(\tilde{x}_i^n, N^{n/(S+1)})) \geq c_{4.6}n^{-c_{4.5}c^{1/(d_w^i-1)}}, \tag{4.2}$$

and for all $z\in B_i'(\tilde{x}_i^{n-1}, N^{(n-1)/(S+1)})$,

$$P^z(X_{N^n}\in B_i(\tilde{x}_i^n, N^{n/(S+1)})) \geq c_{4.8}n^{-c_{4.7}c^{1/(d_w^i-1)}}. \tag{4.3}$$

PROOF. For (4.2) we substitute $\lambda = h_i^c(N^n)$ in Lemma 4.3.

For (4.3), for any vertex pair $z\in B_i'(\tilde{x}_i^{n-1}, N^{(n-1)/(S+1)}), y\in B_i(\tilde{x}_i^n, N^{n/(S+1)})$, there is a constant c_1 such that $d_i(z,y)\leq h_i^c(N^{n-1}) + c_1N^{(n-1)/(S+1)}$. Hence applying the same approach as the proof of Lemma 4.3 will give the result. □

Finally we state a triviality result for tail events associated with our random walk. This is a simple extension to the random walk case of Theorem 8.4 of [**BB3**].

LEMMA 4.5. *Let $A\in\cap_k\sigma(X_n : n\geq k)$ be a tail event, then $P^x(A)=0$ for all x or $P^x(A)=1$ for all x.*

PROOF OF THEOREM 4.1. We need to modify the usual proof of the law of the iterated logarithm to allow for the existence of different types. Firstly, by Lemma 4.5, it is enough to show that there are constants c_1, c_2 such that

$$c_1 \leq \limsup_{k\to\infty}\frac{d_i(0,X_k)}{h_i(k)} \leq c_2, \quad P^0-a.s.$$

The upper bound is standard as we apply the upper estimate of Lemma 4.2 and the first Borel-Cantelli lemma.

For the lower bound we use a slight variant of the second Borel-Cantelli Lemma. In the case of a Markov process we can use a simple modification of the proof to show that if A_n is a sequence of events, each in $\sigma(X_k : k\leq N^n)$, then, if $\sum_n P(A_n\cap A_{n-1}^c)=\infty$, then $P(A_n \;\; i.o.)=1$, for a discussion see [**HHH**] Lemma 4.5. Let $A_n = \{X_{N^n}\in B(\tilde{x}_i^n, N^{n/(S+1)})\}$, so that $A_{n-1}^c = \{X_{N^{n-1}}\notin B(\tilde{x}_i^{n-1}, N^{(n-1)/(S+1)})\}$. Now $B_{n-1} := \{X_{N^{n-1}}\in B_i'(\tilde{x}_i^{n-1}, N^{(n-1)/(S+1)})\}\subset A_{n-1}^c$ and hence

$$\begin{aligned} P^0(A_n\cap A_{n-1}^c) &\geq P^0(A_n\cap B_{n-1}) \\ &= P^0(A_n|B_{n-1})P^0(B_{n-1}) \end{aligned}$$

We can now apply both estimates of Corollary 4.4, to get

$$P^0(A_n\cap A_{n-1}^c) \geq c_{4.6}c_{4.8}n^{-(c_{4.5}+c_{4.7})c^{1/(d_w^i-1)}}.$$

We can now choose $c=(c_{4.5}+c_{4.7})^{-(d_w^i-1)}$ to obtain

$$P^0(A_n\cap A_{n-1}^c) \geq c_{4.6}c_{4.8}n^{-1}.$$

Thus the sum diverges and we have that the event A_n occurs infinitely often. Thus

$$X_{N^n} \in B(\tilde{x}_i^n, N^{n/(S+1)}) \ \ i.o. \ \ P^0 - a.s.$$

By the construction of the set $B(\tilde{x}_i^n, N^{n/(S+1)})$ we have that there is a constant c' such that $d_i(0, X_{N^n}) \geq c' h_i^c(N^n)$ infinitely often. Hence

$$\limsup_{k\to\infty} \frac{d_i(0,X_k)}{h_i(k)} \geq \limsup_{n\to\infty} \frac{d_i(0,X_{N^n})}{h_i^1(N^n)} \geq c'c, \ \ P^0 - a.s.$$

This completes the proof of our LIL. □

We note that this shows that the size of fluctuations in the different directions is different. The construction of $d_i(.,.)$ ensures that if X_k has moved in the direction of type i, then $d_i(0,X_k) \asymp R(0,X_k)^{d_c^i}$ and hence

$$\frac{d_i(0,X_k)^{d_w^i}}{k} \asymp \frac{R(0,X_k)^{S+1}}{k}.$$

Thus the bulk spread of the random walk is determined by the resistance. However if we look at the large fluctuations we see that the direction has an effect as

$$\frac{R(0,X_k)^{S+1}}{k} \asymp (\log\log k)^{d_w^i - 1} \ \ \text{infinitely often.}$$

Thus the size of fluctuations in the different directions is different and we would see that the process makes larger extreme movements in the direction with the largest d_w^i (that is, the smallest α_i).

5. Diffusion case

In this section, we will mention diffusion processes on u.f.r. fractals and their heat kernel estimates. For $u,v \in l(\hat{V}_n)$, define

$$(5.1) \qquad \hat{\mathcal{E}}^n_{Q_0}(u,v) = \rho^n \sum_{i_1,\cdots,i_n \in I} S_{Q_0}(u \circ \Psi_{i_1,\dots,i_n}, v \circ \Psi_{i_1,\dots,i_n}),$$

where $S_{Q_0}(\xi,\xi) = \sum_{i,j\in I_F} (Q_0)_{ij}(\xi_i - \xi_j)^2/2$ for each $\xi \in l(I_F)$. Let $\hat{\nu}$ be the normalized Hausdorff measure on $\hat{K}$. Then, the following is known (see for example, [**HK3, Kig**]).

THEOREM 5.1. *Let $Q_0 \in Int(\mathcal{Q}_M)$ be as in Assumption 2.2 2), i.e. $F(Q_0) = \rho_{Q_0}^{-1} Q_0$. Then, there is a local regular Dirichlet form $(\hat{\mathcal{E}}, \hat{\mathcal{F}})$ in $\mathbb{L}^2(\hat{K}, d\hat{\nu})$ satisfying the following,*

$$\begin{aligned} \hat{\mathcal{F}} &= \{u \in C(\hat{K},\mathbb{R}) : \sup_n \hat{\mathcal{E}}^n_{Q_0}(u,u) < \infty\}, \\ \hat{\mathcal{E}}(u,v) &= \lim_{n\to\infty} \hat{\mathcal{E}}^n_{Q_0}(u,v) \qquad \text{for } \ u,v \in \hat{\mathcal{F}}. \end{aligned}$$

For each $m \in \mathbb{N}$, let $K_m = \alpha^m \hat{K}$ and define $\sigma_m : C(K_m,\mathbb{R}) \to C(\hat{K},\mathbb{R})$ by $\sigma_m u(x) = u(\alpha^m x)$ for $x \in \hat{K}$. Set $\mathcal{F}_{<m>} = \sigma_{-m}\hat{\mathcal{F}}$, $\mathcal{E}_{<m>}(u,v) = \rho^{-m}\hat{\mathcal{E}}(\sigma_m u, \sigma_m v)$ for $u,v \in \mathcal{F}_{<m>}$. Let ν be a Hausdorff measure on K such that $\nu|_{\hat{K}} = \hat{\nu}$ and $N\nu = \nu \circ \Psi_1^{-1}$. Now let

$$\begin{aligned} \mathcal{F}_K &= \{u \in l(K) : u|_{K_m} \in \mathcal{F}_{<m>} \text{ for all } \ m \in \mathbb{N}, \\ &\qquad \lim_{m\to\infty} \mathcal{E}_{<m>}(u|_{K_m}, u|_{K_m}) < \infty\} \cap \mathbb{L}^2(K, d\nu), \\ \mathcal{E}_K(u,v) &= \lim_{m\to\infty} \mathcal{E}_{<m>}(u|_{K_m}, v|_{K_m}) \qquad \text{for } \ u,v \in \mathcal{F}. \end{aligned}$$

Then, $(\mathcal{E}_K, \mathcal{F}_K)$ is a local regular Dirichlet form on $\mathbb{L}^2(K, d\nu)$, $\mathcal{F}_K \subset C(K, \mathbb{R})$ and the following scaling property holds,

$$\mathcal{E}_K(u,v) = \rho\mathcal{E}_K(u \circ \Psi_1, v \circ \Psi_1) \qquad \text{for } \ u, v \in \mathcal{F}_K.$$

Let $\{X_t\}_{t\geq 0}$ (resp. $\{\hat{X}_t\}_{t\geq 0}$) be the diffusion process corresponding to $(\mathcal{E}_K, \mathcal{F}_K)$ (resp. $(\hat{\mathcal{E}}, \hat{\mathcal{F}})$). It can be shown that there is a jointly continuous version of the heat kernel $p_t^K(x,y)$ (resp. $p_t^{\hat{K}}(x,y)$) such that $E^x[f(X_t)] = \int_K p_t^K(x,y)f(y)\nu(dy)$ for all $f : K \to \mathbb{R}$ (resp. $E^x[f(\hat{X}_t)] = \int_{\hat{K}} p_t^{\hat{K}}(x,y)f(y)\hat{\nu}(dy)$ for all $f : \hat{K} \to \mathbb{R}$). In the following, we write $p_t(x,y)$ for both $p_t^K(x,y)$ and $p_t^{\hat{K}}(x,y)$ and state the results for K and $\hat{K}$ simultaneously. Let

$$k(n,t) := \inf\{j : N_{n+j}(x,y)\tau^{n+j} \geq t\}.$$

Then the following holds (see [**HK3**] for the case of K; recall that our labelling of V_n is the opposite to the one given in [**HK3**]).

THEOREM 5.2. *There exist constants $c_{5.1}, c_{5.2} > 0$ such that for all $x, y \in K$ (resp. $\hat{K}$) and $t > 0$ (resp. $0 < t < 1$), if $\rho^{n-1} \leq R(x,y) \leq \rho^n$, then*

$$p_t(x,y) \asymp c_{5.1} t^{-\frac{S}{S+1}} \exp(-c_{5.2} N_{k(n,t)}(x,y)).$$

Now, we consider the case of two step types. For each $x, y \in K$ (or $\hat{K}$), define

$$D_i(x,y) := \limsup_{m\to\infty} \frac{N_{-m}^i(x,y)}{\alpha_i^m}, \ i = 1,2, \qquad D(x,y) := \limsup_{m\to\infty} \frac{N_{-m}(x,y)}{\lambda_{min}^m}.$$

Note that if we take liminf instead of limsup, there is only a (uniform) constant time difference (due to the results in Section 3). By similar arguments to those in Section 3, we can translate Theorem 3.5 to this setting as follows.

THEOREM 5.3. *There exists constants $c_{5.3}, c_{5.4} > 0$ such that for all $x, y \in K$ (resp. $\hat{K}$) and $t > 0$ (resp. $0 < t < 1$), the following holds.*
Case A: In the cases (2-1), (3-1) *and* (3-3):

$$p_t(x,y) \asymp c_{5.3} t^{-S/(S+1)} \exp\left\{ -c_{5.4} \left(\left(\frac{D_1(x,y)^{d_w^1}}{t} \right)^{\frac{1}{d_w^1 - 1}} + \left(\frac{D_2(x,y)^{d_w^2}}{t} \right)^{\frac{1}{d_w^2-1}} \right) \right\}.$$

Case B: In the case (2-2):

$$p_t(x,y) \asymp c_{5.3} t^{-S/(S+1)}$$
$$\times \quad \exp\left\{ -c_{5.4} \left(\left(\frac{(D_1(x,y)\log(\frac{D_1(x,y)}{t}))^{d_w}}{t} \right)^{\frac{1}{d_w-1}} + \left(\frac{D_2(x,y)^{d_w}}{t} \right)^{\frac{1}{d_w-1}} \right) \right\}.$$

Case C: In the remaining cases (i.e., (0), (1), (2-3) *and* (3-2)):

$$p_t(x,y) \asymp c_{5.3} t^{-S/(S+1)} \exp\left(-c_{5.4} \left(\frac{D(x,y)^{d_w}}{t} \right)^{\frac{1}{d_w-1}} \right).$$

By definition, $d_w^i = (S+1)/d_c^i$ is the box dimension of the type i path with respect to the resistance metric, where $d_c^i = \log_\rho \alpha_i$ is the chemical exponent of the type i path (with respect to the resistance metric). Thus Theorem 5.3 gives an affirmative answer to Conjecture 6.9 and 6.10 in [**HK3**] for the case of two step types.

We can also obtain a law of the iterated logarithm for the long and short time behaviour of the diffusion process on $\hat{K}$ which corresponds to that of Theorem 4.1.

References

[Bar] M.T. Barlow, *Diffusions on fractals*, Lectures in Probability Theory and Statistics: Ecole d'été de probabilités de Saint-Flour XXV, Springer, New York, 1998.

[BB1] M.T. Barlow and R.F. Bass, *Stability of parabolic Harnack inequalities,* Preprint 2002.

[BB2] M.T. Barlow and R.F. Bass, *Random walks on graphical Sierpinski carpets,* Random walks and discrete potential theory (Cortona, 1997), 26–55, Sympos. Math., XXXIX, Cambridge Univ. Press, Cambridge, 1999.

[BB3] M.T. Barlow and R.F. Bass, *Brownian motion and harmonic analysis on Sierpinski carpets,* Canad. J. Math., **51** (1999), 673–744.

[BCG] M.T. Barlow, T. Coulhon and A. Grigor'yan, *Manifolds and graphs with slow heat kernel decay,* Invent. Math., **144** (2001), 609–649.

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

[GT1] A. Grigor'yan and A. Telcs, *Harnack inequalities and sub-Gaussian estimates for random walks,* Math. Annalen, **324** (2002), 521–556.

[GT2] A. Grigor'yan and A. Telcs, *Sub-Gaussian estimates of heat kernels on infinite graphs,* Duke Math. J., **109** (2001), 451–510.

[HHH] B.M. Hambly, K. Hattori and T. Hattori, *Self-repelling walk on the Sierpinski gasket,* Probab. Theory Related Fields, **124** (2002), 1–25.

[HK1] B.M. Hambly and T. Kumagai, *Diffusion processes on fractal fields: heat kernel estimates and large deviations,* Probab. Theory Relat. Fields, to appear.

[HK2] B.M. Hambly and T. Kumagai, *Heat kernel estimates for symmetric random walks on a class of fractal graphs and stability under rough isometries,* Proc. of Symposia in Pure Math. (AMS), to appear.

[HK3] B.M. Hambly and T. Kumagai, *Transition density estimates for diffusions on p.c.f. fractals,* Proc. London Math. Soc., **78** (1999), 431–458.

[Jo] O.D. Jones, *Transition probabilities for the simple random walk on the Sierpiński graph,* Stochastic Process. Appl., **61** (1996), 45–69.

[Kig] J. Kigami, *Analysis on Fractals,* Cambridge Univ. Press, Cambridge, 2001.

[Kum1] T. Kumagai, *Homogenization on finitely ramified fractals,* Proc. of the RIMS Project '02 "Stochastic Analysis and Related Topics", Advanced Studies in Pure Math. (MSJ), to appear.

[Kum2] T. Kumagai, *Estimates of transition densities for Brownian motion on nested fractals,* Probab. Theory Relat. Fields, **96** (1993), 205–224.

[Kus] S. Kusuoka, *Diffusion processes on nested fractals,* Statistical Mechanics and Fractals (Lect. Notes Math. **1567**), Springer, 1993.

[Lind] T. Lindstrøm, *Brownian motion on nested fractals,* Mem. Amer. Math. Soc. 420 **83**, 1990.

Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK

E-mail address: hambly@maths.ox.ac.uk

Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan

E-mail address: kumagai@kurims.kyoto-u.ac.jp