Heat Kernel Estimates for Symmetric Random Walks on a Class of Fractal Graphs and Stability under Rough Isometries

Ben M. Hambly and Takashi Kumagai

ABSTRACT. We examine a class of fractal graphs which arise from a subclass of finitely ramified fractals. The two-sided heat kernel estimates for these graphs are obtained in terms of an effective resistance metric and they are best possible up to constants. If the graph has symmetry, these estimates can be expressed as the usual Gaussian or sub-Gaussian estimates. However, without symmetry, the off-diagonal terms show different decay in different directions. We also discuss the stability of the sub-Gaussian heat kernel estimates under rough isometries.

1. Introduction

In the past two decades, there has been a lot of work on the properties of heat kernels on graphs (see [**32**] etc.). One avenue of research has been to explore graphs whose heat kernels admit so-called Gaussian estimates and, for instance, to determine analytic inequalities which are equivalent to the estimates (see [**12, 31, 9**] etc.). On the other hand, the development of diffusion processes on fractals has given many examples where the corresponding heat kernels admit sub-Gaussian estimates (see [**2, 23**] etc.). Similar results have been obtained for some fractal graphs such as the Sierpinski gasket graph [**20**], the Sierpinski carpet graph [**4**] and tree-like graphs [**1**]. For these graphs, the heat kernel $p_n(x, y)$ of the simple random walk satisfies the following sub-Gaussian estimates:

$$\begin{aligned} p_n(x,y) &\leq c_{1.1} n^{-\alpha/\beta} \exp[-(\frac{d(x,y)^\beta}{c_{1.2} n})^{1/(\beta-1)}], \\ p_n(x,y) + p_{n+1}(x,y) &\geq c_{1.3} n^{-\alpha/\beta} \exp[-(\frac{d(x,y)^\beta}{c_{1.4} n})^{1/(\beta-1)}], \end{aligned}$$

where $c_{1.1}, \ldots, c_{1.4} > 0$ and $\alpha \geq 1, 2 \leq \beta \leq \alpha + 1$ are constants. (Note that the Gaussian estimate is the particular case when $\beta = 2$.) These results have

1991 *Mathematics Subject Classification.* Primary 60J10, 60J35; Secondary 31B05, 31C25.

Key words and phrases. Random walks on graphs, sub-Gaussian heat kernel bounds, parabolic Harnack inequality, Poincaré inequality, rough isometry, anomalous diffusion.

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

motivated further work on graphs (and manifolds) whose heat kernels admit sub-Gaussian estimates. In [**13, 14**], it has been shown that the sub-Gaussian heat kernel estimates, with respect to the usual graph distance, are equivalent to various geometric and analytic properties such as a parabolic Harnack inequality of order β (given as (5.1)) for the graph (see Theorem 5.3, and see [**18**] for the manifold setting). In [**3**], for graphs the parabolic Harnack inequality of order β is proved to be equivalent to a volume doubling condition (5.2), a Poincaré inequality (5.3) and a Sobolev type inequality (5.4) (see Theorem 5.7). This gives the stability of the parabolic Harnack inequality of order β (thus stability of the sub-Gaussian estimates) on graphs under changes in the edge weights that are comparable to the original weights. The techniques developed for heat kernel estimates on fractals are also used in [**6**] to answer the question of how slow the decay of the heat kernel can be, given the volume growth.

In this paper, we study further the sub-Gaussian heat kernel estimates on graphs. The aims of this paper are threefold:
(I) To give a broad class of examples which satisfy sub-Gaussian heat kernel estimates.
(II) To discuss what metric is the "intrinsic" metric for the heat kernel.
(III) To show that such heat kernel estimates are stable under "local perturbations" of the graphs.

In order to see the full range of behaviour of graphs with sub-Gaussian heat kernel estimates, the current concrete examples in [**20, 4, 1**] are not enough. We thus translate the results on heat kernels for post-critically finite (p.c.f.) fractals in [**17**] into the graph setting. The framework of p.c.f. fractals allows us to construct a large class of sets which are self-similar but not necessarily spatially symmetric. It is possible to compute the Hausdorff dimension of such fractals and to establish the existence of a spectral dimension under a fixed point assumption [**23**]. This spectral dimension governs the decay of the on-diagonal heat kernel. Due to the lack of spatial symmetry it is possible for the off-diagonal term to have different decay depending on the direction.

The main result of our first aim (I) is to provide a class of fractal graphs (based on p.c.f. fractals) for which there is a subclass with uniform sub-Gaussian heat kernel estimates and there is also a subclass where these uniform sub-Gaussian estimates breakdown (see Theorem 4.10 and Corollary 4.11– 4.16) in that the off-diagonal term depends on direction. In order to express these estimates we require a choice of metric and we adopt the resistance metric as defined in (2.4).

The directional dependence of the off-diagonal bounds follows from the fact that the resistance metric and the graph distance are not necessarily comparable and indeed, it appears that, in general for these graphs, the parabolic Harnack inequality of order β (for some β) holds with respect to the resistance metric but not with respect to the graph distance. Thus we believe that the resistance metric is more suitable than the graph distance for analysis. However we should note that the resistance metric is not a geodesic metric which makes for various technical problems. We also remark that, for the class of fractals we treat, the associated random walks are recurrent, which makes the resistance metric meaningful. For graphs which carry non-recurrent random walks, the resistance metric may be defined, but will no longer be relevant.

For our aim (III), we use a notion of rough isometry between two graphs (see Definition 5.9). Essentially this is a correspondence up to local modifications of the graphs and similar ideas have been developed in other contexts ([**21, 22, 8**] etc.). Another of our main results is to show that the parabolic Harnack inequality of order β (thus the sub-Gaussian estimates) is stable under rough isometry of the graphs (see Theorem 5.11). For the proof, we apply the results in [**3**] and show that the (weak) Poincaré inequality and the Sobolev type inequality (5.7) are stable under rough isometries of graphs, assuming the volume doubling condition (see Proposition 5.15). This type of stability should have an important role to play for global analysis on graphs.

The outline of the paper is as follows. We begin by defining our class of graphs which are based on a subclass of p.c.f. fractals. For a given graph we define a quadratic form and use this to define a resistance metric on the graph. The self-similarity of the graph allows us to define scale factors which determine the resistance, time and mass scalings. In Section 3 we introduce some preliminary results about the geometry of the graphs with respect to the resistance metric. In Section 4 we discuss the heat kernel estimates, giving upper and lower bounds in terms of a shortest path counting function. In the spatially symmetric case, these can be translated into two sided sub-Gaussian estimates, however in the spatially non-symmetric case we have different bounds in general. We give an example to illustrate what can occur. Finally in Section 5, we discuss the relationship between our results and those in [**13, 14**], as well as the stability of the parabolic Harnack inequality as in [**3**].

Throughout the paper, we write $f(x) \asymp g(x)$ for all x if there exist $c_1, c_2 > 0$ such that $c_1 g(x) \le f(x) \le c_2 g(x)$ for all x.

2. Uniform finitely ramified graphs and their quadratic forms

For $\alpha > 1$, let $\{\Psi_i\}_{i=1}^N$ be a collection 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 assume the *open set condition* for $\{\Psi_i\}_{i=1}^N$, that there is a non-empty, bounded open set W such that $\{\Psi_i(W)\}_{i=1}^N$ are disjoint and $\cup_{i=1}^N \Psi_i(W) \subset W$. As $\{\Psi_i\}_{i=1}^N$ is a family of contraction maps, there exists a unique non-void compact set $\hat{K}$ such that $\hat{K} = \cup_{i=1}^N \Psi_i(\hat{K})$. We will consider the case where $\hat{K}$ is connected. Denote $\Psi_{i_1,\dots,i_n} = \Psi_{i_1} \circ \cdots \circ \Psi_{i_n}$ and $S = \{1, 2, \cdots, N\}$.

Let $\hat{F}$ be the set of fixed points of the Ψ_i's, $i \in S$. A point $x \in \hat{F}$ is called an essential fixed point if there exist $i, j \in S,\ i \neq j$ and $y \in \hat{F}$ such that $\Psi_i(x) = \Psi_j(y)$. We write $\hat{V}_0$ for the set of essential fixed points. We assume that the number of elements in $\hat{V}_0$ is greater than 1. Then, $\hat{K}$ is called a (compact) uniform finitely ramified fractal (u.f.r. fractal for short) if it satisfies the following finitely ramified property in addition to the above properties:
If $\{i_1, \dots, i_n\}, \{j_1, \dots, j_n\}$ are distinct sequences, then

$$\Psi_{i_1,\dots,i_n}(\hat{K}) \bigcap \Psi_{j_1,\dots,j_n}(\hat{K}) = \Psi_{i_1,\dots,i_n}(\hat{V}_0) \bigcap \Psi_{j_1,\dots,j_n}(\hat{V}_0). \tag{2.1}$$

U.f.r. fractals form a class of fractals that is wider than nested fractals, whose definition will be given below, and is included in the class of p.c.f. self-similar sets ([**23, 24**]).

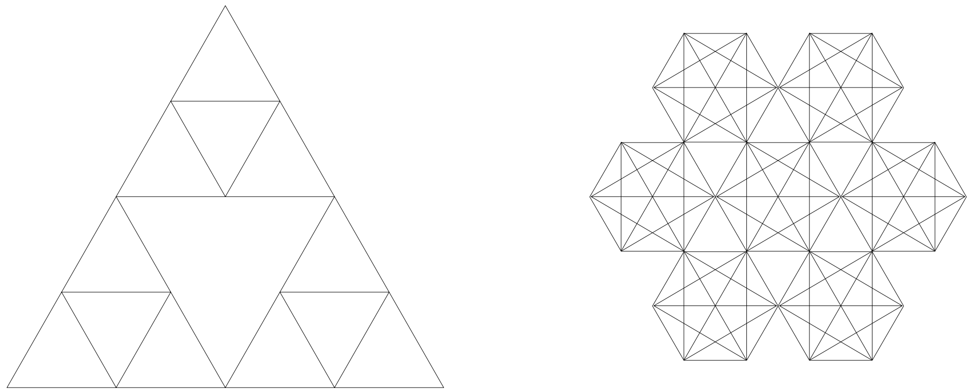

FIGURE 1. The Sierpinski gasket and Lindstrøm snowflake fractal graphs

The class of nested fractals was introduced in [**27**]. Let $F_n = \cup_{i_1,\cdots,i_n=1}^{N} \Psi_{i_1} \circ \cdots \circ \Psi_{i_n}(\hat{V}_0)$. $\hat{K}$ is called a (compact) nested fractal if it satisfies the following symmetry property in addition to the above properties and (2.1):
If $x, y \in \hat{V}_0$, then reflection in the hyperplane $H_{xy} = \{z \in \mathbb{R}^D : |z - x| = |z - y|\}$ maps F_n to itself.

Next we define unbounded u.f.r. fractals. We assume without loss of generality that $\Psi_1(\mathbf{x}) = \alpha_1^{-1}\mathbf{x}$ and $\mathbf{0}$ belongs to $\hat{V}_0$. Set $K = \cup_{n=1}^{\infty} \alpha^n \hat{K}$. Then, clearly $\Psi_1(K) = K$. We call K an unbounded uniform finitely ramified fractal. Let $V = V_0 = \cup_{n=0}^{\infty} \alpha^n F_n$ and $V_n = \alpha^n V$ for $n \in \mathbb{Z}$. (Note that our labelling is the opposite of the usual one given in say [**23**]. As n gets bigger, the graph distance between each vertex of V_n gets larger and $V_n \subset V_{n-1}$.) Then, $K = Cl(\cup_{n \in \mathbb{Z}} V_n)$. For each $l, n \geq 0$ and $i_1, \cdots, i_l \in S$, 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 $\hat{B}_0 := \{\{x, y\} : x \neq y \in \hat{V}_0\}$. 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. Examples of the types of fractal graph that the reader should think about are shown in Figures 1 and 2.

An *electrical network* (V, C) on an arbitrary infinite connected graph with vertices V is an assignment to each edge $\{x, y\}$ of a positive number $C_{xy} = C_{yx}$, the conductance between x and y. We will discuss such networks in Section 5. 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 $\{x, y\} \in \hat{B}_0$. 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 $\{x, y\} \notin B$. Note that, for our fractal electrical networks, there exists $c_{2.1}, c_{2.2} > 0$ such that

$$c_{2.1} \leq C_{xy} \leq c_{2.2} \qquad \text{for all } \{x, y\} \in B. \tag{2.2}$$

REMARK 2.1. Note that we can also consider the case where the contraction rate is different for each contraction (or more generally, graphs which are constructed from p.c.f. self-similar sets). In that case, the natural choice for the conductance of an edge will depend on the length of the edge, and hence (2.2) will not hold. However, by choosing a contraction map with the smallest contraction rate as Ψ_1 and extending in the same way as above, the lower bound of (2.2) is guaranteed and we can obtain most of the results in this paper in more or less the same way. In this paper, we only consider the case where the contraction rates are all the same for simplicity.

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 $\{x,y\} \in B$. For each $f,g \in l(V) := \{h : h \text{ is a function on } V\}$, we define

$$\mathcal{E}(f,g) = \frac{1}{2} \sum_{\substack{x,y \in V \\ \{x,y\} \in B}} (f(x) - f(y))(g(x) - g(y))C_{xy}. \tag{2.3}$$

We sometimes abbreviate $\mathcal{E}(f,f)$ as $\mathcal{E}(f)$. 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 $\{x,y\} \in B$, define $P_{xy} = C_{xy}/\mu_x$, which is then the transition probability matrix of the Markov chain corresponding to $\mathcal{E}$. To be precise, the Markov chain corresponding to $\mathcal{E}$ is a continuous time one in which jumps occur along edge $\{x,y\}$ at rate C_{xy}. In this paper, we will consider instead the induced random walk, the discrete time Markov chain which moves at unit time intervals to any vertex y in the neighbourhood of x with probabilities given by $\{P_{xy}\}$, since its long time asymptotic behaviour is similar to that of the continuous time Markov chain. We denote the induced random walk by $\{X_n\}_{n \geq 0}$. 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 $q_{ii} = 0, q_{ij} = q_{ji} \geq 0$ for any $i,j \in \hat{V}_0\}$. Also, let $\mathrm{Int}(\mathcal{Q}_M)$ be the subset of $\mathcal{Q}_M$ such that every Q has $q_{ii} = 0, q_{ij} = q_{ji} > 0$ for all $i \sim j \in \hat{V}_0$. As our quadratic form on the graph V is constructed from a basic electrical network, the conductances are determined by placing $C \in \mathrm{Int}(\mathcal{Q}_M)$ within each 0-cell, and we denote the form by $\mathcal{E}_C(\cdot,\cdot)$. Given a graph on V_1 which is similar to V_0 and $C' \in \mathrm{Int}(\mathcal{Q}_M(\alpha\hat{V}_0))$, we can define a form on V_1 in the same way:

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

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

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

see [**2, 23, 24**] 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, in that $F(\theta C) = \theta F(C)$ 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 [**30, 28, 29**] etc.). 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 $F(Q) = \rho_Q^{-1} Q_0$. We assume the following.

Assumption 2.2. (Non-degeneracy)*:*
There exists $Q_0 \in Int(\mathcal{Q}_M)$ and $\rho_{Q_0}^{-1} > 0$ such that $F(Q_0) = \rho_{Q_0}^{-1} Q_0$.

By Corollary 6.20 of [**2**], $\rho_{Q_0} > 0$ is uniquely determined, i.e. if $Q_1, Q_2 \in \mathrm{Int}(\mathcal{Q}_M)$ satisfies $F(Q_j) = \rho_{Q_j}^{-1} Q_j$ $(j = 1, 2)$ with $\rho_{Q_1}, \rho_{Q_2} > 0$, then $\rho_{Q_1} = \rho_{Q_2} = \rho_{Q_0}$. We thus denote $\bar{F} = \rho_{Q_0} F$. In the class of fractal graphs we consider, we can prove $\rho_{Q_0} > 1$ using Theorem 4.10 of [**24**]. 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 [**2, 23**] for details.

If the conductance of (V, C) is given by this non-degenerate fixed point Q_0, then all the results in the rest of the paper hold. Otherwise, we make an additional assumption. Let $\bar{F}^n$ be the n-th iteration of $\bar{F}$.

Assumption 2.3. *For all $Q \in Int(\mathcal{Q}_M)$, there exist $c_{1,Q}, c_{2,Q} > 0$ such that,*

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

Note that various examples, including nested fractal graphs, satisfy this assumption. For the rest of the paper we work under Assumption 2.3 unless otherwise stated.

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

We now define an effective resistance between $x \neq y \in V$ as follows.

$$R(x, y)^{-1} = \inf\{\mathcal{E}(f, f) : f \in l(V), f(x) = 1, f(y) = 0\}. \tag{2.4}$$

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. Indeed, by taking $h \in l(V)$ so that $h(x) = 1$ and $h(z) = 0$ for all $z \neq x$, we have

$$R(x, y)^{-1} \leq \sum_{\{x,p\} \in B} C_{xp}/2 = \mu_x/2 < \infty.$$

Also, by taking a shortest path $x = x_0, x_1 \cdots, x_{n-1}, x_n = y$, we have

$$R(x, y)^{-1} \geq \inf\{\frac{1}{2} \sum_{i=0}^{n-1} (f(x_i) - f(x_{i+1}))^2 C_{x_i x_{i+1}} : f(x) = 1, f(y) = 0\} \geq \frac{c_{2.1}}{2n} > 0.$$

Thus, we can follow the same argument as Proposition 2.9 in [**17**], Section 2.3 of [**23**] etc. so that the following holds.

Lemma 2.4. *1) The function $R(\cdot, \cdot)$ is a metric on V.*
2) For all $f \in l(V)$ and $x, y \in V$,

$$|f(x) - f(y)|^2 \leq R(x, y)\mathcal{E}(f, f). \tag{2.5}$$

In Section 3 and 4, we will use this resistance metric as a metric on V unless otherwise stated. Note also that the definition of R can be generalized for arbitrary disjoint $A, B \subset V$ as follows.

$$R(A,B)^{-1} = \inf\{\mathcal{E}(f,f) : f \in l(V), f|_A = 1, f|_B = 0\}. \tag{2.6}$$

REMARK 2.5. As we will discuss in the last part of Section 5, when

$$c_{2.3}R(x,y) \le d(x,y)^\eta \le c_{2.4}R(x,y) \qquad \text{for all } x,y \in V \tag{2.7}$$

holds for some $c_{2.3}, c_{2.4}, \eta > 0$, we only need to assume (2.2) and Assumption 2.2 for the conductance.

3. Preliminary results

In this section, we will give two results on the resistance metric which will be used later in establishing our heat kernel estimates. We begin by defining a sequence of graphs V_n such that the distance between points is α^n when the graph is embedded in a Euclidean space. From Section 2 we have a set of vertices V_n. Let $\hat{B}_n = \{\{x,y\} : x \neq y \in \alpha^n \hat{V}_0\}$. We define edges B_n by placing a copy of $\hat{B}_n$ in each of the n-cells in V_n. We will call the sub graph of $V = V_0$ in an n-cell an n-graph complex. We will write $x \sim_n y$ if $x, y \in V_n$ and $\{x,y\} \in B_n$.

By Assumption 2.3 we know that there is a resistance scale factor ρ such that

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

Hence, on the graph V_n, the resistance along each edge induced by the graph V is controlled by the resistance scale factor.

LEMMA 3.1. *Under Assumption 2.3 there exist constants $c_{3.1}, c_{3.2}$ such that for $x \sim_n y$,*

$$c_{3.1}\rho^n \le R(x,y) \le c_{3.2}\rho^n, \ \ \forall n \in \mathbb{N}.$$

An n-neighbourhood of the vertex x is defined to be

$$\begin{aligned} D_n^0(x) &= \{C : \text{C is an } n\text{-graph complex containing } x\} \\ D_n^1(x) &= D_n^0(x) \cup \{C : \text{C is an } n\text{-graph complex connected to } D_n^0(x)\}. \end{aligned}$$

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 $N_m(x,y) = \max_{x_1 \in \partial D_m^0(x), y_1 \in \partial D_m^0(y)} N_m(x_1,y_1)$.

Let $B_R(x,r)$ be a 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 [**2**]). For the non-degenerate case, we can prove that the volume growth of the ball is a power law. Let $S = \log N / \log \rho$, the Hausdorff dimension of the fractal in the resistance metric.

LEMMA 3.2. *There exists $c_{3.3}, c_{3.4}, r_0 > 0$ such that*

$$c_{3.3}r^S \le \mu(B_R(x,r)) \le c_{3.4}r^S \qquad \text{for all } r \ge r_0. \tag{3.1}$$

PROOF. By Lemma 3.1 there is a constant c_1 such that $D_{n-c_1}^1(x) \subset B_R(x,\rho^n) \subset D_{n+c_1}^1(x)$. For the upper bound we then observe that

$$\mu(B_R(x,\rho^n)) \le \mu(D_{n+c_1}^1(x)) \le c_2 N^{n+c_1}.$$

For a given r we choose n such that $\rho^n \le r \le \rho^{n+1}$ and hence

$$\mu(B_R(x,r)) \le c_3 r^S,$$

where $S = \log N / \log \rho$. The lower bound follows in exactly the same way. □

4. Heat kernel estimates for u.f.r. graphs

In this section we will obtain kernel estimates for the u.f.r. graphs under Assumption 2.3. Denote by $P_k(x, y)$ the transition function after k steps, i.e.

$$P_k(x, y) = P(X_k = y | X_0 = x),$$

where $\{X_k; k \in \mathbb{N}\}$ is the random walk on V associated with the Dirichlet form $\mathcal{E}$. 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 $p_k(x, y) = p_k(y, x)$.

We begin with the on-diagonal upper bound.

PROPOSITION 4.1. *There exists $c_{4.1} > 0$ such that*

$$p_k(x, x) \le c_{4.1} k^{-S/(S+1)}, \qquad \text{for all } x \in V, k \in \mathbb{N}.$$

We will follow the proof of Theorem 2.1 in [**6**]. As we are working in the resistance metric, it is necessary to modify the proof. The part which requires modification is Lemma 2.4 of [**6**]. To state the lemma here, we introduce some more notation. For a non-empty finite set $\Omega \subset V$, denote by $C_0(\Omega)$ the set of all real-valued functions on Ω extended by 0 outside Ω. Denote by $\mathcal{L}_\Omega$ the restriction of $\mathcal{L}$ to $C_0(\Omega)$, that is,

$$\mathcal{L}_\Omega f(x) = \begin{cases} \mathcal{L}f(x) & \text{if } x \in \Omega \\ 0 & \text{if } x \notin \Omega. \end{cases}$$

We can identify $C_0(\Omega)$ with $\mathbb{L}^2(\Omega, \mu)$ and then $-\mathcal{L}_\Omega$ acting on $\mathbb{L}^2(\Omega, \mu)$ is self-adjoint, positive definite and has discrete positive spectrum. The smallest positive eigenvalue, denoted by $\lambda_1(\Omega)$, admits the following variational characterization:

$$\lambda_1(\Omega) = \inf_{f \in C_0(\Omega)} \frac{\mathcal{E}(f, f)}{\sum_{x \in V} f(x)^2 \mu(x)}. \tag{4.1}$$

Let $r(\Omega)$ denote the inradius of Ω, defined by

$$r(\Omega) = \sup\{r \ge 0 : \exists x \in \Omega \text{ such that } B_R(x, r) \subset \Omega\}.$$

where we remark that the ball is with respect to the resistance metric. Note that in [**6**], $r(\Omega)$ is a natural number, but here it can take any non-negative value.

LEMMA 4.2. *For any set $\Omega \subset V$, $\lambda_1(\Omega) \ge 1/(r(\Omega)\mu(\Omega))$.*

PROOF. Take $f \in C_0(\Omega)$ and normalize so that $\max_{x \in \Omega} |f(x)| = 1$. Then $\sum_x f^2(x)\mu(x) \le \mu(\Omega)$. Now consider a point $x_0 \in \Omega$ such that $|f(x_0)| = 1$. If we choose $y_0 \notin \Omega$ so that $R(x_0, y_0) = \min_{y \notin \Omega} R(x_0, y)$, then $R(x_0, y_0) \le r(\Omega)$. We now use (2.5) and obtain

$$\frac{\mathcal{E}(f, f)}{\sum_{x \in V} f(x)^2 \mu(x)} \ge \frac{|f(x_0) - f(y_0)|^2}{R(x_0, y_0)\mu(\Omega)} \ge \frac{1}{r(\Omega)\mu(\Omega)}.$$

This holds for all f and hence we have the result by (4.1). □

Using this Lemma, we can prove Proposition 4.1 in the same way as [**6**]. For completeness, we will sketch the proof. The key fact is that the Faber-Krahn inequality (4.2) implies heat kernel upper bounds, a fact that was established for the case of graphs in [**7**] Proposition V.1, which we state here.

PROPOSITION 4.3. *Assume that* $v_0 := \inf_{x\in V}\mu(x) > 0$. *Suppose also that, for all non-empty finite sets* Ω,

$$\lambda_1(\Omega) \geq h(\mu(\Omega)), \tag{4.2}$$

where $h : [v_0,\infty) \to (0,\infty)$ *is a decreasing function. Then, for all* $x \in V$ *and for all* $k \in \mathbb{N}$, *there are constants* C, c *such that*

$$p_k(x,x) \leq C/\gamma(ck), \tag{4.3}$$

where γ *is defined by*

$$t = \int_{v_0}^{\gamma(t)} \frac{ds}{sh(s)}.$$

PROOF OF PROPOSITION 4.1. We first show that we have

$$\lambda_1(\Omega) \geq c_1\mu(\Omega)^{-(S+1)/S}. \tag{4.4}$$

Indeed, putting $r = r(\Omega)$, we have by definition of $r(\Omega)$ and by (3.1) that $c_{3.3}r^S \leq \mu(\Omega)$. Thus, $r(\Omega) \leq (\mu(\Omega)/c_{3.3})^{1/S}$. So (4.4) follows by Lemma 4.2. Now, by (4.4), we have the Faber-Krahn inequality (4.2) with $h(x) = c_2x^{-(S+1)/S}$. Applying Proposition 4.3, we have (4.3) with $\gamma(t) = (c_3t + v_0^{(S+1)/S})^{S/(S+1)} \geq c_4t^{S/(S+1)}$. We thus obtain the desired inequality. □

Note that we only use the first inequality of (2.2), which is used to guarantee $v_0 > 0$ in Proposition 4.3, and (3.1).

We now consider the off-diagonal upper bound and follow a similar argument to that of [**17**]. We begin with an estimate on the tail of the crossing time for an n-cell.

Let $T_0^n = \inf\{m : X_m \in V_n\}$ and, for $i \geq 1$, let

$$T_i^n = \inf\{m : X_m \in V_n \backslash \{X_{T_{i-1}^n}\}\}.$$

These are the successive hits on the large scale vertex set V_n by the random walk. Using these stopping times we can define a random walk $\{X_k^n; k \in \mathbb{N}\}$ on the graph V_n by setting $X_k^n = X_{T_k^n}$ for $k \in \mathbb{N}$. We let $W_i^n = T_i^n - T_{i-1}^n$ be the crossing time of an n-cell. We also define the exit time from a set A as

$$T_A = \inf\{m : X_m \notin A\}.$$

We need some estimates on the mean hitting times. For this we use the excursion ideas of [**2**] and the fact that there is a fixed point for the decimation map. We first define the time scale factor $\tau = \rho N$.

LEMMA 4.4. *There exist constants* $c_{4.2}, c_{4.3}$ *such that for* $x \in V_n$

$$c_{4.2}\tau^n \leq E^x(W_i^n) \leq c_{4.3}\tau^n, \tag{4.5}$$

and

$$\sup_{z\in D_n^0(x)} E^z(T_0^n) \leq c_{4.4}\tau^n. \tag{4.6}$$

PROOF. For $x \in V_n$, by construction and Assumption 2.3 there is a constant c_1 independent of n such that

$$c_1 \leq P^x(X_{W_1^n} = y) \leq 1 - c_1, \quad \forall y \in \partial D_n^0(x).$$

Let $U = \min\{k > 0 : X_k \in V_n\}$. Then

$$E^x W_1^n = E^x U + P^x(X_U = x) E^x W_1^n.$$

Now there are constants c_2, c_3 such that, $c_2 \rho^{-n} \leq 1 - P^x(X_U = x) \leq c_3 \rho^{-n}$ by construction of the random walk and Assumption 2.3 and hence there are constants c_4, c_5 such that

$$c_4 \rho^n E^x U \leq E^x W_1^n \leq c_5 \rho^n E^x U.$$

The mean return time for a random walk can be expressed in terms of the stationary distribution and is given by $E^x U = 1/\mu(x) \asymp N^n$ and hence we have the first result.

The second follows by considering the path of the random walk from any point z in $D_n^0(x)$. First, conditioning on the exit place on $\partial D_{n-1}^0(z)$, we have

$$\begin{aligned} E^z T_0^n &= E^z T_0^{n-1} + \sum_{x \in \partial D_{n-1}^0(z)} P^z(X_{\partial D_{n-1}^0(z)} = x) E^x T_0^n \\ &\leq E^z T_0^{n-1} + \max_{x \in \partial D_{n-1}^0(z)} E^x T_0^n. \end{aligned}$$

Repeating this procedure gives

$$E^z T_0^n \leq \sum_{k=1}^n \max_{x_k \in \partial D_{k-1}^0(z)} E^{x_k} T_0^k,$$

for all $z \in D_n^0(x)$ where $D_0^0(y) = y$. From (4.5) and an easy Markov chain computation, we have $E^{x_k} T_0^k \leq c\tau^k$. Using this we obtain the result. □

We can now apply the argument of [**2, 5, 17**] to prove the following.

PROPOSITION 4.5. *There exist constants* $0 < c_{4.5} < 1, 0 < c_{4.6}$ *such that*

$$P^x(W_1^n \leq k) \leq 1 - c_{4.5} + c_{4.6} k \tau^{-n}, \quad \forall x \in V, n, k > 0. \tag{4.7}$$

PROOF. Note first that it is enough to prove (4.7) for $x \in V_n$. By the Markov property, for each $x \in V_n$ we have

$$E^x T_{\partial D_n^0(x)} \leq k + E^x[1_{\{T_{\partial D_n^0(x)} > k\}} E^{X_k} T_{\partial D_n^0(x)}].$$

Since $x \in V_n$, $E^y T_{\partial D_n^0(x)} = E^y T_0^n$ if $y \in D_n^0(x) \setminus V_n$ and $T_{\partial D_n^0(x)} = W_1^n$ under P^x. Thus, applying Lemma 4.4, we have

$$c_{4.2}\tau^n \leq k + c_{4.4}\tau^n P^x(T_{\partial D_n^0(x)} > k) = k + c_{4.4}\tau^n(1 - P^x(W_1^n \leq k)).$$

Rearranging gives the result. □

This estimate can be used to obtain the heat kernel upper bound (cf. [**2, 5, 17**] etc.).

THEOREM 4.6. *There exist constants* $c_{4.7}, c_{4.8}$ *such that*

$$p_k(x,y) \leq c_{4.7} k^{-S/(S+1)} \exp(-c_{4.8} N_{l(k,n)}(x,y)), \quad \forall x, y \in V, \quad k > 0,$$

where

$$\rho^n \leq R(x,y) \leq \rho^{n+1},$$

and

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

PROOF. Let $A(x) = \{z \in V : N_{l(k,n)}(x,z) \geq \frac{1}{2}N_{l(k,n)}(x,y)\}$ and $A^c(x) = V\backslash A(x)$. Then

$$\begin{aligned} p_k(x,y) &= P^x(X_k=y)/\mu_y \\ &= \frac{1}{\mu_y}\left(P^x(X_k=y, X_{[k/2]}\in A(x)) + P^x(X_k=y, X_{[k/2]}\in A^c(x))\right) \\ &= \frac{1}{\mu_y}\left(I_1+I_2\right). \end{aligned}$$

We also set for some constant c_1 to be chosen later,

$$l'(k,n)=\inf\{j: N_j(x,y)\tau^j \geq c_1 k\} \wedge n.$$

We first consider the case $l'(k,n) < n$. Observe that by construction of the set $A(x)$,

$$T_{A(x)} \geq \sum_{i=1}^{\frac{1}{2}N_{l'(k,n)}(x,y)} W_i^{l'}. \tag{4.8}$$

Now, using (4.7), (4.8) with Lemma 3.14 of [**2**] and the definition of $l'(k,n)$, we have

$$\begin{aligned} P^x(X_{[k/2]}\in A(x)) &= P^x(T_{A(x)} \leq [k/2]) \\ &\leq P^x(\sum_{i=1}^{\frac{1}{2}N_{l'(k,n)}(x,y)} W_i^{l'} \leq [k/2]) \\ &\leq \exp(-c_2 N_{l'(k,n)}(x,y) + c_3\left(N_{l'(k,n)}(x,y)\tau^{-l}k\right)^{1/2}) \\ &\leq \exp(-c_4 N_{l'(k,n)}(x,y)). \end{aligned}$$

We can ensure that $c_4 > 0$ by choosing c_1 large enough in the definition of l'. Now observe that there is a constant c_5 such that $N_{l'(k,n)}(x,y) \geq c_5 N_{l(k,n)}(x,y)$ and hence

$$P^x(X_{[k/2]}\in A(x)) \leq c_6\exp(-c_7 N_{l(k,n)}(x,y)). \tag{4.9}$$

When $l'(k,n)=n$, then $N_n(x,y)\leq c_8$ so that (4.9) clearly holds.

Finally, by the on-diagonal upper bound and (4.9), we have

$$\begin{aligned} I_1 &= \sum_{z\in A(x)} p_{[k/2]}(x,z)p_{k-[k/2]}(z,y)\mu_y\mu_z \\ &\leq c_{4.1}(k-[k/2])^{-S/S+1}P^x(X_{[k/2]}\in A(x))\mu_y \\ &\leq c_{4.1}(k-[k/2])^{-S/S+1}\exp(-c_6 N_{l(k,n)}(x,y))\mu_y. \end{aligned}$$

By the symmetry of $p_k(x,y)$ we can establish the same result for I_2 and hence obtain the result. □

We now discuss the lower bound for the heat kernel. We note that, if the graph is bipartite, it may be the case that $p_{2k+1}(x,x) = 0$ and hence we would have no non-trivial lower bound. For this reason we work with the sum $u_k(x,y) = p_k(x,y) + p_{k+1}(x,y)$.

LEMMA 4.7. *There exists a constant $c_{4.9}$ such that for all $k \in \mathbb{N}$*

$$p_k(x,x) + p_{k+1}(x,x) \geq c_{4.9} k^{-S/(S+1)}.$$

PROOF. Using (4.7) we have that

$$P^x(X_k \notin D_n^1(x)) \leq P(W_1^n \leq k) \leq 1 - c_{4.5} + c_{4.6} k \tau^{-n}.$$

Hence by choosing a large enough n we have for $c_1 \tau^{n-1} \leq k \leq c_1 \tau^n$,

$$P^x(X_k \notin D_n^1(x)) \leq c_2 < 1.$$

Thus $P^x(X_k \in D_n^1(x)) \geq 1 - c_2 > 0$ and, by Cauchy-Schwarz,

$$(1-c_2)^2 \leq \mu(D_n^1(x)) p_{2k}(x,x).$$

As $\mu(D_n^1(x)) \asymp N^n$, we use the lower bound on our choice of k to get

$$p_{2k}(x,x) \geq c_3 k^{-S/(S+1)}.$$

If we rewrite in terms of k and adjust for parity we have the result. □

For the lower estimate in the finitely ramified fractal literature, it is standard to use the Sobolev type inequality (2.5) to obtain a near diagonal lower bound (cf. [**2, 17**]). This approach allows us to control the heat kernel with a Hölder continuity estimate of the form

$$|p_k(x,y) - p_k(x,y')|^2 \leq R(y,y') \mathcal{E}(p_k(x,.), p_k(x,.)). \tag{4.10}$$

We need a slight modification of this technique here.

A straightforward computation using the definition of the Dirichlet form shows that

$$\mathcal{E}(p_m(x,.), p_n(x,.)) = p_{n+m}(x,x) - p_{n+m+1}(x,x), \quad \forall n,m \in \mathbb{N}, x \in V. \tag{4.11}$$

LEMMA 4.8. *There exist constants $c_{4.10}, c_{4.11}$ such that*

$$u_k(x,y) \geq c_{4.10} k^{-S/(S+1)} \text{ if } R(x,y) \leq c_{4.11} k^{1/(S+1)}.$$

PROOF. We begin by controlling the Hölder continuity of u_k. We compute using (4.11) that

$$\begin{aligned} \mathcal{E}(u_k(x,.)) &= \mathcal{E}(p_k(x,.)) + 2\mathcal{E}(p_k(x,.), p_{k+1}(x,.)) + \mathcal{E}(p_{k+1}(x,.)) \\ &= p_{2k}(x,x) - p_{2k+2}(x,x) + p_{2k+1}(x,x) - p_{2k+3}(x,x). \end{aligned}$$

Note that, as discussed in [**13**], $p_{2k+2}(x,x) - p_{2k}(x,x)$ is the time derivative of the heat kernel. We recall from [**13**] Proposition 12.3 that the time derivative of the discrete heat kernel can be estimated from above. The proof given there transfers to our graphs and hence there are constants c_1, c_2 such that if $p_k(x,x) \leq c_1 k^{-\nu}$, then $|p_{k+2}(x,x) - p_k(x,x)| \leq c_2 k^{-\nu-1}$. Thus, putting u into (2.5), as in (4.10), we have

$$|u_k(x,y) - u_k(x,y')|^2 \leq c_3 R(y,y') k^{-S/(S+1)-1}.$$

Now,

$$\begin{aligned} u_k(x,y) &\geq u_k(x,x) - |u_k(x,x) - u_k(x,y)| \\ &\geq c_4 k^{-S/S+1} - c_3 \left(R(x,y)k^{-S/(S+1)-1}\right)^{1/2} \\ &= c_4 k^{-S/S+1}\left(1 - c_5\left(R(x,y)k^{-1/(S+1)}\right)^{1/2}\right) \\ &\geq \frac{1}{2}c_4 k^{-S/(S+1)}, \end{aligned}$$

by the choice of y such that $R(x,y) \leq (\frac{1}{2c_5})^2 k^{1/(S+1)}$, as required. □

With this result we can proceed to the usual chain argument to derive the off diagonal lower bound. Recall that $d(x,y)$ is the graph distance between x and y.

PROPOSITION 4.9. *There exist constants $c_{4.12}, c_{4.13}$ such that*

$$u_k(x,y) \geq c_{4.12} k^{-S/(S+1)} \exp(-c_{4.13} N_{l(k,n)}(x,y)), \quad \forall x,y \in V, k \geq d(x,y),$$

where $\rho^{n-1} \leq R(x,y) \leq \rho^n$ and

$$l(k,n) = \inf\{j : N_j(x,y)\tau^j \geq k\} \wedge n. \tag{4.12}$$

PROOF. We first note that we require $k \geq d(x,y)$ as otherwise $p_k(x,y) = 0$. We next observe that the bound is satisfied if $R(x,y)^{S+1} < c_{4.11}^{S+1} k$. Hence we take $\Theta = R(x,y)^{S+1}/k \geq c_{4.11}^{S+1}$. As in the upper bound case we define

$$l'(k,n) = \inf\{j : N_j(x,y)\tau^j \geq c_1 k\} \wedge n,$$

where c_1 will be chosen later. We first consider the case $l'(k,n) < n$. By this definition there are constants c_2, c_3 such that

$$c_2 N_{l'(k,n)}(x,y)^{-1}\tau^{n-l'} \leq \Theta \leq c_3 N_{l'(k,n)}(x,y)^{-1}\tau^{n-l'}.$$

Note that the upper bound still holds when $l'(k,n) = 0$ as in that case $N_0(x,y) \geq c_1 k \geq c_1 d(x,y) = c_1 N_0(x,y)$ and hence $\Theta \leq c_4 \tau^n / N_0(x,y)$.

Thus

$$\frac{R(x,y)}{\rho^{n-1-l'}} \leq c_5 \left(\frac{k}{N_{l'}(x,y)}\right)^{1/(S+1)}.$$

We can now choose a minimal path $\{x_i : i = 1, \ldots, N_{l'}(x,y)\}$ on $V_{l'}$ from x to y. Then

$$R(x_i, x_{i+1}) \leq \rho^{l'} \leq \frac{R(x,y)}{\rho^{n-1-l'}} \leq c_5 \left(\frac{k}{N_{l'}(x,y)}\right)^{1/(S+1)}. \tag{4.13}$$

When $l'(k,n) = n$, it is easy to check (4.13) holds.

We now take a ball $B_\epsilon(x_i)$ of radius $\epsilon = \rho^{l'}$ about each point in the path. Then, by construction, for $z_i \in B_\epsilon(x_i), z_{i+1} \in B_\epsilon(x_{i+1})$ we have

$$R(z_i, z_{i+1}) \leq c_6 \left(\frac{k}{N_{l'}(x,y)}\right)^{1/(S+1)}.$$

Let $N = N_{l'}(x,y)$, $m = [k/N] - 1$ and define $\tilde{P}_n = u_n(x,y)\mu_y$. Then, as in [**13**] Lemma 13.4–13.6, we have a constant $C > 1$ such that

$$(\tilde{P}_m)^N \leq C^{N-1}\tilde{P}_{N(m+1)-1} \leq C^{N-1} C^{k-N(m+1)+1}\tilde{P}_k \leq C^{2N}\tilde{P}_k.$$

By choosing c_1 small enough, we can ensure that $c_6 < c_{4.11}^{S+1}$ and hence the near diagonal bound can be used along our path, to give

$$\begin{aligned} u_k(x,y) &\geq C^{-2N}(\tilde{P}_m)^N/\mu_y \\ &\geq C^{-2N}\sum_{B_\epsilon(x_1)}\cdots\sum_{B_\epsilon(x_{N-1})} u_m(x,z_1)\dots u_m(z_{N-1},y)\mu_{z_1}\dots\mu_{z_{N-1}} \\ &\geq c_7^N\prod_{i=1}^{N-1}\mu(B_\epsilon(x_i))(c_8 m^{-S/(S+1)})^N \\ &\geq c_9 m^{-S/(S+1)}\exp(-c_{10}N). \end{aligned}$$

As in the upper bound we can change $N_{l'}$ to N_l by changing the constant c_{10}. □

We can now state our main theorem as follows.

THEOREM 4.10. *There exist constants $c_{4.14},\cdots,c_{4.17} > 0$ such that for all $x,y\in V$ and $k\geq d(x,y)$, if $\rho^{n-1}\leq R(x,y)\leq\rho^n$, then*

$$\begin{aligned} p_k(x,y) &\leq c_{4.14}k^{-\frac{S}{S+1}}\exp(-c_{4.15}N_{l(k,n)}(x,y)), \\ p_k(x,y)+p_{k+1}(x,y) &\geq c_{4.16}k^{-\frac{S}{S+1}}\exp(-c_{4.17}N_{l(k,n)}(x,y)), \end{aligned}$$

where $l(k,n)$ is given in (4.12).

Note that this theorem contains the diagonal estimate as $0\leq N_{l(k,n)}(x,x)\leq 1$ for all $n,k\in\mathbb{N}$ and $x\in V$. For $x,y\in V$ such that $R(x,y)\asymp\rho^n$, define a chemical exponent with respect to the resistance metric, for $0\leq l<n$ as

$$d_l^c(x,y) = \frac{1}{n-l}\log_\rho N_l(x,y).$$

For $l=n$ we can choose it arbitrarily and thus define $d_n^c(x,y)=1$. Using the definition of $l(k,n)$ we can write Theorem 4.10 in terms of the resistance metric and this chemical exponent as follows.

COROLLARY 4.11. *There exist constants $c_{4.18},\cdots,c_{4.21}>0$ such that for $x,y\in V$ and $k\geq d(x,y)$, with n,l as above, then*

$$\begin{aligned} p_k(x,y) &\leq c_{4.18}k^{-\frac{S}{S+1}}\exp\left(-c_{4.19}\left(\frac{R(x,y)^{S+1}}{k}\right)^{\frac{d^c_{l(k,n)}(x,y)}{S+1-d^c_{l(k,n)}(x,y)}}\right), \\ p_k(x,y)+p_{k+1}(x,y) &\geq c_{4.20}k^{-\frac{S}{S+1}}\exp\left(-c_{4.21}\left(\frac{R(x,y)^{S+1}}{k}\right)^{\frac{d^c_{l(k,n)}(x,y)}{S+1-d^c_{l(k,n)}(x,y)}}\right). \end{aligned}$$

Now, using self-similarity, Lemma 3.3, Lemma 3.4 and (3.11) of [**17**], we have for $\{x,y\}\in B_n$,

$$c_1\rho^l\leq N_{n-l}(x,y)\leq c_2\tau^{l/2}.$$

(From this we see that $S\geq 1$.) Thus from the definition of $d_l^c(x,y)$ we have for $0\leq l\leq n$,

$$1-\frac{c_3}{n-l}\leq d_l^c(x,y)\leq\frac{S+1}{2}+\frac{c_4}{n-l},$$

so that the following holds.

COROLLARY 4.12. *There exist constants* $c_{4.22}, \cdots, c_{4.25} > 0$ *and* $0 < \gamma_1 < \gamma_2$ *such that for* $x, y \in V$ *and* $k \geq d(x,y)$, *then*

$$p_k(x,y) \leq c_{4.22} k^{-\frac{S}{S+1}} \exp\left(-c_{4.23}\left(\frac{R(x,y)^{S+1}}{k}\right)^{\gamma_1}\right),$$

$$p_k(x,y) + p_{k+1}(x,y) \geq c_{4.24} k^{-\frac{S}{S+1}} \exp\left(-c_{4.25}\left(\frac{R(x,y)^{S+1}}{k}\right)^{\gamma_2}\right).$$

We note that in general we do not have the usual sub Gaussian estimates on the heat kernel in the resistance metric, and we will give an explicit example later where $\gamma_1 \neq \gamma_2$. We also note that this result is not altered by moving to the graph metric and shows that we are not in the framework of [**13, 14**].

For the nested fractal graphs, when the fractals have spatial symmetry, we can prove in the same way as [**25**] and [**11**] that (2.7) holds and there are constants $d_c, c_1, c_2 > 0$ (d_c is called a chemical exponent) such that for $x, y \in V_n$, with $x \sim_n y$,

$$c_1 \rho^{(n-m)d_c} \leq N_m(x,y) \leq c_2 \rho^{(n-m)d_c}.$$

It is known that Assumption 2.2 and Assumption 2.3 hold in this case (see [**30, 28, 29**] etc.). Hence, writing $d_w = (S+1)/d_c$, we can express our estimates in the following form.

COROLLARY 4.13. *Let* (V, C) *be an electrical network on a nested fractal graph. Then there exist constants* $c_{4.26}, \ldots, c_{4.29}$ *such that for all* $x, y \in V, k > 0$

$$p_k(x,y) \leq c_{4.26} k^{-S/(S+1)} \exp\left(-c_{4.27}\left(\frac{d(x,y)^{d_w}}{k}\right)^{1/(d_w-1)}\right),$$

and for $k > d(x,y)$,

$$p_k(x,y) + p_{k+1}(x,y) \geq c_{4.28} k^{-S/(S+1)} \exp\left(-c_{4.29}\left(\frac{d(x,y)^{d_w}}{k}\right)^{1/(d_w-1)}\right).$$

In [**17**] it was observed that for p.c.f. fractals, if there is no spatial symmetry, there may be dependence of the off-diagonal terms on direction. A corresponding example for u.f.r. fractal graphs is shown in Figure 2. To see that there is a dependency in the exponents on direction we must compute the behaviour of $N_m(x,y)$ for horizontal or vertical moves in the fractal.

In the case of the left fractal in Figure 2 we see that there is a unique exponent for the scaling of the shortest path. It is given by the scaling in the vertical direction as any horizontal move can be made from a vertical move up, followed by a vertical move down. Thus for all $x \sim_n y$ we have $N_m(x,y) \asymp 12^{n-m}$. However in the fractal on the right, where it is the horizontal move which has minimal length, we note that we cannot make a vertical move from a horizontal one. Hence we see that there are two exponents, one for vertical and one for horizontal moves. If $x, y \in V_n$ are connected by a horizontal line, then $N_m(x,y) \asymp 12^{n-m}$, while if $x, y \in V_n$ and they lie on a diagonal line in a triangle, then $N_m(x,y) \asymp 14^{n-m}$.

In the left figure it is possible to calculate the map F explicitly as a rational function of polynomials in one variable and hence deduce the existence of a fixed point at (1,1,1.157) (where we are writing the resistance of each edge in the triangle, with the third component corresponding to the horizontal edge) with resistance scale factor $\rho = 8.2934$. Thus we can compute all the exponents as $S = \log 34/\log 8.2934 = 1.6669$, $d_c^v = \log 12/\log 8.2934 = 1.1746$ and $d_w^v = 2.2704$.

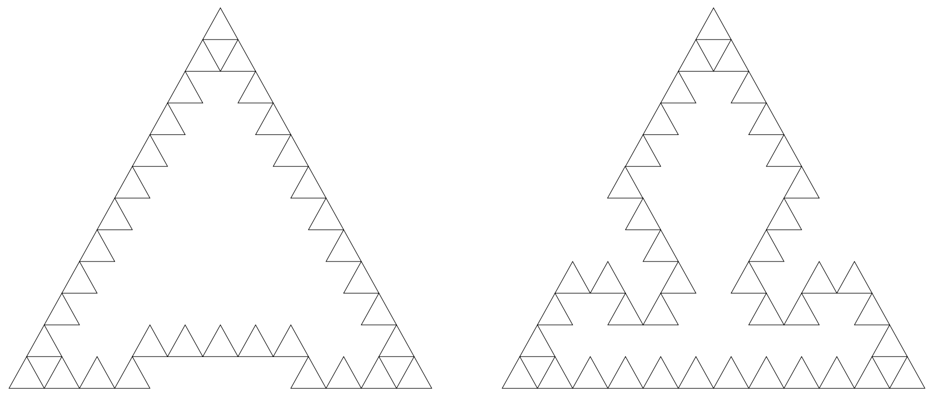

FIGURE 2. Two u.f.r. fractal graphs with different directional dependence

COROLLARY 4.14. *In the left figure of Figure 2, there exist constants $c_{4.30}, \dots, c_{4.33}$ such that for all $x, y \in V, k > d(x,y)$,*

$$p_k(x,y) \le c_{4.30} k^{-S/(S+1)} \exp\left(-c_{4.31}\left(\frac{d(x,y)^{d_w^v}}{k}\right)^{1/(d_w^v-1)}\right),$$

and

$$p_k(x,y) + p_{k+1}(x,y) \ge c_{4.32} k^{-S/(S+1)} \exp\left(-c_{4.33}\left(\frac{d(x,y)^{d_w^v}}{k}\right)^{1/(d_w^v-1)}\right).$$

In the right figure we have a fractal graph for which the analogue of the p.c.f. result obtained in [**17**] Theorem 6.3 and Theorem 6.5 holds. For this graph the off diagonal heat kernel estimate has different exponents in the two different directions. We can once again calculate the map F explicitly as a one dimensional map and hence deduce the existence of a fixed point at (1,1,0.6256) with resistance scale factor $\rho = 9.1593$. Note there is also a degenerate fixed point at (1,1,0) but, as we have the map explicitly, it is easy to check that this fixed point is repulsive and the non-degenerate one is attractive and hence Assumption 2.3 is satisfied. The exponents are given by $S = 1.6767$, $d_c^v = \log 14/\log 9.1593 = 1.1916$ and $d_w^v = 2.2464$ and $d_c^h = \log 12/\log 9.1593 = 1.122$, $d_w^h = 2.3857$.

COROLLARY 4.15. *In the right figure of Figure 2, there exist constants $c_{4.34}, \dots, c_{4.41}$ such that for $x, y \in V$ connected by a horizontal line and $k > d(x,y)$, then*

$$p_k(x,y) \le c_{4.34} k^{-S/(S+1)} \exp\left(-c_{4.35}\left(\frac{d(x,y)^{d_w^h}}{k}\right)^{1/(d_w^h-1)}\right),$$

and

$$p_k(x,y) + p_{k+1}(x,y) \ge c_{4.36} k^{-S/(S+1)} \exp\left(-c_{4.37}\left(\frac{d(x,y)^{d_w^h}}{k}\right)^{1/(d_w^h-1)}\right).$$

If y lies on the 60° diagonal through the origin and $k > d(0,y)$, then

$$p_k(0,y) \le c_{4.38} k^{-S/(S+1)} \exp\left(-c_{4.39}\left(\frac{d(0,y)^{d_w^v}}{k}\right)^{1/(d_w^v-1)}\right),$$

and

$$p_k(0,y) + p_{k+1}(0,y) \ge c_{4.40} k^{-S/(S+1)} \exp\left(-c_{4.41}\left(\frac{d(0,y)^{d_w^v}}{k}\right)^{1/(d_w^v-1)}\right).$$

We can express these estimates in a more general form, that is similar to that used for fractal fields in [**16**]. It is clear that, using the embedding of the graph in $\mathbb{R}^2$, that each edge in this graph can be labelled as either horizontal or vertical. Hence we can write the graph distance as

$$d(x,y) = d_h(x,y) + d_v(x,y),$$

where $d_h(x,y)$ and $d_v(x,y)$ denote the number of horizontal and vertical edges in the shortest path from x to y. We can also define two shortest path counting functions, one for each step type. By considering the paths in the fractal to consist of these two types of steps we can reprove our Theorem 4.10 to obtain off-diagonal terms expressed as exponentials of the sum of these two counting functions, at possibly different levels in the fractal. By rewriting these in terms of the two shortest path metrics we have the following, where S, d_w^h, d_w^v are the same as those in Corollary 4.15.

COROLLARY 4.16. *In the right figure of Figure 2 there exist constants $c_{4.42}, \dots, c_{4.45}$ such that for $x, y \in V$ and $k > d(x,y)$, then*

$$p_k(x,y) \le c_{4.42} k^{-S/(S+1)} \exp\left(-c_{4.43} \sum_{i=h,v} \left(\frac{d_i(x,y)^{d_w^i}}{k}\right)^{1/(d_w^i-1)}\right),$$

and

$$p_k(x,y) + p_{k+1}(x,y) \ge c_{4.44} k^{-S/(S+1)} \exp\left(-c_{4.45} \sum_{i=h,v} \left(\frac{d_i(x,y)^{d_w^i}}{k}\right)^{1/(d_w^i-1)}\right).$$

We believe that for p.c.f. fractal graphs it will be possible to express the heat kernel estimates in general in this form.

5. Harnack inequality and its Stability

5.1. Parabolic Harnack inequality. Let $\beta > 0$. We say (V, C) satisfies $(PHI(\beta))$, a parabolic Harnack inequality of order β if whenever $u(n,x) \ge 0$ is defined on $[0,4N] \times \bar{B}(y,2R)$ and satisfies

$$u(n+1,x) - u(n,x) = \mathcal{L}u(n,x), \qquad (n,x) \in [0,4N] \times B(y,2R),$$

then

$$\max_{\substack{N \le n \le 2N \\ x \in B(y,R)}} u(n,x) \le c_{5.1} \min_{\substack{3N \le n \le 4N \\ x \in B(y,R)}} (u(n,x) + u(n+1,x)), \tag{5.1}$$

where $N \ge 2R$ and $c_{5.2} R^\beta \le N \le c_{5.3} R^\beta$.

PROPOSITION 5.1. *Let (V, C) be a basic electrical network on a u.f.r. graph which satisfies Assumption 2.3. Then, (V, C) satisfies $(PHI(S+1))$ with respect to the resistance metric.*

The proof follows from the estimates on $p_k(x, y)$ in Corollary 4.12 by the same argument as that used in [**10**] Section 3. Note that by Corollary 4.15, we see that $(PHI(\beta))$ does not hold in general with respect to the graph distance for any choice of β.

5.2. The results of Grigor'yan-Telcs and Barlow-Bass. In this subsection, we will explain the results in [**14, 3**]. Throughout this and the next subsections, we use the graph distance $d(\cdot, \cdot)$ and the ball $B(x, r)$ is in the graph distance.

DEFINITION 5.2. Let (V, C) be an electrical network on a connected infinite graph. Let μ be a measure on V such that $\mu(\{x\}) = \mu_x := \sum_{y\in V} C_{xy}$ for $x \in V$.
(1) We say (V, C) satisfies the (p_0) condition if there exists $p_0 > 0$ such that

$$p_{xy} = C_{xy}/\mu_x \geq p_0 \quad \text{for all } \{x, y\} \in B.$$

(2) We say (V, C) satisfies the volume doubling condition (VD) if there exists $c_{5.4} > 1$ such that

$$\mu(B(x, 2R)) \leq c_{5.4}\mu(B(x, R)) \qquad \text{for all } x \in V, R \geq 1. \tag{5.2}$$

(3) We say that (V, C) has walk dimension β and that (V, C) satisfies (E_β) if

$$E^x[\tau_{B(x,r)}] \asymp r^\beta, \qquad \text{for all } r \in [1, \infty),\ x \in V.$$

(4) We say (V, C) satisfies an elliptic Harnack inequality (EHI) if there exists $c_{5.5} > 0$ such that, whenever $x \in V, R \geq 1$ and $h : V \to \mathbb{R}$ is non-negative and harmonic in $B(x, 2R)$,

$$\sup_{y\in B(x,R)} h(y) \leq c_{5.5} \inf_{y\in B(x,R)} h(y).$$

(5) We say (V, C) satisfies (R_β) and has resistance exponent β if

$$R(B(x_0, r), B(x_0, 2r)^c) \asymp \frac{r^\beta}{\mu(B(x_0, r))}, \qquad \text{for all } r \geq 1.$$

THEOREM 5.3. ([**14**]) *Suppose (V, C) is an electrical network on a connected infinite graph satisfying the (p_0) condition. Then the following are equivalent:*
(a) *(V, C) satisfies $PHI(\beta)$.*
(b) *(V, C) satisfies (VD), (E_β) and (EHI).*
(c) *For $x, y \in V, n \geq d(x, y)$, the transition density of $\{X_n\}_n$ on (V, C) satisfies*

$$\begin{aligned} p_n(x, y) &\leq \frac{c_{5.6}}{\mu(B(x, n^{1/\beta}))} \exp[-(\frac{d(x, y)^\beta}{c_{5.6}n})^{1/(\beta-1)}], \\ p_n(x, y) + p_{n+1}(x, y) &\geq \frac{c_{5.7}}{\mu(B(x, n^{1/\beta}))} \exp[-(\frac{d(x, y)^\beta}{c_{5.7}n})^{1/(\beta-1)}]. \end{aligned}$$

(d) *(V, C) satisfies (VD), (EHI) and (R_β).*

REMARK 5.4. When any (thus all) of the above conditions hold, then $\beta \geq 2$ (see [**1**] Lemma 1.1).

In our work the resistance metric is not a geodesic metric and differs from the graph distance. The example we studied (Corollary 4.15) shows that the class of fractal graphs we treat is not within the framework of [**14**]. If we make the additional assumption that there is spatial symmetry in the fractal, then it is within the framework of [**14**] as we have observed in Corollary 4.13. We will discuss this further in the next subsection.

DEFINITION 5.5. (1) We say (V, C) satisfies $(PI(\beta))$, a scaled Poincaré inequality with parameter $\beta \geq 2$, if there exists a constants $c_{5.8} > 0$ such that for any ball $B = B(x_0, R) \subset V$ with $R \geq 1$ and $f : B \to \mathbb{R}$,

$$\sum_{x\in B}(f(x) - \bar{f}_B)^2\mu_x \leq c_{5.8}R^\beta \sum_{x,y\in B} C_{xy}(f(x) - f(y))^2, \tag{5.3}$$

where $\bar{f}_B = \mu(B)^{-1}\sum_{y\in B} f(y)\mu_y$.
(2) Let $\beta \geq 2$ and $\theta \in (0, 1]$. We say (V, C) satisfies $(CS(\beta, \theta))$, the cut-off Sobolev inequality with exponents β and θ, if there exist constants $c_{5.9}, c_{5.10} > 0$ such that for every $x_0 \in V, R \geq 1$, there exists a cut-off function $\varphi(= \varphi_{x_0,R})$ satisfying the following properties.
(a) $\varphi(x) \geq 1$ for $x \in B(x_0, R/2)$.
(b) $\varphi(x) = 0$ for $x \in B(x_0, R)^c$.
(c) $|\varphi(x) - \varphi(y)| \leq c_{5.9}(d(x, y)/R)^\theta$ for all $x, y \in V$.
(d) For any ball $B(x_0, s)$ with $1 \leq s \leq R$ and $f : B(x_0, 2s) \to \mathbb{R}$,

$$\begin{aligned}\sum_{x\in B(x_0,s)} f(x)^2\sum_{y\in V} C_{xy}|\varphi(x) - \varphi(y)|^2 \\ \leq\ c_{5.10}(s/R)^{2\theta}\Big(\sum_{x,y\in B(x_0,2s)} C_{xy}|f(x) - f(y)|^2 + s^{-\beta}\sum_{y\in B(x_0,2s)} f(y)^2\mu_y\Big).\end{aligned} \tag{5.4}$$

REMARK 5.6. Suppose (V, C) satisfies (VD) and $(CS(\beta, \theta))$, but with (a) replaced by
(a') $\varphi(x) \geq 1$ for $x \in B(x_0, \delta R)$, for some $\delta < 1/2$.
Then an easy covering argument (using (VD); see for example, Lemma 2.7 of [**3**]) gives $(CS(\beta, \theta))$ with $\delta = 1/2$.

By the same reasoning, suppose (V, C) satisfies (VD) and $(CS(\beta, \theta))$, except that each $B(x_0, 2s)$ on the right hand side of (5.4) is replaced by $B(x_0, \lambda s)$ for some $\lambda > 1$, then $(CS(\beta, \theta))$ holds with $\lambda = 2$. See Remark 3.2 in [**3**]. The same fact is true for $(WPI(\beta))$, which is defined in Definition 5.12.

THEOREM 5.7. ([**3**]) *Suppose (V, C) is an electrical network on a connected infinite graph satisfying the (p_0) condition. Then the following are equivalent:*
(1) There exists $\theta \in (0, 1]$ such that (V, C) satisfies (VD), $(PI(\beta))$ and $(CS(\beta, \theta))$.
(2) (V, C) satisfies $(PHI(\beta))$.

REMARK 5.8. Note that $(CS(2, 1))$ always holds. Indeed, essentially one can take $\varphi(x) = 2d(x, B(x_0, R)^c)/R$, then $|\varphi(x) - \varphi(y)| \leq 2/R$, if $C_{xy} > 0$, and (5.4) follows easily. Thus Theorem 5.7 is a nice extension of the characterization of $(PHI(2))$ due to [**12**] and [**31**] (for certain graphs, due to [**9**]) to the $\beta > 2$ case on graphs.

5.3. Stability under rough isometries. In this subsection, we will discuss the stability of the parabolic Harnack inequality under rough isometries.

DEFINITION 5.9. Let $(V^{(1)}, C^{(1)})$, $(V^{(2)}, C^{(2)})$ be electrical networks on connected infinite graphs satisfying the (p_0) condition.
(1) A map $T : V^{(1)} \to V^{(2)}$ is called a rough isometry if the following holds. There exist positive constants $a, c > 1, b > 0$ and $M > 0$ such that

$$a^{-1}d^{(1)}(x,y) - b \le d^{(2)}(T(x),T(y)) \le ad^{(1)}(x,y) + b \quad \forall x,y \in V^{(1)}, \tag{5.5}$$

$$d^{(2)}(T(V^{(1)}),y') \le M \quad \forall y' \in V^{(2)}, \tag{5.6}$$

$$c^{-1}\mu_x^{(1)} \le \mu_{T(x)}^{(2)} \le c\mu_x^{(1)} \quad \forall x \in V^{(1)}, \tag{5.7}$$

where $\mu^{(i)}$ and $d^{(i)}(\cdot,\cdot)$ are the measure and the graph distance of $(V^{(i)}, C^{(i)})$ respectively for $i = 1, 2$.
(2) $(V^{(1)}, C^{(1)})$, $(V^{(2)}, C^{(2)})$ are said to be rough isometric if there is a rough isometry between them.

The notion of rough isometry was first introduced by M. Kanai ([**21, 22**]). As this work was mainly concerned with Riemannian manifolds, a definition of rough isometry included only (5.5), (5.6). The definition equivalent to ours is given in [**8**].

We first give a lemma which lists some easy consequences of the definition.

LEMMA 5.10.
(1) There exists $K > 0$ *such that* $\sup_{x' \in T(V^{(1)})} \sharp\{x \in V^{(1)} : T(x) = x'\} \le K$.
(2) $\mu_x^{(i)} \asymp \mu_y^{(i)}$ *for all* $x \sim y \in V^{(i)}$.
(3) $\mu_x^{(1)} \asymp \mu_y^{(1)}$ *for all* $x, y \in V^{(1)}$ *such that* $T(x) = T(y)$.
(4) $\mu_{x'}^{(2)} \asymp \sum_{\substack{x \in V^{(1)} \\ T(x)=x'}} \mu_x^{(1)}$ *for all* $x' \in T(V^{(1)})$.
(5) $C_{xy}^{(1)} \asymp C_{T(x)y'}^{(2)}$ *for all* $x, y \in V^{(1)}$, $y' \in V^{(2)}$ *such that* $x \sim y$, $T(x) \sim y'$.
(6) There exists $h > 1$ *such that*

$$T(B^{(1)}(x,h^{-1}r)) \subset B^{(2)}(x',r) \subset \{M\text{-neighbourhood of } T(B^{(1)}(x,hr))\},$$

for all $r \ge 2(b+2M)$ *and all* $x' \in V^{(2)}$ *where* $x \in V^{(1)}$ *satisfies* $d^{(2)}(T(x),x') \le M$. *Here,* $B^{(i)}(x,r) := \{y \in V^{(i)} : d^{(i)}(x,y) < r\}$ *for* $x \in V^{(i)}$, $i = 1, 2$.
(7) Rough isometry is an equivalence relation.

PROOF. (1)-(6) can be deduced easily from the (p_0) condition and the definition of the rough isometry. For (7), let $(V^{(i)}, C^{(i)})$, $i = 1, 2, 3$, be electrical networks on connected infinite graphs satisfying the (p_0) condition. Clearly, by the identity map, $(V^{(1)}, C^{(1)})$ is rough isometric to itself. If $T_1 : V^{(1)} \to V^{(2)}$ and $T_2 : V^{(2)} \to V^{(3)}$ are rough isometries, then so is the composition $T_2 \circ T_1 : V^{(1)} \to V^{(3)}$. Finally, if $T : V^{(1)} \to V^{(2)}$ is a rough isometry, then we define $T^- : V^{(2)} \to V^{(1)}$ as follows; for each $x' \in V^{(2)}$, choose some $x \in V^{(1)}$ so that $d^{(2)}(T(x),x') \le M$ and put $T^-(x') = x$. T^- is called a rough inverse of T. It is easy to check that T^- is a rough isometry. Thus (7) holds. □

Our main result in this subsection is the stability of $PHI(\beta)$ under rough isometry.

THEOREM 5.11. *Let* $(V^{(1)}, C^{(1)})$, $(V^{(2)}, C^{(2)})$ *be electric networks of connected infinite graphs satisfying the* (p_0) *condition. If* $(V^{(1)}, C^{(1)})$ *satisfies* $(PHI(\beta))$ *with respect to the graph distance and* $(V^{(1)}, C^{(1)})$, $(V^{(2)}, C^{(2)})$ *are rough isometric, then* $(V^{(2)}, C^{(2)})$ *also satisfies* $(PHI(\beta))$ *with respect to the graph distance.*

For the proof we require a weak Poincaré inequality.

DEFINITION 5.12. We say (V,C) satisfies $(WPI(\beta))$, a scaled weak Poincaré inequality with parameter $\beta \geq 2$, if there exists a constants $c_{5.8} > 0$ such that for all $x_0 \in V,\ R \geq 1$ and $f : B(x_0, 2R) \to \mathbb{R}$,

$$\sum_{x \in B(x_0,R)} (f(x) - \bar{f}_B)^2 \mu_x \leq c_{5.8} R^\beta \sum_{x,y \in B(x_0,2R)} C_{xy}(f(x) - f(y))^2, \tag{5.8}$$

where $\bar{f}_B = \mu(B(x_0,R))^{-1} \sum_{y \in B(x_0,R)} f(y)\mu_y$.

REMARK 5.13. For $(WPI(\beta))$ to hold, it is enough to check (5.8) for all non-negative $f : B(x_0, 2R) \to \mathbb{R}_+$. (The same is true for $(PI(\beta))$ and $(CS(\beta,\theta))$.)

To show this, denote the left hand side of (5.8) as $\mathrm{Var}_{B(x_0,R)}(f)$ and the sum in the right hand side of (5.8) as $\mathcal{E}_{B(x_0,2R)}(f)$. Then, by a simple computation, we have

$$\mathrm{Var}_{B(x_0,R)}(f+\alpha) = \mathrm{Var}_{B(x_0,R)}(f),\ \mathcal{E}_{B(x_0,2R)}(f+\alpha) = \mathcal{E}_{B(x_0,2R)}(f),$$

for all constant functions $\alpha \in \mathbb{R}$. Thus, if we assume (5.8) for all non-negative functions, then it holds for all bounded functions. We then see that (5.8) holds for all $f : B(x_0, 2R) \to \mathbb{R}$, by the usual approximation using $f_n := (f \wedge n) \vee (-n)$. The proof for $(CS(\beta,\theta))$ is much simpler and we omit it.

In the definition of $(WPI(\beta))$, the sum on the right hand side is over all $x, y \in B(x_0, 2R)$ whereas in $(PI(\beta))$, it is over all $x, y \in B(x_0, R)$. Clearly, $(PI(\beta))$ implies $(WPI(\beta))$. The next proposition asserts that if the volume doubling condition holds, the converse is true.

PROPOSITION 5.14. *Let (V,C) be an electrical network on a connected infinite graph satisfying the (p_0) condition. If (V,C) satisfies (VD) and $(WPI(\beta))$ with respect to the graph distance, then it also satisfies $(PI(\beta))$.*

This result was first proved by Jerison in [**19**] for Euclidean vector fields satisfying Hörmander's condition. One can easily modify that proof line by line to obtain the same result for this situation. Since the argument is well-known to specialists, we skip the proof here. See also [**15**] for closely related results on geodesic metric spaces.

We now give the key proposition.

PROPOSITION 5.15. *Let $(V^{(1)}, C^{(1)})$, $(V^{(2)}, C^{(2)})$ be electrical networks on connected infinite graphs satisfying the (p_0) condition and assume that they are rough isometric.*
(1) If $(V^{(1)}, C^{(1)})$ satisfies (VD), then so does $(V^{(2)}, C^{(2)})$.
(2) If $(V^{(1)}, C^{(1)})$ satisfies (VD) and $(WPI(\beta))$, then so does $(V^{(2)}, C^{(2)})$.
(3) If $(V^{(1)}, C^{(1)})$ satisfies (VD) and $(CS(\beta,\theta))$, then so does $(V^{(2)}, C^{(2)})$.

PROOF. (1) is relatively easy. First, note that using Lemma 5.10 (2) and the (p_0) condition, $\mu^{(2)}$ satisfies (5.2) for small R. We thus consider the case when R is large. For $x' \in V^{(2)}$, take $x \in V^{(1)}$ so that $d^{(2)}(T(x), x') \leq M$. Then, by Lemma

5.10 (4) and (6), we have

$$\begin{aligned}\mu^{(2)}(B^{(2)}(x',R)) &\ge c_1\mu^{(2)}(T(B^{(1)}(x,h^{-1}R))) = c_1 \sum_{y'\in T(B^{(1)}(x,h^{-1}R))} \mu^{(2)}_{y'} \\ &\ge c_2 \sum_{y\in B^{(1)}(x,h^{-1}R)} \mu^{(1)}_y = c_2\mu^{(1)}(B^{(1)}(x,h^{-1}R)).\end{aligned}$$

By the same observation using Lemma 5.10 (2), (4) and (6), we have

$$\mu^{(2)}(B^{(2)}(x',2R)) \le c_3\mu^{(1)}(B^{(1)}(x,2h^{-1}R)).$$

Thus, using (VD) of $\mu^{(1)}$ iteratively, we obtain (VD) of $\mu^{(2)}$ for large R.

Before proving (2) and (3), we note that it is enough to prove them when $M=1$ (M is defined in (5.6)). To show this, for $2\le k\le M$, let $(V^{(2)},C^{(2)}_k)$ be graphs obtained from $(V^{(2)},C^{(2)})$ by adding an edge $\{x,y\}$ if it is possible to go from x to y in at most k-steps. The conductance of the edge $\{x,y\}$ is given by $(\mu^{(2)}_x+\mu^{(2)}_y)/2$ if $\{x,y\}\notin B^{(2)}$ and by $C^{(2)}_{xy}$ if $\{x,y\}\in B^{(2)}$. Then, clearly $(V^{(2)},C^{(2)}_k)$ satisfies the (p_0) condition (with a different p_0). The identity map $\mathrm{Id}_k : V^{(2)}\to V^{(2)}$ is a rough isometry from $(V^{(2)},C^{(2)}_k)$ to $(V^{(2)},C^{(2)}_{k-1})$ for each $2\le k\le M$ (here we let $(V^{(2)},C^{(2)}_1)=(V^{(2)},C^{(2)})$). Also, T can be regarded as a rough isometry from $(V^{(1)},C^{(1)})$ to $(V^{(2)},C^{(2)}_M)$ (we denote this map $\hat T$) and (5.6) holds with $M=1$. Then, we have $T=\mathrm{Id}_2\circ\cdots\circ\mathrm{Id}_M\circ\hat T$ as maps of networks. We thus consider the case $M=1$ in the following.

We now prove (2). As in Remark 5.13, we will prove (5.8) for all $g: V^{(2)}\to\mathbb{R}_+$. Let

$$\begin{aligned}\|g\|^2_{i,B(x_0,R)} &:= \left(\sum_{y\in B(x_0,R)} g(y)^i\mu_y\right)^{1/i} \quad \text{for } i=1,2,\\ \mathcal{E}_{B(x_0,R)}(g) &:= \sum_{x,y\in B(x_0,R)} C_{xy}(g(x)-g(y))^2.\end{aligned}$$

Take $R_0\ge 1$ large enough. We note that (5.3) and thus (5.8) holds for all $R\le R_0$ where $c_{5.8}$ could depend on R_0. Indeed, if we let $M_{x_0}:=\max\{|g(x')-g(y')| : x',y'\in B^{(2)}(x_0,R_0)\}$, then, as $\min_{x\in B} f(x)\le \bar f_B\le \max_{x\in B} f(x)$, we have

$$(\text{LHS of (5.3)}) \le 2M_{x_0}\sum_{x\in B^{(2)}(x_0,R)} \mu^{(2)}_x \le c_{R_0}\mathcal{E}_{B^{(2)}(x_0,R)}(g) \le (\text{RHS of (5.3)}),$$

so that (5.3), (5.8) hold. Thus it is enough to prove (5.8) for large $R\ge R_0$.

Since $g\circ T: V^{(1)}\to\mathbb{R}_+$, by $(WPI(\beta))$ for $(V^{(1)},C^{(1)})$, we have

$$\|g\circ T-\overline{g\circ T}_{B^{(1)}}\|^2_{2,B^{(1)}(x_0,R)} \le c_{5.8}R^\beta\mathcal{E}_{B^{(1)}(x_0,2R)}(g\circ T), \tag{5.9}$$

where $\overline{g\circ T}_{B^{(1)}} = \|g\circ T\|_{1,B^{(1)}(x_0,R)}/\mu^{(1)}(B^{(1)}(x_0,R))$. Let $x_0' := T(x_0)$. Then, by the (p_0) condition and Lemma 5.10 (2), (5), (6), we can easily obtain

$$\mathcal{E}_{B^{(1)}(x_0,2R)}(g\circ T) \le c_4\mathcal{E}_{B^{(2)}(x_0',2hR)}(g). \tag{5.10}$$

We will next show

$$\|g\|^2_{2,B^{(2)}(x_0',R/h)} \le c_5\|g\circ T\|^2_{2,B^{(1)}(x_0,R)} + c_6\mathcal{E}_{B^{(2)}(x_0',hR+1)}(g). \tag{5.11}$$

To prove this, first note that, by definition of a rough isometry, we have

$$\|g\circ T\|^2_{2,B^{(1)}(x_0,R)} \ge c_7 \sum_{x'\in T(B^{(1)}(x_0,R))} g(x')^2\mu^{(2)}_{x'}. \tag{5.12}$$

Let $A_{x_0,R} := \{\text{1-neighbourhood of } T(B^{(1)}(x_0,R))\}$ and let

$$\begin{aligned} G_1 &:= \{y'\in A_{x_0,R} : g(x')/2 \le g(y') \le 2g(x') \ \text{ for some } \ x'\in T(B^{(1)}(x_0,R)) \\ &\qquad \text{such that } \ x'\sim y'\} \cup T(B^{(1)}(x_0,R)), \\ G_2 &:= A_{x_0,R}\setminus G_1. \end{aligned}$$

If $y'\in G_2$, then an elementary computation shows

$$g(x')^2+g(y')^2 \le 5|g(x')-g(y')|^2, \tag{5.13}$$

where $x'\in T(B^{(1)}(x_0,R))$ and $x'\sim y'$. Using this, the (p_0) condition and Lemma 5.10 (2), we have $g(y')^2\mu^{(2)}_{y'} \le c_8 C^{(2)}_{x'y'}|g(x')-g(y')|^2$. We thus obtain

$$\sum_{y'\in G_2} g(y')^2\mu^{(2)}_{y'} \le c_9\mathcal{E}_{A_{x_0,R}}(g) \le c_9\mathcal{E}_{B^{(2)}(x_0',hR+1)}(g), \tag{5.14}$$

where we apply Lemma 5.10 (6) in the last inequality. Now, if $y'\in G_1$, then there exists $x'\in T(B^{(1)}(x_0,R))$, $x'\sim y'$ such that $g(y')\le 2g(x')$. Thus, using the (p_0) condition and Lemma 5.10 (2), we have

$$\begin{aligned} \sum_{y'\in G_1\setminus T(B^{(1)}(x_0,R))} g(y')^2\mu^{(2)}_{y'} &\le c_{10}\sum_{x'\in T(B^{(1)}(x_0,R))} g(x')^2\mu^{(2)}_{x'} \\ &\le c_{11}\|g\circ T\|^2_{2,B^{(1)}(x_0,R)}, \end{aligned} \tag{5.15}$$

where we use (5.12) in the last inequality. Combining (5.12), (5.14), (5.15) and using the fact $A_{x_0,R}\supset B^{(2)}(x_0',R/h)$ which is due to Lemma 5.10 (6), we obtain (5.11). Let $c_g := \overline{g\circ T}_{B^{(1)}}$. Now, substituting (5.10) and (5.11) in (5.9), where we use $g-c_g$ instead of g in (5.11), we obtain

$$\|g-c_g\|^2_{2,B^{(2)}(x_0',R/h)} \le c_{12}\mathcal{E}_{B^{(2)}(x_0',2hR)}(g). \tag{5.16}$$

Note that in general the minimum of $\|g-\alpha\|^2_{2,B^{(2)}(x_0',R/h)}$ (as a function of $\alpha\in\mathbb{R}$) is attained when $\alpha=\|g\|_{2,B^{(2)}(x_0',R/h)}/\mu^{(2)}(B^{(2)}(x_0',R/h))$. Thus, in view of Remark 5.6 (where we apply (VD)), we obtain (5.8) for $(V^{(2)},C^{(2)})$ when $x_0\in T(V^{(1)})$. For $x_0\notin T(V^{(1)})$, we can find $\hat{x}_0\in T(V^{(1)})$ such that $d^{(2)}(x_0,\hat{x}_0)\le 1$, so that the same result holds by changing h suitably. We have thus obtained $(WPI(\beta))$ for $(V^{(2)},C^{(2)})$.

We finally prove (3). Take $R_0\ge 1$ large enough. Note first that (5.4) holds for all $R\le R_0$ where $c_{5.10}$ could depend on R_0. Indeed, in view of Remark 5.8, we know $(CS(2,1))$ always holds. As in (5.4) the exponents β and θ are irrelevant for small R (as long as $c_{5.10}$ could depend on R_0), we see that (5.4) holds for all $R\le R_0$. Thus it is enough to prove (5.4) for large $R\ge R_0$.
Now, take $g: V^{(2)}\to\mathbb{R}$ arbitrary. Since $g\circ T: V^{(1)}\to\mathbb{R}$, by $(CS(\beta,\theta))$ for $(V^{(1)},C^{(1)})$, we have

$$\begin{aligned} &\sum_{x\in B^{(1)}(x_0,s)} g\circ T(x)^2 \sum_{y\in V^{(1)}} C^{(1)}_{xy}|\varphi(x)-\varphi(y)|^2 \\ &\le c_{5.10}\Big(\frac{s}{R}\Big)^{2\theta}\big(\mathcal{E}_{B^{(1)}(x_0,2s)}(g\circ T)+s^{-\beta}\|g\circ T\|^2_{2,B^{(1)}(x_0,2s)}\big), \end{aligned} \tag{5.17}$$

where φ is a cut-off function which satisfies (a)-(c) in Definition 5.5 (2). (We assume that (a) holds on $B(x_0, \delta R)$ instead of $B(x_0, R/2)$ where $\delta > 0$ is chosen sufficiently small, cf. Remark 5.6.) Let $x_0' = T(x_0)$. Then, by definition of a rough isometry and Lemma 5.10 (6), we have

$$\|g \circ T\|_{2,B^{(1)}(x_0,2s)} \le c\|g\|_{2,B^{(2)}(x_0',2hs)}. \tag{5.18}$$

We can use (5.10) to estimate the first term of the right hand side of (5.17). Let T^- be a rough inverse given in the proof of Lemma 5.10 and set $\tilde{\varphi} := \varphi \circ T^-$. Note that by definition of a rough isometry, $\tilde{\varphi}$ also satisfies (a)-(c) in Definition 5.5 (2) (with $\delta' R$ instead of $R/2$ in (a) and $\delta'' R$ instead of R in (b) where $0 < \delta' < \delta''$). We will prove

$$\begin{aligned} &\sum_{x' \in B^{(2)}(x_0', s/h)} g(x')^2 \sum_{y' \in V^{(2)}} C^{(2)}_{x'y'} |\tilde{\varphi}(x') - \tilde{\varphi}(y')|^2 \\ \le\ & c_{13}\Big(\sum_{x \in B^{(1)}(x_0,s)} g \circ T(x)^2 \sum_{y \in V^{(1)}} C^{(1)}_{xy} |\varphi(x) - \varphi(y)|^2 + \frac{1}{R^{2\theta}} \mathcal{E}_{B^{(2)}(x_0', c_{14}s)}(g)\Big), \end{aligned} \tag{5.19}$$

where $c_{13}, c_{14} > 1$ are some constants. Once we prove (5.19), by substituting (5.10), (5.18) and (5.19) in (5.17), we obtain (5.4) for $(V^{(2)}, C^{(2)})$ in view of Remark 5.6 (where we apply (VD)) when $x_0 \in T(V^{(1)})$. By the same observation as in the last part of the proof of (2), we have the same result for $x_0 \notin T(V^{(1)})$, and we obtain $(CS(\beta, \theta))$ for $(V^{(2)}, C^{(2)})$. Thus, all we need is to prove (5.19).

To prove (5.19), set $F(x', y') := g(x')^2 C^{(2)}_{x'y'} |\tilde{\varphi}(x') - \tilde{\varphi}(y')|^2$ for $x' \sim y'$. We first consider the case $x' = T(x)$ for some $x \in B^{(1)}(x_0, s)$. Denoting $\bar{x} := T^-(x')$ and $\bar{y} := T^-(y')$, we have $F(x', y') = g \circ T(x)^2 C^{(2)}_{x'y'} |\varphi(\bar{x}) - \varphi(\bar{y})|^2$. Note that, by definition of a rough isometry and a rough inverse, there exists M_0 (independent of $x' \in V^{(2)}$) so that $\bar{x}, \bar{y} \in B^{(1)}(x, M_0)$. Now take a chain $\{x_i\}_{i=0}^L$ so that $x_0 = \bar{x}, x_L = \bar{y}$, $x_i \sim x_{i+1}$ and $L \le 2M_0$. Using Lemma 5.10 (2) and (5), we have

$$F(x', y') \le c_{15} g \circ T(x)^2 \sum_{i=0}^{L-1} C^{(1)}_{x_i, x_{i+1}} |\varphi(x_i) - \varphi(x_{i+1})|^2. \tag{5.20}$$

Set $G_1^x := \{y \in V^{(1)} : g \circ T(y)/2 \le g \circ T(x) \le 2g \circ T(y)\}$. By using (5.13) and (c) of Definition 5.5 (2), we see that the right hand side of (5.20) is smaller than

$$\begin{aligned} c_{16}\Big(&\sum_{\substack{i=0,\cdots,L-1 \\ x_i \in G_1^x}} g \circ T(x_i)^2 C^{(1)}_{x_i, x_{i+1}} |\varphi(x_i) - \varphi(x_{i+1})|^2 \\ &+ \frac{1}{R^{2\theta}} \sum_{\substack{i=0,\cdots,L-1 \\ x_i \notin G_1^x}} |g \circ T(x) - g \circ T(x_i)|^2 \mu^{(1)}_{x_i}\Big). \end{aligned} \tag{5.21}$$

We make another chain $\{y_j\}_{j=0}^{L'}$ from x to x_i (note $L' \le M_0$), so that

$$|g \circ T(x) - g \circ T(x_i)|^2 \mu^{(1)}_{x_i} \le c_{17} \sum_{j=0}^{L'-1} |g \circ T(y_j) - g \circ T(y_{j+1})|^2 \mu^{(1)}_{y_j}. \tag{5.22}$$

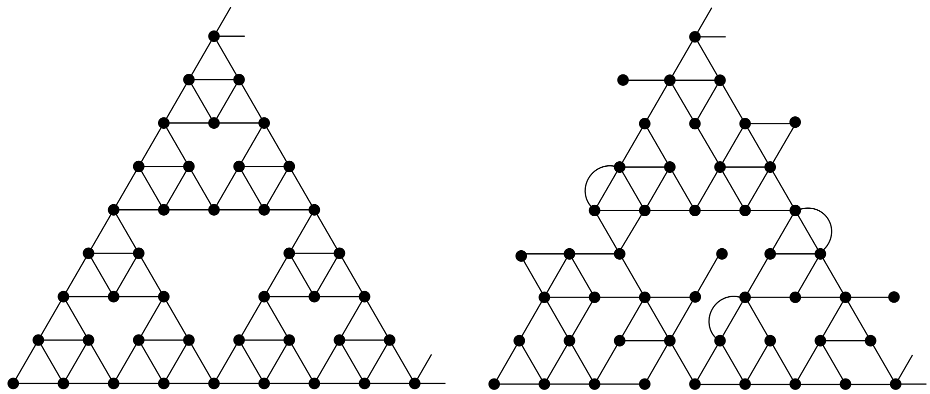

FIGURE 3. S.G. graph and modified S.G. graph

Combining (5.20), (5.21) and (5.22), for each x' we have

$$(5.23)\quad \sum_{y'\sim x'} F(x',y')$$
$$\le c_{18}\Big(\sum_{y\in B^{(1)}(x,M_0)} g\circ T(y)^2\Big(\sum_{z\sim y} C^{(1)}_{yz}|\varphi(y)-\varphi(z)|^2\Big)+\frac{1}{R^{2\theta}}\mathcal{E}_{B^{(1)}(x,M_0)}(g\circ T)\Big)$$
$$\le c_{19}\Big(\sum_{y\in B^{(1)}(x,M_0)} g\circ T(y)^2\Big(\sum_{z\sim y} C^{(1)}_{yz}|\varphi(y)-\varphi(z)|^2\Big)+\frac{1}{R^{2\theta}}\mathcal{E}_{B^{(2)}(x',hM_0)}(g)\Big),$$

where we apply (5.10) in the last inequality.

We next consider the case $x'\in B^{(2)}(x_0',s/h)\setminus T(B^{(1)}(x_0,s))$. In view of Lemma 5.10 (6), there exists $z\in B^{(1)}(x_0,s)$ so that $z':=T(z)\sim x'$. If $g(z')/2\le g(x')\le 2g(z')$, then by changing $g(x')$ as $g(z')$ we can apply the same argument as above and obtain (5.23). Otherwise, using (5.13) and (c) of Definition 5.5 (2) again, for each x' we have

$$\sum_{y'\sim x'} F(x',y')\le \frac{c_{20}}{R^{2\theta}}\mathcal{E}_{B^{(2)}(x',1)}(g),$$

and obtain (5.23). Altogether, we have proved (5.23) for all $x'\in B^{(2)}(x_0',s/h)$. We note again that M_0 is independent of the choice of x', s and R. Then, summing (5.23) for $x'\in B^{(2)}(x_0',s/h)$, using the (p_0) condition and Lemma 5.10 (6), we obtain (5.19). □

Now the proof of Theorem 5.11 is immediate by combining the results of Theorem 5.7, Proposition 5.14 and Proposition 5.15.

For the rest of the paper, we consider (V,C) to be an electrical network on a u.f.r. graph which satisfies (2.7). In particular we consider nested fractal graphs as Assumption 2.2 holds (see [**27**]). If the conductance on the basic electrical network (V,C) is derived from the non-degenerate fixed point, we have Corollary 4.13, that is the network satisfies two sided sub-Gaussian estimates with respect to the graph distance. Thus Theorem 5.3 and Theorem 5.7 can be applied and we see that the same results hold if we change the conductance (even randomly) as long as it satisfies (2.2). We note that related results on homogenization are discussed in [**26**].

Finally we consider a graph which is rough isometric to a nested fractal graph. The graph on the right of Figure 3 is made by locally modifying the 2-dimensional Sierpinski gasket graph so that (5.5) and (5.6) are satisfied. Note that the global gasket structure is preserved by the local modification of the graph. The two networks in Figure 3 are rough isometric if the conductance on each edge is bounded from above and below by positive constants. Thus, using Theorem 5.11, the network on the right of Figure 3 also satisfies two sided sub-Gaussian estimates and thus satisfies $(PHI(\log 5/\log 2))$ with respect to the graph distance.

ACKNOWLEDGEMENTS. The second author is grateful to Professor A. Grigor'yan for motivating him to work on this problem. He also thanks Professor M.T. Barlow and Professor R.F. Bass for helpful comments concerning Section 5.

References

[1] M.T. Barlow, *Which values of the volume growth and escape time exponents are possible for a graph?* In: Proc. of Conference held in Graz 2001, Birkhäuser, to appear.

[2] 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.

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

[4] 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.

[5] M.T. Barlow and R.F. Bass, *Transition density estimates for Brownian motion on the Sierpinski carpet,* Probab. Theory Relat. Fields, **91** (1992), 307–330.

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

[7] T. Coulhon, *Ultracontractivity and Nash type inequalities,* J. Func. Anal., **141** (1996), 510–539.

[8] T. Coulhon and L. Saloff-Coste, *Variétés riemanniennes isométriques à l'infini,* Rev. Mat. Iberoamericana, **11** (1995), 687–726.

[9] T. Delmotte, *Parabolic Harnack inequality and estimates of Markov chains on graphs,* Rev. Mat. Iberoamericana, **15** (1999), 181–232.

[10] E.B. Fabes and D.W. Stroock, *A new proof of Moser's parabolic Harnack inequality via the old idea of Nash,* Arch. Rational Mech. Anal., **96** (1986), 327–338.

[11] 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.

[12] A. Grigoryan, *The heat equation on non-compact Riemannian manifolds,* (in Russian) Matem. Sbornik., **182** (1991), 55–87. (English transl.) Math. USSR Sb., **72** (1992), 47–77.

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

[14] A. Grigor'yan and A. Telcs, *Harnack inequalities and sub-Gaussian estimates for random walks,* Math. Ann., to appear.

[15] P. Hajłasz and P. Koskela, *Sobolev meets Poincaré,* C. R. Acad. Sci. Paris Sér. I Math., **320** (1995), 1211–1215.

[16] B.M. Hambly and T. Kumagai, *Diffusion processes on fractal fields: heat kernel estimates and large deviations,* Preprint 2002.

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

[18] W. Hebisch and L. Saloff-Coste, *On the relation between elliptic and parabolic Harnack inequalities,* Ann. Inst. Fourier (Grenoble), **51** (2001), 1437–1481.

[19] D. Jerison, *The Poincaré inequality for vector fields satisfying Hörmander's condition,* Duke Math. J., **53** (1986), 503–523.

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

[21] M. Kanai, *Rough isometries, and combinatorial approximations of geometries of non-compact riemannian manifolds,* J. Math. Soc. Japan, **37** (1985), 391–413.
[22] M. Kanai, *Analytic inequalities, and rough isometries between non-compact riemannian manifolds,* Lect. Notes Math. **1201**, Springer, 1986, 122–137.
[23] J. Kigami, *Analysis on Fractals,* Cambridge Univ. Press, Cambridge, 2001.
[24] J. Kigami, *Harmonic calculus on P.C.F. self-similar sets,* Trans. Amer. Math. Soc., **335** (1993), 721–755.
[25] T. Kumagai, *Estimates of transition densities for Brownian motion on nested fractals,* Probab. Theory Relat. Fields, **96** (1993), 205–224.
[26] T. Kumagai and S. Kusuoka, *Homogenization on nested fractals,* Probab. Theory Relat. Fields, **104** (1996), 375–398.
[27] T. Lindstrøm, *Brownian motion on nested fractals,* Memoirs Amer. Math. Soc. 420 **83**, 1990.
[28] V. Metz, *Renormalization contracts on nested fractals,* J. Reine Angew. Math., **480** (1996), 161–175.
[29] R. Peirone, *Convergence and uniqueness problems for Dirichlet forms on fractals,* Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8), **3** (2000), 431–460.
[30] C. Sabot, *Existence and uniqueness of diffusions on finitely ramified self-similar fractals,* Ann. Sci. École Norm. Sup. (4), **30** (1997), 605–673.
[31] L. Saloff-Coste, *A note on Poincaré, Sobolev and Harnack inequality,* Internat. Math. Res. Notices, **2** (1992), 27–38.
[32] W. Woess, *Random Walks on Infinite Graphs and Groups,* Cambridge University Press, Cambridge, 2000.

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