Diffusion processes on fractal fields: heat kernel estimates and large deviations

B.M. Hambly* and T. Kumagai†

Dedicated to Professor Tokuzo Shiga on his 60th birthday.

May 2, 2003

Abstract

A fractal field is a collection of fractals with, in general, different Hausdorff dimensions, embedded in $\mathbb{R}^2$. We will construct diffusion processes on such fields which behave as Brownian motion in $\mathbb{R}^2$ outside the fractals and as the appropriate fractal diffusion within each fractal component of the field. We will discuss the properties of the diffusion process in the case where the fractal components tile $\mathbb{R}^2$. By working in a suitable shortest path metric we will establish heat kernel bounds and large deviation estimates which determine the trajectories followed by the diffusion over short times.

1 Introduction

Let $\{K_i\}_{i=1}^M$ $(1 \le M \le \infty)$ be a countable family of sets containing different disordered media on $\mathbb{R}^d$. Is it possible to construct a diffusion process which moves over the whole of $\mathbb{R}^d$, whose behaviour is like that of the (in general singular) Brownian motion inside each of the different media K_i, and also like that of Brownian motion on $\mathbb{R}^d$ outside? If so, how does such a diffusion behave? In particular, what is asymptotically the "most probable path" for the diffusion in the short time limit? In this paper we will give some answers to these questions when the sets K_i are fractals.

The initial interest in this area came from the study of the range of behaviour for diffusion processes in more than one dimension. In [10] a class of diffusions was constructed on $\mathbb{R}^d$ with singular symmetrizing measure and generators that were not necessarily differential operators. It was shown by [24] that some of these processes could be regarded as superpositions of Dirichlet forms and this has led to work on understanding the semigroups and Markov processes which correspond to superpositions of Dirichlet forms, (for recent results see [22] and the references therein). All this work has been done for the case where the sets K_i are Euclidean subspaces or more generally submanifolds in $\mathbb{R}^d$. Note that, even in this setting, little is known about the detailed properties of the corresponding processes.

In the fractal context this type of problem was first considered by Lindstrøm [21], in the case of the Sierpinski gasket embedded in $\mathbb{R}^2$. He was able to prove an existence theorem for

*Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK.
E-mail: hambly@maths.ox.ac.uk
†Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan.
E-mail: kumagai@kurims.kyoto-u.ac.jp

a process which behaves like a diffusion on the Sierpinski gasket within the unit triangle, yet standard Brownian motion in $\mathbb{R}^2$ outside. This problem was then viewed as the superposition of two Dirichlet forms, which led to constructions of 'penetrating' diffusions for other fractals embedded in $\mathbb{R}^d$, [17], [12].

We wish to consider the more general situation where there may be many (non-overlapping) fractals embedded in $\mathbb{R}^d$, with different Hausdorff dimensions and which may touch each other or not. Our work will extend the existence theorem to the case of many different nested fractals, which may be embedded in $\mathbb{R}^2$ with gaps between them, but also may tile the space. We will call spaces of either type fractal fields and two examples are shown in Figures 1 and 2.

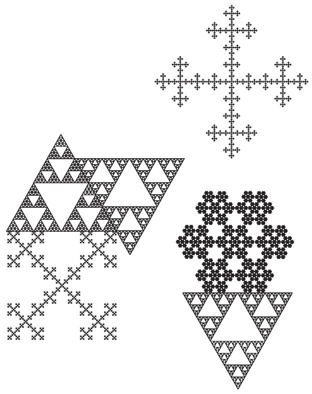

Figure 1: A part of a fractal field

A key family of examples that we would like readers to bear in mind throughout the paper is what we will call the gasket tiling in $\mathbb{R}^2$. Consider the triangular lattice on $\mathbb{R}^2$ where each edge is of length 1. We will fill each triangle with a version of the Sierpinski gasket in a periodic way. More precisely, let $SG(l)$ be the family of 2-dimensional Sierpinski gaskets from [7] with sidelength 1, each constructed from a set of contraction maps with contraction factor $1/l$. Now, take a set of triangles (we let L be the number of triangles in the set) from the triangular lattice so that their union is a connected closed set. In each triangle we place $\{SG(l_k)\}_{k=1}^L$ and denote the union of these fractals by G_0. Without loss of generality, we can assume that one of the vertices of G_0 is $(0,0)$. We take $i_x \in \mathbb{N}$ so that $G_0 \cap (G_0 + (i_x, 0)) \neq \emptyset$ and Int $G_0 \cap$ Int $(G_0 + (i_x, 0)) = \emptyset$. We also take $i_y \in \mathbb{N}$ in the same way by taking $(0, i_y)$ instead of $(i_x, 0)$. Then, $G \equiv \cup_{l,m\in\mathbb{Z}}(G_0 + (li_x, mi_y))$ is the space we will consider. Figure 2 illustrates the case when G_0 is a parallelogram filled with SG(2) and SG(4).

In this particular example, as the boundaries of the constituent fractals fit together in a natural way, the construction of the Dirichlet form is quite a straightforward extension of the approach of [14] as a limit of a sequence of Dirichlet forms on the natural approximating lattices. However, in general we will not have such a simple boundary structure between the different components and the construction problem requires more work.

Throughout we will only consider our field to consist of nested fractals in $\mathbb{R}^2$ (The definition and key properties of nested fractals are given in an Appendix). Indeed we will assume that

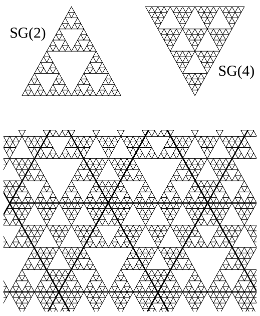

Figure 2: A part of a fractal field made from SG(2) and SG(4)

the number of different nested fractals used in the construction of the field is finite. We restrict ourselves to nested fractals because we do not know how to prove the regularity of the Dirichlet form when two (or more) fractals intersect at a particular boundary in general (see Remark 2.7 (1)). The restriction to $\mathbb{R}^2$ is because Assumption 2.1 (3) and Assumption 2.2 seldom hold for fractals embedded in higher dimensions (see Remark 2.3 (4)). For the construction problem we work in the more general setting where our diffusion may move through $\mathbb{R}^2$ as well as the fractals (as in Figure 1). For the study of heat kernels and large deviations we restrict ourselves further to the case of fractal tilings (as in Figure 2). Despite these restrictions we still have a rich class of examples.

We now give an outline of the ingredients of this paper. Our initial aim will be to consider the general construction problem. Here we incorporate the ideas of Kumagai [17], for the embedding into a Euclidean space, with an idea originally due to Kusuoka, [18] which shows how to extend a Lipschitz function from the boundary of a fractal to the interior while controlling its energy. This will allow us to build up a Dirichlet form and establish various properties, such as a Nash inequality and the existence of a resistance metric in the tiling case.

Once we have the existence of a regular local Dirichlet form and hence a continuous strong Markov process on the field we turn to establishing some heat kernel estimates and determining the short time asymptotics for the process when the fractals tile $\mathbb{R}^2$. The heat kernel on fractals has been considered in a number of papers, see [1], [2], [3], [8], [15] and it is typically of a sub-Gaussian form in that for short times there are constants c, c' such that for all x, y in the fractal,

$$p_t(x,y) \le ct^{-d_s/2} \exp\left(-c' \left(\frac{d(x,y)^{d_w}}{t}\right)^{1/(d_w-1)}\right), \tag{1.1}$$

with a corresponding lower bound (with different constants) and where d_s, d_w are the spectral and walk dimensions respectively. The metric $d(.,.)$ is an intrinsic shortest path metric considered for instance in [15].

In the fractal tiling setting we are able to build a family of shortest path metrics $d^{(j)}(.,.)$, based on counting shortest paths in a resistance metric, each of which lives on $G^{(j)}$, the component of the field of type j. The resistance approximation is used as the Harnack inequality holds with respect to this metric, and though we do not explicitly state such an inequality, we require this type of control to get at the off-diagonal heat kernel. In the case where the fractals do not tile the space we cannot construct a resistance metric and for this reason we restrict our attention to the tiling case. We will extend the usual probabilistic techniques for heat kernel estimation on fractals to this setting. It is on the boundaries between components that control is difficult and, as a result, we are not able to obtain uniform upper and lower estimates of the same functional form as in (1.1).

We state the heat kernel estimates for our gasket tiling example to show what we are able to deduce. Let $l_1 < l_2 < \ldots < l_{M'}$ where $M' < \infty$ denote the different fractal types and write $d_w^{(l_i)}, d^{(l_i)}(.,.)$ for the walk dimension and metric on $G^{(l_i)}$ respectively. Also, for each $x \in G$, let $d_s(x)$ be the local spectral dimension at $x \in G$ whose definition will be given at the beginning of Section 4.

Theorem 1.1 *For the gasket tiling, there are constants $c_{1.1}, c_{1.2}, c_{1.3}, c_{1.4}$ so that for each $x, y \in G$ there are constants $t_0(x), t_0(y) << 1$ such that for $0 < t < t_0(x) \wedge t_0(y)$,*

$$p_t(x,y) \le c_{1.1} t^{-(d_s(x) \vee d_s(y))/2} \Phi(d^{(l_1)}(x,y), \ldots, d^{(l_{M'})}(x,y), c_{1.2}t),$$

and there are further constants $\rho(x,y), \eta_{l_i}(x,y) \ge 0, \theta(x,y) \in \mathbb{R}$ such that for $0 < t < 1$,

$$p_t(x,y) \ge c_{1.3} t^{\theta(x,y)} \Xi(x,y,t) \Phi(d^{(l_1)}(x,y), \ldots, d^{(l_{M'})}(x,y), c_{1.4}t),$$

where

$$\begin{aligned} \Phi(u_1, \ldots, u_{M'}, t) &= \exp\Big(-\sum_{i=1}^{M'} \Big(\frac{u_i^{d_w^{(l_i)}}}{t}\Big)^{1/(d_w^{(l_i)}-1)}\Big), \\ \Xi(x,y,t) &= \prod_{i=1}^{M'} d^{(l_i)}(x,y)^{-\eta_{l_i}(x,y)} e^{-\rho(x,y) \sum_{i=1}^{M'} (\log(\frac{d^{(l_i)}(x,y)}{t}))^2}. \end{aligned}$$

Note that if $x = y$ in the lower bound we can take $\rho = \eta_{l_i} = 0$, so that $\Xi(x,y,t) = 1$ and $\theta(x,x) = -\min_i d_s^{(l_i)}/2$. The correction factor for the lower bound $\Xi(x,y,t)$ arises from the technique of proof and has the property that $\Xi(x,y,t) \to 0$ as $\hat{d}(x,y) \to \infty$ or $t \to 0$. However the rate is sufficiently slow that, from the exponential term, we are able to obtain a Schilder type large deviation estimate in terms of the metric $d^{(1)}$ and demonstrate the shape of the shortest paths through our fractal tiling. Our main observation is that it is the component fractals with smallest d_w which take the longest time to cross (in the short time limit) and which determine the shortest paths followed by the diffusion. We cannot obtain precise large deviation asymptotics as there are small oscillations which do not allow the existence of the usual large deviation principle, even in the case of the Sierpinski gasket [5].

In particular we will see that in the gasket tiling case consisting of $l = 2, l = 4$, as $d_w^{(4)} > d_w^{(2)}$ the asymptotic shortest paths for our diffusion through this tiling avoid SG(2) and move through SG(4). By [9], for SG(l) we have $d_w^{(l)} \to 2$ as $l \to \infty$ and hence by choosing a large enough l we have that the asymptotic shortest paths now prefer to move through SG(2) and avoid SG(l), as shown in Figure 3 (where $l = 16$). We note that the shortest paths which achieve the infimum in the two fields shown in Figure 3 are not unique.

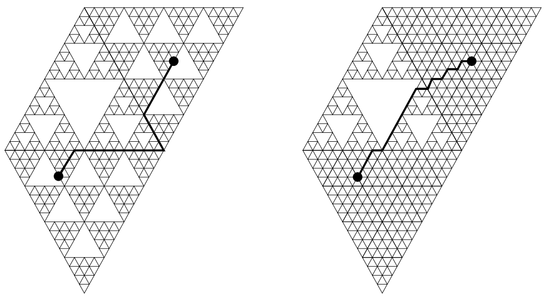

Figure 3: Shortest paths across two fractal tilings

2 Fractal fields and their Dirichlet forms

In this section we will introduce the framework of fractal fields. Our initial aim is to construct local regular Dirichlet forms on these spaces.

2.1 Framework and construction of Dirichlet forms

Let $\{K_i\}_{i=1}^M \subset \mathbb{R}^2$ $(1 \le M \le \infty)$ be a family of (bounded or unbounded) nested fractals (the definition can be found in the Appendix). When K_i is unbounded, we denote by $\hat{K}_i$ the corresponding bounded nested fractal (when K_i is bounded, $\hat{K}_i = K_i$) and denote by $\{\Psi_j^{(i)}\}_{j\in S_i}$ the family of contractions which determine $\hat{K}_i$ $(S_i = \{1, 2, \cdots, N_i\})$. Let $V_0^{(i)}$ be the set of essential fixed points for $\hat{K}_i$.

For each closed set A, let Cov (A) be the set of points covered by A. That is, if we decompose $\mathbb{R}^2 \backslash A$ into connected components $\{D_j\}_{j=1}^\infty$ and denote by $\{D_j\}_{j\in U(A)}$ the unbounded components, Cov $(A) = \mathbb{R}^2 \setminus \cup_{j\in U(A)} D_j$. We note that if the set A has holes, these are contained in Cov(A).

Define $G = \cup_{i=1}^M K_i$ and $D = \mathbb{R}^2 \setminus \mathrm{Cov}(G)$. Clearly, $D = \cup_{j\in U(G)} D_j$. We make the following assumption on the family of fractals $\{K_i\}_i$.

Assumption 2.1 *(1) For each* $1 \le i \ne j \le M$,

$$\mathit{Int}\ (\mathit{Cov}\ (K_i)) \cap \mathit{Int}\ (\mathit{Cov}\ (K_j)) = \emptyset,$$

where, for a set A, Int (A) is the interior of A, that is Int$(A) = Cl(A)\backslash(Cl(A) \cap Cl(\mathbb{R}^2 \setminus A))$.
(2) For each compact set $C \subset \mathbb{R}^2$, $|\{i : C \cap K_i \ne \emptyset\}| < \infty$, where for a set A we denote by $|A|$ the number of elements in A.
(3) $\{x \in \mathbb{R}^2 : x \in K_i \cap K_j \cap Cl(D)$ for some $1 \le i \ne j \le M\}$ is a discrete set.

We will make some comments on this assumption in Remark 2.7. Note that G is a closed set by Assumption 2.1 (2). We define $\tilde{G} = G \cup D$ and call it a *fractal field* generated by $\{K_i\}_{i=1}^M$. (When G is connected as in the introduction, we also call G a *fractal field* or a *fractal tiling.*) Note that we can define fractal fields on $\mathbb{R}^d$ in the same way using nested fractals on $\mathbb{R}^d$, but as our Assumptions 2.1 (3) and 2.2, which will be introduced later, seldom hold for nested fractals in $\mathbb{R}^d$ for $d \ge 3$, we will restrict our attention to the case $d = 2$.

Let $\partial_{ex} G$ be the topological boundary of $\mathrm{Cov}(G)$ as a subset of $\mathbb{R}^2$. For $1 \le i \ne j \le M$, let

$$\Gamma_{ij} = \mathrm{Cov}\ (K_i) \cap \mathrm{Cov}\ (K_j), \qquad \partial_{in} G = \cup_{1\le i\ne j\le M} \Gamma_{ij}. \tag{2.1}$$

Set $\partial G = \partial_{ex} G \cup \partial_{in} G$. Let μ_i be a Hausdorff measure on K_i, and set $\mu = \sum_{i=1}^M \mu_i$, $\tilde{\mu} = m|_D + \mu$ where m is the Lebesgue measure on $\mathbb{R}^2$.

Our next task is to define a Dirichlet form on $\tilde{G}$. First, for each i, the local regular Dirichlet form $(\mathcal{E}_{K_i}, \mathcal{F}_{K_i})$ on $\mathbb{L}^2(K_i, \mu_i)$ is given as in Theorem A.2 and Theorem A.5 in the Appendix. We denote by $d_f(K_i), d_s(K_i)$, $d_w(K_i), d_c(K_i)$ the Hausdorff, spectral, walk dimension and chemical exponent respectively with respect to the shortest path metric introduced in the Appendix.

Before we define a Dirichlet form on the fractal field we need a technical assumption which allows us to prove regularity. Let $K \subset K_i$ be a compact nested fractal which is congruent to $\hat{K}_i$ (thus, when K_i is bounded, $K = K_i$). For each Γ_{ij} in (2.1) where $1 \le i \ne j \le M$ and for $\omega \in \Sigma_i \equiv (S_i)^{\mathbb{N}}$, let $d_{\Gamma_{ij},K}(\omega) = \min\{n \ge 1 : \Gamma_{ij} \cap \Psi^{(K)}_{\omega_1\cdots\omega_n}(K) = \emptyset\}$ where $\{\Psi^{(K)}_j\}_{j\in S_i}$ is the family of α_i-contractions which determine K, and define

$$\begin{aligned} \kappa(\Gamma_{ij}, K) &= -\frac{1}{\log N_i} \limsup_{n\to\infty} \frac{1}{n} \log \nu_i(d_{\Gamma_{ij},K}(\omega) > n), \\ \kappa(\Gamma_{ij}, K_i) &= \inf_K \kappa(\Gamma_{ij}, K). \end{aligned}$$

Here ν_i is the Bernoulli measure on Σ_i such that $\nu_i(\{\omega \in \Sigma_i : \omega_1 = l\}) = 1/N_i$ for each $l \in S_i$ and the infimum in the second definition is taken over all compact sets $K \subset K_i$ which are congruent to $\hat{K}_i$. We adopt the convention that $-\log 0 = \infty$.

Assumption 2.2 *For each $1 \le i \le M$ and each $j \ne i$, the following holds,*

$$\frac{2}{d_s(K_i)} - \frac{2}{d_f(K_i) d_c(K_i)} < \kappa(\Gamma_{ij}, K_i). \tag{2.2}$$

Remark 2.3 (1) As mentioned in the Appendix, $d_{f,E}(K) = d_f(K) d_c(K)$ is the Hausdorff dimension of K with respect to the Euclidean metric.
(2) When $\Gamma_{ij} = \emptyset$, then $\kappa(\Gamma_{ij}, K_i) = \infty$ and (2.2) always holds. Further, when $|\Gamma_{ij}| < \infty$, (in that K_i and K_j intersect at only a finite number of points) and, writing ρ_i for the conductance scale factor of K_i, $\rho_i < \alpha_i^2$, (2.2) always holds. Indeed, in such a case we see that for any $K \subset K_i$ which is congruent to $\hat{K}_i$, $\nu_i(d_{\Gamma_{ij},K}(\omega) > n) \le c_1 N_i^{-n}$ for large n so that $\kappa(\Gamma_{ij}, K_i) \ge 1$. Using the fact that $d_f(K_i) d_c(K_i) = \log N_i / \log \alpha_i$ and $d_s(K_i) = 2 \log N_i / \log(\rho_i N_i)$, we see that

$$\frac{2}{d_s(K_i)} - \frac{2}{d_f(K_i) d_c(K_i)} = 1 + \frac{\log(\rho_i/\alpha_i^2)}{\log N_i} < 1,$$

where we use $\rho_i < \alpha_i^2$ in the last inequality, so that (2.2) holds. Note further that $\rho_i \le \alpha_i < \alpha_i^2$ always holds when $d_c(K_i) = 1$, as $d_w(K_i) \le d_f(K_i) + 1$ (see Theorem 3.20 in [1]; the exponents are with respect to the shortest path metric).
(3) For the gasket tiling introduced in the Introduction, (2.2) always holds. Indeed, let $K = SG(l)$ $l \ge 2$ and $\Gamma = \Gamma_{ij}$ be the bottom line of K. As there are l^n n-cells which intersect with Γ, we see that $\nu(d_{\Gamma,K}(\omega) > n) = l^n / N^n$ where $N = l(l+1)/2$. Thus, $\kappa(\Gamma, K) = 1 - \log l / \log N$ and (2.2) is equivalent to

$$\frac{\log(\rho N) - 2 \log l}{\log N} < 1 - \frac{\log l}{\log N},$$

which is equivalent to $\rho < l$. The last inequality can easily be checked, either by using $d_w(K) \leq d_f(K) + 1$ as mentioned above (equality holds only in the case where the fractal is a tree), or by simple electrical network arguments.

(4) For fractals embedded in $\mathbb{R}^d$ ($d \geq 3$), it seems that (2.2) is a very restrictive condition. In fact, even for the 3-dimensional Sierpinski gasket with the contraction rate $1/2$, (2.2) does not hold when $\Gamma = \Gamma_{ij}$ is the bottom triangle of the gasket. Indeed, by similar computations to the above, we have

$$\frac{2}{d_s(K)} - \frac{2}{d_f(K)d_c(K)} = \frac{\log 3}{2\log 2} - \frac{1}{2}$$

whereas

$$\kappa(\Gamma, K) = 1 - \frac{\log 3}{2\log 2} < \frac{\log 3}{2\log 2} - \frac{1}{2}.$$

As mentioned in the introduction, this is part of the reason why we restrict our framework to 2-dimensions.

Assumption 2.1 and Assumption 2.2 will be in force throughout the paper. We define a bilinear form $(\tilde{\mathcal{E}}, \mathcal{D}(\tilde{\mathcal{E}}))$ on $\mathbb{L}^2(\tilde{G}, \tilde{\mu})$ as follows,

$$\begin{aligned} \tilde{\mathcal{E}}(u,v) &= \sum_{i=1}^{M} \mathcal{E}_{K_i}(u|_{K_i}, v|_{K_i}) + \frac{1}{2}\sum_{j \in U(G)} \int_{D_j} \nabla u(x) \nabla v(x) dx \quad \text{for all } u, v \in \mathcal{D}(\tilde{\mathcal{E}}), \\ \mathcal{D}(\tilde{\mathcal{E}}) &= \{u \in C_0(\tilde{G}) : u|_{K_i} \in \mathcal{F}_{K_i}\ \forall\, i,\ \ u|_{D_j} \in W^{1,2}(D_j)\ \forall\, j,\ \ \tilde{\mathcal{E}}(u,u) < \infty\}, \end{aligned}$$

where $D = \cup_{j \in U(G)} D_j$ is a decomposition of D into open connected components, $C_0(\tilde{G})$ is the space of continuous functions on $\tilde{G}$ with compact support and $W^{1,2}(D_j)$ is the Sobolev space of functions on D_j with first derivatives in $\mathbb{L}^2(D_j, m|_{D_j})$. Then, it is easy to check the following.

Lemma 2.4 *(1) $(\tilde{\mathcal{E}}, \mathcal{D}(\tilde{\mathcal{E}}))$ is closable in $\mathbb{L}^2(\tilde{G}, \tilde{\mu})$.*
(2) $\mathcal{D}(\tilde{\mathcal{E}})$ is an algebra.
(3) For each j, $x \in K_j$ and each $U(x)$ which is a neighbourhood of x, there exists $f \in \mathcal{F}_{K_j} \cap C_0(K_j)$ such that $f(x) > 0$ and Supp $f \subset U(x) \cap K_j$, where Supp f denotes the support of f.
(4) $C_0^\infty(D) \subset \mathcal{D}(\tilde{\mathcal{E}})$.

Now, let $\tilde{\mathcal{F}} = \overline{\mathcal{D}(\tilde{\mathcal{E}})}^{\tilde{\mathcal{E}}_{(1)}}$ so that $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$ is the smallest extension of $(\tilde{\mathcal{E}}, \mathcal{D}(\tilde{\mathcal{E}}))$, where $\tilde{\mathcal{E}}_{(1)}(f,f) = \tilde{\mathcal{E}}(f,f) + \|f\|^2_{\mathbb{L}^2(\tilde{G},\tilde{\mu})}$. We then have the following.

Theorem 2.5 *$(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$ is a local regular Dirichlet form on $\mathbb{L}^2(\tilde{G}, \tilde{\mu})$.*

By the general theory of Dirichlet forms ([6]), there is a one to one correspondence between a local regular Dirichlet form on $\mathbb{L}^2(\tilde{G}, \tilde{\mu})$ and a $\tilde{\mu}$-symmetric diffusion process on $\tilde{G}$ up to an exceptional set of starting points. We will denote by $\{\tilde{X}_t\}_{t \geq 0}$ the diffusion process corresponding to $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$. Note that, as the original forms on $\{K_i\}_i$ and $\{D_j\}_j$ are strong local, $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$ is also strong local.

For the proof of Theorem 2.5, the key part is to prove the following.

Proposition 2.6 *(1) For each $x \neq y \in \tilde{G}$, there exists $g \in \mathcal{D}(\tilde{\mathcal{E}})$ such that $g(x) \neq g(y)$.*
(2) For any compact set L in $\tilde{G}$, there exists $f \in \mathcal{D}(\tilde{\mathcal{E}})$ such that $f|_L = 1$.

Indeed, using this proposition, we can prove Theorem 2.5 as follows. It is easy to see that $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$ is a local Dirichlet form. Also, as $\tilde{\mathcal{F}} = \overline{\mathcal{D}(\tilde{\mathcal{E}})}^{\tilde{\mathcal{E}}_{(1)}}$, it is clear that $\mathcal{D}(\tilde{\mathcal{E}})$ is dense in $\tilde{\mathcal{F}}$ with respect to the $\tilde{\mathcal{E}}_{(1)}$-norm. Thus, all we need for the regularity of the form is to show that $\mathcal{D}(\tilde{\mathcal{E}})$ is dense in $C_0(\tilde{G})$ with respect to the $||\cdot||_\infty$-norm. Now, as $\mathcal{D}(\tilde{\mathcal{E}})$ is an algebra (Lemma 2.4 (2)), we see that for each compact set L in $\tilde{G}$, $\mathcal{D}(\tilde{\mathcal{E}})|_L$ is dense in $C(L)$ by using Proposition 2.6 and applying the Stone-Weierstrass theorem. This establishes regularity and we have completed the proof.

Remark 2.7 (1) In this paper, we assume that each K_i is a nested fractal. We need this technical condition as we could only prove Proposition 2.6 for $x \neq y \in \partial_{in}G \setminus \partial_{ex}G$ under this condition and Assumption 2.2 (see Proposition 2.13). If we assume the following strong separation condition instead of Assumption 2.1 and 2.2, we can prove Theorem 2.5 even if some of the K_i are fractals such as the Sierpinski carpet, where the domains of the Dirichlet forms are Lipschitz (Besov) spaces (see [17], [23]):
There exists $\delta_0 > 0$ such that $d^E(K_i, \cup_{j\neq i} K_j) > \delta_0$ for all $1 \le i \le M$ where d^E is the Euclidean distance.
(2) E.B. Davies pointed out to us that Assumption 2.1 (2) is not necessary for the construction of the Dirichlet form. Indeed, when $\{K_i\}_i$ does not satisfy this assumption, we can always approximate G by a sequence $\{G_j\}_j$ which does satisfy the assumption. Then, we can construct a monotone sequence of regular local Dirichlet forms using $\{G_j\}_j$ so that the desired form is obtained as a limit of this monotone sequence of forms. However we will leave the assumption in force as we will need it for the heat kernel estimates discussed in Section 4.
(3) It is not hard to extend our results to the cases where (a) Dirichlet integrals are added on each domain inside $\mathrm{Cov}(G)$; (b) the reference measure μ is weighted so that $\mu = m|_D + \sum_i a_i\mu_i$ with a suitable positive sequence $\{a_i\}_i$, or the mixture of the two cases (in [17], this possibility is discussed for the case $M = 1$). In this paper, we do not choose to include these extensions for simplicity.
(4) Assumption 2.1 (3) is used to prove Proposition 2.6 when $x, y \in \partial_{in}G \cap \partial_{ex}G$ (see the last part of Section 2.2).

For each $B \subset \mathbb{R}^2$, define $\tau_B = \inf\{t \ge 0 : \tilde{X}_t \in B\}$. We can then prove that $\tilde{X}_t$ penetrates into each K_i. More precisely we have the following.

Proposition 2.8 *Assume that $m(G) = 0$ where m is the Lebesgue measure on $\mathbb{R}^2$. Then, for any nearly Borel set B with positive 1-capacity (w.r.t. $\tilde{\mathcal{E}}$),*

$$\tilde{P}^x(\tau_B < \infty) > 0 \qquad \text{for quasi-every } x \in \mathbb{R}^2. \tag{2.3}$$

Especially, when B is either a subset of K_i whose 1-capacity w.r.t. $\mathcal{E}_{K_i}$ is positive or a subset of $\mathbb{R}^2$ whose 1-capacity w.r.t. the Dirichlet integral is positive, then (2.3) holds.

Noting that $d_{f,E}(K_i) > 2-2=0$ and points in K_i have positive capacities w.r.t. $\mathcal{E}_{K_i}$ for all K_i, the proof is the same as Proposition 2.9 in [17].

In order to study the properties of $\{\tilde{X}_t\}_{t\geq 0}$, we will make the following additional assumption. Firstly we say that two fractal components K_i are equivalent if there exists a similitude mapping one to the other, and we say that a fractal which belongs to such an equivalence class is of a particular type.

Assumption 2.9 *There are only a finite number M' of types in $\{K_i\}_{i=1}^M$. Further, $d_w(K_i) \neq d_w(K_j)$ if K_i and K_j are different types.*

In the same way as Theorem 2.11 in [17], we can prove a Nash inequality and a corresponding upper estimate for the heat semigroup. If $P_t^{\tilde{\mathcal{E}}}$ $(t>0)$ denotes the semigroup corresponding to $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$, then the following holds (see [17] for the proof).

Proposition 2.10 *Under Assumption 2.1, 2.2 and 2.9, set $d_s^{\min} = \min_{i=1}^M d_s(K_i)$. Then, there exists $c_{2.1} > 0$ such that the following holds*

$$\|P_t^{\tilde{\mathcal{E}}}\|_{1\to\infty} \leq \begin{cases} c_{2.1}t^{-1}, & \text{for all } \ t\in(0,1], \\ c_{2.1}t^{-d_s^{\min}/2}, & \text{for all } \ t\in[1,\infty). \end{cases} \tag{2.4}$$

2.2 Proof of Proposition 2.6

In this subsection, we will give a proof of Proposition 2.6. The crucial part is to show (1) for the case $x\vee y \in \partial_{in}G$ and $x\vee y\in\partial_{ex}G$, where $x \vee y$ means x or y. We adopt completely different methods for the two cases; we use self-similarity and the nesting property for the first case and, for the second case, we apply the extension operator used in the trace theory of Besov spaces.

We will first prove (1) for the case where $x\vee y\in\partial_{in}G$. Assumption 2.2 will be used here. For each $f\in C(\mathbb{R}^2)$, let $\|f\|_{\text{Lip}} = \sup\{|f(x)-f(y)|/\|x-y\| : \ x,y\in\mathbb{R}^2\}$ and let Lip $(\mathbb{R}^2) = \{f\in C(\mathbb{R}^2) : \|f\|_{\text{Lip}} < \infty\}$ where $\|\cdot\|$ is the Euclidean metric.

We now give an important lemma due essentially to Kusuoka ([18]) about extension of functions on fractals. We give a more general version than we need as the result is of interest in its own right. Let $d_{A,K}(\omega) = \min\{n\geq 1 : A\cap\Psi_{\omega_1\cdots\omega_n}(K) = \emptyset\}$.

Lemma 2.11 *Let K be a compact nested fractal defined by the family of α-contractions $\{\Psi_j\}_{j=1}^N$ and let $A\subset K$ be a subset of K. Let $H_{A,K}$: Lip $(\mathbb{R}^d)\to C(K)$ be a linear operator given by*

$$H_{A,K}g(x) = E^x[g(X_{\tau_A})] \qquad \text{for all } \ x\in K, \ g\in \text{Lip } (\mathbb{R}^d) \tag{2.5}$$

where $\{X_t\}$ is the Brownian motion on K and $\tau_A = \inf\{t\geq 0 : X_t\in A\}$. Then there exists $c_{2.2} = c_{2.2}(K) > 0$ such that

$$\mathcal{E}_K(H_{A,K}g, H_{A,K}g) \leq c_{2.2}\{\int_\Sigma (\rho N\alpha^{-2})^{d_{A,K}(\omega)}\nu(d\omega)\}\|g\|_{\text{Lip}}^2 \tag{2.6}$$

holds for any $g\in$ Lip $(\mathbb{R}^d)$ where $\Sigma = \{1,\cdots,N\}^{\mathbb{N}}$ and ρ is the conductance scale factor of K.

Proof: For each $g \in C(K)$, define $h_A(\cdot : g) : K \to \mathbb{R}$ as follows,

$$h_A(\pi(\omega) : g) = \begin{cases} E^{\pi(\sigma^m \omega)}[g \circ \Psi_{\omega_1 \cdots \omega_m}(X_{\tau_{V_0}})] & \text{if } \ d_{A,K}(\omega) = m, \\ g(\pi(\omega)) & \text{if } \ d_{A,K}(\omega) = \infty, \end{cases} \tag{2.7}$$

for each $\omega \in \Sigma = S^{\mathbb{N}}$ (see the Appendix for the notation). It is easy to see that $h_A(\cdot : g)$ is a well-defined continuous map which is harmonic inside $\Psi_{\omega_1 \cdots \omega_m}(K)$ if $d_{A,K}(\omega) = m$, and $h_A(\cdot : g)|_A = g|_A$. Moreover, noting that

$$\mathcal{E}_n(g) = \rho^n \sum_{w \in S^n} \mathcal{E}_0(g \circ \Psi_w) \qquad \text{for all } \ g \in C(V_n),$$

where we abbreviate $\mathcal{E}_n(g,g)$ to $\mathcal{E}_n(g)$, we can easily see that

$$\mathcal{E}_n(h_A(\cdot : g)|_{V_n}) = \int_\Sigma \rho^{d_{A,K}(\omega) \wedge n} \cdot N^{d_{A,K}(\omega) \wedge n} \mathcal{E}_0(\{g(\pi([\omega, i]_{d_{A,K}(\omega) \wedge n})); \pi(\dot{i}) \in V_0\}) \nu(d\omega), \tag{2.8}$$

where we set $[\omega, i]_l = \omega_1 \cdots \omega_l i i \cdots$. Note also that there exists $c_1 > 0$ such that

$$c_1^{-1} \mathcal{E}_0(u) \le \max\{|u(x) - u(y)|^2 : x, y \in V_0\} \le c_1 \mathcal{E}_0(u), \tag{2.9}$$

for any $u \in C(V_0)$. Using (2.8), (2.9) and the fact $\rho N \alpha^{-2} > 1$, (which is shown in [1], Proposition 6.30), we have for each $g \in C(K)$ that

$$\begin{aligned} \mathcal{E}_n(h_A(\cdot : g)|_{V_n}) \quad &\le \quad c_1 \cdot \{\int_\Sigma (\rho N \alpha^{-2})^{d_{A,K}(\omega)} \nu(d\omega)\} \\ &\qquad \times \sup_m \{\alpha^m \cdot \max\{|g(x) - g(y)| : x, y \in \Psi_\xi(V_0)\}; m \ge 0, \xi \in S^m\}^2. \end{aligned}$$

We thus obtain

$$\mathcal{E}(h_A(\cdot : g), h_A(\cdot : g)) \le c_1 \cdot \int_\Sigma (\rho N \alpha^{-2})^{d_{A,K}(\omega)} \nu(d\omega) \cdot \{\text{diam } K\}^2 \cdot \|g\|_{\text{Lip}}^2,$$

for each $g \in \text{Lip}(\mathbb{R}^d)$. As

$$\mathcal{E}(H_{A,K} g, H_{A,K} g) = \inf\{\mathcal{E}(u,u) : \ u \in \mathcal{F}, \ u|_A = g\} \le \mathcal{E}(h_A(\cdot : g), h_A(\cdot : g)) \quad \text{for all } \ g \in \text{Lip}(\mathbb{R}^d),$$

we have completed the proof. ∎

As a corollary we have the following result which we can apply to our fractal fields. This shows that a Lipschitz function on the boundary of the fractal as a set in $\mathbb{R}^2$ can be extended to the interior, provided there is not too much mass of the fractal near the boundary.

Corollary 2.12 *For each Γ_{ij} in (2.1) where $1 \le i \ne j \le M$ and for each compact set $K \subset K_i$ which is congruent to $\hat{K}_i$ and $K \cap \Gamma_{ij} \ne \emptyset$, let $H_{\Gamma_{ij},K} : \text{Lip } (\mathbb{R}^2) \to C(K)$ be a linear operator given by*

$$H_{\Gamma_{ij},K} g(x) = E^x[g(X_{\tau_{\Gamma_{ij}}})] \qquad \textit{for all } \ x \in K, \ g \in \text{Lip } (\mathbb{R}^2). \tag{2.10}$$

Then, under Assumption 2.2, $H_{\Gamma_{ij},K} g \in \mathcal{F}_K$ and

$$\mathcal{E}_K(H_{\Gamma_{ij},K} g, H_{\Gamma_{ij},K} g) \le c_{2.2} \{\int_{\Sigma_i} (\rho_i N_i \alpha_i^{-2})^{d_{\Gamma_{ij},K}(\omega)} \nu(d\omega)\} \|g\|_{\text{Lip}}^2 < \infty \tag{2.11}$$

holds for any $g \in \text{Lip } (\mathbb{R}^2)$.

Note that Assumption 2.2 is used to guarantee $\int_{\Sigma_i} (\rho_i N_i \alpha_i^{-2})^{d_{\Gamma_{ij},K}(\omega)} \nu(d\omega) < \infty$.

Using this we can now show (1) of Proposition 2.6 for the case $x \vee y \in \partial_{in} G \setminus \partial_{ex} G$.

Proposition 2.13 *For each $x \neq y \in \tilde{G}$ where $x \in \partial_{in} G \setminus \partial_{ex} G$, there exists $f \in \mathcal{D}(\tilde{\mathcal{E}})$ such that $f(x) = 1, f(y) = 0$.*

Proof: For $x \in \partial_{in} G \setminus \partial_{ex} G$, denote by $\{K_i\}_{i \in I(x)}$ the set of all K_i such that $x \in K_i$. For each K_i, $i \in I(x)$, take $m_i \in \mathbb{N}$ such that $\alpha_i^{-m_i-1} \leq e^{-m} < \alpha_i^{-m_i}$ and define $D_m(x)$ as a union of the m_i-complexes which contain x for each $i \in I(x)$. We also define $D_m^1(x)$ as the union of the m_i-complexes which intersect with $D_m(x)$. We take m suitably large so that $D_m^1(x) \cap \tilde{G} \subset \cup_{i \in I(x)} K_i$, $(\cup_{i \in I(x)} V_0^{(i)}) \cap (D_m^1(x) \setminus D_m(x)) = \emptyset$ and $y \notin D_m^1(x)$. It is enough to prove that there exists $g \in \mathcal{D}(\tilde{\mathcal{E}})$ such that

$$g|_{D_m(x)} = 1, \quad \text{Supp } g \subset D_m^1(x). \tag{2.12}$$

We will now construct $g \in \mathcal{D}(\tilde{\mathcal{E}})$ which satisfies (2.12). Set $g|_{D_m(x)} = 1$ and take an arbitrary connected component of $\Gamma_{ij} \cap (D_m^1(x) \setminus D_m(x))$, $i, j \in I(x)$ which we denote by Γ. Denote by $a_0 \in D_m(x), a_1 \notin D_m(x)$ the end vertices of Γ. Take $f \in \text{Lip}(\mathbb{R}^2)$ so that $f(a_0) = 1, f(a_1) = 0$. Also, for each $i \in I(x)$, take a compact fractal $\tilde{K}_i$ congruent to $\hat{K}_i$ such that $\tilde{K}_i \cap D_m^1(x) = K_i \cap D_m^1(x)$. Then, by Corollary 2.12, we can construct continuous functions $H_{\Gamma, \tilde{K}_i} f$ and $H_{\Gamma, \tilde{K}_j} f$ on the m_i-complexes on each side of Γ such that $H_{\Gamma, \tilde{K}_l} f|_\Gamma = f|_\Gamma$ and $\mathcal{E}_{(1)}(H_{\Gamma, \tilde{K}_l} f) < \infty$ for $l = i, j$. We apply the same procedure to each connected component of $\Gamma_{ij} \cap (D_m^1(x) \setminus D_m(x))$, $i, j \in I(x)$. Then, using the m-harmonic extension (A.2) for the rest of $D_m^1(x) \setminus D_m(x)$, we can easily extend $\{H_{\Gamma, \tilde{K}_l} f\}_{\Gamma, \tilde{K}_l}$ $(l \in I(x))$ m-harmonically and construct g which satisfies (2.12). Thus, by construction, we see that $g \in \mathcal{D}(\tilde{\mathcal{E}})$. ∎

We next consider the case where $x \vee y \in \partial_{ex} G$. As we mentioned before, we will apply the extension operator used in the theory of Besov spaces (see [13] for details of this theory). For this purpose, we will briefly explain the construction of an extension operator. It is a slight modification of the operator which extends a function in the Lipschitz (Besov) space on the set K_i to a function in a Besov space on $\mathbb{R}^d$ ($d = 2$ for our case, but we can argue in the same way for all $d \in \mathbb{N}$).

We begin by setting up the Whitney decomposition of the complement of K_i. It consists of a collection of closed cubes $\{Q_j^{(i)}\}_{j \in \mathbb{N}}$, with mutually disjoint interiors and sides parallel to the axes so that $\mathbb{R}^d \setminus K_i = \cup_j Q_j^{(i)}$. We assume that the sidelength of the cubes is of the form $2^{-L}, L \in \mathbb{Z}$. Denote the centre of $Q_j^{(i)}$ by $x_j^{(i)}$, its diameter by $l_j^{(i)}$ and its sidelength by $s_j^{(i)}$. Then $s_j^{(i)} = l_j^{(i)}/\sqrt{d} \in \{2^{-L} : L \in \mathbb{Z}\}$. (In the following, we may omit the superscript $^{(i)}$ when there is no confusion.) This decomposition has the following properties,

$$l_j \leq d(Q_j, K_i) \leq 4l_j, \quad Q_j \cap Q_k \neq \emptyset \Rightarrow l_j/4 \leq l_k \leq 4l_j. \tag{2.13}$$

Let $0 < \epsilon < 1/4$ and put $Q_j^* = (1+\epsilon)Q_j$. Note that by the above properties of $\{Q_j\}_j$, each point in $\mathbb{R}^d \setminus K_i$ is contained in at most $M_0(d)$ (which depends only on the Euclidean dimension) cubes Q_j^* and, $Q_j^* \cap Q_k \neq \emptyset$ if and only if $Q_j \cap Q_k \neq \emptyset$. To this decomposition, we associate a partition of unity, consisting of nonnegative functions $\{\varphi_j\}_{j\in\mathbb{N}}$ such that $\varphi_j|_{(Q_j^*)^c}=0$, $\sum_j \varphi_j(x)=1$ for all $x\in\mathbb{R}^d\setminus K_i$, and

$$|D^k\varphi_j(x)|\le A_k(l_j)^{-|k|} \quad \text{for all } x\in\mathbb{R}^d, j\in\mathbb{N}, k\in(\mathbb{N}\cup\{0\})^{\,n}, \tag{2.14}$$

for some constant $A_k>0$ depending only on k. Here, for $k=(k_1,\cdots,k_d)$, we set $D^k=\frac{\partial^{k_1}}{\partial x_1^{k_1}}\cdots\frac{\partial^{k_d}}{\partial x_d^{k_d}}$ and $|k|=k_1+\cdots k_d$.

We now define the extension operator ξ_{δ_0}. Set $m_j=\mu_i(B(x_j,6l_j))^{-1}$. Note that when $l_j=\sqrt{d}2^{-\nu}$ for $\nu\in\mathbb{N}$, then $m_j\le c_1 2^{\nu d_i}$ (where d_i is the Hausdorff dimension of K_i). Now, for $f\in\mathbb{L}^2(K_i,\mu_i)$, define

$$\xi_{\delta_0}f(x)=\sum_{j\in I_{\delta_0}}\varphi_j(x)m_j\int_{||t-x_j||\le 6l_j} f(t)d\mu_i(t) \quad \text{for all } x\in\mathbb{R}^d\setminus K_i, \tag{2.15}$$

where $\delta_0>0$ and

$$I_{\delta_0}\equiv\{j\in\mathbb{N}: s_j\le c_2\delta_0\}. \tag{2.16}$$

We note that in the definition of the usual extension operator, $I\equiv\{j\in\mathbb{N}: s_j\le 1\}$ is used instead of I_{δ_0}. The concrete value 6 used in (2.15) is not important; it is enough to choose a sufficiently large number α_0 so that $\mu_i(\{t:||t-x_j||\le\alpha_0 l_j\}\cap K_i)$ is bounded away from 0. Take $f\in C_0(K_i)$. For each fixed $x\in\mathbb{R}^d\setminus K_i$, there are only a finite number of φ_j where $\varphi_j(x)\ne 0$ so that $\xi_{\delta_0}f$ is well defined and in $C^\infty(\mathbb{R}^d\setminus K_i)$. Further, by (2.13) and by the definition of I_{δ_0}, $\xi_{\delta_0}f(x)=0$ if $x\in Q_j, s_j>c_3(\delta_0)$ for some $c_3(\delta_0)$ which depends on c_2 and δ_0. We will take c_2 (which depends only on the dimension of the Euclidean space) small enough so that Supp $\xi_{\delta_0}f$ is in the δ_0-neighbourhood of K_i. Thus we see that $\xi_{\delta_0}f\in C_b^\infty(\mathbb{R}^d\setminus K_i)$ for $f\in C_0(K_i)$, where $C_b^\infty(\mathbb{R}^d\setminus K_i)$ is the space of infinitely differentiable functions with bounded support on $\mathbb{R}^d\setminus K_i$. In this case, $\xi_{\delta_0}f$ is uniformly continuous on $\mathbb{R}^d\setminus K_i$ and $\lim_{x\to x_0\in\partial K_i}\xi_{\delta_0}f(x)=f(x_0)$, which can be proved in the same way as in [17] p78, p80. Thus, by defining $\xi_{\delta_0}f(x)=f(x)$ for $x\in K_i$, it holds that $\xi_{\delta_0}f\in C_0(\mathbb{R}^d)$ for each $f\in C_0(K_i)$. It can be also proved by the general trace theory of Besov spaces (or, for this case, as in [17] p79) that $\int_{(K_i)^c}|\nabla(\xi_{\delta_0}f)(x)|^2dx<\infty$. Noting that Supp $\xi_{\delta_0}f$ is in the δ_0-neighbourhood of K_i, we obtain that $\xi_{\delta_0}f\in\mathcal{D}(\tilde{\mathcal{E}})$ for each $f\in C_0(K_i)$. Using ξ_{δ_0}, we now show (1) of Proposition 2.6 for the case where $x\vee y\in\partial_{ex}G\setminus\partial_{in}G$.

Proposition 2.14 *For each $x\ne y\in\tilde{G}$ where $x\in\partial_{ex}G\setminus\partial_{in}G$, there exists $f\in\mathcal{D}(\tilde{\mathcal{E}})$ such that $f(x)=1, f(y)=0$.*

Proof: As $x\in\partial_{ex}G\setminus\partial_{in}G$, there is a unique K_i such that $x\in K_i$. Denote by $B(x,r)$ a ball in $\mathbb{R}^2$ centred at x and radius r. We take $r,\delta_0>0$ small enough so that $B(x,r+\delta_0)\cap G\subset K_i$ and $y\notin B(x,r+\delta_0)$. Using Lemma 2.4 (3), we see that there exists $f\in\mathcal{F}_{K_i}\cap C_0(K_i)$ such that $f(x)=1$ and Supp $f\subset B(x,r)\cap K_i$. Now using the above extension operator, $\xi_{\delta_0}f\in\mathcal{D}(\tilde{\mathcal{E}})$, $(\xi_{\delta_0}f)|_{K_i}=f$ and Supp $f\subset B(x,r+\delta_0)$. Thus $\xi_{\delta_0}f(x)=1$, $\xi_{\delta_0}f(y)=0$ and the proof is completed. ∎

<u>End of the proof of Proposition 2.6</u>

We first complete the proof of (1). When $x \vee y \in \tilde{G} \setminus \partial G$, (1) is clear using Lemma 2.4 (3) and (4). When x and y are both in ∂G, there are three cases: (a) $x \vee y \in \partial_{in} G \setminus \partial_{ex} G$, (b) $x \vee y \in \partial_{ex} G \setminus \partial_{in} G$, (c) $x, y \in \partial_{in} G \cap \partial_{ex} G$. For cases (a) and (b), (1) is proved in Proposition 2.13 and Proposition 2.14 respectively. For the case (c), denote by $\{K_i\}_{i \in I(x)}$ the set of all K_i such that $x \in K_i$. In the same way as Proposition 2.13 (using Corollary 2.12 repeatedly), we can construct non-negative $f \in C_0(G)$ such that $f|_{K_i} \in \mathcal{F}_{K_i}$ for all $i \in I(x)$, $f|_{U(x)} = 1$ for some small neighbourhood $U(x)$ of x, Supp $f \subset \cup_{i \in I(x)} K_i$ and $d^E(\text{Supp } f, \partial_{in} G \cap \partial_{ex} G \setminus \{x\}) > \epsilon$ for some $\epsilon > 0$ where d^E is the Euclidean metric (the last claim is guaranteed by Assumption 2.1 (3)). Next we take the Whitney decomposition $\{Q_j\}$ of $(\cup_{i \in I(x)} K_i)^c$, with its associated partition of unity $\{\varphi_j\}$, and define $\xi_{\delta_0} f$ in the same way as (2.15) using this $\{Q_j\}$, $\{\varphi_j\}$ and $\mu \equiv \sum_{i \in I(x)} \mu_i$. For $y \in \cup_{i \in I(x)} K_i$, we set $\xi_{\delta_0} f(y) = f(y)$. Now consider the balls $\{t \in G : \|t - x_j\| < 6 l_j\}$, for $j \in I_{\delta_0}$ in the definition of $\xi_{\delta_0} f$, which intersect more than one K_i. By taking δ_0 small, either $f(t) = 1$ or $f(t) = 0$ for all t which belong to the ball, depending on the distance of the ball to x. Thus, either $\xi_{\delta_0} f(z) = \sum_{j \in I_{\delta_0}} \varphi(z)$, or $\xi_{\delta_0} f(z) = 0$ for $z \in Q_j$, where Q_j is the cube of the Whitney decomposition centred at x_j above. For the other $z \in Q_{j'}$ (i.e., $Q_{j'}$ where $\{t \in G : \|t - x_{j'}\| < 6 l_{j'}\}$ intersects only one K_i), $\xi_{\delta_0} f(z)$ equals $\xi_{\delta_0}^{(i)}(f|_{K_i})(z)$, as constructed through the Whitney decomposition of $(K_i)^c$. Using these facts, we can prove $\xi_{\delta_0} f \in \mathcal{D}(\tilde{\mathcal{E}})$ in the same way as before so that $\xi_{\delta_0} f$ is the desired function.

We next prove (2). For each compact set $L \subset \tilde{G}$, define $I_L = \{i : L \cap K_i \neq \emptyset\}$. Note that $|I_L| < \infty$, by Assumption 2.1 (2). As each K_i is closed, we can take $\delta_0'(L) > 0$ so that the set of the indices of the K_i which intersect with $\{y : d^E(L, y) \leq \delta_0'(L)\}$ is equal to I_L. Now, in a similar way to the proof of (1), there exists $f \in \mathcal{D}(\tilde{\mathcal{E}})$ such that $f|_{L \cap G} = 1$. Let $L' = L \setminus \{x \in L : f(x) \geq 1/2\}$. Then there exists $g \in C_0^\infty(\mathbb{R}^d)$ so that $g|_{L'} = 1$ and the support of g is in $\{x \in \tilde{G} : d^E(L, x) \leq \delta_0'(L)\} \setminus G$. Clearly $g \in \mathcal{D}(\tilde{\mathcal{E}})$. Define $h = 2f + g \in \mathcal{D}(\tilde{\mathcal{E}})$. Then, $h|_L \geq 1$. Thus, $\bar{h} \equiv (h \vee 0) \wedge 1$ (which is in $\mathcal{D}(\tilde{\mathcal{E}})$ by the Markovian property of $\tilde{\mathcal{F}}$) is the desired function. ∎

2.3 More properties of the forms on G

When G is connected, we can define a bilinear form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on $\mathbb{L}^2(G, \mu)$ in the same way as before:

$$
\begin{aligned}
\mathcal{E}(u, v) &= \sum_{i=1}^{M} \mathcal{E}_{K_i}(u|_{K_i}, v|_{K_i}) \quad \text{for all } u, v \in \mathcal{D}(\mathcal{E}), \\
\mathcal{D}(\mathcal{E}) &= \{u \in C_0(G) : u|_{K_i} \in \mathcal{F}_{K_i} \ \forall\, i, \ \ \mathcal{E}(u, u) < \infty\}.
\end{aligned}
$$

The conclusions of Lemma 2.4 (1),(2),(3) hold for $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ and we can define $\mathcal{F} = \overline{\mathcal{D}(\mathcal{E})}^{\mathcal{E}_{(1)}}$ so that $(\mathcal{E}, \mathcal{F})$ is the smallest extension of $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$. As in the last subsection we can also prove that $(\mathcal{E}, \mathcal{F})$ is a local regular Dirichlet form on $\mathbb{L}^2(G, \mu)$. In this subsection, we will describe more properties of $(\mathcal{E}, \mathcal{F})$ in the case where G is connected.

Let $R_{K_i}(.,.)$ be the resistance metric on K_i with respect to $(\mathcal{E}_{K_i}, \mathcal{F}_{K_i})$ (see the Appendix for details). For each $x, y \in G,\ x \neq y$, define

$$R(x,y)^{-1} = \inf\{\mathcal{E}(f,f) : f \in \mathcal{D}(\mathcal{E}), f(x) = 1, f(y) = 0\}, \tag{2.17}$$

and $R(x,x) = 0$ for all $x \in G$. Using Proposition 2.6 (1), we see that $R(x,y) \neq 0$ for $x \neq y$. Further, for each $f \in \mathcal{D}(\mathcal{E})$, we have that $f|_{K_i} \in \mathcal{F}_{K_i}$ and $\mathcal{E}(f,f) \geq \mathcal{E}_{K_i}(f|_{K_i}, f|_{K_i})$, and hence

$$R(x,y) \leq R_{K_i}(x,y) \qquad \text{for all } x, y \in K_i. \tag{2.18}$$

Using these facts, Proposition A.1, (where $\hat{K}$ should be changed to G, $\hat{\mathcal{E}}$ to $\mathcal{E}$ and $l(V_*)$ to $\mathcal{D}(\mathcal{E})$) can be proved in the same way and we have

$$|f(x) - f(y)|^2 \leq R(x,y)\mathcal{E}(f,f) \quad \text{for all} \ \ f \in \mathcal{D}(\mathcal{E}),\ x, y \in G. \tag{2.19}$$

We can easily extend (2.19) to all $f \in \mathcal{F}$ so that $\mathcal{F} \subset C(G)$. Thus we can take $f \in \mathcal{F}$ in the infimum of the definition of $R(x,y)$ in (2.17). Using (2.19), if we set $\mathcal{E}_{(\beta)}(.,.) = \mathcal{E}(.,.) + \beta(.,.)_{\mathbb{L}^2(G,\mu)}$ for $\beta > 0$, then by routine arguments we can prove $\mathcal{E}_{(\beta)}$ admits a positive symmetric continuous reproducing kernel. Summarizing the results,

Theorem 2.15 *If G is connected, then under Assumption 2.1 and 2.2, $(\mathcal{E}, \mathcal{F})$ is a local regular Dirichlet form on $\mathbb{L}^2(G,\mu)$ which has the following property:*

$$|f(x) - f(y)|^2 \leq R(x,y)\mathcal{E}(f,f) \quad \textit{for all} \ \ f \in \mathcal{F},\ x, y \in G. \tag{2.20}$$

Further, $\mathcal{E}_{(\beta)}$ admits a positive symmetric continuous reproducing kernel.

From this theorem, we can obtain several properties of the Dirichlet form by using the same arguments as those for nested fractals. For instance, using (2.20), we can prove that there exists a jointly continuous heat kernel $p_t(.,.)$ and therefore the corresponding diffusion is a Feller diffusion (so there is no ambiguity in the starting points for the correspondence with the Dirichlet form). Also, using the same proof as Theorem 2.11 in [17], Proposition 2.10 corresponds to the following.

Proposition 2.16 *If G is connected, then under Assumption 2.1, 2.2 and 2.9, with $d_s^{\min} = \min_{i=1}^M d_s(K_i)$ and $d_s^{\max} = \min_{i=1}^M d_s(K_i)$, there exists $c_{2.3} > 0$ such that the following holds*

$$\|P_t^{\mathcal{E}}\|_{1\to\infty} \leq \begin{cases} c_{2.3} t^{-d_s^{\max}/2}, & \textit{for all} \ \ t \in (0,1], \\ c_{2.3} t^{-d_s^{\min}/2}, & \textit{for all} \ \ t \in [1,\infty). \end{cases} \tag{2.21}$$

3 Shortest path metric and hitting time estimates

For the rest of the paper we will work on G, the fractal tiling of $\mathbb{R}^2$ and make Assumptions 2.1, 2.2 and 2.9. This restriction is because our arguments rely heavily on the existence of a resistance metric. (We believe that similar results hold for $\tilde{G}$, though we do not have a proof as yet.) We will also, without loss of generality, assume that the types are labelled in order of increasing d_w in the shortest path metric and denote by $\tau(j)$ the type of fractal K_j. We decompose the tiling as $G = \cup_{i=1}^{M'} K^{(j)}$, where $K^{(j)} = \cup_{i=1,\tau(i)=j}^{M} K_i$. We begin this section by defining shortest paths on our fractal fields. Once we have obtained such paths it is straightforward to obtain a hierarchy of shortest path metrics on the field.

To begin we introduce a sequence of subsets which approximate the field using conductivity coordinates. Firstly, for a compact fractal K_i of unit side, we can define conductivity coordinates as usual, by ensuring that each edge in the n-th lattice is of resistance roughly e^{-n},

$$\Lambda_n^i = \{w = w_1 \dots w_m \in \Sigma_i : \rho_i^{m-1} \leq e^n < \rho_i^m\}.$$

This can be extended to larger and unbounded fractals by allowing a two sided shift space Σ_i. For each fractal subset K_i we can define its conductivity coordinate approximation $V_{\Lambda_n^i}$ as the vertices of the Λ_n^i-cells, [8]. We refer to a Λ_n^i-cell, with address w as the vertices in $V_{\Lambda_n^i} \cap \Psi_w^{(i)}(K_i)$. We denote by G_n^i the graph which has an edge from x to y in K_i if $P_{\Lambda_n^i}(x,y) > 0$ for the transition probability on $V_{\Lambda_n^i}$ induced by the Brownian motion on the fractal K_i.

Let $\Lambda_n = \oplus_{i=1}^{M} \Lambda_n^i$. The conductivity coordinate approximation H_{Λ_n} to the field will be different from the approximation to the individual fractals as the components of the boundaries between fractals incorporated in the approximation will not be just points. We begin by considering the interiors of the constituent fractals. Let $x \in H_{\Lambda_n}$ if $x \in V_{\Lambda_n^i} \setminus \cup_j \Gamma_{ij}$ and we consider such an x to be connected to $y \in V_{\Lambda_n^i} \setminus \cup_j \Gamma_{ij}$ if $\{x,y\} \in G_n^i$. We now define the boundary and its connections. Consider Γ_{ij}. We assume without loss of generality that $\rho_i > \rho_j$ and hence there are fewer vertices of $V_{\Lambda_n^i}$ on Γ_{ij} compared to vertices of $V_{\Lambda_n^j}$. We now include in the set H_{Λ_n} the component of the boundary which joins each vertex of $V_{\Lambda_n^i} \cap \Gamma_{ij}$ with the closest vertex of $V_{\Lambda_n^j} \cap \Gamma_{ij}$. If these two vertices coincide, and are not at a corner where two or more boundary components meet, we only take the point itself. In the case of a corner, we include the boundary components from the corner to all the nearest points in $V_{\Lambda_n^j}$ as shown in Figure 4.

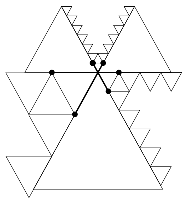

Figure 4: The connected component of H_{Λ_n} at a corner where 6 fractals meet

The possible connections between elements of H_{Λ_n} are then inherited from the constituent fractals in that we regard each point in H_{Λ_n} as connected to any other point for which there is a path in the fractal to that point which does not pass through any other point of H_{Λ_n}. A boundary complex in the gasket tiling case is shown in Figure 5. Note that the thick line segments on the boundary are elements of H_{Λ_n} and hence all points in the segment are connected to the other elements on H_{Λ_n} in the complex.

We introduce a little more notation. As in [20] we call the set in K_i corresponding to the word $w \in \Lambda_n^i$ a Λ_n^i-complex. For the field G the Λ_n-complexes consist of all the Λ_n^i-complexes for each i which do not intersect the boundary while, if the Λ_n-complex intersects the boundary Γ_{ij},

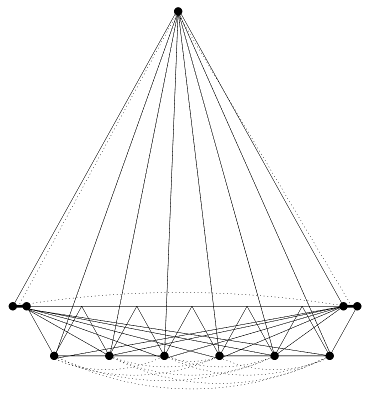

Figure 5: The H_{Λ_n} boundary complex and its internal connections

where $\rho_i > \rho_j$, there will be one Λ_n-complex for each Λ_n^i-complex which intersects the boundary. It consists of the Λ_n^i-complex A and all the Λ_n^j-complexes B such that $B \cap \Gamma_{ij} \subset A$ as well as those with at least half of $B \cap \Gamma_{ij} \subset A$. As we take components of the boundary between fractals in our Λ_n-cell (the boundary of the Λ_n-complex) any path from the interior of one Λ_n-complex to another must pass through an element of H_{Λ_n}.

Let

$$\begin{aligned} D^0_{\Lambda_n}(x) &= \{A : A \text{ is a } \Lambda_n\text{-complex containing } x\}, \\ D^1_{\Lambda_n}(x) &= D^0_{\Lambda_n} \cup \{A : A \text{ is a } \Lambda_n\text{-complex connected to } D^0_{\Lambda_n}\}. \end{aligned}$$

We write $\partial D^i_{\Lambda_n}(x) = Cl(G \backslash D^i_{\Lambda_n}(x)) \cap D^i_{\Lambda_n}(x)$ for the components of the boundary of the set $D^i_{\Lambda_n}(x)$ which connect the neighbourhood of x to the rest of the field. We denote a path from x to y on H_{Λ_n} as

$$\pi_n(x,y) = \{(x_0, x_1, \ldots, x_N) : x_0 \in \partial D^0_{\Lambda_n}(x), x_N \in \partial D^0_{\Lambda_n}(y), x_i \in H_{\Lambda_n}, i = 1, \ldots, N-1\}.$$

Note that if the path runs along a boundary between two fractals it will only contain the points from the side with larger spacing between points. We define the set of all such paths to be $\Pi_n(x,y)$.

An intrinsic shortest path metric can be defined on the fractal and will be used to express the heat kernel bounds and the large deviation estimates in the next section. Firstly we note that as our fractal tiling is composed of nested fractals there exists a shortest path scaling factor for each type (see the Appendix). That is, if we write $N_n^{(j)}(x,y)$ for the number of steps in the shortest path on $V_{\Lambda_n^j}$ from x to $y \in K_j$, then $N_n^{(j)}$ satisfies $c^{(j)} N_n^{(j)} N_m^{(j)} \leq N_{n+m}^{(j)} \leq N_n^{(j)} N_m^{(j)}$ and hence there exists a $b_j > 1$ such that $c_1^{(j)} b_j^n \leq N_n^{(j)}(x,y) \leq c_2^{(j)} b_j^n$. Note that, as we work on the resistance approximation, the scale factor b_j depends on the resistance.

We introduce the following metric to make our later construction of a hierarchy of metrics clearer. We begin by defining a distance function on H_{Λ_n} as,

$$\hat{d}_n(x,y) = \inf\{\sum_{j=1}^{M'} \frac{|\{x_i \in \pi_n(x,y) \cap H_{\Lambda_n} \cap K^{(j)}\}|}{b_j^n} : \pi_n(x,y) \in \Pi_n(x,y)\}.$$

The convergence of $\hat{d}_n(x,y) \to \hat{d}(x,y)$ as $n \to \infty$ can be established using the convergence of the shortest path counting function on each type. The path (or paths) that achieves this infimum is the shortest path from x to y in G.

The metric $\hat{d}$ on G can be decomposed into a family of metrics, one on each of the different $K^{(i)}$, for $i = 1, \dots, M'$. We begin by defining $G^{(1)}$ to be the field G where we identify all points in each connected component of the closure of the complement of $K^{(1)}$ in G which are not of type 1. Now we define, for $x, y \in G^{(1)}$,

$$d_n^{(1)}(x,y) = \frac{N_n^{(1)}(x,y)}{b_i^n}, \tag{3.1}$$

where $N_n^{(1)}(x,y)$ denotes the number of steps in the shortest path on the $H_{\Lambda_n} \cap K^{(1)}$ approximation to $G^{(1)}$. Then the shortest path metric is defined as in [1] page 78, by taking a limit along a suitable subsequence, $d^{(1)}(x,y) = \lim_{k\to\infty} d_{n_k}^{(1)}(x,y)$.

We now define in a hierarchical fashion our family of metrics. Let $G^{(j)}$ be the field G in which the interior of all components of type i for $i < j$ are removed completely, while all points are identified in the closure of the connected components of $K^{(i)}$ with $i > j$. We can then count the number of steps between x and y on $H_{\Lambda_n} \cap G^{(j)}$, where if the fractal $G^{(j)}$ is disconnected we fix the end points of the path which achieves the infimum on $G^{(i)}$ for $i < j$. Thus if we define

$$d_n^{(j)}(x,y) = \frac{N_n^{(j)}(x,y)}{b_j^n},$$

the metric $d^{(j)}$ can be defined as before by taking a limit (along a subsequence if necessary). The consequences of the above discussion are summarized in the following.

Proposition 3.1 *(1) The distance function $\hat{d}(x,y)$ on the field G can be written as*

$$\hat{d}(x,y) = \sum_{i=1}^{M'} d^{(i)}(x,y).$$

(2) There exist constants $c_{3.1}, c_{3.2}$ such that

$$c_{3.1} d^{(j)}(x,y) b_j^n \leq N_n^{(j)}(x,y) \leq c_{3.2} d^{(j)}(x,y) b_j^n. \tag{3.2}$$

We recall that the walk dimension in the effective resistance metric is given by $d_{w,R}(K_i) = S_i + 1$ where S_i is the Hausdorff dimension of the fractal K_i in the effective resistance metric. We will also define the chemical exponent for type j fractals as $d_c^{(j)} = \log b_j$. For $x \in G$ we write $I_n^l(x) = \{\tau(i) : D_{\Lambda_n}^l(x) \cap Int(K_i) \neq \emptyset\}$, for $l = 0, 1$ for the indices of the different types of fractal which are in the neighbourhood of the point x.

The key is to show that our approximating sets are compatible with the resistance metric.

Lemma 3.2 *There exist constants $c_{3.3}, c_{3.4}$ such that if $x, y \in H_{\Lambda_m}$ and $x, y \in D_{\Lambda_m}^0$ and x, y not in the same connected component of H_{Λ_m}, then*

$$c_{3.3} e^{-m} \leq R(x,y) \leq c_{3.4} e^{-m}.$$

Proof: We first obtain the upper bound for $x, y \in H_{\Lambda_m} \setminus \cup_j \Gamma_{ij}$ by observing that if $x, y \in V_{\Lambda^i_m}$, then $\{x, y\} \in G^i_m$ for some i. Hence

$$R(x,y) \leq R_{K_i}(x,y) \leq c_1 e^{-m},$$

using [8] Lemma 3.2. For $x, y \in \Gamma_{ij}$, (where $\rho_i > \rho_j$) then, as they are not in the same connected component, we can take $x_i \in V_{\Lambda^i_m}$ in the component containing x, so that, $R(x, x_i) \leq c_2 e^{-m}$ and similarly for y_i. Putting these estimates together

$$R(x,y) \leq R(x,x_i) + R(x_i,y_i) + R(y_i,y) \leq c_3 e^{-m}.$$

For a point $x \in \Gamma_{ij}$ and $y \in V_{\Lambda^i_m}$,

$$R(x,y) \leq R(x,x_i) + R(x_i,y) \leq c_4 e^{-m}.$$

Finally for $y \in V_{\Lambda^j_m}$, as it is adjacent to the boundary, there is a point $z_j \in \Gamma_{ij} \cap V_{\Lambda^j_m}$ such that $R(z_j, y) \leq c_5 e^{-m}$. Now, as there is a c_6 such that $R(x_i, z) \leq c_6 e^{-m}$ for all $z \in D^0_{\Lambda^i_n}(x)$, we have

$$R(x,y) \leq R(x,x_i) + R(x_i,z_j) + R(z_j,y) \leq c_7 e^{-m}.$$

For the lower bound, we begin by considering $x \in V_{\Lambda^i_m} \setminus \cup_j \Gamma_{ij}$ for some i. For any $y \in G$,

$$\begin{aligned} R(x,y)^{-1} &= \inf\{\mathcal{E}(f,f) : f(x) = 0, f(y) = 1\} \\ &= \inf\{\sum_{i=1}^{M} \mathcal{E}_{K_i}(f|_{K_i}, f|_{K_i}) : f(x) = 0, f(y) = 1\}. \end{aligned}$$

We can upper bound this by choosing a particular function g. Let $g(y) = 0$ for $y \in K_j, j \neq i$ and $g(x) = 1$. We also set $g(y) = 0$ for all $y \neq x \in V_{\Lambda^i_m}$. In a similar way to [8] Lemma 3.2, though with a little more thought if $y \in \Gamma_{ij}$, this gives $R(x,y)^{-1} \leq c_8 e^m$ and hence the required lower bound.

Finally for the lower bound on the boundary we use Corollary 2.12. Let $x, y \in \Gamma_{ij}$. We now consider the restriction to Γ_{ij} of a function $g \in \mathrm{Lip}(\mathbb{R}^2)$ which has $g(x) = 0$ and $g(y) = 1$. We then extend this to the neighbouring cells using Corollary 2.12 in the same way as in Proposition 2.13, where instead of the support being in $D^1_m(x)$ it lies in $D^1_{\Lambda_m}(x)$. We keep the notation g for this new function. Now we consider a unit cell of K_i, then, writing $|w|$ for the length of the word w,

$$\mathcal{E}_{K_i}(g,g) = \sum_{w \in \Lambda^i_m} \mathcal{E}_{K_i}(g \circ \Psi^{(i)}_w, g \circ \Psi^{(i)}_w) \rho_i^{|w|},$$

and using the fact that g is only non-zero on a bounded number of Λ_m-cells, there are constants C_i, C'_i, C''_i such that

$$\mathcal{E}_{K_i}(g,g) \leq \sum_{w : g \circ \Psi^{(i)}_w \neq 0} C_i e^m \mathcal{E}_{K_i}(g \circ \Psi^{(i)}_w, g \circ \Psi^{(i)}_w) \leq \sum_{w : g \circ \Psi^{(i)}_w \neq 0} C'_i e^m \|g \circ \Psi^{(i)}_w\|_{\mathrm{Lip}} \leq C''_i e^m.$$

Now, by Assumption 2.9, there are only a finite number of C''_i, we can sum them all to get

$$R(x,y)^{-1} \leq \mathcal{E}(g,g) = \sum_{i=1}^{M} \mathcal{E}_{K_i}(g|_{K_i}, g|_{K_i}) \leq c_9 e^m,$$

which gives the result. ∎

It is straightforward to establish an estimate on the measure of Λ_n-complexes.

Lemma 3.3 *There exist constants $c_{3.5}, c_{3.6}$ such that for all $x \in G, n \in \mathbb{N}, i = 0,1,$*

$$c_{3.5} e^{-n \min_i \{S_i : i \in I_n^i(x)\}} \le \mu(D^i_{\Lambda_n}(x)) \le c_{3.6} e^{-n \min_i \{S_i : i \in I_n^i(x)\}}.$$

For $i = 0,1$, let $d^i_{w,R}(n,x) = \min\{S_j + 1 : j \in I_n^i(x)\}$, the smallest walk dimension with respect to the resistance metric among the Λ_n-cells in $D^i_{\Lambda_n}(x)$. We will abbreviate $d_{w,R}(n,x) = d^1_{w,R}(n,x)$. Let $T_0^n = \inf\{t \ge 0 : X_t \in H_{\Lambda_n}\}$ and for $i > 0$ let

$$T_i^n = \inf\{t > T_{i-1}^n : X_t \in H_{\Lambda_n} \backslash \{X_{T_{i-1}^n}\}\}.$$

We then define the crossing time $W_i^n = T_i^n - T_{i-1}^n$.

Lemma 3.4 *There exist constants $c_{3.7}, c_{3.8}$ such that*

$$\sup_{y \in D^0_{\Lambda_n}(x)} E^y T_{D^0_{\Lambda_n}(x)} \le c_{3.7} e^{-n d^0_{w,R}(n,x)}, \quad \forall x \in G, \quad n \ge 0, \tag{3.3}$$

$$E^y T_{D^0_{\Lambda_n}(y)} \ge c_{3.8} e^{-n d^0_{w,R}(n,y)}, \quad \forall y \in H_{\Lambda_n}, \quad n \ge 0. \tag{3.4}$$

Proof: We first observe that in a ball B about the point x we can define the process killed when it exits the ball. The green kernel for this process g_B is a reproducing kernel for the associated Dirichlet form $(\mathcal{E}_B, \mathcal{F}_B)$. As the form is obtained by taking the trace, the resistance metric on B is compatible with the resistance metric on the field. The reproducing kernel $g_B(x,y) = E^x(L^y_{T_B})$ ([1]), the expected local time at y for the process started from x before first exit from B. Thus we have the following estimate

$$g_B(x,y) = g_B(y,x) = P^y(T_x < T_B) g_B(x,x) \le g_B(x,x). \tag{3.5}$$

For the upper bound in (3.3) we have for any $z \in D^0_{\Lambda_n}(x)$,

$$E^z T_{D^0_{\Lambda_n}(x)} = \int_{D^0_{\Lambda_n}(x)} g_{D^0_{\Lambda_n}(x)}(z,y)\mu(dy) \le g_{D^0_{\Lambda_n}(x)}(z,z)\mu(D^0_{\Lambda_n}(x)).$$

We note that, for our tiling, points have positive capacity and hence

$$g_{D^0_{\Lambda_n}(x)}(z,z) = R(z, D^0_{\Lambda_n}(x)^c) \le c_1 e^{-n}. \tag{3.6}$$

Combining this with Lemma 3.3 we have

$$E^z T_{D^0_{\Lambda_n}(x)} \le c_2 e^{-n d^0_{w,R}(n,x)}.$$

Note that as the constant c_2 is independent of the starting point within $D^0_{\Lambda_n}(x)$ we have (3.3).

For the lower bound (3.4) we follow the same argument as Barlow [1] Proposition 8.11. The first step is to show that for $y \in H_{\Lambda_n}$, $g_{D^0_{\Lambda_n}(y)}(y,z)$ is bounded below over $D^0_{\Lambda_{n+1}}(y)$. To establish this we first show that $R(y, D^0_{\Lambda_n}(y)^c) \ge c_3 e^{-n}$. Consider

$$R(y, D^0_{\Lambda_n}(y)^c)^{-1} = \inf\{\mathcal{E}(f,f) : f(y) = 1, f(z) = 0, z \in D^0_{\Lambda_n}(y)^c\}.$$

If $D^0_{\Lambda_n}(y)$ is contained within one fractal type, this result is standard. We only need to consider the case where $D^0_{\Lambda_n}(y)$ overlaps the boundary between two or more types. In this case we can define a function h by setting $h(a) = 1$, for $a \in D^0_{\Lambda_{n+1}}(y)$ and $h(a) = 0$ for $a \in D^0_{\Lambda_n}(y)^c$ and we take h to be Lipschitz on the boundary. The function h can then be extended away from the boundary using Corollary 2.12, as in Proposition 2.13, which ensures that the energy of this extension is controlled. To do this we observe that, as the support of h within fractal K_i is contained within a few Λ_n-cells, $\mathcal{E}_{K_i}(h,h) \leq c_3 e^n \mathcal{E}_{K_i}(h\circ\Psi_w, h\circ\Psi_w)$, where Ψ_w is the composition of similitudes that leads to the Λ_n cell supporting h. As the Lipschitz norm of $h\circ\Psi_w \leq c_4$, we have $\mathcal{E}(h,h) \leq c_5 e^n$. For other parts of the fractal within a single component we can use the harmonic extension. Thus we can control the energy of h over the whole field and hence

$$R(y, D^0_{\Lambda_n}(y)^c) \geq c_3 e^{-n}. \tag{3.7}$$

Let $f(z) = g_{D^0_{\Lambda_n}(y)}(y,z)/g_{D^0_{\Lambda_n}(y)}(y,y)$, so that $f \in \mathcal{F}$, $f(y) = 1$ and, by (3.5), $f \leq 1$. By (3.7) we have

$$\mathcal{E}(f,f) = g_{D^0_{\Lambda_n}(y)}(y,y)^{-1} = R(y, D^0_{\Lambda_n}(y)^c)^{-1} \leq c_4 e^n.$$

Hence, by (2.20), we can choose a c_5 such that

$$|f(y) - f(z)|^2 \leq 1/2, \text{ if } R(y,z) \leq c_5 e^{-n}.$$

Thus $f(z) \geq c_6$ on $D^0_{\Lambda_{m+\hat{c}}}(y)$ for some $\hat{c}$, and hence

$$E^y T_{D^0_{\Lambda_n}(y)} = \int_{D^0_{\Lambda_n}(y)} g_{D^0_{\Lambda_n}(y)}(y,z)\mu(dz) \geq c_7 g_{D^0_{\Lambda_n}(y)}(y,y)\mu(D^0_{\Lambda_{n+\hat{c}}}(y)),$$

which gives the required result. ∎

We now give an extension of [2] Lemma 1.1 which will be useful for the next result.

Lemma 3.5 *Let $U_1,\ldots,U_V, V \in \mathbb{N}$ and let $Z_{ij}, i = 1,\ldots,U_j, j = 1,\ldots,V$ be non-negative random variables. If a random variable X satisfies $X \geq \sum_{j=1}^V Y_j$, where $Y_j \geq \sum_{i=1}^{U_j} Z_{ij}$, and*

$$P(Z_{ij} \leq x|\sigma(Z_{kj}, k < i)) \leq p_j + q_j x, \quad i = 1,\ldots,U_j, x \geq 0,$$

where $0 < p_j < 1, q_j > 0$ for $j = 1,\ldots,V$, then,

$$P(X \leq x) \leq \exp\left(2\sum_{j=1}^V \left((\frac{q_j U_j x}{p_j})^{1/2} - U_j \log\frac{1}{p_j}\right)\right).$$

Proof: We recall from the proof of Lemma 1.1 in [2] that $E(e^{-\theta Z_{ij}}|\sigma(Z_{kj}, k < i)) \leq p_j + q_j\theta^{-1}$. Now for any positive vector $(\theta_1,\ldots,\theta_M)$ we have

$$\begin{aligned}
P(X \leq x) &\leq P(\sum_{j=1}^V \theta_j Y_j \leq \sum_{j=1}^V \theta_j x) \\
&\leq e^{\sum_{j=1}^V \theta_j x} E(e^{-\sum_{j=1}^V \theta_j Y_j}) \\
&= e^{\sum_{j=1}^V \theta_j x} \prod_{j=1}^V E e^{-\theta_j Y_j} \\
&\leq e^{\sum_{j=1}^V \theta_j x} \prod_{j=1}^V (p_j + q_j\theta_j^{-1})^{U_j}.
\end{aligned}$$

We now optimise over each θ_j, by setting $\theta_j = (q_j U_j / p_j x)^{1/2}$, to get the result. ∎

With this estimate we can now extend the standard machinery to deduce an estimate on the tail of the crossing time distribution in our setting.

For $x, y \in G$ such that $e^{-n-1} \leq R(x,y) \leq e^{-n}$, let $\pi_{n,k}(x,y)$ be a potential minimal path from x to y on $H_{\Lambda_{n+k}}$. That is, if we write $\hat{N}^{j,\pi}_{n,k}(x,y) = |\{x_i \in \pi_{n,k}(x,y) \cap V_{\Lambda^j_n}\}|$, the number of steps in the path π from x to y on a fractal of type j, then $(\hat{N}^{1,\pi}_{n,k}(x,y), \ldots, \hat{N}^{M',\pi}_{n,k}(x,y))$ attains the lower boundary of this set of vectors in the usual partial order in $\mathbb{N}^{M'}$. If the path runs along the boundary between two cells we count it as in the fractal with the smallest j (that is the smallest d_w in the shortest path metric). We write $N^j_{n,k}(x,y)$ for the number of type j steps in the shortest path from x to y. We note that, as in [8] Lemma 3.3, $\hat{N}^{j,\pi}_{n,r}(x,y) e^{-(n+r)(S_j+1)} \to 0$ as $r \to \infty$ for $j = 1, \ldots, M'$ and for all such $\pi_{n,k}(x,y)$.

Lemma 3.6 *There exist constants $c_{3.9}, c_{3.10}$ such that for each $x \in G$, there exists a $t_0(x) < 1$ such that*

$$P^x(T_y \leq t) \leq c_{3.9} \exp(-c_{3.10} \sum_{j=1}^{M'} \left(\frac{d^{(j)}(x,y)^{d_w^{(j)}}}{t} \right)^{1/(d_w^{(j)}-1)}), \;\; 0 < t < t_1(x).$$

Proof: To obtain a preliminary estimate, we use the argument of [2] Lemma 4.3, to show that for $y_0 \in H_{\Lambda_n}$ and $0 < t < 1$,

$$P^{y_0}(T_{D^0_{\Lambda_n}(y_0)} \leq t) \leq 1 - \frac{E^{y_0} T_{D^0_{\Lambda_n}(y_0)}}{\sup_{z \in D^0_{\Lambda_n}(y_0)} E^z T_{D^0_{\Lambda_n}(y_0)}} + \frac{t}{\sup_{z \in D^0_{\Lambda_n}(y_0)} E^z T_{D^0_{\Lambda_n}(y_0)}}.$$

We now substitute in the estimates on the mean exit times from Lemma 3.4 to obtain constants $0 < c_1 < 1, c_2$ such that

$$P^{y_0}(T_{D^0_{\Lambda_n}(y_0)} \leq t) \leq 1 - c_1 + c_2 t e^{n d^0_{w,R}(n,y_0)}, \;\; \forall y_0 \in H_{\Lambda_n}, \;\; 0 < t < 1. \tag{3.8}$$

From the definition of W_1^n we have

$$P^x(W_1^n \leq t) = \int_{\partial D^0_{\Lambda_n}(x)} P^y(W_1^n \leq t) P^x(X_{T_0^n} \in dy).$$

As $W_1^n = T_{D^0_{\Lambda_n}(y_0)}, \;\; P^{y_0} - a.s.$ if $y_0 \in H_{\Lambda_n}$, by (3.8) and the fact $d^0_{w,R}(n,x) \geq d^0_{w,R}(n,y_0)$ for all $y_0 \in \partial D^0_{\Lambda_n}(x)$ if $x \notin H_{\Lambda_n}$, we have

$$P^x(W_1^n \leq t) \leq 1 - c_1 + c_2 t e^{n d^0_{w,R}(n,x)}, \;\; 0 < t < 1. \tag{3.9}$$

In order to apply Lemma 3.5 we now need to consider the shortest path from x to y on finer levels $H_{\Lambda_{n+k}}$. The path will exit each cell at a particular boundary point but this path and exit point may vary as we refine the level $n+k$ on which they are viewed. Indeed there may be small fractals in the field which are only seen when $n+k$ gets large. However, by Assumption 2.1 (2), there is a level $\hat{k}_n$ where the shortest path to each boundary point is fixed. For the moment we will assume that we are working on finer levels than $\hat{k}_n$.

We begin by considering the crossing time random variable T_y. As in the discussion before the Lemma, $\pi_{n,k}(x,y)$ denotes a minimal path on $H_{\Lambda_{n+k}}$ from x to y where $e^{-n-1} \leq R(x,y) \leq e^{-n}$. For the minimal path $\{a_0 = x, a_1, \cdots, a_{l_0} = y\}$, when the component of the boundary of the Λ_n-cell containing a_l $(1 \leq l < l_0)$ is not a point, we take $a_l \in V_{\Lambda_n^i}$ for one of the possible types i containing a_l. Given such a path $\pi = \pi_{n,k}(x,y)$ we have the following P^x almost sure inequality for any k_j,

$$T_y \geq \sum_{j=1}^{M'} \sum_{i=1}^{\hat{N}^{j,\pi}_{n,k_j}(x,y)} W_i^{n+k_j}(\pi), \tag{3.10}$$

where $W_i^r(\pi)$ denotes the crossing time random variable for the i-th step in the path π. We will condition on the path $\pi_{n,k}(x,y)$ and write $P^x(\pi_{n,k}(x,y))$ for the probability that the process hits the points in $\pi_{n,k}(x,y)$ successively. If we label the i-th step in the path π on $G^{(j)}$ as x_i^j, then by the definition of $\hat{N}^{j,\pi}_{n,k}$ for such a potential minimal path (that the steps lie in the fractal with smaller d_w if the path runs along the boundary), $d^0_{w,R}(n+k_j, x_i^j) = S_{\tau(j)}+1$, for $i = 1, \ldots, \hat{N}^{j,\pi}_{n,k_j}$. Using this and the fact that $k_j > 0$, we have, by Lemma 3.5, (3.9) and (3.10), that

$$\begin{aligned}
P^x(T_y < t) &\leq \sum_{\pi_{n,k}} P^x\Big(\sum_{j=1}^{M'} \sum_{i=1}^{\hat{N}^{j,\pi}_{n,k_j}} W_i^{n+k_j}(\pi) < t | \pi_{n,k}(x,y)\Big) P^x(\pi_{n,k}(x,y)) \\
&\leq \sum_{\pi_{n,k}} \exp\left(c_3 \sum_{j=1}^{M'} \left((\hat{N}^{j,\pi}_{n,k_j} e^{(n+k_j)(S_j+1)} t)^{1/2} - c_4 \hat{N}^{j,\pi}_{n,k_j} \right) \right) P^x(\pi_{n,k}(x,y)) \\
&\leq \sum_{\pi_{n,k}} \exp\left(-c_3 c_4 \sum_{j=1}^{M'} \hat{N}^{j,\pi}_{n,k_j} \left(1 - (c_4^{-2} (\hat{N}^{j,\pi}_{n,k_j})^{-1} e^{(n+k_j)(S_j+1)} t)^{1/2} \right) \right) P^x(\pi_{n,k}(x,y)).
\end{aligned}$$

For this particular path we now choose the level $k_j(\pi)$ for each $j = 1, \ldots, M'$ so that

$$c_4^{-2} (\hat{N}^{j,\pi}_{n,k_j(\pi)})^{-1} e^{(n+k_j(\pi))(S_j+1)} t \leq 1/4.$$

Hence if we take a constant $c_{3.11}$ large enough and take

$$k_j(\pi) = k_j \equiv \inf\{r : \frac{\hat{N}^{j,\pi}_{n,r}(x,y)}{e^{(n+r)(S_j+1)}} \leq c_{3.11} t\}, \quad j = 1, \ldots, M', \tag{3.11}$$

then

$$\begin{aligned}
P^x(T_y \leq t) &\leq \sum_{\pi_{n,k}} c_5 \exp(-c_6 \sum_{j=1}^{M} \hat{N}^{j,\pi}_{n,k_j(\pi)}(x,y)) P^x(\pi_{n,k}(x,y)), \\
&\leq c_5 \exp(-c_6 \min_{\pi_{n,k}} \sum_{j=1}^{M'} \hat{N}^{j,\pi}_{n,k_j(\pi)}(x,y)).
\end{aligned} \tag{3.12}$$

Set $\bar{d}^{(j)}_{n+k_j}(x,y) = \hat{N}^{j,\pi}_{n,k_j}(x,y)/b_j^{n+k_j}$. Then, by our choice of k_j and the fact that $d_c^{(j)} = \log b_j$, for small $t > 0$ we have

$$b_j^{k_j} \asymp \Big(\frac{t e^{n(S_j+1-d_c^{(j)})}}{\bar{d}^{(j)}_{n+k_j}(x,y)}\Big)^{-d_c^{(j)}/(S_j+1-d_c^{(j)})},$$

where we write $a_m \asymp b_m$ if there exist constants C_1, C_2 such that $C_1 b_m \le a_m \le C_2 b_m$ for all m. Then, as $d_w^{(j)} = (S_j+1)/d_c^{(j)}$,

$$\begin{aligned}
\hat{N}^{j,\pi}_{n,k_j}(x,y) &= \bar{d}^{(j)}_{n+k_j}(x,y) b_j^{n+k_j} \\
&\asymp \bar{d}^{(j)}_{n+k_j}(x,y) b_j^{n} \left(\frac{t e^{n(S_j+1-d_c^{(j)})}}{\bar{d}^{(j)}_{n+k_j}(x,y)} \right)^{-d_c^{(j)}/(S_j+1-d_c^{(j)})} \\
&\asymp \left(\frac{\bar{d}^{(j)}_{n+k_j}(x,y)^{d_w^{(j)}}}{t} \right)^{1/(d_w^{(j)}-1)} .
\end{aligned} \tag{3.13}$$

Now, there is a $t_0'(x) << 1$ such that, in order to minimize (3.12) for $t < t_0'(x)$, we should first take $\bar{d}^{(1)}_{n+k_j}(x,y)$ as small as possible, because $t^{-1/(d_w^{(1)}-1)} > t^{-1/(d_w^{(j)}-1)}$ for all $j > 1$ and for $t \le 1$. The minimum of this quantity is attained by the $d^{(1)}_{n+k_j}(x,y)$ given in (3.1). (Note there is a small difference as we count the smallest d_w on the boundary, however the difference only affects the constants and hence does not change our estimate.) We can inductively choose the minimum $\bar{d}^{(j)}_{n+k_j}(x,y)$, given an i-type path for $i < j$, and obtain

$$(\text{RHS of (3.13)}) \asymp \left(\frac{d^{(j)}_{n+k_j}(x,y)^{d_w^{(j)}}}{t} \right)^{1/(d_w^{(j)}-1)} \asymp \left(\frac{d^{(j)}(x,y)^{d_w^{(j)}}}{t} \right)^{1/(d_w^{(j)}-1)},$$

and the desired result is obtained. We note that we have an alternative expression

$$P^x(T_y \le t) \le c_3 \exp(-c_4 \sum_{j=1}^{M'} N^j_{n,k_j}(x,y)). \tag{3.14}$$

Finally, in order to carry through the above calculations we must ensure that $k_j(\pi) > \hat{k}_n$. Using the definition of k_j this can be done by choosing a sufficiently small t_0'', and thus for $t < t_0(x) = t_0'(x) \wedge t_0''$ we have the upper bound. ∎

Corollary 3.7 *There exists constants $c_{3.12}$ and $c_{3.13} < 1$ such that, for all $x \in G$,*

$$P^x(T_{D^1_{\Lambda_n}(x)} < t) \le c_{3.13}, \quad \textit{for} \quad t \le c_{3.12} e^{-n d^0_{w,R}(n,x)}.$$

Proof: We observe that $T_{D^1_{\Lambda_n}(x)} \ge W_1^n$ and thus from (3.9), we have

$$P^x(T_{D^1_{\Lambda_n}(x)} \le t) \le 1 - c_1 + c_2 t e^{n d^0_{w,R}(n,x)},$$

for all $x \in G$ and all $t < 1$. As $c_1 > 0$, the result holds by taking, for instance, $c_{3.12} = c_1/(2c_2)$ and $c_{3.13} = 1 - c_1/2$. ∎

4 Heat kernel bounds

With the estimate on the tail of the crossing time distribution and a local on-diagonal estimate we can obtain a heat kernel upper estimate in terms of our hierarchy of shortest path metrics. The lower bound requires a more detailed analysis and yields a slightly weaker bound, though still expressed in terms of our shortest path metrics.

Firstly, as noted in Section 2.3, there exists a jointly continuous heat kernel on our fractal field. We also note that, by definition of $d^i_{w,R}(n,x)$ and Assumption 2.1 (2), for each $x \in G$, there exists $N_0(x) \in \mathbb{N}$ such that

$$\begin{aligned} &\lim_{m\to\infty} d^0_{w,R}(m,x) = d^0_{w,R}(n,x) \\ =\ &\lim_{m\to\infty} d^1_{w,R}(m,x) = d^1_{w,R}(n,x) \qquad \text{for all } n \geq N_0(x). \end{aligned} \tag{4.1}$$

We denote this limit by $d_w(x)$.

An upper bound on the norm of the heat semigroup was stated in Proposition 2.16, and hence there is an on-diagonal upper bound for the heat kernel. We begin by giving a local version of this result. On our fractal tilings, in the same way as (3.6), we can obtain a uniform upper estimate for the on-diagonal Green density,

$$g_{D^1_{\Lambda_n}(x)}(y,y) \leq c_{4.1} e^{-n}, \ \ \forall y \in D^0_{\Lambda_n}(x), \ \ \forall x \in G.$$

Using the tail of the crossing time distribution we can convert this into an estimate on the λ-Green density as, by [3] Lemma 4.3,

$$g_\lambda(y,y) \leq P^y(T_{D^1_{\Lambda_n}(x)} \geq \zeta_\lambda)^{-1} g_{D^1_{\Lambda_n}(x)}(y,y),$$

where ζ_λ is an independent exponentially distributed random variable with mean λ^{-1}. Following the proof of [3] Proposition 4.4, Corollary 3.5, we have from the crossing time estimate of Corollary 3.7, that

$$P^y(T_{D^1_{\Lambda_n}(x)} \geq \zeta_\lambda) \geq 1/2 \quad \text{if} \ \ \lambda e^{-n d^0_{w,R}(n,x)} \geq c_{4.2}.$$

Thus, if we choose λ large so that there exists $n \geq N_0(x)$ with $c_{4.2} e^{n d_w(x)} < \lambda \leq c_{4.2} e^{(n+1) d_w(x)}$, then

$$g_\lambda(x,x) \leq c_{4.3} \lambda^{-1/d_w(x)}.$$

Denote $d_s(x)/2 = (d_w(x)-1)/d_w(x)$. Using Tauberian ideas in the same way as in the proof of [3] Lemma 5.2 we get the following.

Lemma 4.1 *There exists a constant $c_{4.4}$ such that for each $x \in G$, $n \geq N_0(x)$ and for $t \leq c_{4.2} e^{-n d_w(x)}$,*

$$p_t(x,x) \leq c_{4.4} t^{-d_s(x)/2}.$$

We note that here we obtain a sharper result than that arising from Proposition 2.16 as we use the local structure of the field.

By extending the standard approach, see [8] Theorem 4.2, we can convert our bound on the tail of the hitting time distribution into an off diagonal upper bound on the heat kernel. Recall that $d_w^{(i)}, d^{(i)}(.,.)$ are the walk dimension and metric on $G^{(i)}$ respectively where $i = \{1, \dots, M'\}$ and $M' < \infty$ denotes the different fractal types. We will also write $d_s^{(i)}$ for the spectral dimension of a fractal of type i.

Theorem 4.2 *There exist constants $c_{4.5}, c_{4.6}$ such that, for each x, y, there are constants $t_1(x), t_1(y)$ < 1 such that for all $0 < t < t_1(x) \wedge t_1(y)$,*

$$p_t(x,y) \leq c_{4.5} t^{-(d_s(x) \vee d_s(y))/2} \exp\left(-c_{4.6} \sum_{j=1}^{M'} \left(\frac{d^{(j)}(x,y)^{d_w^{(j)}}}{t} \right)^{1/(d_w^{(j)}-1)} \right). \tag{4.2}$$

Proof: Recall that $N_{n,k}^j(x,y)$ denotes the number of steps of type j in the shortest path from x to y on $H_{\Lambda_{n+k}}$, where $e^{-n-1} \leq R(x,y) \leq e^{-n}$. It is enough to prove

$$p_t(x,y) \leq c_1 t^{-(d_s(x) \vee d_s(y))/2} \exp(-c_2 \sum_{j=1}^{M'} N_{n,k_j}^j(x,y)), \tag{4.3}$$

where $k_j = k_j(t,x,y) \equiv \inf\{r : N_{n,r}^j(x,y)/\exp((n+r)(S_j+1)) \leq c_{3.11} t\}$. Indeed, we can then use a version of (3.13) to rewrite this estimate in terms of the metric and obtain (4.2).

To show (4.3), we will apply the technique of [8] Theorem 4.2 as follows. Fix $x \neq y$ and let ϵ be sufficiently small. Let $\nu_z = \mu|_{B_\epsilon(z)}$ for any $z \in G$, where $B_\epsilon(z)$ denote a ball of radius ϵ about z. Now, fix $j \leq M'$ and define

$$A_x^j = \{z : N_{n,k_j}^j(x,z) \geq \frac{1}{2} N_{n,k_j}^j(x,y)\}, \quad A_y^j = \{z : N_{n,k_j}^j(y,z) \geq \frac{1}{2} N_{n,k_j}^j(x,y)\}$$

(note that we take the same $k_j = k_j(t,x,y)$ for all the terms). Then,

$$\begin{aligned} P^{\nu_x}(X_t \in B_\epsilon(y)) &= P^{\nu_x}(X_t \in B_\epsilon(y), X_{t/2} \in A_x^j) + P^{\nu_x}(X_t \in B_\epsilon(y), X_{t/2} \in (A_x^j)^c) \\ &= I_1 + I_2. \end{aligned}$$

Take ϵ small enough. Using (3.14) for this particular j,

$$\begin{aligned} P^{\nu_x}(X_{t/2} \in A_x^j) &\leq \max_{z \in B_\epsilon(x)} P^z(X_{t/2} \in A_x^j) \mu(B_\epsilon(x)) \\ &\leq \max_{z \in B_\epsilon(x)} P^z(T_{\partial A_x^j} \leq t/2) \mu(B_\epsilon(x)) \\ &\leq c_3 \exp(-c_4 N_{n,k_j}^j(x,y)) \mu(B_\epsilon(x)). \end{aligned}$$

Let $q(z) = P(X_t \in B_\epsilon(y) | X_{t/2} = z)$, then by the on-diagonal upper bound

$$q(z) = \int_{B_\epsilon(y)} p_{t/2}(z,w) \mu(dw) \leq c_3 t^{-d_s(y)/2} \mu(B_\epsilon(y)).$$

Putting these two estimates together we obtain

$$\begin{aligned} I_1 &= E^{\nu_x}[q(X_{t/2}) : X_{t/2} \in A_x^j] \\ &\leq c_5 \mu(B_\epsilon(x)) \mu(B_\epsilon(y)) t^{-d_s(y)/2} \exp(-c_4 N_{n,k_j}^j(x,y)). \end{aligned} \tag{4.4}$$

For I_2 we use the reversibility to write

$$P^{\nu_x}(X_t \in B_\epsilon(y), X_{t/2} \in (A_x^j)^c) = P^{\nu_y}(X_t \in B_\epsilon(x), X_{t/2} \in (A_x^j)^c).$$

Using the fact that $(A_x^j)^c \subset A_y^j$, we can obtain a similar estimate to (4.4), with the only difference being that we use $d_s(x)$ instead of $d_s(y)$. We can now add the bounds to arrive at

$$P^{\nu_x}(X_t \in B_\epsilon(y)) \leq c_6 \mu(B_\epsilon(x))\mu(B_\epsilon(y)) t^{-(d_s(x)\vee d_s(y))/2} \exp(-c_7 N^j_{n,k_j}(x,y)).$$

Since this estimate holds for all $j \leq M'$, we have

$$\begin{aligned} P^{\nu_x}(X_t \in B_\epsilon(y)) &\leq c_8 \mu(B_\epsilon(x))\mu(B_\epsilon(y)) t^{-(d_s(x)\vee d_s(y))/2} \min_{j=1}^{M'}\{\exp(-c_9 N^j_{n,k_j}(x,y))\} \\ &= c_8 \mu(B_\epsilon(x))\mu(B_\epsilon(y)) t^{-(d_s(x)\vee d_s(y))/2} \exp(-c_9 \max_{j=1}^{M'} N^j_{n,k_j}(x,y)) \\ &\leq c_8 \mu(B_\epsilon(x))\mu(B_\epsilon(y)) t^{-(d_s(x)\vee d_s(y))/2} \exp(-c_{10} \sum_{j=1}^{M'} N^j_{n,k_j}(x,y)). \end{aligned}$$

Taking $\epsilon \to 0$, we obtain (4.3). ∎

For the lower bound we follow a standard route through a sequence of estimates, see [1] Section 3, [8] Section 5, making appropriate extensions but only obtaining weaker results. Firstly, to obtain the on-diagonal lower bound, we use our estimate on the tail of the crossing time distribution.

Lemma 4.3 *There exists a constant $c_{4.7}$ such that for each $x \in G$, $n \geq N_0(x)$ and for $t \leq e^{-nd_w(x)}$,*

$$p_t(x,x) \geq c_{4.7} t^{-d_s(x)/2}.$$

Proof: As in [8] Lemma 5.1, we use our local estimate from Corollary 3.7 to get that for each $x \in G$ and for $t \leq e^{-nd^0_{w,R}(n,x)}$,

$$p_t(x,x) \geq c_{4.7} e^{n(d^1_{w,R}(n,x)-1)}.$$

Using the fact $n \geq N_0(x)$, we can obtain the result. ∎

For the lower bound we require an estimate on the Hölder continuity of the heat kernel. As in the proof of [8] Lemma 5.2, using (2.19), we have

$$|p_t(x,y) - p_t(x,z)|^2 \leq R(y,z) t^{-1} p_t(x,x). \tag{4.5}$$

Lemma 4.4 *There exist constants $c_{4.8}, c_{4.9}$ such that for each $x \in G$, $n \geq N_0(x)$ and $t \leq e^{-nd_w(x)}$,*

$$p_t(x,y) \geq c_{4.8} t^{-d_s(x)/2} \quad \textit{if } R(x,y) \leq c_{4.9} t^{1/d_w(x)}.$$

Remark 4.5 *As this result is symmetric in x and y, we can state it for $n \geq N_0(y)$; then $d_w(x), d_s(x)$ are changed to $d_w(y), d_s(y)$ accordingly.*

Proof: Using (4.5),

$$\begin{aligned} p_t(x,y) &\geq p_t(x,x) - |p_t(x,x) - p_t(x,y)| \\ &\geq p_t(x,x)\left(1 - \sqrt{\frac{R(x,y)}{tp_t(x,x)}}\right) \\ &\geq \frac{1}{2} p_t(x,x) \end{aligned}$$

if $R(x,y) \leq tp_t(x,x)/4$. Thus, by Lemma 4.3,

$$tp_t(x,x)/4 \geq c_{4.7} t t^{-(d_w(x)-1)/d_w(x)} \geq c_{4.9} t^{1/d_w(x)},$$

as required. ∎

We are now ready to use a chaining argument to get a lower bound. This is based on the following chaining Lemma.

Lemma 4.6 *Let $\pi = \{x_i : i = 1, \ldots, N\}$ be a path from $x = x_0$ to $y = x_{N+1}$, $f_i(t), i = 0, \ldots, N$ be decreasing functions of t and assume $p_{t_i}(z_i, z_{i+1}) \geq f_i(t_i)$ for $i = 0, \ldots, N$ and $z_i \in B(x_i, r_i)$, then, if $\sum_{i=0}^N t_i = t$,*

$$p_t(x,y) \geq f_0(t_0) \prod_{i=1}^N f_i(t_i) \mu(B(x_i, r_i)).$$

Proof: The transition density can be bounded below by

$$p_t(x,y) \geq \int_{B_1} \cdots \int_{B_N} p_{t_0}(x, z_1) \ldots p_{t_N}(z_N, y) \mu(dz_1) \ldots \mu(dz_N).$$

The result follows by applying the lower bound on the heat kernel given in the assumption. ∎

A key lemma for tackling the lower bound is the following extension of the near diagonal bound Lemma 4.4. It allows us to control the heat kernel as we move across a boundary where the walk dimension drops.

Lemma 4.7 *There exist constants $c_{4.10}, c_{4.11}$ such that for $y \in K^{(j)} \cap K^{(j')}$, $x \in K^{(j)} \cap D^1_{\Lambda_l}(y)$ with $B(x, e^{-l}) \subset K_j$ and $d_{w,R}(l', y) = S' + 1 < S_j + 1$, where l is given by $e^{-l(S_j+1)} \leq c_{4.9}^{S_j+1} t \leq e^{-(l-1)(S_j+1)}$, and l' is given by $e^{-l'(S'+1)} \leq t \leq e^{-(l'-1)(S'+1)}$, then*

$$p_t(z_0, z_1^b) \geq c_{4.10} t^{-d_s^{(j)}/2} e^{-c_{4.11}(l'-l)},$$

where $z_0 \in B(x, e^{-l}), z_1^b \in B(y, e^{-l'})$.

Proof: We let $x_0 = x$ and note that by the assumptions, y is a point on the boundary with $R(x,y) \leq e^{-l} \leq c_{4.9} t^{1/(S_j+1)}$ and x is such that $e^{-N_0(x) d_w(x)} \leq t < e^{-(N_0(x)-1)d_w(x)}$ ($N_0(x)$ is defined in (4.1)). Thus, for this time, at level $N_0(x)$, we cannot use our near diagonal bound Lemma 4.4. In order to use the near diagonal bound we construct a path from x to y consisting of points $\{x_i \in H_{\Lambda_{l+k}}, i = 1, \ldots, i_*(k), k = 1, \ldots, l' - l + c_1\}$ and times $t(k)$ for $k = 0, \ldots, l' - l + c_1$ where $i_*(0) = 0$ and

$$i_*(k) = \inf\{i : i \geq i_*(k-1), t(k) > e^{-N_0(x_{i+1}) d_w(x_{i+1})}\}, \quad k = 1, \ldots, l' - l + c_1,$$

and $t(k) = e^{-k(S_j+1)} t$. The constant c_1 will be chosen later and we note that there is an $\bar{i} < \infty$ such that $i_*(k) - i_*(k-1) \leq \bar{i}$ for all $k = 1, \ldots, l' - l + c_1$, and hence that there is a c_2 such that

$$c_2 t \leq \sum_{k=1}^{l'-l+c_1} (i_*(k) - i_*(k-1)) t(k) \leq \frac{\bar{i}}{1 - e^{-(S_j+1)}} t.$$

Let $B(x_i, c_3e^{-l-k})$ denote a ball at x_i for $i \in i_*(k-1)+1,\dots,i_*(k)$ with radius c_3e^{-l-k} in the resistance metric. Then, for $i \in i_*(k-1)+1,\dots i_*(k)-1$, we can choose c_3, in order to have

$$R(z_i, z_{i+1}) \le 2c_3e^{-l-k} + R(x_i, x_{i+1}) \le c_{4.9}t(k)^{1/(S_j+1)},$$

and, by our construction of the path and times, we are able to apply Lemma 4.4 to get

$$p_{t(k)}(z_i, z_{i+1}) \ge c_{4.8}t(k)^{-d_s^{(j)}/2},$$

for $z_i \in B(x_i, c_3e^{-l-k})$. For $i = i_*(k)$ we have

$$R(z_i, z_{i+1}) \le c_3(1+e^{-1})e^{-l-k} + R(x_i, x_{i+1}) \le c_{4.9}t(k)^{1/(S_j+1)},$$

which also gives

$$p_{t(k)}(z_i, z_{i+1}) \ge c_{4.8}t(k)^{-d_s^{(j)}/2},$$

for $z_i \in B(x_{i_*(k)}, c_3e^{-l-k}), z_{i+1} \in B(x_{i_*(k)+1}, c_3e^{-l-k-1})$ for $k = 1,\dots,l'-l+c_1$.

Finally, we can choose c_1, to ensure that the point $x_{i_*(l'-l+c_1)} \in H_{\Lambda_{l'+c_1}}$ is close enough to the boundary such that

$$R(z_{i_*(l'-l+c_1)}, z_1^b) \le c_{4.9}t(l'-l+c_1)^{1/(S_j+1)},$$

in order to obtain

$$p_{t(l'-l+c_1)}(z_{i_*(l'-l+c_1)}, z^b) \ge c_{4.8}t(l'-l+c_1)^{-d_s^{(j)}/2},$$

for $z_{i_*(l'-l+c_1)} \in B(x_{i_*(l'-l+c_1)}, c_3e^{-l'})$ and $z_b \in B(y, e^{-l'})$.

We write $B_i(k) = B(x_i, e^{-l-k})$. For the path in the vicinity of the boundary point y, but with $B_i(k) \subset K^{(j)}$, there are constants c_4, c_5 such that

$$c_4t(k)^{d_s^{(j)}/2} \le \mu(B_i(k)) \le c_5t(k)^{d_s^{(j)}/2}, \quad i = i_*(k-1)+1,\dots,i_*(k),$$

for $k = 1,\dots,l'-l+c_1$. Using these estimates in Lemma 4.6, there is a constant $c_8 < 1$ such that

$$\begin{aligned} p_t(z_0, z_1^b) &\ge c_6t(0)^{-d_s^{(j)}/2} \prod_{k=1}^{l'-l+c_1} \prod_{i=1}^{i_*(k)} c_{4.8}t(k)^{-d_s^{(j)}/2}\mu(B_i(k)), \\ &\ge c_7t^{-d_s^{(j)}/2} \prod_{k=1}^{l'-l+c_1} c_8 \end{aligned}$$

as required. ∎

Theorem 4.8 *There exist constants $c_{4.12}, c_{4.13}, c_{4.14}$ such that, for each $x, y \in G$, there are constants $\theta(x,y) \in \mathbb{R}, \rho(x,y), \eta_i(x,y) \ge 0,\ i = 1,\dots,M'$ such that, for $0 < t < 1$,*

$$\begin{aligned} p_t(x,y) &\ge c_{4.12}t^{\theta(x,y)} \prod_{i=1}^{M'} d^{(i)}(x,y)^{-\eta_i(x,y)} \exp\left(-c_{4.13}\rho(x,y)\sum_{i=1}^{M'}\left(\log\left(\frac{d^{(i)}(x,y)}{t}\right)\right)^2\right) \\ &\quad \times \exp\left(-c_{4.14}\sum_{i=1}^{M'}\left(\frac{d^{(i)}(x,y)^{d_w^{(i)}}}{t}\right)^{1/(d_w^{(i)}-1)}\right). \end{aligned} \tag{4.6}$$

Remark 4.9 (1) If $R(x,y) \leq c_{4.9} t^{1/d_w(x)}$ and $t \leq e^{-N_0(x) d_w(x)}$ we will have the near diagonal bound by choice of $\theta(x,y) = -d_s(x)/2$ and $\eta_i(x,y) = 0, \rho(x,y) = 0$.
(2) The dependence of θ, ρ, η_i on x, y is only through the number and type of path segments over different fractals in the shortest path from x to y.

Proof: Fix x, y. The shortest path between the two points $x, y \in G$ can be constructed recursively as follows. Firstly we fix the path on $G^{(1)}$ using the $d^{(1)}$ metric. Once this part of the path is fixed we determine the shortest path on $G^{(2)}$ given the path on $G^{(1)}$, using the $d^{(2)}$ metric, and continue recursively until all constituent fractals have been exhausted. Once we have determined the shortest path we can approximate it, where it may be on different Λ_n-coordinates within different types, by

$$\pi'(x,y) = \{x = x_1^1, \ldots, x^1_{N^1_{l_1}}, x_1^2, \ldots, x^m_{N^m_{l_m}}, x_1^{m+1} = y\},$$

where $x_1^j \subset \partial_{in} G$ denotes a point in a connected component of the boundary for $j = 2, \ldots, m$ and $x_i^j \in H_{\Lambda_{l_j}}$ for some l_j, for $i = 2, \ldots, N^j_{l_j}$. The number of steps in the path will depend upon the scale and we write $N^j_{l_j}(x_1^j, x_1^{j+1})$ for the number of steps in the shortest path from x_1^j to x_1^{j+1} on $H_{\Lambda_{l_j}}$ (we may drop the (x_1^j, x_1^{j+1}) when it is clear). The j-th segment of the path, where it lies in a particular fractal on level l_j, is $\{x_1^j, \ldots, x^j_{N^j_{l_j}}, x_1^{j+1}\}$. If $x = x_1^1$ is on a boundary, we regard it as a separate segment with $N^1_{l_1} = 0$ only if $d_w^{(\tau(1))} < d_w^{(\tau(2))}$ (similarly for $y = x_1^{m+1}$). Note that in this proof our notation differs from the exit time distribution upper bound proof, as here the j refers to the j-th segment of path, not the j-th type of fractal. We will also modify the notation $\tau(j)$ and use it to denote the type of the j-th segment. If the shortest path runs along the boundary $\Gamma_{kk'}$, then we consider our path π' to run in the fractal K_k on the side where $d_w^{(k)} > d_w^{(k')}$.

Let $a_j = d^{(\tau(j))}(x_1^j, x_1^{j+1})/(d^{(\tau(j))}(x,y)M')$. The key idea is to split the shortest path up in such a way that over a fractal of a particular type we examine the fractal at a particular depth. Fix a segment j, and find n_j, m_j such that

$$e^{-(n_j+1)(S_{\tau(j)}+1)} \leq a_j t \leq e^{-n_j(S_{\tau(j)}+1)},$$

and

$$e^{-m_j-1} \leq R(x_1^j, x_1^{j+1}) \leq e^{-m_j}.$$

Now let

$$D_j = \frac{R(x_1^j, x_1^{j+1})^{S_{\tau(j)}+1}}{a_j t}.$$

We first assume that $D_j \geq c_1 = c_{4.9}^{S_{\tau(j)}+1}$ and hence there is a $c_2 \geq 0$ such that $n_j \geq m_j + c_2$. Thus, for $\hat{c} > 0$, we can choose a $k_j > 0$ such that

$$k_j = \inf\{k : \frac{N^j_{m_j+k}(x_1^j, x_1^{j+1})}{e^{(m_j+k)(S_{\tau(j)}+1)}} \leq \hat{c} a_j t\}.$$

With this choice we treat the j-th path segment on level $l_j = m_j + k_j$. As $k_j > 0$, there is a constant c_3 such that

$$c_3 N^j_{l_j}(x_1^j, x_1^{j+1})^{-1} e^{(S_{\tau(j)}+1)k_j} \leq D_j \leq \hat{c} N^j_{l_j}(x_1^j, x_1^{j+1})^{-1} e^{(S_{\tau(j)}+1)k_j}. \quad (4.7)$$

Thus, letting $s_j = a_j t / N^j_{l_j}(x^j_1, x^{j+1}_1)$, we have

$$R(x^j_1, x^{j+1}_1)e^{-k_j} \le (\hat{c}s_j)^{1/(S_{\tau(j)}+1)}, \tag{4.8}$$

and observe that by the definition of D_j, (4.7), and the fact that $N^j_{l_j} \asymp b_j^{k_j} \asymp e^{-k_j d_c^{(\tau(j))}}$, where $d_c^{(i)} = \log b_i$ is the chemical exponent, there are constants c_4, c_5 such that

$$c_4 N^j_{l_j}(x^j_1, x^{j+1}_1) \le \left(\frac{R(x^j_1, x^{j+1}_1)^{S_{\tau(j)}+1}}{a_j t} \right)^{d_c^{(\tau(j))}/(S_{\tau(j)}+1-d_c^{(\tau(j))})} \le c_5 N^j_{l_j}(x^j_1, x^{j+1}_1). \tag{4.9}$$

Also, by rearranging in the definition of k_j and m_j, we have

$$0 < l_j \asymp \frac{\log (M' d^{(\tau(j))}(x,y)/t)}{S_j + 1 - d_c^j}. \tag{4.10}$$

For our segment j, under the assumption that $D_j \ge c_1$, there are three cases to consider depending on the size of $d_w^{(\tau(j))}$ relative to that for the neighbouring fractals. In Figure 6 a tiling of SG(2) and SG(4), where $d_w^{(SG(2))} < d_w^{(SG(4))}$, is shown. The three cases for segments are the pieces x to x^2_1, x^* to y and x^2_1 to x^3_1 respectively.

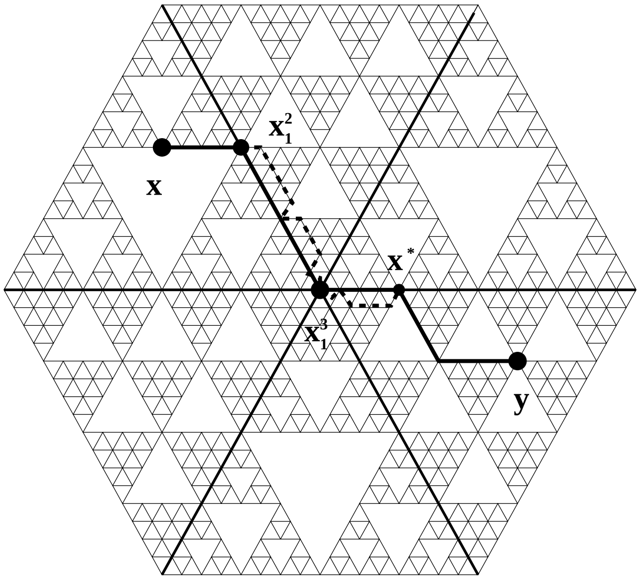

Figure 6: A short path from x to y showing the different cases for short path segments

(1) $S_{\tau(j)} + 1 = d^1_{w,R}(l_j, x^j_i)$ *for all* x^j_i *in the* j*-th segment.*
The shortest path along the j-th segment can be realized as $\{x^j_i : i = 1, \dots, N^j_{l_j}(x^j_1, x^{j+1}_1)\}$. By our construction so far we have

$$R(x^j_i, x^j_{i+1}) \le c_6 e^{-m_j - k_j} \le c_7 R(x^j_1, x^{j+1}_1)e^{-k_j} \le c_7 (\hat{c}s_j)^{1/(S_{\tau(j)}+1)}.$$

We set $\epsilon_j = e^{-l_j}$ and write $B^j_i = B_{\epsilon_j}(x^j_i)$ for the resistance ball of radius ϵ_j about x^j_i. As $d^1_{w,R}(l_j, x^j_i) = S_{\tau(j)} + 1$ for $1 \le i \le N^j_{l_j}$, we have that if $z^j_i \in B^j_i$ and $z^j_{i+1} \in B^j_{i+1}$ for $i = 1, \dots, N^j_{l_j}(x^j_1, x^{j+1}_1)$, then, by (4.8),

$$R(z^j_i, z^j_{i+1}) \le 2\epsilon_j + R(x^j_i, x^j_{i+1}) \le c_8 (\hat{c}s_j)^{1/(S_{\tau(j)}+1)}.$$

By choice of $\hat{c}$ we have

$$R(z_i^j, z_{i+1}^j) \leq c_{4.9} s_j^{1/(S_{\tau(j)}+1)}.$$

As $d_w(x_i^j) = S_{\tau(j)}$ and $t < e^{-N_0(x_i^j)d_w(x_i^j)}$, for all $x_i^j, i = 1, \ldots, N_{l_j}^j$, we can apply Lemma 4.4 to obtain

$$p_{s_j}(z_i^j, z_{i+1}^j) \geq c_{4.8} s_j^{d_s^{(\tau(j))}/2}, \tag{4.11}$$

for all $z_i^j \in B_i^j, z_{i+1}^j \in B_{i+1}^j$.

We note that in the segment j, by the choice of ϵ_j, the measure of the ball can be controlled by

$$c_9 s_j^{d_s^{(\tau(j))}/2} \leq \mu(B_i^j) \leq c_{10} s_j^{d_s^{(\tau(j))}/2}, \quad i = 1, \ldots, N^j.$$

and we have that $c_{11}(j) = \mu(B_i^j) s_j^{d_s^{(\tau(j))}/2}$ is bounded above and below by constants for $i = 1, \ldots, N^j$. Thus, from Lemma 4.6, we have the following lower bound on the transition density for crossing between two balls at each end of a segment of this type,

$$p_{a_j t}(z_1^j, z_1^{j+1}) \geq c_{12}(a_j t)^{-d_s^{(\tau(j))}/2} \exp(-c_{13} N_{l_j}^j), \tag{4.12}$$

for all $z_1^a \in B(x_1^a, e^{-l_j})$, for $a = j, j+1$.

(2) $S_{\tau(j)} + 1 > d_{w,R}^1(l_j, x_i^j)$ for a set of x_i^j at one end of the j-th segment.

We will assume that $d_{w,R}^1(l_j, x_1^j) = S_{\tau(j)} + 1$, but if not, we can apply the following proof at the beginning of the path segment. In this situation, we may not be able to use Lemma 4.4 right up to the boundary but we can use Lemma 4.7 at the boundary. We assume here that x_1^{j+1} is the first point in the j-th segment after x_1^j which is on a boundary. If this is not the case (as in Figure 6), we can follow the same idea up to the first point on a boundary (the point x^*) and then use case (3).

First note that there exists $i_*^j(0) = \inf\{i - 1 : s_j > e^{-N_0(x_i^j)d_w(x_i^j)}\}$ and $i_*^j(0) \leq N_{l_j}^j$ and by Lemma 4.4 we will have (4.11) for $1 \leq i \leq i_*^j(0)$.

We determine the appropriate time scale at the boundary and need l_j' in place of l_j, where

$$e^{-(l_j'+1)(S'+1)} \leq s_j \leq e^{-l_j'(S'+1)},$$

where $S' + 1 = d_w(x_1^{j+1})$. Note that $S' \neq S_{\tau(j+1)}$ and hence $l_j' \neq l_{j+1}$ in general. By definition of case (2) we have $l_j < l_j'$.

For $i = 1, \ldots, N^1 - i_*^1(0)$ we have $c_{11}(j)$ as in case (1). For the final part of the path use Lemma 4.7. We note that from the definition of l_j' and the estimate on l_j in (4.10),

$$l_j' - l_j \asymp \frac{S_j - S'}{(S'+1)(S_j + 1 - d_c^j)} \log \frac{M' d^{(\tau(j))}(x,y)}{t}.$$

Using these estimates in Lemma 4.6 and Lemma 4.7 we have the following lower bound for a segment of type (2),

$$\begin{aligned} p_{a_j t}(z_1^j, z_1^{j+1}) &\geq \prod_{i=1}^{N^j - i_*^j(0)} s_j^{-d_s^{(\tau(j))}/2} \prod_{i=2}^{N^j - i_*^j(0)} \mu(B_i^j) c_{4.10} s_j^{-d_s^{(\tau(j))}/2} \exp(-c_{4.11}(l_j - l_j')) & (4.13) \\ &\geq c_{14} s_j^{-d_s^{(\tau(j))}/2} \left(\frac{d^{(\tau(j))}(x,y)}{t}\right)^{\eta} e^{-c_{15} N_{l_j}^j} & (4.14) \end{aligned}$$

for some $\eta = \eta(x, y) > 0$ for all $z_1^j \in B(x_1^j, e^{-l_j}), z_1^{j+1} \in B(x_1^{j+1}, e^{-l'_j})$.

(3) $S_{\tau(j)} + 1 > d^1_{w,R}(l_j, x_i^j)$ *for a subset of* x *in the* j*-th segment.*

The worst case occurs when the shortest path runs along the boundary between two components and we concentrate on this situation. If it is only for a portion of the segment (as in Figure 6 from x_1^3 to x^*) we can incorporate this with case (2) to get a suitable lower bound for the path segment. We can define a new path $\bar{\pi}$ across this segment (the dotted line in Figure 6) with the property that $|\bar{\pi}(x_1^j, x_1^{j+1})| \leq c_{16} N^j_{l_j}(x_1^j, x_1^{j+1})$ and such that there is a sequence $\bar{x}_i^j$ of at most $c_{17} k_j = c_{17}(l_j - m_j)$ boundary points at which the path must be at the boundary. In between these points our path consists of a sequence of points $\{y_i^j \in H_{\Lambda_{l_j}}\}$ where y_i^j are chosen such that $B(y_i^j, e^{-l_j}) \subset K_{\tau(j)}$. We note that this can be done by the choice of $k_j > 0$ which ensures that the steps are on a finer level than the size of the component $K_{\tau(j)}$. The case where we are near the intersection with other fractals will be considered later.

We note that there is an l'_j such that $(\hat{c} s_j)^{1/d_w(x)} \asymp e^{-l'_j}$ and hence at $\bar{x}_i^j$ we can only apply Lemma 4.4 on the level l'_j. We choose the points of our new path to use Lemma 4.7 to move from the resistance ball of radius $e^{-l'_j}$ at $\bar{x}_i^j$ on the boundary to that of radius e^{-l_j} centred at the neighbouring point $\bar{y}_i^j$ and then along inside the fractal $K_{\tau(j)}$ until it must return to $\bar{x}_{i+1}^j$.

We note that by the definition of l'_j we have

$$l'_j - l_j \asymp \left(\frac{1}{d_w(\bar{x}_i^j)} - \frac{1}{S_{\tau(j)} + 1}\right) \log(s_j). \tag{4.15}$$

The total number of steps on the modified path is controlled by $N^j_{l_j}$.

Finally we note that it may not be possible to choose $y_i^j \in H_{\Lambda_{l_j}}$ if the path approaches a vertex. We follow a similar approach to case (2), with a path that moves inside the fractal $K_{\tau(j)}$ but with the points $\bar{y}_i^j$ chosen on finer levels than l_j in the same way as the proof of Lemma 4.7 (see the dotted path near to x_1^3 in Figure 6). The path must move into the interior and out again and the transition probability can be controlled using Lemma 4.6 and Lemma 4.7 again.

Using the chaining Lemma 4.6, (4.15) and the definition of l_j we have

$$\begin{aligned} p_{a_j t}(z_1^j, z_1^{j+1}) &\geq c_{18}(a_j t)^{-d_s^{(\tau(j+1))}/2}\left(\prod_{k=1}^{l'_j - l_j} c_{19}(k)\right)^{c_{17} k_j} e^{-c_{20} N^j_{l_j}} \\ &\geq c_{21} e^{-c_{21} l_j^2}(a_j t)^{-d_s^{(\tau(j+1))}/2} e^{-c_{20} N^j_{l_j}}, \\ &\geq c_{21}(a_j t)^{-d_s^{(\tau(j+1))}/2} e^{-c_{22}(\log \frac{d^{(\tau(j))}(x,y)}{t})^2} e^{-c_{20} N^j_{l_j}}, \end{aligned} \tag{4.16}$$

for $\bar{z}_1^a \in B(x_1^a, e^{-l'_j})$ for $a = j, j+1$.

By combining cases (2) and (3) we can control the path which meets the boundary at some internal point of the shortest path on the j-th segment and moves along it after that. Again this follows from the chaining Lemma 4.6,

$$p_{a_j t}(z_1^j, z_1^{j+1}) \geq c_{23}(a_j t)^{\theta} d^{(\tau(j))}(x, y)^{\eta} e^{-c_{24}(\log \frac{d^{(\tau(j))}(x,y)}{t})^2} e^{-c_{25} N^j_{l_j}},$$

for $\bar{z}_1^a \in B(x_1^a, e^{-l'_j})$ for $a = j, j+1$.

Finally we return to our initial assumption, that $D_j \geq c_1$, and consider the case where $D_j < c_1$. In this case we choose $k_j = 0$ and split into two cases.

Firstly, if $S_{\tau(j)}+1 = d^1_{w,R}(n_j,x)$ for either $x = x_1^j$ or $x = x_1^{j+1}$, by the choice of c_1 we can apply Lemma 4.4 to get that

$$p_{a_j t}(z_1^j, z_1^{j+1}) \geq c_{4.8}(a_j t)^{-S_{\tau(j)}/(S_{\tau(j)}+1)}, \;\; z_1^a \in B(x_1^a, e^{-m_a}), \; a = j, j+1.$$

The other case to consider is where $S_{\tau(j)}+1 > \max\{d^1_{w,R}(n_j,x) : x = x_1^j, x_1^{j+1}\}$. Let $S_j'+1 = \min\{d^1_{w,R}(n_j,x) : x = x_1^j, x_1^{j+1}\}$, and find n_j' such that

$$c_{26} e^{-(n_j'+1)(S_j'+1)} \leq a_j t \leq c_{26} e^{-n_j'(S_j'+1)}.$$

Thus if $n_j' < m_j$, by choice of c_{26} we can apply Lemma 4.4 to obtain

$$p_{a_j t}(z_1^j, z_1^{j+1}) \geq c_{4.8}(a_j t)^{-S_j'/(S_j'+1)}, \;\; z_1^a \in B(x_1^a, e^{-m_a}), \; a = j, j+1.$$

Alternatively we have $n_j < m_j \leq n_j'$, in which case we use the approach for cases (2) and (3) above. We allow a time $e^{-n_j'(S_j'+1)}$ to move from x_1^j to a neighbouring point at resistance distance $e^{-n_j'}$ on a path which has asymptotically the shortest number of steps. We use Lemma 4.7 to make the move. In this case we have to move from n_j' to $m_j + c_{27}$ and back again. We omit the details which lead to

$$p_{a_j t}(z_1^j, z_1^{j+1}) \geq c_{28}(a_j t)^{\frac{-S_j'}{S_j'+1}} \prod_{k=1}^{n_j'-m_j} c_{29}(k).$$

Using the fact that $m_j < n_j'$, there is a constant χ such that

$$p_t(z_1^j, z_1^{j+1}) \geq c_{30}(a_j t)^{\chi},$$

for $z_1^a \in B(x_1^a, e^{-m_a}), \; a = j, j+1$.

Finally we put these two cases together and use the fact that $k_j = 0$, so that $N_{l_j}(x_1^j, x_1^{j+1}) = 1$, to write

$$p_t(z_1^j, z_1^{j+1}) \geq c_{30}(a_j t)^{\chi} e^{-c_{31} N_{l_j}(x_1^j, x_1^{j+1})}, \tag{4.17}$$

for $z_1^a \in B(x_1^a, e^{-m_a}), \; a = j, j+1$.

We are now ready to chain along the whole path. Note that in the case where x, y are both in the interior of one fractal component we just use the standard chaining argument.

We now proceed inductively starting from the case where our path moves between 2 different fractals, $m = 2$ and assuming $S_1 > S_2$ and $l_1 < l_2$ and the path does not touch the boundary except at the final point (we are in cases (1) and (2)). Writing $N^j = N^j_{l_j}(x_1^j, x_1^{j+1})$, and $s_j = a_j t / N^j$, we have by the chaining Lemma 4.6,

$$p_t(x,y) \geq c_{14} s_1^{-d_s^{(\tau(1))}/2} \left(\frac{d^{(\tau(1))}(x,y)}{t}\right)^{\eta} e^{-c_{15}N^1} \mu(B(x_1^2, e^{-l_2})) c_{12} s_2^{-d_s^{(\tau(2))}/2} e^{-c_{13}N^2}.$$

Note if x was on the boundary we would be in the $m = 1$ case, and if y is on the boundary we are in the $m = 2$ case but with $N^2 = 0$.

If $D_1 > c_1$, then by (4.10), we have $d^{(\tau(1))}(x,y) \geq t/M'$, if not we can use to (4.17), to obtain

$$\begin{aligned} p_t(x,y) &\geq c_{32} \min\{t^{-d_s^{(\tau(1))}/2}, t^{\chi}\} e^{-c_{33}\sum_{j=1}^2 N^j}. \\ &\geq c_{34} t^{\theta'} e^{-c_{33}\sum_{j=1}^2 N^j}. \end{aligned} \tag{4.18}$$

It is straightforward to show that if we had $l_1 > l_2$, the exponent for leading term in t in (4.18) would be $\theta' = \max\{-d_s^{(\tau(2))}/2, \chi\}$.

If the path is along the boundary for a segment

$$p_t(x,y) \geq c_{36} t^{\theta'} e^{-c_{35}\sum_{j=1}^{2}(\log \frac{d^{(\tau(j))}(x,y)}{t})^2} e^{-c_{33}\sum_{j=1}^{2} N^j}. \tag{4.19}$$

We now incorporate more segments using the chaining Lemma 4.6. To apply this we need bounds on the transition probabilities across each different segment $p_{a_j t}(z_1^z, z_1^{j+1}) \leq f_j(a_j t)$ and on the measure of the boundary ball. For the bounds on the transition probabilities between boundary points we use the bounds from cases (1), (2) and (3) when $D_j \geq c_1$, otherwise we use (4.17). Each time we arrive at a boundary we may need to adjust the level from l'_j to an l'_{j+1}. This is done by considering the boundary ball to be in the medium with smaller d_w and using a near diagonal estimate. A lower bound on the measure of the boundary ball can be written $\mu(B_1^j) \geq c_{37} s_j^{\eta'} = c_{38}(t/d^{(\tau(j))}(x,y))^{\eta'}$ for some $\eta' > 0$. Combining these two sets of bounds we have constants $\theta \in \mathbb{R}, \rho, \eta_i \geq 0$, for $i = 1, \ldots, M'$ (which may depend on m, the number of different segments in the path), such that

$$p_t(x,y) \geq c_{39} t^{\theta} \prod_{j=1}^{m} d^{(\tau(j))}(x,y)^{-\eta_j} e^{-c_{40}\sum_{j=1}^{m}(\log \frac{d^{(\tau(j))}(x,y)}{t})^2} e^{-c_{41}\sum_{j=1}^{m} N^j}.$$

Using (4.9), we can write this lower estimate as

$$\begin{aligned} p_t(x,y) \quad \geq \quad & c_{39} t^{\theta(x,y)} \prod_{i=1}^{M'} d^{(i)}(x,y)^{-\eta_i(x,y)} e^{-c_{40}\rho(x,y)\sum_{i=1}^{M'}(\log \frac{d^{(i)}(x,y)}{t})^2} \\ & \times \exp\left(-c_{42}\sum_{j=1}^{m}\left(\frac{R(x_1^j, x_1^{j+1})^{S_{\tau(j)}+1}}{a_j t}\right)^{d_c^{(\tau(j))}/(S_{\tau(j)}+1-d_c^{(\tau(j))})}\right). \end{aligned} \tag{4.20}$$

Now, using the construction of the shortest path metrics, and the fact that x_1^j, x_1^{j+1} are in this shortest path

$$d^{(\tau(j))}(x_1^j, x_1^{j+1}) \asymp R(x_1^j, x_1^{j+1})^{d_c^{(\tau(j))}}.$$

Using this we have

$$\left(\frac{R(x_1^j, x_1^{j+1})^{S_{\tau(j)}+1}}{a_j t}\right)^{d_c^{(\tau(j))}/(S_{\tau(j)}+1-d_c^{(\tau(j))})} \asymp \left(\frac{d^{(\tau(j))}(x_1^j, x_1^{j+1})^{d_w^{(\tau(j))}}}{a_j t}\right)^{1/(d_w^{(\tau(j))}-1)}.$$

From the definition of a_j we have $d^{(\tau(j))}(x_1^j, x_1^{j+1}) = a_j d^{(\tau(j))}(x,y) M'$ and hence

$$\left(\frac{d^{(\tau(j))}(x_1^j, x_1^{j+1})^{d_w^{(\tau(j))}}}{a_j t}\right)^{1/(d_w^{(\tau(j))}-1)} \asymp a_j \left(\frac{d^{(\tau(j))}(x,y)^{d_w^{(\tau(j))}} M'^{d_w^{(\tau(j))}}}{t}\right)^{1/(d_w^{(\tau(j))}-1)}.$$

Thus the exponent in the exponential term becomes

$$\begin{aligned} \sum_{j=1}^{m}\left(\frac{R(x_1^j, x_1^{j+1})^{S_{\tau(j)}+1}}{a_j t}\right)^{\frac{d_c^{(\tau(j))}}{S_{\tau(j)}+1-d_c^{(\tau(j))}}} \quad & \asymp \quad \sum_{i=1}^{M'}\sum_{j=1}^{m} a_j M'^{\frac{d_w^{(\tau(j))}}{d_w^{(\tau(j))}-1}} I_{\{\tau(j)=i\}} \left(\frac{d^{(\tau(j))}(x,y)^{d_w^{(\tau(j))}}}{t}\right)^{\frac{1}{d_w^{(\tau(j))}-1}} \\ & \asymp \quad \sum_{i=1}^{M'} a'_i M'^{\frac{d_w^{(i)}}{d_w^{(i)}-1}} \left(\frac{d^{(i)}(x,y)^{d_w^{(i)}}}{t}\right)^{\frac{1}{d_w^{(i)}-1}}, \end{aligned}$$

where $a'_i = \sum_{j=1}^m a_j I_{\{\tau(j)=i\}} = 1/M'$. Thus $a'_i M'^{d_w^{(i)}/(d_w^{(i)}-1)} = M'^{1/(d_w^{(i)}-1)} \leq M'$, giving the result. ∎

Remark 4.10 (1) We believe that the on-diagonal part of the above estimate can be improved but we do not know how to do this via the usual chaining argument.
(2) The term $e^{-c_{4.13}(\log(d^{(i)}(x,y)/t))^2}$ arises from shortest paths which move along a boundary. If the points x and y are such that the shortest path between them only intersects component boundaries at single points it is not required.

5 Large deviations

Now that we have estimates for the heat kernel (4.2) and (4.6), it is reasonably straight forward, using the well-known results of [26] (see also [4]), to obtain large deviation results for the process. In this section we will briefly explain these results. Throughout this section we will write $d(.,.) = d^{(1)}(.,.)$ for the shortest path metric defined in Section 3 and $d_w = d_w^{(1)}$, for the smallest value of the walk dimension. We note that by construction $d(.,.)$ is a pseudo-metric over the whole field as points in connected components of the closure of the complement of $K^{(1)}$ in G are identified. The large deviation results presented here hold for all points in the field but are determined by the component of the fractal field with smallest d_w and this associated pseudo-metric.

For $T > 0$ and $x \in G$, let $\Omega_x = \{\phi \in C([0,T] \to G) : \phi(0) = x\}$ with uniformly continuous topology. For $\phi \in \Omega_x$, we say ϕ is absolutely continuous if for each $\epsilon > 0$, there exists $\delta > 0$ such that $\sum_{i=1}^n d(\phi(t_i), \phi(t_{i-1})) < \epsilon$ for any n and any disjoint collection of intervals $\{(t_{i-1}, t_i)\}_{i=1}^n$ in $[0,T]$ whose lengths satisfy $\sum_i (t_i - t_{i-1}) < \delta$. It can be proved by routine arguments that if ϕ is absolutely continuous, then $\dot\phi(t) \equiv \lim_{s\to t} \frac{d(\phi(s),\phi(t))}{|s-t|}$ exists for a.e. $t \in [0,T]$, $\dot\phi \in \mathbb{L}^1([0,T], dt)$ and $\int_0^T \dot\phi(t)dt$ is the length of the path $\{\phi(s) : 0 \leq s \leq T\}$. Now, for $\phi \in \Omega_x$, define $I_x(\phi)$ as

$$I_x(\phi) = \begin{cases} \int_0^T (\dot\phi(t))^{d_w/(d_w-1)} dt, & \phi \text{ is absolutely continuous,} \\ \infty, & \text{otherwise.} \end{cases} \tag{5.1}$$

When $\phi \in C([0,T] \to G)$ (no restriction for $\phi(0)$), we denote the corresponding rate function by $I(\phi)$.

For $\Delta : 0 = t_0 < t_1 < t_2 \cdots < t_m = T$ and $\phi \in \Omega_x$, we set $\Pi_\Delta \phi = \{\phi(t_1), \cdots, \phi(t_m)\}$. Also, define $\phi_\Delta \in \Omega_x$ by taking points $\{\phi(t_j)\}$ and joining them successively by geodesic paths. If there is more than one geodesic path between two such points, it is immaterial which one is chosen. Thus, ϕ_Δ is a piecewise geodesic path and $\phi_\Delta(t_j) = \phi(t_j)$ $(0 \leq j \leq m)$. We then have the following (see [4] Lemma 2.4 or [5] Lemma 3.2 for the proof).

Lemma 5.1 *(a) On $C([0,T] \to G)$ we have*

$$\inf_{\substack{\phi(\alpha)=a,\\ \phi(\beta)=b}} I(\phi) = \Big(\frac{d(a,b)^{d_w}}{\beta - \alpha}\Big)^{1/(d_w-1)},$$

where the infimum is attained by the geodesic path on G.
(b) On $C_x([0,T] \to G)$ we have

$$\inf_{\substack{\phi(t_i)=x_i,\\ i=1,\cdots,m}} I_x(\phi) = I_x(\phi_\Delta) = \sum_{i=1}^{m} \Big(\frac{d(x_i,x_{i-1})^{d_w}}{t_i - t_{i-1}}\Big)^{1/(d_w-1)}$$

where $\Delta : 0 = t_0 \le t_1 \le \cdots \le t_m \le T$, $x_0 = x, x_1, \cdots, x_m \in G$ and ϕ_Δ is a piecewise geodesic path with $\phi_\Delta(t_j) = x_j$ $(0 \le j \le m)$.

Now by a standard argument, we have the following (see, for example, [4] Lemma 2.5).

Lemma 5.2 *(1) The function $I_x(\phi)$ is lower semi-continuous. Further, for every $N > 0$, $\{\phi : I_x(\phi) \le N\}$ is compact.*
(2) If $C \subset \Omega_x$ is closed in Ω_x, then

$$\lim_{\delta\to 0} \inf_{\phi\in C_\delta} I_x(\phi) = \inf_{\phi\in C} I_x(\phi),$$

where $C_\delta = \{\phi \in \Omega_x : \|\phi - \psi\| < \delta \quad \text{for some } \psi \in C\}$.

Here $\|\phi - \psi\| = \sup_{t\le T} d(\phi(t), \psi(t))$.

We have the following estimates.

Proposition 5.3 *(1) For each $x, y \in G$,*

$$\begin{aligned} c_{5.1} d(x,y)^{\frac{d_w}{d_w-1}} &\le -\limsup_{\epsilon\to 0} \epsilon^{\frac{1}{d_w-1}} \log p_\epsilon(x,y) \\ &\le -\liminf_{\epsilon\to 0} \epsilon^{\frac{1}{d_w-1}} \log p_\epsilon(x,y) \le c_{5.2} d(x,y)^{\frac{d_w}{d_w-1}}. \end{aligned}$$

(2) For each $K_0 > 0$ and $T_0 < 1$, there exist $c_{5.3}, c_{5.4} > 0$ such that the following holds,

$$0 < -t^{\frac{1}{d_w-1}} \log p_t(x,y) \le c_{5.3} T_0^{\frac{1}{d_w-1}} ((\log 1/T_0)^2 + c_{5.4}) \qquad \text{for all } t \le T_0, \hat{d}(x,y) \le K_0,$$

where $\hat{d}(.,.)$ is a distance on G given in Section 3.
(3) There exist $c_{5.5}, c_{5.6} > 0$ which depend only on $0 < T_0 < 1$ such that

$$P^x(\sigma_{B_d(x,r)} \le t) \le c_{5.5} \exp\Big(- c_{5.6} \Big(\frac{r^{d_w}}{t}\Big)^{\frac{1}{(d_w-1)}} \Big),$$

for all $0 < r, 0 < t \le T_0, x \in G$, where $B_d(x,r)$ is a ball centred x radius r with respect to the $d^{(1)}$ metric and $\sigma_A = \inf\{t \ge 0 : \tilde{X}_t \notin A\}$.

Proof: Using (4.2), (4.6), (1) is straightforward. (2) is a consequence of our uniform lower bound Theorem 4.8. (3) is an easy consequence of Lemma 3.6. ∎

Let P_ϵ^x be the law for $X^x(\epsilon t)$ where X^x is the process starting at x. We now state our large deviation theorem.

Theorem 5.4 *There exist $c_{5.7}, c_{5.8} > 0$ such that for each $A \subset \Omega_x$,*

$$\begin{aligned} -c_{5.7} \inf_{\phi\in IntA} I_x(\phi) &\le \liminf_{\epsilon\to 0} \epsilon^{1/(d_w-1)} \log P_\epsilon^x(A) \\ &\le \limsup_{\epsilon\to 0} \epsilon^{1/(d_w-1)} \log P_\epsilon^x(A) \le -c_{5.8} \inf_{\phi\in ClA} I_x(\phi). \end{aligned}$$

Using Lemma 5.1, Lemma 5.2 and Proposition 5.3, the proof of this theorem is almost the same as that of the classical case (see, for instance, [26]) though we do not necessarily have a limit in Proposition 5.3. In [4] Section 2, the proof of this type of large deviation theorem is given and thus we omit the proof here. Note that, even though we have not proved compact uniform convergence in Proposition 5.3 (1), we can change limits and integrals in the proof of Theorem 5.4 using Proposition 5.3 (2) (by the dominated convergence theorem).

Remark 5.5 (1) We cannot take $c_{5.7} = c_{5.8}$ in general; as we explain in the Appendix (Theorem A.7), even for Brownian motion on the Sierpinski gasket, there are oscillations which mean that we cannot take $c_{5.7} = c_{5.8}$.
(2) The results hold for all $x \in G$, though the probabilities only have a non-zero limit when the paths must pass through components of type 1.

Finally, we mention that we can "observe" the following asymptotic expansion of $\log P_\epsilon^x(A)$ using (4.2), (4.6). Define $I_x^{(l)}(\cdot)$ in the same way as (5.1) using $d^{(l)}(.,.)$ and $d_w^{(l)}$ (thus $I_x^{(1)}(\phi) = I_x(\phi)$). Then, for each $A \subset \Omega_x$,

$$\log P_\epsilon^x(A) \asymp -\sum_{l=1}^{M'} \epsilon^{-1/(d_w^{(l)}-1)} \inf_{\phi\in A} I_x^{(l)}(\phi).$$

As $\epsilon^{1/(d_w-1)} < \epsilon^{1/(d_w^{(i)}-1)}$ for $\epsilon < 1$ and $i > 1$, only the rate function $I_x = I_x^{(1)}$ appears in the asymptotic behaviour of $\epsilon^{1/(d_w-1)} \log P_\epsilon^x(A)$ in Theorem 5.4.

A Appendix

In this appendix, we will briefly summarize properties of nested fractals and their Brownian motions, as introduced by Lindstrøm ([20]). See [1], [14], [15], [19], etc. for details.
Let $S = \{1, 2, \cdots, N\}$ $(N < \infty)$ and let $\{\Psi_i\}_{i\in S}$ be *similitudes* on $\mathbb{R}^d$, i.e., $\Psi_i(x) = \alpha^{-1}U_i x + \beta_i$, $x \in \mathbb{R}^d$ where U_i is a unitary map, $\alpha > 1$ and $\beta_i \in \mathbb{R}^d$. We assume the *open set condition* for $\{\Psi_i\}_{i\in S}$, that is, there is a non-empty, bounded open set V such that $\{\Psi_i(V)\}_{i\in S}$ are disjoint and $\cup_{i\in S}\Psi_i(V) \subset V$. As $\{\Psi_i\}_{i\in S}$ is a family of contraction maps, there exists a unique non-void compact set $\hat{K}$ such that $\hat{K} = \cup_{i\in S}\Psi_i(\hat{K})$. Before defining nested fractals, we give some more definitions and notation. Let F be the set of fixed points of the maps Ψ_i, $i \in S$ (thus $|F| = N$). A point $x \in F$ is called an essential fixed point if there exist i, j $(i \neq j)$ and $y \in F$ such that $\Psi_i(x) = \Psi_j(y)$. Let V_0 be the set of essential fixed points. Set $V_n = \cup_{x\in V_0} \cup_{i_1,\cdots,i_n\in S} \Psi_{i_1\cdots i_n}(x)$ where $\Psi_{i_1\cdots i_n} \equiv \Psi_{i_1} \circ \cdots \circ \Psi_{i_n}$ and $V_* = \cup_{n\geq 0} V_n$; then $\hat{K} = cl(V_*)$. For $i_1, \cdots, i_n \in S$, we call $\Psi_{i_1\cdots i_n}(V_0)$ an n-cell and $\Psi_{i_1\cdots i_n}(\hat{K})$ an n-complex. For $x, y \in \mathbb{R}^d (x \neq y)$, set $H_{xy} = \{z \in \mathbb{R}^d : |z-x| = |z-y|\}$ and let $U_{xy} : \mathbb{R}^d \to \mathbb{R}^d$ be a symmetric transformation with respect to H_{xy}. Then $\hat{K}$ is called a (compact) nested fractal if the following holds in addition to the above conditions:
(1) $\hat{K}$ is connected, $|V_0| \geq 2$.
(2) (Nesting) If $(i_1, \cdots, i_n)$ and $(j_1, \cdots, j_n)$ are distinct elements of S^n, then

$$\Psi_{i_1\cdots i_n}(\hat{K}) \cap \Psi_{j_1\cdots j_n}(\hat{K}) = \Psi_{i_1\cdots i_n}(V_0) \cap \Psi_{j_1\cdots j_n}(V_0).$$

(3) (Symmetry) For $x, y \in V_0$ $(x \neq y)$, U_{xy} maps n-cells to n-cells, and it maps any n-cell which contains elements in both sides of H_{xy} to itself for each $n \geq 0$.

From (2), we know that every nested fractal is a finitely ramified fractal. It is known that for each nested fractal, V_0 should be vertices of a regular planar polygon, a d-dimensional tetrahedron or a d-dimensional simplex (see [1], page 71). As in [17],[15], we also make the following assumption on nested fractals.

(NF*) There exists $k_0 > 0$ satisfying the following for all $m \geq 0$:

If $x, y \in \hat{F}$ satisfy $\|x - y\| \leq k_0 \alpha^{-m}$, then x, y join either in the same m-complex or adjacent m-complexes.

Set $\Sigma = S^{\mathbb{N}}$ and define a continuous surjective map $\pi : \Sigma \to \hat{K}$ as $\pi(\omega) = \lim_{m\to\infty} \Psi_{\omega_1 \cdots \omega_m}(x_0)$ where $x_0 \in V_0$. Let $\sigma : \Sigma \to \Sigma$ be the shift map, i.e. $\sigma w = w_2 w_3 \cdots$ for $w = w_1 w_2 \cdots$.

The Hausdorff dimension of $\hat{K}$ with respect to the Euclidean metric is $d_{f,E}(\hat{K}) \equiv \log N / \log \alpha$. A Bernoulli measure $\hat{\mu}$ on $\hat{K}$ with the property that $\hat{\mu}(\Psi_{i_1 \cdots i_n}(\hat{K})) = N^{-n}$ is a normalized Hausdorff measure.

We will next summarize how to construct a Dirichlet form on $\hat{K}$. Let $\{l_1, \cdots, l_r\} = \{|x - y| : x, y \in V_0, x \neq y\}$ (where $l_1 < \cdots < l_r$). Set $m_i = |\{y \in V_0 : |x - y| = l_i\}|$ (note that m_i is independent of $x \in V_0$) and let $\mathcal{P} = \{(p_1, \cdots, p_r) : p_1, \cdots, p_r > 0, \sum_{i=1}^r m_i p_i = 1\}$. Now, for $f, g \in l(V_n) \equiv \{f : V_n \to \mathbb{R}\}$ and $(p_1, \cdots, p_r) \in \mathcal{P}$, set

$$\begin{aligned} B_n(f,g) &= \sum_{i_1,\cdots,i_n \in S} \sum_{x,y \in V_0} (f \circ \Psi_{i_1 \cdots i_n}(x) - f \circ \Psi_{i_1 \cdots i_n}(y)) \\ &\qquad \times (g \circ \Psi_{i_1 \cdots i_n}(x) - g \circ \Psi_{i_1 \cdots i_n}(y)) q_{xy} \end{aligned}$$

(where $q_{xy} = p_i$ if $|x - y| = l_i$, 0 otherwise). Then, it is known that there exists a unique vector $(p_1, \cdots, p_r) \in \mathcal{P}$ and a unique $\rho > 1$ such that

$$\rho \cdot \inf\{B_1(g,g) : g|_{V_0} = v\} = B_0(v,v) \qquad \text{for all } v \in l(V_0). \tag{A.1}$$

In the following we use the vector $(p_1, \cdots, p_r)$ to define the form. For $f, g \in l(V_n)$, set

$$\hat{\mathcal{E}}_n(f,g) = \rho^n B_n(f,f).$$

Using (A.1) and the nesting property of $\hat{K}$,

$$\hat{\mathcal{E}}_n(f,f) \leq \hat{\mathcal{E}}_{n+1}(f,f) \quad \text{for all } f \in l(V_{n+1})$$

(equality holds when f is harmonic on $V_{n+1} \setminus V_n$). Define

$$\hat{\mathcal{F}} = \{f \in l(V_*) : \lim_{n\to\infty} \hat{\mathcal{E}}_n(f,f) < \infty\}, \quad \hat{\mathcal{E}}(f,g) = \lim_{n\to\infty} \hat{\mathcal{E}}_n(f,g) \quad \text{for all } f, g \in \hat{\mathcal{F}}.$$

Then, for each $f \in \hat{\mathcal{F}}$, there exists a unique $P_m f \in \hat{\mathcal{F}}$ such that

$$\hat{\mathcal{E}}(P_m f, P_m f) = \hat{\mathcal{E}}_m(f|_{V_m}, f|_{V_m}), \tag{A.2}$$

which is called the *m-harmonic extension* of $f|_{V_m}$. In order to embed this closed form in $\mathbb{L}^2(\hat{K}, \mu)$, we prepare the following. Let

$$R(p,q)^{-1} = \inf\{\hat{\mathcal{E}}(f,f) : f \in V_*, f(p) = 1, f(q) = 0\} \quad \text{for all } p, q \in V_*, \ p \neq q.$$

This function $R(p,q)$ is the effective resistance between p and q over the network V_*. We set $R(p,p) = 0$ for each $p \in V_*$.

Proposition A.1 *(1) $R(.,.)$ is a metric on V_*. It can be extended to a metric on $\hat{K}$, (which will be denoted by the same symbol R) and it gives the same topology on $\hat{K}$ as that inherited from the Euclidean metric.*
(2) For $p \neq q \in V_*$, $R(p,q) = \sup\{|f(p) - f(q)|^2/\hat{\mathcal{E}}(f,f) : f \in \hat{\mathcal{F}}, f(p) \neq f(q)\}$.

Note that $\rho > 1$ is important for $R(.,.)$ to be a metric on $\hat{K}$. From (2), we have that $|f(p) - f(q)|^2 \leq R(p,q)\hat{\mathcal{E}}(f,f)$ for $f \in \hat{\mathcal{F}}, p, q \in V_*$. Therefore $f \in \hat{\mathcal{F}}$ can be extended continuously to $\hat{K}$ and hence we can regard $\hat{\mathcal{F}} \subset C(\hat{K}, \mathbb{R}) \subset \mathbb{L}^2(\hat{K}, \hat{\mu})$.

Theorem A.2 *$(\hat{\mathcal{E}}, \hat{\mathcal{F}})$ is a local regular Dirichlet form on $\mathbb{L}^2(\hat{K}, \hat{\mu})$ with the following property,*

$$|f(p) - f(q)|^2 \leq R(p,q)\hat{\mathcal{E}}(f,f) \qquad \text{for all } \ f \in \hat{\mathcal{F}}, \text{and } p, q \in \hat{K}, \tag{A.3}$$

$$\hat{\mathcal{E}}(f,g) = \rho \sum_{i \in S} \hat{\mathcal{E}}(f \circ \Psi_i, g \circ \Psi_i) \qquad \text{for all } \ f, g \in \hat{\mathcal{F}}. \tag{A.4}$$

Further, for $\beta > 0$, $\hat{\mathcal{E}}_{(\beta)}$ admits a positive symmetric continuous reproducing kernel, where $\hat{\mathcal{E}}_{(\beta)}(\cdot,\cdot) = \hat{\mathcal{E}}(.,.) + \beta(.,.)_{\mathbb{L}^2(\hat{K},\hat{\mu})}$.

By the general theory of Dirichlet forms ([6]), there is a one to one correspondence between a local regular Dirichlet form on $\mathbb{L}^2(\hat{K}, \hat{\mu})$ and a $\hat{\mu}$-symmetric diffusion process on $\hat{K}$ up to some exceptional set of starting points. In this case, thanks to (A.3), we can prove the Feller property of the process to show that this one to one correspondence holds without any ambiguity in the starting points. We will denote by $\{\hat{X}_t\}_{t \geq 0}$ the diffusion process corresponding to $(\hat{\mathcal{E}}, \hat{\mathcal{F}})$. Roughly speaking, this process is constructed from the random walk $\hat{X}_n$ on V_n (whose transition probability is given by $(p_1, \cdots, p_r)$) by multiplying the time by t_K^n (that is $\hat{X}_n([t_K^n t])$; where $t_K \equiv \rho N$ is the time scaling factor) and letting $n \to \infty$. It is known that any self-similar Feller diffusion process which is invariant under locally symmetric transformations on $\hat{K}$ is a constant time change of this process, so that we call this process Brownian motion on $\hat{K}$.

On nested fractals, there is another important metric $d(.,.)$ called the shortest path metric (see [15], [1]). For each $x, y \in \hat{K}$, let $N_n(x,y)$ be the number of steps in the shortest path on V_n from x_0 to y_0 where $x_0 \in V_n$ (respectively $y_0 \in V_n$) is chosen so that x and x_0 (respectively y and y_0) are in the same n-complex. Then, there exists $b > 1$ such that $N_n(x,y)/b^n$ is bounded from above and below by some positive constants for all n. Taking a subsequence if necessary, we can define the metric $d(x,y) = \lim_{k \to \infty} N_{n_k}(x,y)/b^{n_k}$. d is a geodesic metric and each similitude Ψ_i has a contraction rate b^{-1} with respect to $d(.,.)$. Set $d_c = \log b / \log \alpha \geq 1$ (d_c is called the chemical exponent). We also define the walk dimension $d_w = \log t_K / \log b$ and Hausdorff dimension $d_f = \log N / \log b$, with respect to the shortest path metric. Note that $d_{w,E} = d_w d_c$, $d_{f,E} = d_f d_c$ are the walk dimension and the Hausdorff dimension with respect to the Euclidean metric. Then, for each $p, q \in \hat{K}$, we have the following relation,

$$d(.,.) \asymp \|p - q\|^{d_c} \asymp R(p,q)^{d_c/(d_w - d_f)}.$$

Here $\|\cdot\|$ is a Euclidean metric and $f(x) \asymp g(x)$ means $f(x)/g(x)$ are bounded from above and below by some positive constants.

Define $d_s = 2\log N/\log t_K$ which is called the spectral dimension. We then have the following heat kernel estimates.

Theorem A.3 *Brownian motion on $\hat{K}$ has a jointly continuous transition density (heat kernel) $\hat{p}_t(x,y)$ $t>0, x,y \in \hat{K}$. Further, there exist $c_{A.1},\cdots,c_{A.4}$ such that*

$$\begin{aligned} c_{A.1}t^{-d_s/2}\exp(-c_{A.2}(\frac{d(x,y)^{d_w}}{t})^{\frac{1}{d_w-1}}) &\leq \hat{p}(t,x,y) \\ &\leq c_{A.3}t^{-d_s/2}\exp(-c_{A.4}(\frac{d(x,y)^{d_w}}{t})^{\frac{1}{d_w-1}}), \end{aligned}$$

for all $0<t<1$ and all $x,y \in \hat{K}$.

The domain of the form is characterized as follows.

Theorem A.4 *([11], [17], [23])*

$$\hat{\mathcal{F}} = Lip(\frac{d_w d_c}{2}, 2, \infty)(\hat{K}), \tag{A.5}$$

where the Lipschitz space $Lip(d_w d_c/2, 2, \infty)(\hat{K})$ is the set of $f \in \mathbb{L}^2(\hat{K},\hat{\mu})$ such that

$$\sup_{\nu \in \mathbb{N}\cup\{0\}} \alpha_0^{\nu d_c(d_w+d_f)} \int\int_{||x-y||<c_0\alpha_0^{-\nu}} |f(x)-f(y)|^2 d\hat{\mu}(x)d\hat{\mu}(y) < \infty \tag{A.6}$$

for some $\alpha_0 > 1, c_0 > 0$.

Note that it is easy to see that in (A.6), different values of the constants c_0 and α_0 give equivalent spaces as long as the former is positive and the latter is greater than 1. For this theorem, the exponents are simpler if we express them with respect to the Euclidean metric; $d_{w,E} = d_w d_c, d_{f,E} = d_f d_c$. It is known that when $d_{w,E}/2 = d_w d_c/2 \notin \mathbb{Z}$, this Lipschitz space corresponds to (a subspace of) the Besov space $B^{2,\infty}_{d_{w,E}/2}(\hat{K})$ (see [13] Chapter V Proposition 3 and [11] Proposition 1).

If we assume, without loss of generality, that $\Psi_1(x) = \alpha^{-1}x$, then an unbounded nested fractal K can be constructed as $K = \cup_{n=1}^{\infty}\alpha^n\hat{K}$. The local regular Dirichlet form $(\mathcal{E},\mathcal{F})$ on K, whose restriction to $\hat{K}$ is $\hat{\mathcal{E}}$, can be constructed on $\mathbb{L}^2(K,\mu)$ (where μ is a Bernoulli measure on K so that $\mu|_{\hat{K}} = \hat{\mu}$) as follows. Set $\hat{K}_{<l>} = \alpha^l\hat{K}$ and define $\sigma_l : l(\hat{K}_{<l>}) \to l(\hat{K})$ by $\sigma_l f(x) = f(\alpha^l x) = f\circ\Psi_1^{-l}(x)$ for all $x \in \hat{K}$. Set $\hat{\mathcal{F}}_{\hat{K}_{<l>}} = \sigma_{-l}\hat{\mathcal{F}}$ and $\hat{\mathcal{E}}_{\hat{K}_{<l>}}(f,g) = \rho^{-l}\hat{\mathcal{E}}(\sigma_l f, \sigma_l g)$ for all $f,g \in \hat{\mathcal{F}}_{\hat{K}_{<l>}}$. It is easy to see

$$\hat{\mathcal{E}}_{\hat{K}_{<l-1>}}(f|_{\hat{K}_{<l-1>}}, f|_{\hat{K}_{<l-1>}}) \leq \hat{\mathcal{E}}_{\hat{K}_{<l>}}(f,f) \qquad \text{for all } \; f \in \hat{\mathcal{F}}_{\hat{K}_{<l>}}. \tag{A.7}$$

Define

$$\begin{aligned} \mathcal{D}_K = \{f \in C_0(K) : f|_{\hat{K}_{<l>}} \in \mathcal{F}_{\hat{K}_{<l>}} \forall l \in \mathbb{N}, \; \lim_{l\to\infty}\hat{\mathcal{E}}_{\hat{K}_{<l>}}(f|_{\hat{K}_{<l>}}, f|_{\hat{K}_{<l>}}) < \infty\}, \\ \mathcal{E}(f,g) = \lim_{l\to\infty}\hat{\mathcal{E}}_{\hat{K}_{<l>}}(f|_{\hat{K}_{<l>}}, g|_{\hat{K}_{<l>}}) \qquad \text{for all } \; f,g \in \mathcal{D}_K. \end{aligned}$$

It is easy to show that $(\mathcal{E},\mathcal{D}_K)$ is closable in $\mathbb{L}^2(K,\mu)$ by using (A.7). Let $\mathcal{F}=\overline{\mathcal{D}_K}^{\mathcal{E}_{(1)}}$ so that $(\mathcal{E},\mathcal{F})$ is the smallest extension of $(\mathcal{E},\mathcal{D}_K)$. Then we can define the resistance metric $R(.,.)$ in the same way as for the compact case and we have the following.

Theorem A.5 *$(\mathcal{E},\mathcal{F})$ is a local regular Dirichlet form on $\mathbb{L}^2(K,\mu)$ which satisfies (A.3) and the following scaling property,*

$$\mathcal{E}(f,g)=\lambda\mathcal{E}(f\circ\Psi_1,g\circ\Psi_1)\ \textit{ for all }\ f,g\in\mathcal{F}.$$

Further, for $\beta>0$, $\mathcal{E}_{(\beta)}$ admits a positive symmetric continuous reproducing kernel.

We call the corresponding diffusion process Brownian motion on K. Theorem A.3 holds for the heat kernel on K for $0<t<\infty$. Similarly to Theorem A.4, we have $\mathcal{F}=\tilde{\mathrm{Lip}}(\frac{d_wd_c}{2},2,\infty)(K)$, where $\tilde{\mathrm{Lip}}(d_wd_c/2,2,\infty)(K)$ is a set of $f\in\mathbb{L}^2(K,\mu)$ such that

$$\sup_{\nu\in\mathbb{Z}}\alpha^{\nu(d_c(d_w+d_f))}\int\int_{||x-y||<c_0\alpha_0^{-\nu}}|f(x)-f(y)|^2d\mu(x)d\mu(y)<\infty \tag{A.8}$$

for some $\alpha_0>1, c_0>0$.

Finally, we will mention the detailed short time asymptotics (the so called Varadhan type estimate) and (Schilder type) large deviations on fractals. As we will see, these results do not hold in their original forms on $\mathbb{R}^d$, due to oscillations which arise from the fractal structure. So far, these results are only proved for the two-dimensional Sierpinski gasket, but it is believed that such behaviour is "typical" for Brownian motion on nested fractals.

Let K be a (either compact or unbounded) 2-dimensional Sierpinski gasket.

Theorem A.6 *([16]) There exists a periodic non-constant positive continuous function F with period $5/2$ such that the following holds for all $z>0,x,y\in E$.*

$$-\lim_{n\to\infty}((2/5)^nz)^{\frac{1}{d_w-1}}\log p_{(2/5)^nz}(x,y)=d(x,y)^{\frac{d_w}{d_w-1}}F(\frac{z}{d(x,y)}).$$

In particular, $-\lim_{t\to 0}t^{\frac{1}{d_w-1}}\log p_t(x,y)$ does not exists.

This F is defined as the Legendre transform of a limiting function arising from the Laplace transform of the hitting time distribution of the Brownian motion X and a "tiny" oscillation in the tail of the hitting time distribution makes F non-constant (this is also related to a "tiny" oscillation in the generating function of an associated branching process).

For the same reason, we have an oscillation in the large deviations. For fixed $T>0$, let $\Omega_x\equiv C_x([0,T]\to K)=\{\phi\in C([0,T]\to K):\phi(0)=x\}$ with uniformly continuous topology. Let P_ϵ^x be the law for $X^x(\epsilon t)$. Then the following holds.

Theorem A.7 *([5]) For each $z>0$, $A\subset\Omega_x$,*

$$\begin{aligned}-\inf_{\phi\in IntA}I_x^z(\phi)\ &\le\ \liminf_{n\to\infty}((2/5)^nz)^{1/(d_w-1)}\log P^x_{(2/5)^nz}(A)\\ &\le\ \limsup_{n\to\infty}((2/5)^nz)^{1/(d_w-1)}\log P^x_{(2/5)^nz}(A)\le-\inf_{\phi\in ClA}I_x^z(\phi).\end{aligned}$$

Here $\{I_x^z\}_{z\in[2/5,1)}$ is a sequence of rate functions defined as follows for each $\phi \in \Omega_x$,

$$I_x^z(\phi) = \begin{cases} \int_0^T (\dot\phi(t))^{d_w/(d_w-1)} F(z/\dot\phi(t))dt, & \phi \text{ is absolutely continuous,} \\ \infty, & \text{otherwise,} \end{cases}$$

where F is as before and $\dot\phi(t) \equiv \lim_{s\to t} \frac{d(\phi(s),\phi(t))}{|s-t|}$ for $t \in [0,T]$. This result tells us that the usual (Schilder type) large deviation principle does not hold when one takes the time scaling ϵ to 0. When $A = \{f \in \Omega_x : f(T) = y\}$, $\inf\{I_x^z(\phi) : \phi \in A\}$ is attained independently of z by the path(s) which moves on the geodesic(s) between x and y homogeneously. Thus "the most probable path" should be this path, but the energy (action functional) of the path depends on the time sequence determined by z.

Acknowledgement. The authors thank Professor S. Kusuoka for allowing us to use his unpublished result. This paper was completed during a visit by the second author to the University of Oxford. He thanks the University of Oxford for its hospitality. This research was partly supported by the Grant-in-Aid for Scientific Research of Japan.

References

[1] M.T. Barlow, Diffusions on fractals, Lectures in Probability Theory and Statistics: Ecole d'été de probabilités de Saint-Flour XXV (Lect. Notes Math., vol. 1690), Springer, New York, 1998.

[2] M.T. Barlow and R.F. Bass, Construction of Brownian motion on the Sierpinski carpet, *Ann Inst H. Poincare*, **25** (1989), 225–257.

[3] 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.

[4] R.F. Bass and T. Kumagai, Laws of the iterated logarithm for some symmetric diffusion processes, *Osaka J. Math.*, **37** (2000), 625–650.

[5] G. Ben Arous and T. Kumagai, Large deviations for Brownian motion on the Sierpinski gasket, *Stoch. Proc. Their Appl.*, **85** (2000), 225–235.

[6] M. Fukushima, Y. Oshima and M. Takeda, Dirichlet forms and symmetric Markov processes, de Gruyter, Berlin, 1994.

[7] B.M. Hambly, Brownian motion on a random recursive Sierpinski gasket, *Ann. Probab.*, **25** (1997), 1059–1102.

[8] B.M. Hambly and T. Kumagai, Transition density estimates for diffusion processes on post critically finite self-similar fractals, *Proc. London Math. Soc.*, **78** (1999), 431–458.

[9] B.M. Hambly and T. Kumagai, Asymptotic behaviour of the spectral and walk dimensions as fractals approach Euclidean space, *Fractals*, **10** (2002), 403–413.

[10] N. Ikeda and S. Watanabe, The local structure of a class of diffusions and related problems, In: Proceedings of the Second Japan-USSR Symposium on Probability Theory (Kyoto, 1972), pp. 124–169, Lect. Notes Math., vol. 330, Springer, Berlin, 1973.

[11] A. Jonsson, Brownian motion on fractals and function spaces, *Math. Z.*, **222** (1996), 496–504.

[12] A. Jonsson, Dirichlet forms and Brownian motion penetrating fractals, *Potential Analysis*, **13** (2000), 69–80.

[13] A. Jonsson and H. Wallin, Function spaces on subsets of $\mathbf{R}^n$, Mathematical Reports Vol. 2, Part 1, Acad. Publ., Harwood, 1984.

[14] J. Kigami, Analysis on fractals, Cambridge Univ. Press, Cambridge, 2001.

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

[16] T. Kumagai, Short time asymptotic behavior and large deviations for Brownian motion on some affine nested fractals, *Publ. RIMS. Kyoto Univ.*, **33** (1997), 223–240.

[17] T. Kumagai, Brownian motion penetrating fractals -An application of the trace theorem of Besov spaces-, *J. Func. Anal.*, **170** (2000), 69–92.

[18] S. Kusuoka, Unpublished work (1993).

[19] S. Kusuoka, Diffusion processes on nested fractals, In: R.L. Dobrushin and S. Kusuoka: Statistical Mechanics and Fractals, Lect. Notes Math., vol. 1567, Springer, New York, 1993.

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

[21] T. Lindstrøm, Brownian motion penetrating the Sierpinski gasket, In: Asymptotic Problems in Probability Theory, stochastic models and diffusions on fractals (K. D. Elworthy and N. Ikeda (eds.)), pp. 248–278, Longman Scientific, Harlow UK, 1993.

[22] Y. Ogura, M. Tomisaki and M. Tsuchiya, Superposition of diffusion processes—Feller property, In: Trends in probability and related analysis (Taipei, 1998), pp. 113–128, World Sci. Publishing, River Edge, NJ, 1999.

[23] K. Pietruska-Paluba, On function spaces related to fractional diffusions on d-sets, *Stochastics Stochastics Rep.*, **70** (2000), 153–164.

[24] M. Tomisaki, Superposition of diffusion processes, *J. Math. Soc. Japan*, **32** (1980), 671–696.

[25] H. Triebel, Fractals and spectra -related to Fourier analysis and function spaces-, Monographs in Math. Vol. 91, Birkhäuser, Basel, 1997.

[26] S.R.S. Varadhan, Diffusion processes in a small time interval, *Commun. Pure Appl. Math.*, **20** (1967), 659–685.