Electronic Journal of Probability

Vol. 6 (2001) Paper no. 9, pages 1–23

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol6/paper9.abs.html

TRANSITION DENSITY ASYMPTOTICS FOR SOME DIFFUSION PROCESSES WITH MULTI-FRACTAL STRUCTURES

Martin T. Barlow
Department of Mathematics, University of British Columbia, Vancouver, V6T 1Z2, Canada
barlow@math.ubc.ca

Takashi Kumagai
Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan
kumagai@kurims.kyoto-u.ac.jp

Abstract We study the asymptotics as $t \to 0$ of the transition density of a class of μ-symmetric diffusions in the case when the measure μ has a multi-fractal structure. These diffusions include singular time changes of Brownian motion on the unit cube.

Keywords diffusion process, heat equation, transition density, spectral dimension, multi-fractal

AMS subject classification 60J60, 31C25, 60J65.

Research partially supported by a NSERC (Canada) grant and Grant-in-Aid for Scientific Research (B)(2) 10440029 of Japan.

Submitted to EJP on June 19, 2000. Final version accepted on March 3, 2001.

1 Introduction

Let $\widehat{X}_t$ be reflecting Brownian motion on $K = [0,1]$. It is well known that $\widehat{X}$ has generator $\widehat{\mathcal{L}} = \frac{1}{2}\Delta$, and Dirichlet form given by

$$\mathcal{E}(f,f) = \frac{1}{2}\int_{[0,1]} |\nabla f(x)|^2 dx \quad \text{on } L^2([0,1],dx).$$

The transition density $\widehat{p}_t(x,y)$ of X satisfies the heat equation $\partial\widehat{p}_t/\partial t = \widehat{\mathcal{L}}\widehat{p}_t$, and for $x \in [0,1]$ the short time asymptotics of $\widehat{p}_t(x,x)$ are given by

$$\widehat{p}_t(x,x) \sim (2\pi t)^{-1/2}, \quad t \to 0.$$

Now let μ be a measure on K, with closed support K, and consider the Dirichlet form $\mathcal{E}(f,f)$ on $L^2(K,\mu)$. In probabilistic terms the associated process X can be obtained by a time change of $\widehat{X}$. Set $A_t = \int_K L_t^a \mu(da)$, where (L_t^a) are the (jointly continuous) local times of $\widehat{X}$, and let $\tau_t = \inf\{s : A_s > t\}$ be the right-continuous inverse of A. Then (see [9], Theorem 6.2.1), $X_t = \widehat{X}_{\tau_t}$. If $d\mu/dx = a(x)$, where a is strictly positive and continuous, then X has a generator

$$\mathcal{L}f(x) = \frac{1}{2}a(x)^{-1}\Delta f(x),$$

and the transition density $p_t(x,y)$ of X satisfies

$$p_t(x,x) \sim (2a(x)\pi t)^{-1/2}, \quad t \to 0.$$

In this paper we wish to study the short time asymptotics of $p_t(x,x)$ in the case when μ is singular with respect to Lebesgue measure, but still has closed support equal to K. For the moment we will just discuss the case $K = [0,1]$, but our results do hold for more general self-similar sets. We will assume that the measure μ is "multi-fractal" or self-similar. For $[0,1]$ examples of measures of this kind are the de Rham p-measures $\mu = \mu^{(p)}$, where $0 < p < 1$. $\mu^{(p)}$ is characterized by the property that, for any $n \geq 1$ and $0 \leq k \leq 2^n - 1$,

$$\mu^{(p)}([k2^{-n}, k2^{-n} + 2^{-(n+1)}]) = p\mu^{(p)}([k2^{-n},(k+1)2^{-n}]).$$

(This is the measure under which the coefficients x_i in the dyadic expansion of x are independent identically distributed random variables with mean $1-p$.)
Define

$$d_s(x) = 2\lim_{t\to 0} \frac{\log p_t(x,x)}{-\log t},$$

for those $x \in [0,1]$ for which this limit exists.

Theorem 1.1 *Let $p_t(x,y)$ be the transition density of the process X associated with $\mathcal{E}$ on $L^2([0,1],\mu^{(p)})$, where $\frac{1}{2} \leq p < 1$. For $0 \leq \theta \leq 1$ let*

$$a(\theta) = \theta\log\frac{1}{p} + (1-\theta)\log\frac{1}{1-p}.$$

1) If x is a dyadic rational then $d_s(x)/2 = a(1)/(\log 2 + a(1)) = \log(1/p)/\log(2/p)$.
2) $\mu^{(\theta)}$ almost everywhere $d_s(x)/2 = a(\theta)/(\log 2 + a(\theta))$.
3) There exist points x at which

$$\liminf_{t\to 0} \frac{\log p_t(x,x)}{-\log t} < \limsup_{t\to 0} \frac{\log p_t(x,x)}{-\log t}.$$

In fact our methods handle more general compact self-similar sets K, and include the following:

1) P.c.f. fractals with a 'regular harmonic structure' – see [16].
2) The unit cube $[0,1]^d$ for $d \geq 2$.
3) P.c.f. fractals with a harmonic structure which is not regular.
4) Sierpinski carpets in dimensions $d \geq 2$ – see [5].

The unit interval is a special case of 1), and we can treat 2) as a special case of 4). In cases 1) and 3) the underlying diffusion is that given by the harmonic structure, while for 2) it is standard Brownian motion on the unit cube, with normal reflection on the boundary. For 4) it is the diffusion constructed in [2]. We restrict ourselves to self-similar (Bernoulli) measures μ for which the topological support is the whole of K. In case 1) this is the only condition on μ, but in the other cases a further condition (see (2.2)) is needed to ensure that μ does not charge sets of capacity zero.

The main results of this paper are Theorem 3.5 and Corollary 3.6, which give upper and lower bounds on the transition density $p_t(x,x)$. Specializing to the case $K = [0,1]$ we obtain Theorem 1.1.

The essential idea of this paper is to decompose K into regions $D_i^{(n)}$ such that the process X takes a time $O(e^{-n})$ to cross each of these sets. The self-similarity of K means that these sets are all the same 'shape', but in general different 'sizes'. We therefore expect that, for most $x \in D_i^{(n)}$, one should have

$$p_t(x,x) \simeq \mu(D_i^{(n)})^{-1}. \tag{1.1}$$

This estimate turns out to be correct whenever, on one hand, t is small enough so that $P^x(X_t \in D_i^{(n)}) > c > 0$, and on the other hand t is large enough so that $p_t(x,\cdot)$ has diffused over a significant proportion of $D_i^{(n)}$. We will see that when x is suitably far from the boundary of $D_i^{(n)}$ then (with a few added constants) (1.1) holds.

We can, however, have adjacent regions $D_i^{(n)}$, $D_j^{(n)}$, with very different measures. For example, in the case of $[0,1]$ with $\mu^{(p)}$ the appropriate sets will be $[\frac{1}{2} - 2^{-n_1}, \frac{1}{2}]$, $[\frac{1}{2}, \frac{1}{2} + 2^{-n_2}]$, where $n_i \simeq n/(\log(2/p_i))$ and $p_1 = 1-p$, $p_2 = p$. Since $p_t(x,x)$ is continuous, (1.1) clearly cannot hold close to $\frac{1}{2}$.

In this paper we do not tackle the problem, which seems in general quite hard, of identifying how $p_t(x,x)$ behaves in these boundary zones. We are, however, able to show that the sets of bad points (where our upper and lower bounds differ significantly) is small, and this enables us to make the kind of estimates given in Theorem 1.1.

If we set

$$J_\gamma = \{x \in K : d_s(x) \text{ exists and equals to } \gamma\}$$

then we have a multi-fractal decomposition of K into $\{J_\gamma\}_\gamma$ and $K \setminus \cup_\gamma J_\gamma$. A forthcoming paper, [12], studies the Hausdorff dimensions of these sets.

2 Dirichlet forms on some self-similar sets with multi-fractal measure

2.1 Self-similar sets

In this section we describe the spaces we consider, and give the properties of the Dirichlet forms on them that we will need. We begin with the definition of a self-similar space: see [1], [18] for more details and examples.

Notation.
1) Let $S = \{1, 2, \cdots, N\}$. The one-sided shift space Σ is defined by $\Sigma = S^{\mathbb{N}}$.
2) For $w \in \Sigma$, we denote the i-th element in the sequence by w_i and write $w = w_1 w_2 w_3 \cdots$.
3) If $w \in S^n$, we define $|w| = n$.
4) Let $\sigma : \Sigma \to \Sigma$ be the left shift map, i.e. $\sigma w = w_2 w_3 \cdots$ if $w = w_1 w_2 \cdots$. Define $\widetilde{\sigma}_s : \Sigma \to \Sigma$ by $\widetilde{\sigma}_s w = sw$ for $s \in S$.

Definition 2.1 *Let K be a compact metrizable space and for each $s \in S$, $F_s : K \to K$ be a continuous injection. Then, $\mathcal{L} = (K, S, \{F_s\}_{s \in S})$ is said to be a self-similar structure on K if there exists a continuous surjection $\pi : \Sigma \to K$ such that $\pi \circ \widetilde{\sigma}_s = F_s \circ \pi$ for every $s \in S$.*

For $w \in S^n$, we denote $F_w = F_{w_1} \circ F_{w_2} \circ \cdots \circ F_{w_n}$ and $K_w = F_w(K)$. In particular, $K_s = F_s(K)$ for $s \in S$. We remark that the unit interval is a simple example of a self-similar structure: take $N = 2$ so that $\Sigma = \{1, 2\}^{\mathbb{N}}$ and let

$$\pi(w) = \sum_{i=1}^{\infty} (w_i - 1) 2^{-i}.$$

Definition 2.2 *Let $\mathcal{L} = (K, S, \{F_s\}_{s \in S})$ be a self-similar structure on K. Then the critical set of $\mathcal{L}$ is defined by*

$$C(\mathcal{L}) = \pi^{-1}(\cup_{s,t \in S, s \neq t}(K_s \cap K_t))$$

and the post critical set of $\mathcal{L}$ is defined by

$$P(\mathcal{L}) = \cup_{n \geq 1} \sigma^n (C(\mathcal{L})).$$

See [1], Section 5, for the computation of $C(\mathcal{L})$ and P for some simple examples. In the case of the unit interval, with the self-similar structure given above, we have $P(\mathcal{L}) = \{0, 1\}$.

For $m \geq 0$, let

$$P^{(m)} = \cup_{w \in S^m} wP, \quad V_m = \pi(P^{(m)}), \quad V_* = \cup_{m \geq 0} V_m \text{ and } \overset{\circ}{V_m} = V_m - V_0.$$

We call V_0 the boundary of K. A Bernoulli (probability) measure on K is a measure μ on K such that $\mu(F_i(K)) = \mu_i > 0$, where $\sum_{i=1}^{N} \mu_i = 1$. For $u \in L^1(K, \mu)$ we write $\bar{u} = \int_K u d\mu$.

In this paper, we will consider connected self-similar sets $(K, S, \{F_s\}_{s \in S})$, with a local regular Dirichlet forms $(\mathcal{E}, \mathcal{F})$ on $L^2(K, \mu)$ which satisfy the following assumption.

Assumption 2.3 *(a) $(\mathcal{E}, \mathcal{F})$ is a closed local regular Dirichlet form on $L^2(K, \mu)$ so that for each $f \in \mathcal{F}$, $f \circ F_i \in \mathcal{F}$ for all $i \in S$. Further, $(\mathcal{E}, \mathcal{F})$ satisfies, for some $\rho_i > 0, i \in S$, the following self-similarity property:*

$$\mathcal{E}(f, g) = \sum_{i=1}^{N} \rho_i \mathcal{E}(f \circ F_i, g \circ F_i) \quad \forall f, g \in \mathcal{F}. \tag{2.1}$$

(b) μ is a Bernoulli measure on K with

$$0 < \mu_i < \rho_i, \quad \forall i \in S. \tag{2.2}$$

(c) There exists $c_{2.1} > 0$ such that

$$\mathcal{E}(f, f) \geq c_{2.1} \int_K |f - \bar{f}|^2 d\mu \quad \forall f \in \mathcal{F}. \tag{2.3}$$

(d) The semigroup $(P_t)_{t \geq 0}$ associated with $\mathcal{E}$ on $L^2(K, \mu)$ has a jointly continuous density $p_t(x, y)$, $t > 0$, $x, y \in K$. (This is the transition density of the associated diffusion process X with respect to μ.)

In the remainder of this section we will discuss the existence of Dirichlet forms satisfying this assumption for the two classes of spaces treated in this paper: p.c.f self-similar sets, and Sierpinski carpets.

First we give some more notation. Set $t_i = \rho_i / \mu_i$, for $1 \leq i \leq N$. We remark that ρ_i can be interpreted as the conductance associated with $F_i(K)$ and that t_i^{-1} is the time scaling factor for the diffusion process on $F_i(K)$. Let Λ_n be defined by

$$\Lambda_n = \{w = w_1 \cdots w_k \in \cup_{i \geq 0} S^i : t_{w_1} \cdots t_{w_{k-1}} \leq e^n < t_{w_1} \cdots t_{w_k}\},$$

with $\Lambda_0 = \emptyset$. We write $t_w = \Pi_{i=1}^k t_{w_i}$, $\rho_w = \Pi_{i=1}^k \rho_{w_i}$ etc. for $w = w_1 \cdots w_k$. Throughout the paper, we denote

$$t^* = \max_i t_i, \quad t_* = \min_i t_i, \quad \mu^* = \max_i \mu_i, \quad \mu_* = \min_i \mu_i.$$

From (2.1) we have

$$\mathcal{E}(f, f) = \sum_{w \in \Lambda_n} \rho_w \mathcal{E}(f \circ F_w, f \circ F_w) \quad \forall f \in \mathcal{F}.$$

We call a set of the form $F_w(K)$, an m-complex if $w \in S^m$ and a Λ_l-complex if $w \in \Lambda_l$. For $A \subset K$ and $m \geq 0$, let

$$D_m(A) = \{C : C \text{ is a } m\text{-complex such that } A \cap C \neq \emptyset\},$$

and let $D_m^1(A) = D_m(D_m(A))$. We define $D_{\Lambda_l}(x)$ and $D_{\Lambda_l}^1(x)$ analogously. Set $\partial D_{\Lambda_l}(x) = cl(K \setminus D_{\Lambda_l}(x)) \cap D_{\Lambda_l}(x)$. For $x \in K - V_*$ let $\Lambda_r(x)$ be the length of the word of the Λ_r-complex to which x belongs: note that $D_{\Lambda_{r(x)}}(x) = D_{\Lambda_r}(x)$.

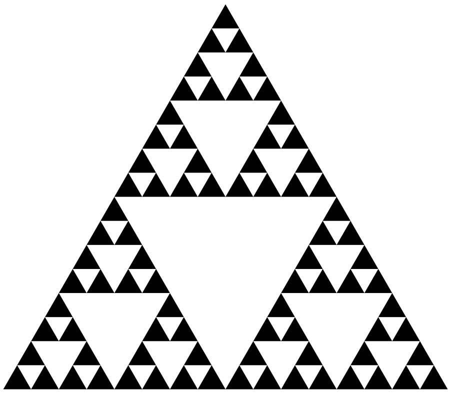

Figure 1: The Sierpinski gasket

2.2 P.c.f. self-similar sets and their Dirichlet forms

We call the self-similar set $(K, S, \{F_s\}_{s\in S})$ a p.c.f fractal set if the post critical set $P(\mathcal{L})$ is a finite set – p.c.f. here stands for 'post critically finite'. This condition implies that K is finitely ramified.

These sets were introduced by Kigami ([16]). In [16], [18], [20] it is shown that, provided a 'non-degenerate harmonic structure' exists, then a closed regular local Dirichlet form satisfying (2.1) exists, with the property that $\mathcal{E}(f,f) = 0$, $f \in \mathcal{F}$, implies that f is constant. (For work on the existence of non-degenerate harmonic structures see [25], [23].) In [16] the additional hypothesis of 'regularity' of the harmonic structure was imposed: in our context this means that the conductivities ρ_i satisfy

$$\rho_i > 1, \qquad \forall i \in S. \tag{2.4}$$

We now summarise how the remainder of Assumption 2.3 is proved in this case. Because the resolvent operator is compact (see [18], [20]) and $P_t f = f$ if and only if f is constant, there is a spectral gap so that (2.3) holds.

Let $\mathcal{L}_\mu$ be the self-adjoint operator on $L^2(K,\mu)$ associated with the Dirichlet form $(\mathcal{E}, \mathcal{F})$, and let $\{\lambda_n\}_n$ be the eigenvalues of $-\mathcal{L}_\mu$ and $\{\varphi_n\}_n$ be the normalized eigenfunctions. In [18], it is proved that φ_n is continuous and

$$\|\varphi_n\|_\infty \le \lambda_n^\kappa \qquad n \ge 1,$$

where κ depends only on the Dirichlet form and K. Thus, by Mercer's theorem,

$$p_t(x,y) = \sum_n e^{-\lambda_n t}\varphi_n(x)\varphi_n(y),$$

and the right hand side converges uniformly. This proves joint continuity of the transition density, and completes the verification of Assumption 2.3.

Let $n^\mu(x) = \#\{\lambda : \lambda \text{ is an eigenvalue of } -\mathcal{L}_\mu \le x.\}$. In [19], [18] it is proved that, if $d_s^e(\mu) > 0$ is the unique positive number satisfying

$$\sum_{i=1}^N (\mu_i/\rho_i)^{d_s^e(\mu)/2} = 1, \tag{2.5}$$

then

$$0 < \liminf_{x\to\infty} n^\mu(x)/x^{d_s^e(\mu)/2} \le \limsup_{x\to\infty} n^\mu(x)/x^{d_s^e(\mu)/2} < \infty.$$

In the case when (2.4) and (2.5) holds, let ν be the Bernoulli measure satisfying

$$\nu_i = \rho_i^{-\sigma} \qquad \forall i \in S, \tag{2.6}$$

where σ is the unique constant which satisfies $\sum_{i=1}^N \rho_i^{-\sigma} = 1$. Then $\max_\mu d_s^e(\mu)/2$ (where μ is taken to be a Bernoulli measure on K) is attained only at ν, and the maximum value is $\sigma/(\sigma+1)$. For this special case, (i.e. $\mu = \nu$) detailed estimates on $p_t(x,y)$ are obtained in [13]. We remark that if (2.4) holds then (2.2) is satisfied for any Bernoulli measure μ (with $\mu_i > 0$), and that $d_s^e(\mu) < 2$. In general, however, it is possible to have $d_s^e(\mu) > 2$.

2.3 Sierpinski carpets and their Dirichlet forms

Let $H_0 = [0,1]^d$, and let $l \in \mathbb{N}$, $l \ge 2$ be fixed. Set $\mathcal{Q} = \{\Pi_{i=1}^d[(k_i-1)/l, k_i/l] : 1 \le k_i \le l,\ k_i \in \mathbb{N}\ (1 \le i \le d)\}$, let $l \le N \le l^d$ and let F_i, $1 \le i \le N$ be orientation preserving affine maps of H_0 onto some element of $\mathcal{Q}$. (We assume that the sets $F_i(H_0)$ are distinct.) Set $H_1 = \cup_{i\in S} F_i(H_0)$. Then, there exists a unique non-empty compact set $K \subset H_0$ such that $K = \cup_{i\in S} F_i(K)$ and $(K, S, \{F_s\}_{s\in S})$ is a self-similar structure. K is called a Sierpinski carpet if the following hold:
(SC1) (Symmetry) H_1 is preserved by all the isometries of the unit cube H_0.
(SC2) (Connected) H_1 is connected.
(SC3) (Non-diagonality) Let B be a cube in H_0 which is the union of 2^d distinct elements of $\mathcal{Q}$. Then if $(H_1 \cap B)^o$ is non-empty, it is connected.
(SC4) (Borders included) H_1 contains the line segment $\{x : 0 \le x_1 \le 1, x_2 = \cdots = x_d = 0\}$.

Here (see [5]) (SC1) and (SC2) are essential, while (SC3) and (SC4) are included for technical convenience. The main difference from p.c.f. self-similar sets is that Sierpinski carpets are infinitely ramified: the critical set $C(\mathcal{L})$ in Definition 2.2 is infinite, and K cannot be disconnected by removing a finite number of points. In fact, for the classical Sierpinski carpet in $\mathbb{R}^d$ with $l = 3$ and $N = 3^d - 1$ we have $V_0 = \partial[0,1]^d$. Write $d_f = d_f(K) = \log N/\log l$ for the Hausdorff dimension of K. Note that the d-dimensional unit cube $[0,1]^d$, $d \ge 2$, can be included as an example of a Sierpinski carpet by taking $N = l^d$.

We write ν for the Bernoulli measure with weights $\nu_i = 1/N$: ν is a multiple of the Hausdorff measure on K. In [2], [21], [5], [14] a non-degenerate Dirichlet form $\mathcal{E}'$ on $L^2(K,\nu)$ is constructed on these spaces, with the property that $\mathcal{E}'$ is invariant under local isometries of K – and in particular $\mathcal{E}'$ is the same on each k-complex. The uniqueness of $\mathcal{E}'$ is an open problem – see [5]. If $\mathcal{E}'$ were unique then (2.1) would follow immediately. However, without requiring uniqueness, in [21] (see also Remark 5.11 of [5] and [17]) a compactness argument is used to construct a Dirichlet form $\mathcal{E}$ with the same invariances as $\mathcal{E}'$ and in addition satisfying (2.1) in the case when, for a constant ρ_K depending on K,

$$\rho_i = \rho_K, \quad 1 \le \forall i \le N.$$

Let $t_K = t_i = N\rho_K$, and let $\widehat{X} = (\widehat{X}_t, t \ge 0)$ be the diffusion associated with $\mathcal{E}$ and $L^2(K,\nu)$. We define $d_w = \log t_K/\log l$, the walk dimension of K, and $d_s = 2\log N/\log t_K$, the spectral

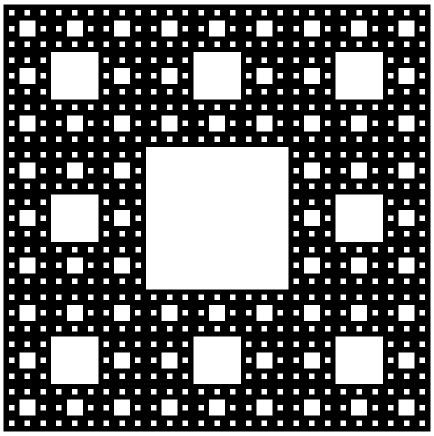

Figure 2: The Sierpinski carpet

dimension of K. When $K = [0,1]^d$, $\widehat{X}$ is just reflecting Brownian motion on $[0,1]^d$, $t_K = l^2$ and $\rho_K = l^{2-d}$. As $\widehat{X}$ satisfies the same local isotropy condition as the processes studied in [2], [5], the techniques of those papers apply to $\widehat{X}$ and lead to the same estimates for the Green's function and transition density of the process.

Let μ be a Bernoulli measure satisfying (2.2). We now verify Assumption 2.3. For functions f, g, write $f \asymp g$ if there exists $c_1, c_2 > 0$ such that $c_1 g(x) \le f(x) \le c_2 g(x)$ for all x. We have the following estimate of the 1-order Green's kernel for the process $\widehat{X}$. The proof follows from the estimates and methods of [1], [4], [5].

Proposition 2.4 *There exists a Green's kernel $\widehat{g}_1(x,y)$ which is continuous on $K \times K \setminus \{x \neq y\}$ (and on $K \times K$ when $d_s < 2$), and satisfies the following:*

$$\begin{aligned} E^x\Big[\int_0^\infty e^{-t} f(\widehat{X}_t)dt\Big] &= \int_K \widehat{g}_1(x,y) f(y) d\nu \quad \forall f \in \mathcal{B}(K), \\ \widehat{g}_1(x,y) &\asymp \begin{cases} c_{2.2}|x-y|^{d_w - d_f} & \text{if } d_s > 2, \\ -c_{2.3}\log|x-y| + c_{2.4} & \text{if } d_s = 2, \\ c_{2.5} & \text{if } d_s < 2. \end{cases} \end{aligned}$$

We now wish to consider $\mathcal{E}$ on the space $L^2(K,\mu)$, and to do this we use an argument due to Osada [24].

For an open set $B \subset K$, define the capacity of B by

$$\mathrm{Cap}(B) = \inf\{\widehat{\mathcal{E}}_1(u,u) : u \in \mathcal{F}, u \ge 1 \text{ on } B\},$$

where for $\beta \ge 0$, $\widehat{\mathcal{E}}_\beta(u,u) = \mathcal{E}(u,u) + \beta\|u\|^2_{L^2(K,\nu)}$. The capacity of any set $F \subset K$ is defined as the infimum of the capacity of open sets which contain F. We say that μ charges no set of zero capacity if the following holds:

$$\mu(A) = 0 \quad \text{for all } A \in \mathcal{B}(K) \text{ such that } \mathrm{Cap}(A) = 0.$$

Lemma 2.5 *Under the condition (2.2), μ charges no sets of zero capacity.*

PROOF. If $d_s < 2$ then points have strictly positive capacity, and the result is immediate. We prove the result for $d_s > 2$: the proof for $d_s = 2$ is similar. It is well-known that for each compact set $M \subset K$,

$$\begin{aligned}\mathrm{Cap}(M) = \sup\{m(M) \quad : \quad & m \text{ is a positive Radon measure, } \mathrm{Supp}[m] \subset M, \\ & \widehat{G}_1 m(x) \equiv \int_M \widehat{g}_1(x,y) m(dy) \le 1, \ \forall x \in K\}.\end{aligned} \tag{2.7}$$

Using Proposition 2.4,

$$\begin{aligned}\int_M \widehat{g}_1(x,y)\mu(dy) &\le \int_K \widehat{g}_1(x,y)\mu(dy) \\ &\le \sum_{n=0}^{\infty} \int_{l^{-n-1} \le |x-y|_{\max} < l^{-n}} \widehat{g}_1(x,y)\mu(dy) \\ &\le c_1 \sum_n l^{n(d_f-d_w)} \mu(l^{-n-1} \le |x-y|_{\max} < l^{-n}) \\ &\le c_2 \sum_n l^{n(d_f-d_w)} (\mu^*)^n \equiv c_3 < \infty,\end{aligned}$$

because of the assumption (2.2) (note that $l^{(d_f-d_w)} = N/t_K = \rho_K^{-1}$). Here $|y|_{\max} = \max_i |y_i|$ for $y = (y_1, \cdots, y_d) \in \mathbb{R}^d$. Thus, setting $\mu_M(\cdot) \equiv \mu(\cdot \cap M)$, we have $G_1 \mu_M \le c_3$. Using (2.7), $\mathrm{Cap}(M) \ge \mu(M)/c_3$ for each compact set M, thus for each Borel set, which completes the proof. □

As μ is a Radon measure which charges no set of zero capacity, it is a smooth measure in the sense of [9] (p. 80). Thus, it is a Revuz measure for some positive continuous additive functional A (see section 5.1 of [9]) and we can time change by the inverse of A – see section 6.2 of [9]. Let $\mathcal{S}$ denote the quasi support of μ (see p. 168 of [9]) for the Dirichlet form on $L^2(K, \nu)$. Note that $\mathcal{S} \subset K$ and $\mu(K \setminus \mathcal{S}) = 0$. For a set A, denote $\sigma_A = \inf\{t > 0 : \widehat{X}_t \in A\}$. Define

$$\begin{aligned}\widetilde{\mathcal{F}} &= \{\phi \in L^2(K,\mu) : \phi = u \ \mu-\text{a.e. for some } u \in \mathcal{F}_e\}, \\ \widetilde{\mathcal{E}}(\phi,\phi) &= \mathcal{E}(H_{\mathcal{S}}u, H_{\mathcal{S}}u), \ \phi \in \widetilde{\mathcal{F}}, \phi = u \ \mu-\text{a.e.}, u \in \mathcal{F}_e,\end{aligned}$$

where $H_{\mathcal{S}}u(x) = E^x u(\widehat{X}_{\sigma_{\mathcal{S}}})$ for $x \in K$ and $\mathcal{F}_e$ is the extended Dirichlet space associated with $(\mathcal{E}, \mathcal{F})$ (p. 35 of [9]). By Theorem 5.1.5 and Theorem 6.2.1 of [9], $(\widetilde{\mathcal{E}}, \widetilde{\mathcal{F}})$ is a closed regular Dirichlet form on $L^2(K,\mu)$. The next proposition shows that $\mathcal{E} = \widetilde{\mathcal{E}}$, so that the Dirichet form is not affected by the time change.

Proposition 2.6 *Assume* (2.2). *Then*

$$P^x(\sigma_{\mathcal{S}} = 0) = 1 \quad \forall x \in K. \tag{2.8}$$

PROOF. If $d_s < 2$, then $\mathcal{S} = K$ and (2.8) is clear. We now prove (2.8) for $d_s > 2$: a similar proof works for $d_s = 2$. Let $A_n(x) = \{y \in K : l^{-n-1} \le |x-y|_{\max} < l^{-n}\}$, and $\mathcal{S}^n(x) = \mathcal{S} \cap A_n(x)$. Note that as $\mu_i > 0$ for all $i \in S$, $\mu(\mathcal{S}^n(x)) \ne 0$ for $n \ge 0$. To prove (2.8) it is enough to prove that

$$\sum_{n=1}^{\infty} l^{n(d_f-d_w)} \ \mathrm{Cap}(\mathcal{S}^n(x_0)) = \infty \quad \forall x_0 \in K. \tag{2.9}$$

If $K=[0,1]^d$ this is just the classical Wiener test (see [15] or [22]); the result used here follows, using Proposition 2.4, by exactly the same arguments.

Using Kelvin's principle (see Section 2.2 of [9]), we have for each compact set $M \subset K$,

$$\{\mathrm{Cap}(M)\}^{-1} = \inf \int_{M\times M} \widehat{g}_1(x,y)m(dx)m(dy), \tag{2.10}$$

where the infimum is taken over the positive Radon measures m with $m(M)=1$. Now, take an arbitrary compact set $M \subset \mathcal{S}^n(x_0)$ such that $\mu(M) \geq \mu(\mathcal{S}^n(x_0))/2$. Then $\mu(M\cap A_k(x)) \leq \mu(\mathcal{S}^n(x_0)\cap A_k(x)) \leq c_1\mu(\mathcal{S}^n(x_0))(\mu^*)^{k-n} \leq 2c_1\mu(M)(\mu^*)^{k-n}$ for $k\geq n$. So, using Proposition 2.4 as before, we have

$$\begin{aligned}
\mu(M)^{-2}\int_{M\times M}\widehat{g}_1(x,y)\mu(dx)\mu(dy) &\leq \mu(M)^{-2}\int_M \mu(dx)\int_M c_2|x-y|_{\max}^{d_w-d_f}\mu(dy)\\
&\leq c_3\mu(M)^{-2}\int_M\mu(dx)\sum_{k=n}^{\infty} l^{k(d_f-d_w)}\mu(M\cap A_k(x))\\
&\leq c_4\mu(M)^{-2}\int_M\mu(dx)\sum_{k=n}^{\infty} l^{k(d_f-d_w)}\mu(M)(\mu^*)^{k-n}\\
&= c_4 l^{n(d_f-d_w)}\sum_{j=0}^{\infty} l^{j(d_f-d_w)}(\mu^*)^j \leq c_5 l^{n(d_f-d_w)}.
\end{aligned}$$

Here we used the fact that $l^{d_f-d_w}\mu^* = \rho_K^{-1}\mu^* < 1$ (due to assumption (2.2)) in the last inequality. Taking $m(\cdot)=\mu(\cdot\cap M)/\mu(M)$ in (2.10), we have $\mathrm{Cap}(M)\geq c_5^{-1}l^{-n(d_f-d_w)}$. As this holds for all compact sets $M\subset\mathcal{S}^n(x_0)$ with $\mu(M)\geq\mu(\mathcal{S}^n(x_0))/2$, we have $\mathrm{Cap}(\mathcal{S}^n(x_0))\geq c_5^{-1}l^{-n(d_f-d_w)}$, which proves (2.9). □

We thus obtain a closed local regular Dirichlet form $(\mathcal{E},\widetilde{\mathcal{F}})$ on $L^2(K,\mu)$ with the property (2.1): write X for the associated diffusion.

We next show that (2.3) holds. If $d_s<2$ then this is easy to verify directly. We omit the argument for $d_s=2$: it follows by similar arguments to those for $d_s>2$.

For $A\subset K$ write $T_A=\inf\{t\geq 0: X_t\in A\}$. Let $x\in K-V_*$, and $\sigma_n=T_{D_n(x)^c}$. Let $g(x,y)$ be the (0th order) Green's function for $\widehat{X}$ killed on hitting $D_n(x)^c$. From [5] we have that $g(x,y)\leq c|x-y|^{d_w-d_f}$. As the 0th-order Green's function is not affected by the time change, by a calculation similar to that in Lemma 2.5

$$\begin{aligned}
E^x\sigma_n &= \int_{D_n(x)} g(x,y)\mu(dy)\leq c_1\int_{D_n(x)}|x-y|^{d_w-d_f}d\mu\\
&\leq c_2\sum_{k=n}^{\infty} l^{k(d_f-d_w)}(\mu^*)^k \leq c_3(t_*)^{-n}\leq c_3e^{-c_4 n}.
\end{aligned}\tag{2.11}$$

Define $U^\lambda f(x)=E^x[\int_0^\infty e^{-\lambda t}f(X_t)dt]$ for $\lambda>0$. Using (2.11), the argument of [5] Proposition 6.14 goes through (with some modification) to prove that there exists $\beta>0$ such that

$$|U^\lambda f(x)-U^\lambda f(y)|\leq c_1(2+\lambda^{-1})|x-y|^\beta\|f\|_\infty \quad \forall x,y\in K. \tag{2.12}$$

Then, by Ascoli-Arzelà's theorem, X has a compact resolvent. It is clear that $P_t f = f$ if and only if f is constant. Thus there is a spectral gap, and (2.3) holds.

Finally, we prove the joint continuity of the transition density. As P_t is a self-adjoint compact operator on $L^2(K,\mu)$, there exist $\widetilde{\varphi}_i$ that form a complete orthonormal system in $L^2(K,\mu)$ with

$$p_t(x,y) = \sum_{i=1}^{\infty} e^{-\lambda_i t}\widetilde{\varphi}_i(x)\widetilde{\varphi}_i(y), \qquad \text{for } \mu^2\text{-a.e. } (x,y).$$

Further, the convergence is absolute and takes place in $L^\infty(K\times K)$ (see [5] Proposition 6.15). As $P_t\widetilde{\varphi}_i = e^{-\lambda_i}\widetilde{\varphi}_i$ a.e., we have $U^\lambda\widetilde{\varphi}_i = (\lambda+\lambda_i)^{-1}\widetilde{\varphi}_i$ a.e.. Defining $\varphi_i = (\lambda+\lambda_i)U^\lambda\widetilde{\varphi}_i$, we have $\varphi_i = \widetilde{\varphi}_i$ a.e. and by (2.12), φ_i is continuous. On the other hand, by a routine argument from (2.3) plus hitting time estimates in Proposition 4.3 which will be proved later, we have

$$p_t(x,x) \le c_1 t^{-\gamma_1} \qquad \forall x \in K, 0<t<1, \tag{2.13}$$

for some $\gamma_1 > 0$ depending only on the Dirichlet form and μ (the detailed argument will be given for the killed process on $D_{\Lambda_m(x)}$ in Proposition 5.1). Note that this upper bound is not sharp, but it is enough to deduce the continuity of $p_t(x,y)$. By this estimate, we see that $\|P_t\|_{2\to\infty} \le c_1 t^{-\gamma_1/2}$. Thus,

$$\|P_t\varphi_i\|_\infty = e^{-\lambda_i t}\|\varphi_i\|_\infty \le c_1 t^{-\gamma_1/2} \qquad \forall i.$$

Taking $t = 1/\lambda_i$, $\|\varphi_i\|_\infty \le c_2\lambda_i^{-\gamma_1/2}$. Thus we can take a version of transition density as

$$p_t(x,y) = \sum_n e^{-\lambda_n t}\varphi_n(x)\varphi_n(y),$$

and the convergence is uniform. This proves joint continuity of the density.

We thus obtain a Dirichlet form on $L^2(K,\mu)$ which satisfies Assumption 2.3.

3 Main Theorems

In the following, we identify S^n and the set of all n-complexes. For $x,y \in K$, we say $\Pi = \{x, x_1, \cdots, x_l, y\}$ is an *m-walk* of length $|\Pi| = l+1$ if $x_1,\cdots,x_l \in S^m, x \in F_{x_1}(K), y \in F_{x_l}(K)$ and F_{x_i}, $F_{x_{i+1}}$ are adjacent m-complexes for $1 \le i \le l-1$. For simplicity, we will assume the following in the p.c.f. self-similar set case.

Assumption 3.1 *Let $\mathcal{L} = (K, S, \{F_s\}_{s\in S})$ be a p.c.f. self-similar set. For each $x,y \in V_0$ with $x \ne y$ and for each $m \ge 0$,*

$$\min\{|\Pi| : \Pi \textit{ is a } m\textit{-walk from } x \textit{ to } y\} \ge 2^m.$$

We remark that if the self-similar structure $\mathcal{L}$ is changed to $\mathcal{L}_l = (K, S^l, \{F_w\}_{w\in S^l})$, then for sufficiently large l, $\mathcal{L}_l$ satisfies Assumption 3.1. For $x \in K\backslash V_*$, let $n_{r,j}(x)$ be the shortest number of steps by a $(\Lambda_r(x)+j)$-walk from x to $\partial D_{\Lambda_r}(x)$. Further, for $x \in K \setminus V_*$, define

$$p_r(x) = \min\{k : x \notin D_{\Lambda_r(x)+k}(\partial D_{\Lambda_r}(x))\}.$$

Since $\cap_k D_k(A) = A$ if $A \subset K$ is closed, we have $p_r(x) < \infty$ for $x \in K \setminus V_*$. Note that $p_{n-1}(x) \le p_n(x) + C$ for some $C > 0$, independent of n and x.

Lemma 3.2 *There exist $c_{3.1} > 0$ such that the following holds for all $r, j \geq 0$ and $x \in K \setminus V_*$.*

$$e^{(j-p_r(x))_+/2} \leq 2^{(j-p_r(x))_+} \leq n_{r,j}(x) \leq c_{3.1} e^{([\log t^*]+1)j/2}, \tag{3.1}$$

where for $x \in \mathbb{R}$, $x_+ = \max(x, 0)$.

PROOF. Using Assumption 3.1 in the p.c.f. case, and the fact that $l \geq 2$ in the carpet case, we have $n_{r,j}(x) \geq 2^{(j-p_r(x))_+}$. Since $e < 4$ the first inequality is immediate.

Noting that if $r \geq n \log t^*$ then $\Lambda_r(x) > n$ for all $x \in K - V_*$, the third inequality can be obtained by an easy modification of the proof of Lemma 3.3 in [13]. □

Now let $\widehat{\mu}$ be any Bernoulli measure on K so that $\widehat{\mu}(V_0) = 0$, $\widehat{\mu}(F_i(K)) = \widehat{\mu}_i$ for all i where $\widehat{\mu}_i \geq 0$ satisfies $\sum_i \widehat{\mu}_i = 1$. We emphasise that we continue to consider the density of X with respect to μ: the role of $\widehat{\mu}$ will be to select subsets of K with different limiting behaviour of $p_t(x, x)$.

As in [11], Lemma 6, we have

Proposition 3.3 *There exists $\alpha > 0$ and $g : K \to [0, \infty)$ such that for $\hat{\mu}$-a.e. $x \in K$,*

$$(rj)^{-\alpha} e^{j/2} \leq n_{r,j}(x) \leq c_{3.1} e^{([\log t^*]+1)j/2} \qquad \forall r \geq 0, \forall j \geq g(x).$$

PROOF. First note that if $x \notin D_k(V_0)$ then any j-walk from x to V_0 requires at least 2^{j-k} steps. Also, there exists $\theta < 1$ such that $\widehat{\mu}(D_k(V_0)) \leq c_0 \theta^k$ for $k \geq 1$.

Since each Λ_n complex is a scaled copy of K it follows that for $2 \leq k \leq j$,

$$\widehat{\mu}(\{x : n_{r,j}(x) \leq 2^{j-k}\}) \leq \; c_0 \theta^k.$$

So, if $\alpha_0 = \log 2 / \log(1/\rho)$,

$$\widehat{\mu}(\{x : 2^{-j} n_{r,j}(x) \leq \lambda) \leq c_1 \lambda^{1/\alpha_0} \qquad \text{for } 0 < \lambda < 1.$$

So

$$\begin{aligned} \widehat{\mu}(\{x : 2^{-j} n_{r,j}(x) \leq \lambda r^{-\alpha_0 - 1} \text{ for some } r \geq 0\}) \;&\leq\; \sum_{r=0}^{\infty} \widehat{\mu}(2^{-j} n_{r,j}(x) \leq \lambda r^{-\alpha_0 - 1}) \\ &\leq\; \sum_{r} c_1 (\lambda r^{-\alpha_0 - 1})^{1/\alpha_0} \leq c_2 \lambda^{1/\alpha_0}. \end{aligned}$$

Taking $\lambda = j^{-\alpha_0 - 1}$ and applying Borel-Cantelli it follows that, $\widehat{\mu}$-a.e., $n_{r,j}(x) \geq (rj)^{-\alpha_0 - 1} 2^j$ for all sufficiently large j. Putting $\alpha = \alpha_0 + 1$, the first inequality follows. The second inequality is the same as in (3.1). □

We now state our main theorems.

Theorem 3.4 *There exists $c_{3.2}, \cdots, c_{3.8} > 0$ such that the following holds.*
1) (Lower estimate) For each $x \in K \setminus V_*$ and $t \le c_{3.2} e^{-n'}$ $(n' = n + c_{3.3} p_n(x))$,

$$c_{3.4}\{\mu(D_{\Lambda_n}(x))\}^{-1} \le p_t(x,x).$$

2) (Upper estimate) For each $x \in K \setminus V_*$ and for each t which satisfies $c_{3.5} e^{-n''} \le t \le c_{3.6} e^{-m' - c_{3.7}\log n'}$ $(n' = n + c_{3.3} p_n(x),\ n'' = n' + c_{3.7}\log n')$ for some $m \le n$,

$$p_t(x,x) \le c_{3.8}\{ \min_{\substack{w \in \Lambda_{n''} \\ F_w(K) \subset D_{\Lambda_m}(x)}} \mu_w \}^{-1}.$$

We note that in general n' is not monotone increasing w.r.t. n. As we do not have a good comparison between $p_t(x,x)$ and $p_{t/2}(x,x)$, we need an extra $\log n'$ in the time interval for the upper estimate. See the proof, which will be given in Section 5, for details.

Remark. For the p.c.f. fractals, we can also obtain the following estimate by the similar (but simpler) argument to the proof of Theorem 3.4.
There exist $c_{3.9}, c_{3.10}, c_{3.11} > 0$ such that for each $x \in K$ and for each $e^{-(n+1)} \le t \le e^{-n}$,

$$c_{3.9}\{\mu(D^1_{\Lambda_{n-c_{3.10}}}(x))\}^{-1} \le p_t(x,x) \le c_{3.11}\{ \min_{\substack{w \in \Lambda_n \\ F_w(K) \subset D^1_{\Lambda_n}(x)}} \mu_w \}^{-1}. \tag{3.2}$$

Note that (3.2) always gives some estimate of the kernel for each fixed t whereas Theorem 3.4 does not (unless t is small enough). On the other hand, when t is in the interval where Theorem 3.4 gives the estimate, it is much sharper than that of (3.2).

Concerning the lower bound for the p.c.f. fractals, (3.2) with $c_{3.10} = 0$, which is sharper than (3.2), is proved in Section 5 of [18]. For the p.c.f. fractals with a 'regular harmonic structure', sharper upper estimate of $p_t(x,x)$ is given in appendix of [12].

We cannot obtain (3.2) for the carpet cases as $D^1_{\Lambda_n}(x)$ could have a very 'bad shape' in general. When the sizes of two adjacent Λ_n-complexes in $D^1_{\Lambda_n}(x)$ are very different, particles could escape from $D^1_{\Lambda_n}(x)$ much faster than e^{-n}.

Using Proposition 3.3, we have the following almost sure result.

Theorem 3.5 *Let $\widehat{\mu}$ be a Bernoulli measure on K with weights $\widehat{\mu}_i$ satisfying the hypotheses of Proposition 3.3. There exist $c_{3.12}, c_{3.13}, c_{3.14} > 0$ and $h : K \to [0,1]$ such that*
1) $h(x) > 0 \qquad \hat{\mu} - a.e.$,
2) For $t < h(x)$, if $e^{-(n+1)} \le t \le e^{-n}$, then

$$c_{3.13}(\log\frac{1}{t})^{-c_{3.12}}\mu(D_{\Lambda_n}(x))^{-1} \le p_t(x,x) \le c_{3.14}(\log\frac{1}{t})^{c_{3.12}}\mu(D_{\Lambda_n}(x))^{-1}. \tag{3.3}$$

The proof is essentially the same as that of Theorem 3.4: we will remark on the necessary modifications in the last part of Section 5.

Now, let

$$d_s(\hat{\mu})/2 = \frac{\sum_i \hat{\mu}_i \log(1/\mu_i)}{\sum_i \hat{\mu}_i \log(\rho_i/\mu_i)}. \tag{3.4}$$

Corollary 3.6 *The following holds for $\hat{\mu}$-a.e. $x \in K$:*

$$-\lim_{t\to 0} \frac{\log p_t(x,x)}{\log t} = d_s(\hat{\mu})/2.$$

PROOF. For $x \in K - V_*$ let $x_1x_2\dots$ be the word in Σ associated with x. Let $M_k(x) = \log(t_{x_1}\dots t_{x_n})$. Since, under $\widehat{\mu}$, x_i are independent identically distributed random variables, using the strong law of large numbers $M_k/k \to \sum_i \hat{\mu}_i \log t_i$ for $\hat{\mu} - a.e.\ x \in K$.

Now if $\Lambda_n(x) = k$ then $M_k(x) > n \geq M_k(x) - \log t^*$, and therefore for $\widehat{\mu}$-a.e. $x \in K$

$$\lim_{n\to\infty} \frac{\Lambda_n(x)}{n} = \left(\lim_{k\to\infty} \frac{M_k(x)}{k}\right)^{-1} = \frac{1}{\sum_i \widehat{\mu}_i \log(\rho_i/\mu_i)}. \tag{3.5}$$

Finally, by Theorem 3.5, we have, $\widehat{\mu}$-a.e.,

$$\lim_{t\to 0} \frac{\log p_t(x,x)}{\log t} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{\Lambda_n(x)} \log \mu_{x_i} = \lim_{n} \frac{\Lambda_n(x)}{n} \cdot \lim_{n} \frac{1}{n} \sum_{i=1}^{n} \log \mu_{x_i},$$

and using (3.5) and the law of large numbers completes the proof. □

REMARKS. 1) The formula for the $\hat{\mu}$-a.e. spectral dimension (3.4) has the same form as that for the (stationary) homogeneous random Sierpinski gasket studied in [7]. However, in that case, $\{\hat{\mu}_i\}_i$ corresponds to the frequency that each random pattern appears.
2) The condition (2.2) implies that $0 < d_s(\hat{\mu}) < \infty$. Also, for the p.c.f. case with regular harmonic structure (2.4), we have $d_s(\hat{\mu})/2 < 1$, as one expects. If $\widehat{\mu}$ is given by (2.6), then $d_s(\hat{\mu}) = \sigma/(\sigma+1)$ and so is independent of $\hat{\mu}$; this has already been proved in [13].

Note that while $d_s(\mu)/2 \geq d_s^e(\mu)/2$ (recall (2.5)), these two numbers are in general not equal. This lack of correspondence between the asymptotic growth of the eigenvalues and transition density emphasises the lack of uniformity in the behaviour of $p_t(x,x)$.
3) Let $\sigma(\hat{\mu}) = (\sum_i \hat{\mu}_i \log(1/\mu_i))/(\sum_i \hat{\mu}_i \log \rho_i)$ when $\sum_i \hat{\mu}_i \log \rho_i \neq 0$. We can then write

$$d_s(\hat{\mu})/2 = \frac{\sigma(\hat{\mu})}{\sigma(\hat{\mu})+1},$$

with the convention that $\sigma(\hat{\mu}) = \infty$, $d_s(\hat{\mu})/2 = 1$ when $\sum_i \hat{\mu}_i \log \rho_i = 0$. Since $0 < d_s(\hat{\mu}) < \infty$, $\sigma(\hat{\mu}) \notin [-1, 0]$ – this can also be proved directly from (2.2). As $d_s(\hat{\mu})/2$ is increasing w.r.t. $\sigma(\hat{\mu})$, one can calculate the region of $d_s(\hat{\mu})/2$ as $\hat{\mu}$ varies.

The case of $[0,1]$ with $\mu_1 = p > 1/2$, $\mu_2 = q < 1/2$, $p + q = 1$, and $\rho_i = 2$ was discussed in the introduction; some properties of this diffusion process were studied in [10]. For this case, $\log_2(1/p) < \sigma(\hat{\mu}) < \log_2(1/q)$ and by Proposition 4.7 in [10] (plus an easy Tauberian argument), the corresponding value for dyadic rational points is $\log_2(1/p)$, the infimum of the interval. Using the continuity of $p_t(x,x)$ and Corollary 3.6 one can show there exist points for which σ takes the maximum value $\log_2(1/q)$.
4) In the carpet case the sign of $\sigma(\hat{\mu})$ is the same as that of $(\rho_K - 1)$, and so, for example, if $\rho_K < 1$ then $d_s(\hat{\mu}) > 2$ for any $\hat{\mu}$.

5) If $\min_i \log \rho_i / \log(1/\mu_i) \neq \max_i \log \rho_i / \log(1/\mu_i)$, using Corollary 3.6 and the continuity of $p_t(x,y)$, one can prove by an elementary argument that there are (uncountably many) x such that $\lim_{t\to 0} \log p_t(x,x)/\log t$ does not exist.
6) Theorem 1.1 now follows from remarks 3) and 5).
7) Using the methods of [6] and [18], together with the (worse) lower bound in (5.2), which will be used for the proof of Theorem 3.4, and a chaining argument we have

$$p_t(x,y) > 0 \qquad \forall t > 0, \forall x, y \in K.$$

4 Hitting time estimates

In this section, we will prove some hitting time estimates for the process which will be needed for the transition density estimates.

We first give some notation. For each $x \in K$ and $m \geq 0$, we fix a m-complex which contains x and denote it as $D_m^*(x)$. Note that $D_m^*(x) = D_m(x)$ if $x \in K \setminus V_*$. For $x \in K$ and $m \geq 0$, let

$$D_m^{1,*}(x) = \{C : C \text{ is a } m\text{-complex which intersects } D_m^*(x).\}$$

Let $\Lambda_r^*(x)$ be the length of the word of $D_{\Lambda_r}^*(x)$, and define $D_{\Lambda_r(x)+l}^{1,*}(x)$ to be the union of $(\Lambda_r^*(x)+l)$-complexes which intersect $D_{\Lambda_r(x)+l}^*(x)$. In the following, we will treat the case where $x_0 \in K,\ r, m \geq 0$ satisfy

$$D_{\Lambda_r(x_0)+m}^{1,*}(x_0) \subset D_{\Lambda_r}^*(x_0). \tag{4.1}$$

Note that when (4.1) holds for the carpet case, all $(\Lambda_r^*(x_0)+m)$-complexes are same size, so that $D_{\Lambda_r(x_0)+m}^{1,*}(x_0)$ is a cube (or square when $d = 2$).

We will consider the process on $D_{\Lambda_r(x_0)+m}^{1,*}(x_0)$ whose Dirichlet form is the same as that of $\{X_t\}$ but whose measure $\widetilde{\mu}$ satisfies $\widetilde{\mu}(K_{wv_1\cdots v_n}) = \mu_w(\mu_*)^m \mu_{v_{m+1}} \cdots \mu_{v_n}$ for each $(\Lambda_r^*(x_0)+n)$-complex $K_{wv_1\cdots v_n} \subset K_w = D_{\Lambda_r}^*(x_0)$ and all $n \geq m$. Let $\widetilde{X}_t$ be the process associated with $\mathcal{E}$ and $L^2(K, \widetilde{\mu})$ killed on $\partial D_{\Lambda_r(x_0)+m}^{1,*}(x_0)$. As before, we define $T_A = \inf\{t \geq 0 : X_t \in A\}$ for $A \subset K$, and write $T_A(\widetilde{X})$ for the analogous hitting time for $\widetilde{X}$. We then have the following.

Lemma 4.1 *Let $x_0 \in K, r, m \geq 0$ satisfy* (4.1).
1) There exist $0 < c_{4.1} < c_{4.2}$ (independent of x_0, r, m) which satisfy the following,

$$c_{4.1} e^{-(r+m\log t^*)} \leq E^{x_0}(T_{\partial D_{\Lambda_r(x_0)+m}^{1,*}(x_0)}(\widetilde{X})),$$
$$E^x(T_{\partial D_{\Lambda_r(x_0)+m}^{1,*}(x_0)}(\widetilde{X})) \leq c_{4.2} e^{-(r+m\log t^*)} \qquad \forall x \in D_{\Lambda_r(x_0)+m}^{1,*}(x_0).$$

2) There exist $0 < c_{4.3},\ 0 < c_{4.4} < 1$ (independent of x_0, r, m) such that for all $0 < t < 1$,

$$P^{x_0}(T_{\partial D_{\Lambda_r(x_0)+m}^{1,*}(x_0)}(\widetilde{X}) \leq t) \leq c_{4.3} e^{r+m\log t^*} t + c_{4.4}.$$

Proof. For the p.c.f. case, 1) can be proved by a simple modification of Lemma 3.5 in [13] (one can also prove it using Green's density killed at $\partial D^{1,*}_{\Lambda_r(x_0)+m}(x_0)$). For the carpet case, 1) is proved in the same way as Proposition 5.5 in [5].

2) now follows directly from 1), as in Lemma 3.6 of [13], or Lemma 3.16 of [1]. $\square$

Note that $\widetilde{X}_t$ is a time change of X_t, so that the trajectory of $\{\widetilde{X}_t\}$ is the same as that of $\{X_t\}$, but that $\widetilde{X}_t$ moves faster than X_t. Therefore, we have

$$P^{x_0}(T_{\partial D^{1,*}_{\Lambda_r(x_0)+m}(x_0)}(\widetilde{X}) \le t) \ge P^{x_0}(T_{\partial D^{1,*}_{\Lambda_r(x_0)+m}(x_0)}(X) \le t).$$

Thus, we have from Lemma 4.1 2) that

$$P^{x_0}(T_{\partial D^{1,*}_{\Lambda_r(x)+m}(x)}(X) \le t) \le c_{4.3} e^{r+m\log t^*} t + c_{4.4}, \tag{4.2}$$

for each $0 < t < 1$ and for each $x_0 \in K, r, m \ge 0$ which satisfies (4.1).

Now, for a process X on K and for $n \ge 0$, we define a sequence of hitting times as follows,

$$\begin{aligned}
\sigma^n_0(X) &= 0, \\
\sigma^n_1(X) &= \inf\{t \ge 0 : X_t \in \partial D^{1,*}_n(X_0)\}, \\
\sigma^n_{k+1}(X) &= \inf\{t \ge \sigma^n_k(X) : X_t \in \partial D^{1,*}_n(X_{\sigma^n_k})\}, \quad \forall k \ge 1, \\
W^n_k &= \sigma^n_k(X) - \sigma^n_{k-1}(X), \quad \forall k \ge 1.
\end{aligned}$$

We have the following estimate of the crossing time distribution.

Lemma 4.2 *There exist $c_{4.5}, c_{4.6}, c_{4.7} > 0$ such that for all $0 \le r$,*

$$P^x(T_{\partial D_{\Lambda_r(x)}} \le c_{4.7} t) \le c_{4.5} \exp\{-c_{4.6} n_{r,k}(x)\} \qquad 0 < \forall t < 1, \forall x \in K \setminus V_*, \tag{4.3}$$

where $k = k(r,z,t) = \inf\{j : \frac{n_{r,j}(z)}{e^{r+j\log t^*}} \le t\}$.

Remark. Note that, from (3.1), $k < \infty$ for each r, t, z.

Proof. As the result is clear when $k = 0$, we consider the case $k \ge 1$. First, for each $n \ge r \ge 0$, set

$$\xi_{r,n} = \sup\{i : \sigma^n_i \le T_{\partial D_{\Lambda_r(x)}}\}.$$

By the structure of K, there exists $c_1 > 0$ such that

$$c_1 n_{r,m} \le \xi_{r,\Lambda_r(x)+m} \qquad \forall r, m \ge 0.$$

We also note that for each $r, m \ge 0$ and $k \le \xi_{r,\Lambda_r(x)+m} - 1$, $\{W^{\Lambda_r(x)+m}_k\}_k$ behave like $\{T_{D^{1,*}_{\Lambda_r(x)+m}(X_{\sigma^{\Lambda_r(x)+m}_{k-1}})}\}_k$ so that

$$P(W^{\Lambda_r(x)+m}_k \le t | W^{\Lambda_r(x)+m}_j, 1 \le j \le k-1) \le c_{4.3} e^{r+m\log t^*} t + c_{4.4} \tag{4.4}$$

holds for $k \leq \xi_{r,\Lambda_r(x)+m} - 1$ by (4.2) ((4.1) clearly holds in these cases). Using these facts, we have for each $x \in K \setminus V_*$,

$$\begin{aligned} P^x(T_{\partial D_{\Lambda_r(x)}} \leq c_{4.7}t) &\leq P^x(\sum_{i=1}^{c_1 n_{r,k-1}(x)} W_i^{\Lambda_r(x)+k-1} \leq c_{4.7}t) \\ &\leq \exp\left(c_2(n_{r,k-1}(x)e^{r+(k-1)\log t^*}c_{4.7}t)^{\frac{1}{2}} - c_3 n_{r,k-1}(x)\right) \\ &= \exp(-c_4 n_{r,k-1}(x)) \leq \exp(-c_5 n_{r,k}(x)), \end{aligned} \tag{4.5}$$

where we use Lemma 1.1 of [2] and (4.4) in the second inequality, and in the last equality we choose $c_{4.7}$ so that $c_{4.7}^{1/2} < c_3/c_2$ ($c_4 = c_3 - c_2 c_{4.7}^{1/2}$). We thus obtain the result. □

Let $\beta \equiv \log t^*$. For each $c > 0$, set

$$k_c = k_c(n, x, t) = \inf\{j : \frac{n_{n,j}(x)}{e^{n+\beta j}} \leq ct/c_{4.7}\}.$$

Then, by Lemma 4.2, we have for each $0 < t < 1$ and $x \in K \setminus V_*$,

$$P^x(T_{\partial D_{\Lambda_n(x)}} \leq ct) \leq c_{4.5}\exp\{-c_{4.6}n_{n,k_c}(x)\}. \tag{4.6}$$

We now have the following exponential decay of the distribution of hitting times.

Proposition 4.3 *There exist $c_{4.8}, \gamma > 0$ such that for each $x \in K \setminus V_*, 0 < s < 1$ and for each $t \leq e^{-n'}$ $(n' = n + \beta p_n(x))$,*

$$P^x(T_{\partial D_{\Lambda_n(x)}} \leq st) \leq c_{4.5}\exp(-c_{4.8}s^{-\gamma}).$$

PROOF. First, take $a = \min\{1, \beta\}/2 > 0$ and take $\bar{m} \in \mathbf{Z}$ such that

$$e^{-(\beta-a)(\bar{m}+1)} \leq s/c_{4.7} < e^{-(\beta-a)\bar{m}}. \tag{4.7}$$

Then, using (3.1),

$$\frac{n_{n,j}}{e^{n+\beta j}} \geq e^{a(j-p_n(x))-n-\beta j} > \frac{s}{c_{4.7}}e^{-\beta p_n(x)-n} \geq st/c_{4.7} \tag{4.8}$$

for all $j \leq \bar{m} + p_n(x), t \leq e^{-n-\beta p_n(x)}$ so that

$$k_s(n, x, t) - 1 \geq \bar{m} + p_n(x). \tag{4.9}$$

Thus, $k_s - 1 - p_n(x) \geq \bar{m}$. Now, from (3.1) and (4.9), we have

$$n_{n,k_s}(z) \geq n_{n,k_s-1}(z) \geq e^{\frac{1}{2}(k_s-1-p_n(x))} \geq e^{\frac{1}{2}\bar{m}}$$

so that the right hand side of (4.6) is less than $c_{4.5}\exp(-c_{4.6}e^{\frac{1}{2}\bar{m}})$. By (4.7), $e^{\frac{1}{2}\bar{m}} \geq c_1 s^{-\gamma}$ where $\gamma = (2(\beta - a))^{-1}$ and the proof is completed. □

5 Transition density estimates

5.1 Lower bound

In this subsection, we will obtain the lower bound of the transition density.
PROOF OF THEOREM 3.4 1). Set $c_{3.3}=\beta$ and let $s>0$ satisfy $c_{4.5}\exp\{-c_{4.8}s^{-\gamma}\}\le 1/2$. Then, by Proposition 4.3, we have

$$P^x(T_{\partial D_{\Lambda_n}(x)}\le st)\le 1/2 \qquad \forall x\in K\setminus V_*,\ \forall t\le e^{-n'}.$$

Thus, $P^x(X_{st}\in D_{\Lambda_n}(x))\ge 1/2$. Now, using Cauchy-Schwarz,

$$\begin{aligned}(1/2)^2\le P^x(X_{st}\in D_{\Lambda_n}(x))^2 &= \Big(\int_{D_{\Lambda_n}(x)} p_{st}(x,y)\mu(dy)\Big)^2 \\ &\le \mu(D_{\Lambda_n}(x))\int_{D_{\Lambda_n}(x)} p_{st}(x,y)^2\mu(dy) \\ &\le \mu(D_{\Lambda_n}(x))p_{2st}(x,x).\end{aligned} \tag{5.1}$$

Hence we deduce that $p_{2st}(x,x)\ge c_1\{\mu(D_{\Lambda_n}(x))\}^{-1}$. □
In Lemma 5.1 of [13], one of the author proves similar results, but the proof is incomplete as it could be carried out only when $k>0$ where k, which appears in the proof, is defined similarly to ours. The proof can be completed following the argument of ours, using Proposition 4.3.
By a similar argument, it is easy to prove the weaker estimate

$$c_1\{\mu(D^1_{[n/\log t^*]}(x))\}^{-1}\le p_t(x,x) \qquad \forall x\in K\setminus V_*,\ \forall t\le c_2e^{-n}.$$

Since $\mu(D^1_{[n/\log t^*]}(x))\le c_3(\mu^*)^{c_4 n}$, we have

$$p_t(x,x)\ge c_5t^{-\gamma_2} \qquad \forall x\in K\setminus V_*,\ 0<\forall t<1, \tag{5.2}$$

where $\gamma_2=c_4\log(1/\mu^*)$. This uniform (worse) lower estimate will be used later.

5.2 Upper bounds

In this subsection, we will obtain the upper bound of the transition density. For the purpose, we will first obtain the upper estimate of the kernel with Dirichlet boundary condition. For each $x\in K$ and $n\ge 0$, let $p_t^{D_{\Lambda_n}(x)}(x,y)$ be the transition density of the process killed at $\partial D_{\Lambda_n}(x)$.

Proposition 5.1 *There exists $c_{5.1},c_{5.2}>0$ such that for each $x\in K, c_{5.1}e^{-n}\le t$ and $m\le n$,*

$$p_t^{D_{\Lambda_m}(x)}(x,x)\le c_{5.2}\{\min_{\substack{w\in\Lambda_n\\ F_w(K)\subset D_{\Lambda_m}(x)}}\mu_w\}^{-1}.$$

PROOF. For $w\in\Lambda_m$ write $f_w=f\circ F_w$ and define

$$\bar f_w=\int_K f_w(x)\mu(dx)=\mu_w^{-1}\int_{F_w(K)} f(x)\mu(dx).$$

Note that for $v \in \mathcal{F}$, $\bar{v} = \int v d\mu = \sum_{w\in\Lambda_m} \bar{v}_w \mu_w$. In the following, we fix x, m and denote $B = D_{\Lambda_m}(x)$. Let $u_0 \in \mathcal{D}(\mathcal{L})$ with $u_0 \geq 0$, Supp $u_0 \subset B$ and $\|u_0\|_1 = 1$. Set $u_t(x) = (P_t^B u_0)(x)$ and $g(t) = \|u_t\|_2^2$. We remark that g is continuous and decreasing. As the semigroup is symmetric and Markov,

$$\|u_t\|_1 = \int P_t^B u_0 d\mu = \int u_0 P_t^B 1 d\mu \leq \|u_0\|_1 = 1.$$

For each $l \geq 0$,

$$\begin{aligned}
\frac{d}{dt} g(t) &= 2(\mathcal{L}u_t, u_t) \\
&= -2\mathcal{E}(u_t, u_t) \\
&= -2 \sum_{\substack{w\in\Lambda_{m+l} \\ F_w(K)\subset B}} \rho_w \mathcal{E}(u_t \circ F_w, u_t \circ F_w) \\
&\leq -2c_1 \sum_w \rho_w \int (u_{t,w} - \bar{u}_{t,w})^2 d\mu \qquad \text{(by (2.3))} \\
&= -2c_1 \sum_w \rho_w \mu_w^{-1} \int_{F_w(K)} (u_t)^2 d\mu + 2c_1 \sum_w \rho_w (\mu_w^{-1} \int_{F_w(K)} u_t d\mu)^2 \\
&\leq -2c_1 e^{m+l} \|u_t\|_2^2 + 2c_1 e^{m+l+1} C_B^l \sum_w (\int_{F_w(K)} u_t d\mu)^2 \\
&\leq -2c_1 e^{m+l} g(t) + 2c_1 e^{m+l+1} C_B^l (\sum_w \int_{F_w(K)} u_t d\mu)^2 \\
&\leq -2c_1 e^{m+l} (g(t) - eC_B^l),
\end{aligned}$$

where $C_B^l \equiv \max_{\substack{w\in\Lambda_{m+l} \\ F_w(K)\subset B}} \mu_w^{-1}$. Therefore

$$-\frac{d}{dt} \log (g(t) - eC_B^l) \geq c_2 e^{m+l}, \quad \text{if } g(t) > eC_B^l. \tag{5.3}$$

Note that $C_B^l \leq C_B^{l+1}$ and $C_B^l \to \infty$ as $l \to \infty$. Let $s_l = \inf\{t \geq 0 : g(t) \leq eC_B^l\}$ for $l \in \mathbb{N}$. Thus (5.3) holds for $0 < t < s_l$. Note that $s_l \to 0$ as $l \to \infty$. Integrating (5.3) from s_{l+2} to s_{l+1} we obtain

$$\begin{aligned}
c_2 e^{m+l} (s_{l+1} - s_{l+2}) &\leq -\log (g(s_{l+1}) - eC_B^l) + \log (g(s_{l+2}) - eC_B^l) \\
&= \log (C_B^{l+2} - C_B^l)/(C_B^{l+1} - C_B^l) \leq c_3.
\end{aligned}$$

Thus $s_{l+1} - s_{l+2} \leq c_4 e^{-(m+l)}$, and iterating this we have

$$s_l \leq c_4 \sum_{k=l-1}^{\infty} e^{-(m+k)} \leq c_5 e^{-(m+l)}.$$

This implies that $g(c_5/e^{m+l}) \leq g(s_l) = eC_B^l$. Taking $l = n - m$, it follows that if $t/c_5 \geq e^{-n}$, then

$$g(t) \leq c_7 \{ \min_{\substack{w\in\Lambda_n \\ F_w(K)\subset B}} \mu_w \}^{-1}.$$

Using the fact that $\|P_t\|_{1\to\infty} \le \|P_t\|_{1\to 2}^2$, we deduce the result. $\square$

Proof of Theorem 3.4 2). The main step in the proof is to compare the transition densities of X killed at $\partial D_{\Lambda_m}(x)$ with those of the unkilled process. Write $\bar{X}$ for X killed at $\partial D_{\Lambda_m}(x)$, and let $\tau = T_{\partial D_{\Lambda_m}(x)}$. First, by Proposition 4.3,

$$P^x(\tau \le s\bar{t}) \le c_{4.5}\exp\{-c_{4.8}s^{-\gamma}\},$$

for $\bar{t} \le e^{-m'}$. Set $t = s\bar{t}$. We first assume $e^{-n'} \le \bar{t} \le e^{-m'}$ for some $m \le n$ and later determine the right value of s (right interval for t) where the comparison of Dirichlet and Neumann boundaries holds. For $k \ge n$ define $B_k = D_{\Lambda_k}(x) \subset D_{\Lambda_n}(x)$. Then

$$\begin{aligned}\int_{B_k}\int_{B_k} p_t(z,y)\mu(dz)\mu(dy) &= \int_{B_k} P^z(X_t \in B_k)\mu(dz)\\ &= \int_{B_k} P^z(\bar{X}_t \in B_k)\mu(dz) + \int_{B_k} P^z(X_t \in B_k, \tau < t)\mu(dz)\end{aligned}$$

The second term above equals

$$\int_{B_k} P^z(X_t \in B_k, \tau \le t/2)\mu(dz) + \int_{B_k} P^z(X_t \in B_k, t/2 < \tau < t)\mu(dz) = J_1 + J_2 \qquad (5.4)$$

Since the process X started with measure μ is symmetric,

$$\begin{aligned}J_2 &= P^\mu(X_0 \in B_k, X_t \in B_k, t/2 < \tau < t)\\ &\le P^\mu(X_0 \in B_k, X_t \in B_k, \exists s \in [t/2, t) : X_s \in \partial D_{\Lambda_m}(x))\\ &= P^\mu(X_t \in B_k, X_0 \in B_k, \exists s \in (0, t/2] : X_s \in \partial D_{\Lambda_m}(x))\\ &= P^\mu(X_0 \in B_k, X_t \in B_k, \tau \le t/2) = J_1\end{aligned}$$

Write $a(t/2) = \sup\{p_s(y,y) : y \in K, t/2 \le s \le t\}$. We have

$$\begin{aligned}P^z(X_t \in B_k, \tau \le t/2) &= E^z 1_{(\tau \le t/2)} P^{X_\tau}(X_{t-\tau} \in B_k)\\ &= E^z 1_{(\tau \le t/2)} \int_{B_k} p_{t-\tau}(X_\tau, y)\mu(dy)\\ &\le P^z(\tau \le t/2)a(t/2)\mu(B_k).\end{aligned}$$

Combining these estimates, letting $k \to \infty$, and using the continuity of $p_t(\cdot,\cdot)$, we deduce that

$$p_t(x,x) \le p_t^{D_{\Lambda_m}(x)}(x,x) + 2a(t/2)P^x(\tau \le t/2).$$

By (2.13) and (5.2), there exists $a > 0$ such that

$$a(t/2) \le c_1 t^{-a} p_t(z,z) \qquad \forall z \in K \setminus V_*,\ 0 < \forall t < 1.$$

Thus by Proposition 4.3

$$p_t(x,x) \le p_t^{D_{\Lambda_m}(x)}(x,x) + c_2 t^{-a} p_t(x,x)\exp\{-c_3 s^{-\gamma}\}. \tag{5.5}$$

Thus the last term in (5.5) is estimated from above by $c_2 s^{-a}\exp\{an' - c_3 s^{-\gamma}\}p_t(x,x)$ when $e^{-n'} \le \bar{t}$. Now, by taking $s = (hn')^{-l} < 1$ with sufficiently large $l, h > 0$, we can take $c_2 s^{-a}\exp\{an' - c_3 s^{-\gamma}\} \le 1/2$ for all $n \ge 1$. We thus have

$$p_t(x,x) \le \frac{1}{2} p_t^{D_{\Lambda_m}(x)}(x,x) \qquad \forall x \in K \setminus V_*,\ c_4(n')^{-l}e^{-n'} \le \forall t \le c_5(n')^{-l}e^{-m'}.$$

We thus obtain the result using Proposition 5.1. □

Note for the proof of Theorem 3.5.

First, it is enough to prove the following:

There exist $C_1, \cdots, C_6 > 0$ and $k(x) < \infty$ such that for $\hat{\mu}$-a.e. $x \in K$ and for each $\bar{n} = n + C_1 \log n$, $\widetilde{n} = n + C_2 \log n$, $n > k(x)$ (note that $\bar{n}$, $\widetilde{n}$ are increasing w.r.t. n),

$$C_3\{\mu(D_{\Lambda_{n-C_4}}(x))\}^{-1} \le p_t(x,x) \qquad \forall t \le e^{-\bar{n}}, \tag{5.6}$$

$$p_t(x,x) \le C_5\{\min_{\substack{w \in \Lambda_{\widetilde{n+C_6}} \\ F_w(K) \subset D_{\Lambda_{(n+C_6)}}(x)}} \mu_w\}^{-1} \qquad e^{-(\widetilde{n}+1)} \le \forall t \le e^{-\widetilde{n}}. \tag{5.7}$$

Indeed, by taking $\theta = \mu_*^{-\log t^*} > 1$, we see

$$\mu(D_{\Lambda_{n-r}}(x)) \le \mu(D_{\Lambda_n}(x))\theta^r \qquad \forall r \ge 0.$$

Thus, noting $\bar{n} - C_1 \log \bar{n} \le n \le \bar{n}$ (as $\bar{n} = n + C_1 \log n \le n + C_1 \log \bar{n}$), we have

$$\begin{aligned} \mu(D_{\Lambda_{n-C_4}}(x)) &\le \mu(D_{\Lambda_{\bar{n}}}(x))\theta^{\bar{n}-n+C_4} \\ &\le \mu(D_{\Lambda_{\bar{n}}}(x))\theta^{c_1 \log \bar{n}} = e^{c_2 \log \bar{n}}\mu(D_{\Lambda_{\bar{n}}}(x)) \\ &\le (\log \frac{1}{t})^{c_2}\mu(D_{\Lambda_{\bar{n}}}(x)), \end{aligned}$$

so that we can replace the left hand side of (5.6) by $(\log \frac{1}{t})^{-c_2}\mu(D_{\Lambda_{\bar{n}}}(x))^{-1}$. By a similar argument for the upper bound, we obtain Theorem 3.5. Note that by taking $c_{3.12}$ large in (3.3), we can state the upper and lower bounds simultaneously.

Now, in order to apply the proof in this section for (5.6) and (5.7), the following modification is needed. First, taking $c' < 1/2$, we have by Proposition 3.3 that for $\hat{\mu}$-a.e $x \in K$, there exists $\bar{g}(x) < \infty$ such that

$$r^{-\alpha}e^{c'j} \le n_{r,j}(x) \qquad \forall r \ge 0,\ \forall j \ge \bar{g}(x).$$

We will change the definition of k_b so that

$$k_b = \inf\{j \ge \bar{g}(x) : \frac{n_{r,j}(x)}{e^{r+\beta j}} \le bt/c_{4.7}, \frac{n_{r,j-1}(x)}{e^{r+\beta(j-1)}} \ge bt/c_{4.7}\}. \tag{5.8}$$

Then (4.3) holds for this k_b (with bt instead of t). Noting that $p_r(x)/2$ in Lemma 3.2 corresponds to $\alpha \log r$ in this case, it is not hard to modify Proposition 4.3 using k_b in (5.8). Then, the lower and upper bounds can be proved in the same way. The extra $\log n'$ term is taken into $C_2 \log n$. □

Acknowledgment. The authors thank B. M. Hambly, J. Kigami and H. Osada for their fruitful comments. Part of this research was done while the authors were visiting the Isaac Newton Institute, Cambridge, during the program on Mathematics and Applications of Fractals.

References

[1] M. T. Barlow, Diffusions on fractals, *Lectures on Probability Theory and Statistics: Ecole d'Eté de Probabilités de Saint-Flour XXV*, Springer, 1998.

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

[3] M. T. Barlow and R. F. Bass, Local times for Brownian motion on the Sierpinski carpet, *Probab. Theory Relat. Fields,* **85** (1990), 91–104.

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

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

[6] M. T. Barlow, R. F. Bass and K. Burdzy, Positivity of Brownian transition densities, *Elect. Comm. in Probab.*, **2** (1997), 43–51.

[7] M. T. Barlow and B. M. Hambly, Transition density estimates for Brownian motion on scale irregular Sierpinski gasket, *Ann. Inst. H. Poincaré*, **33** (1997), 531–557.

[8] M. T. Barlow and E. A. Perkins, Brownian motion on the Sierpinski gasket, *Probab. Theory Relat. Fields*, **79** (1988), 543–624.

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

[10] T. Fujita, Some asymptotic estimates of transition probability densities for generalized diffusion processes with self-similar speed measures, *Publ. Res. Inst. Math. Sci.*, **26** (1990), 819–840.

[11] B. M. Hambly and O. D. Jones, Asymptotically one-dimensional diffusion on the Sierpinski gasket and multi-type branching processes with varying environment, preprint 1998.

[12] B. M. Hambly, J. Kigami and T. Kumagai, Multifractal formalisms for the local spectral and walk dimensions, *Maths. Proc. Cam. Phil. Soc.*, to appear.

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

[14] B. M. Hambly, T. Kumagai, S. Kusuoka and X. Y. Zhou, Transition density estimates for diffusion processes on homogeneous random Sierpinski carpets, J. Math. Soc. Japan, **52** (2000), 373–408.

[15] K. Itô and H. P. McKean, Jr., *Diffusion processes and thier sample paths (Second printing)*, Springer, Berlin, 1974.

[16] J. Kigami, Harmonic calculus on P.C.F. self-similar sets, *Trans. Amer. Math. Soc.*, **335** (1993), 721–755.

[17] J. Kigami, Markov property of Kusuoka-Zhou's Dirichlet forms on self-similar sets, *J. Math. Sci. Univ. Tokyo*, **7** (2000), 27–33.

[18] J. Kigami, Analysis on fractals, *Cambridge Univ. Press*, to appear.

[19] J. Kigami and M. L. Lapidus, Weyl's spectral problem for the spectral distribution of Laplacians on P.C.F. self-similar fractals, *Comm. Math. Phys.*, **158** (1993), 93–125.

[20] T. Kumagai, Regularity, closedness and spectral dimensions of the Dirichlet forms on P.C.F. self-similar sets, *J. Math. Kyoto Univ.*, **33** (1993), 765–786.

[21] S. Kusuoka and X. Y. Zhou, Dirichlet forms on fractals: Poincaré constant and resistance, *Probab. Theory Relat. Fields*, **93** (1992), 169–196.

[22] N. S. Landkof, *Foundations of Modern Potential Theory*, Springer, New York-Heidelberg, 1972.

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

[24] H. Osada, A family of diffusion processes on Sierpinski carpets, *Probab. Theory Relat. Fields*, **119** (2001), 275–310.

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