Uniqueness of Brownian motion on Sierpinski carpets

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

Richard F. Bass †
Department of Mathematics, University of Connecticut
Storrs CT 06269-3009 USA
Email: bass@math.uconn.edu

Takashi Kumagai ‡§
Department of Mathematics, Faculty of Science
Kyoto University, Kyoto 606-8502, Japan
Email: kumagai@math.kyoto-u.ac.jp

and

Alexander Teplyaev¶
Department of Mathematics, University of Connecticut
Storrs CT 06269-3009 USA
Email: teplyaev@math.uconn.edu

31 August 2009

Abstract

We prove that, up to scalar multiples, there exists only one local regular Dirichlet form on a generalized Sierpinski carpet that is invariant with respect to the local symmetries of the carpet. Consequently for each such fractal the law of Brownian motion is uniquely determined and the Laplacian is well defined.

*Research partially supported by NSERC (Canada), and EPSRC (UK).
†Research partially supported by NSF grant DMS-0601783.
‡Corresponding author
§Research partially supported by the Grant-in-Aid for Scientific Research (B) 18340027.
¶Research partially supported by NSF grant DMS-0505622.

1 Introduction

The standard Sierpinski carpet $F_{\rm SC}$ is the fractal that is formed by taking the unit square, dividing it into 9 equal subsquares, removing the central square, dividing each of the 8 remaining subsquares into 9 equal smaller pieces, and continuing. In [3] two of the authors of this paper gave a construction of a Brownian motion on $F_{\rm SC}$. This is a diffusion (that is, a continuous strong Markov process) which takes its values in $F_{\rm SC}$, and which is non-degenerate and invariant under all the local isometries of $F_{\rm SC}$.

Subsequently, Kusuoka and Zhou in [30] gave a different construction of a diffusion on $F_{\rm SC}$, which yielded a process that, as well as having the invariance properties of the Brownian motion constructed in [3], was also scale invariant. The proofs in [3, 30] also work for fractals that are formed in a similar manner to the standard Sierpinski carpet: we call these *generalized Sierpinski carpets* (GSCs). In [5] the results of [3] were extended to GSCs embedded in $\mathbb{R}^d$ for $d \geq 3$. While [3, 5] and [30] both obtained their diffusions as limits of approximating processes, the type of approximation was different: [3, 5] used a sequence of time changed reflecting Brownian motions, while [30] used a sequence of Markov chains.

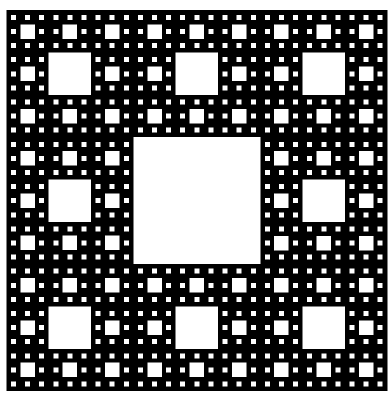

Figure 1: The standard Sierpinski carpet

These papers left open the question of uniqueness of this Brownian motion – in fact it was not even clear whether or not the processes obtained in [3, 5] or [30] were the same. This uniqueness question can also be expressed in analytic terms: one can define a *Laplacian* on a GSC as the infinitesimal generator of a Brownian motion, and one wants to know if there is only one such Laplacian. The main result of this paper is that, up to scalar multiples of the time parameter, there exists only one such Brownian motion; hence, up to scalar multiples, the Laplacian is uniquely defined.

GSCs are examples of spaces with *anomalous diffusion*. For Brownian motion on $\mathbb{R}^d$ one has $\mathbb{E}|X_t - X_0| = ct^{1/2}$. Anomalous diffusion in a space F occurs when instead one has $\mathbb{E}|X_t - X_0| = o(t^{1/2})$, or (in regular enough situations), $\mathbb{E}|X_t - X_0| \approx t^{1/d_w}$, where d_w (called the *walk dimension*) satisfies $d_w > 2$. This phenomena was first observed by mathematical physicists working in the transport properties of disordered media, such as (critical) percolation clusters – see [1, 37]. Since these sets are subsets of the lattice $\mathbb{Z}^d$, they are not true fractals, but their large scale structure still exhibits fractal properties, and the simple random walk is expected to have anomalous diffusion.

For critical percolation clusters (or, more precisely for the incipient infinite cluster) on trees and $\mathbb{Z}^2$, Kesten [23] proved that anomalous diffusion occurs. After this work, little progress was made on critical percolation clusters until the recent papers [7, 8, 27].

As random sets are hard to study, it was natural to begin the study of anomalous diffusion in the more tractable context of regular deterministic fractals. The simplest of these is the Sierpinski gasket. The papers [1, 37] studied discrete random walks on graph approximations to the Sierpinski gasket, and soon after [19, 29, 11] constructed Brownian motions on the limiting set. The special structure of the Sierpinski gasket makes the uniqueness problem quite simple, and uniqueness of this Brownian motion was proved in [11]. These early papers used a probabilistic approach, first constructing the Brownian motion X on the space, and then, having defined the Laplacian $\mathcal{L}_X$ as the infinitesimal generator of the semigroup of X, used the process X to study $\mathcal{L}_X$. Soon after Kigami [24] and Fukushima-Shima [18] introduced more analytical approaches, and in particular [18] gave a very simple construction of X and $\mathcal{L}_X$ using the theory of Dirichlet forms.

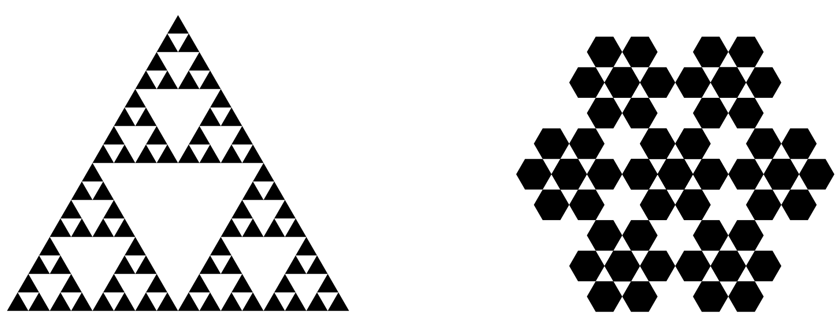

Figure 2: The Sierpinski gasket (left), and a typical nested fractal, the Lindstrøm snowflake (right)

It was natural to ask whether these results were special to the Sierpinski gasket. Lindstrøm [31] and Kigami [25] introduced wider families of fractals (called *nested fractals*, and *p.c.f. self-similar sets* respectively), and gave constructions of diffusions on these spaces. Nested fractals are, like the Sierpinski carpet, highly symmetric, and the uniqueness problem can be formulated in a similar fashion to that for GSCs. Uniqueness for nested fractals was not treated in [31], and for some years remained a significant challenge, before being solved by Sabot [41]. (See also [33, 36] for shorter proofs). For p.c.f. self-similar sets, while some sufficient conditions for uniqueness are given in [41, 21], the general problem is still open.

The study of these various families of fractals (nested fractals, p.c.f self-similar sets, and GSCs) revealed a number of common themes, and showed that analysis on these spaces differs from that in standard Euclidean space in several ways, all ultimately connected with the fact that $d_w > 2$:

- The energy measure ν and the Hausdorff measure μ are mutually singular,
- The domain of the Laplacian is not an algebra,
- If $d(x, y)$ is the shortest path metric, then $d(x, \cdot)$ is not in the domain of the Dirichlet form.

See [2, 26, 43] for further information and references.

The uniqueness proofs in [21, 33, 36, 41] all used in an essential way the fact that nested fractals and p.c.f. self-similar sets are finitely ramified – that is, they can be disconnected by removing a finite number of points. For these sets there is a natural definition of a set V_n of 'boundary points at level n' – for the Sierpinski gasket V_n is the set of vertices of triangles of side 2^{-n}. If one just looks at the process X at the times when it passes through the points in V_n, one sees a finite state Markov chain $X^{(n)}$, which is called the *trace of* X *on* V_n. If $m > n$ then $V_n \subset V_m$ and the trace of $X^{(m)}$ on V_n is also $X^{(n)}$. Using this, and the fact that the limiting processes are known to be scale invariant, the uniqueness problem for X can be reduced to the uniqueness of the fixed point of a non-linear map on a space of finite matrices.

While the boundaries of the squares (or cubes) have an analogous role to the sets V_n in the geometrical construction of a GSC, attempts to follow the same strategy of proof encounter numerous difficulties and have not been successful. We use a different idea in this paper, and rather than studying the restriction of the process X to boundaries, our argument treats the Dirichlet form of the process on the whole space. (This also suggests a new approach to uniqueness on finitely ramified fractals, which will be explored elsewhere.)

Let F be a GSC and μ the usual Hausdorff measure on F. Let $\mathfrak{E}$ be the set of non-zero local regular conservative Dirichlet forms $(\mathcal{E}, \mathcal{F})$ on $L^2(F, \mu)$ which are invariant with respect to all the local symmetries of F. (See Definition 2.15 for a precise definition.) We remark that elements of $\mathfrak{E}$ are not required to be scale invariant – see Definition 2.17. Our first result is that $\mathfrak{E}$ is non-empty.

Proposition 1.1 *The Dirichlet forms associated with the processes constructed in [3, 5] and [30] are in* $\mathfrak{E}$.

Our main result is the following theorem, which is proved in Section 5.

Theorem 1.2 *Let* $F \subset \mathbb{R}^d$ *be a GSC. Then, up to scalar multiples,* $\mathfrak{E}$ *consists of at most one element. Further, this one element of* $\mathfrak{E}$ *satisfies scale invariance.*

An immediate corollary of Proposition 1.1 and Theorem 1.2 is the following.

Corollary 1.3 *The Dirichlet forms constructed in [3, 5] and [30] are (up to a constant) the same.*
(b) The Dirichlet forms constructed in [3, 5] satisfy scale invariance.

A Feller process is one where the semigroup T_t maps continuous functions that vanish at infinity to continuous functions that vanish at infinity, and $\lim_{t\to 0} T_t f(x) = f(x)$ for each $x \in F$ if f is continuous and vanishes at infinity. Our main theorem can be stated in terms of processes as follows.

Corollary 1.4 *If X is a continuous non-degenerate symmetric strong Markov process which is a Feller process, whose state space is F, and whose Dirichlet form is invariant with respect to the local symmetries of F, then the law of X under $\mathbb{P}^x$ is uniquely defined, up to scalar multiples of the time parameter, for each $x \in F$.*

Remark 1.5 Osada [35] constructed diffusion processes on GSCs which are different from the ones considered here. While his processes are invariant with respect to some of the local isometries of the GSC, they are not invariant with respect to the full set of local isometries.

In Section 2 we give precise definitions, introduce the notation we use, and prove some preliminary lemmas. In Section 3 we prove Proposition 1.1. In Section 4 we develop the properties of Dirichlet forms $\mathcal{E} \in \mathfrak{E}$, and in Section 5 we prove Theorem 1.2.

The idea of our proof is the following. The main work is showing that if $\mathcal{A}, \mathcal{B}$ are any two Dirichlet forms in $\mathfrak{E}$, then they are comparable. (This means that $\mathcal{A}$ and $\mathcal{B}$ have the same domain $\mathcal{F}$, and that there exists a constant $c = c(\mathcal{A}, \mathcal{B}) > 0$ such that $c\mathcal{A}(f, f) \le \mathcal{B}(f, f) \le c^{-1}\mathcal{A}(f, f)$ for $f \in \mathcal{F}$.) We then let λ be the largest positive real such that $\mathcal{C} = \mathcal{A} - \lambda\mathcal{B} \ge 0$. If $\mathcal{C}$ were also in $\mathfrak{E}$, then $\mathcal{C}$ would be comparable to $\mathcal{B}$, and so there would exist $\varepsilon > 0$ such that $\mathcal{C} - \varepsilon\mathcal{B} \ge 0$, contradicting the definition of λ. In fact we cannot be sure that $\mathcal{C}$ is closed, so instead we consider $\mathcal{C}_\delta = (1 + \delta)\mathcal{A} - \lambda\mathcal{B}$, which is easily seen to be in $\mathfrak{E}$. We then need uniform estimates in δ to obtain a contradiction.

To show $\mathcal{A}, \mathcal{B} \in \mathfrak{E}$ are comparable requires heat kernel estimates for an arbitrary element of $\mathfrak{E}$. Using symmetry arguments as in [5], we show that the estimates for corner moves and slides and the coupling argument of [5, Section 3] can be modified so as to apply to any element $\mathcal{E} \in \mathfrak{E}$. It follows that the elliptic Harnack inequality holds for any such $\mathcal{E}$. Resistance arguments, as in [4, 34], combined with results in [20] then lead to the desired heat kernel bounds. (Note that the results of [20] that we use are also available in [10].)

A key point here is that the constants in the Harnack inequality, and consequently also the heat kernel bounds, only depend on the GSC F, and not on the particular element of $\mathfrak{E}$. This means that we need to be careful about the dependencies of the constants.

The symmetry arguments are harder than in [5, Section 3]. In [5] the approximating processes were time changed reflecting Brownian motions, and the proofs used the convenient fact that a reflecting Brownian motion in a Lipschitz domain in $\mathbb{R}^d$ does not hit sets of dimension $d-2$. Since we do not have such approximations for the processes corresponding to an arbitrary element $\mathcal{E} \in \mathfrak{E}$, we have to work with the diffusion X associated with $\mathcal{E}$, and this process might hit sets of dimension $d-2$. (See [5, Section 9] for examples of GSCs in dimension 3 for which the process X hits not just lines but also points.)

We use C_i to denote finite positive constants which depend only on the GSC, but which may change between each appearance. Other finite positive constants will be written as c_i.

2 Preliminaries

2.1 Some general properties of Dirichlet forms

We begin with a general result on local Dirichlet forms. For definitions of local and other terms related to Dirichlet forms, see [17]. Let F be a compact metric space and m a Radon (i.e. finite) measure on F. For any Dirichlet form $(\mathcal{E},\mathcal{F})$ on $L^2(F,m)$ we define

$$\mathcal{E}_1(u,u) = \mathcal{E}(u,u) + \|u\|_2^2. \tag{2.1}$$

Functions in $\mathcal{F}$ are only defined up to quasi-everywhere equivalence (see [17] p. 67); we use a quasi-continuous modification of elements of $\mathcal{F}$ throughout the paper. We write $\langle \cdot,\cdot \rangle$ for the inner product in $L^2(F,m)$ and $\langle \cdot,\cdot \rangle_S$ for the inner product in a subset $S \subset F$.

Theorem 2.1 *Suppose that $(\mathcal{A},\mathcal{F})$, $(\mathcal{B},\mathcal{F})$ are local regular conservative irreducible Dirichlet forms on $L^2(F,m)$ and that*

$$\mathcal{A}(u,u) \le \mathcal{B}(u,u) \quad \text{for all } u \in \mathcal{F}. \tag{2.2}$$

Let $\delta > 0$, and $\mathcal{E} = (1+\delta)\mathcal{B} - \mathcal{A}$. Then $(\mathcal{E},\mathcal{F})$ is a regular local conservative irreducible Dirichlet form on $L^2(F,m)$.

Proof. It is clear that $\mathcal{E}$ is a non-negative symmetric form, and is local.

To show that $\mathcal{E}$ is closed, let $\{u_n\}$ be a Cauchy sequence with respect to $\mathcal{E}_1$. Since $\mathcal{E}_1(f,f) \ge (\delta \wedge 1)\mathcal{B}_1(f,f)$, $\{u_n\}$ is a Cauchy sequence with respect to $\mathcal{B}_1$. Since $\mathcal{B}$ is a Dirichlet form and so closed, there exists $u \in \mathcal{F}$ such that $\mathcal{B}_1(u_n - u, u_n - u) \to 0$. As $\mathcal{A} \le \mathcal{B}$ we have $\mathcal{A}(u_n - u, u_n - u) \to 0$ also, and so $\mathcal{E}_1(u_n - u, u_n - u) \to 0$, proving that $(\mathcal{E},\mathcal{F})$ is closed.

Since $\mathcal{A}$ and $\mathcal{B}$ are conservative and F is compact, $1 \in \mathcal{F}$ and $\mathcal{E}(1,h) = 0$ for all $h \in \mathcal{F}$, which shows that $\mathcal{E}$ is conservative by [17, Theorem 1.6.3 and Lemma 1.6.5].

We now show that $\mathcal{E}$ is Markov. By [17, Theorem 1.4.1] it is enough to prove that $\mathcal{E}(\bar{u},\bar{u}) \le \mathcal{E}(u,u)$ for $u \in \mathcal{F}$, where we let $\bar{u} = 0 \vee (u \wedge 1)$. Since $\mathcal{A}$ is local and $u_+ u_- = 0$, we have $\mathcal{A}(u_+,u_-) = 0$ ([42, Proposition 1.4]). Similarly $\mathcal{B}(u_+,u_-) = 0$, giving $\mathcal{E}(u_+,u_-) = 0$. Using this, we have

$$\mathcal{E}(u,u) = \mathcal{E}(u_+,u_+) - 2\mathcal{E}(u_+,u_-) + \mathcal{E}(u_-,u_-) \ge \mathcal{E}(u_+,u_+) \tag{2.3}$$

for $u \in \mathcal{F}$. Now let $v = 1 - u$. Then $\bar{u} = (1 - v_+)_+$, so

$$\begin{aligned}\mathcal{E}(u,u) = \mathcal{E}(v,v) \geq \mathcal{E}(v_+,v_+) = \mathcal{E}(1-v_+,1-v_+)\\ \geq \mathcal{E}((1-v_+)_+,(1-v_+)_+) = \mathcal{E}(\bar{u},\bar{u}),\end{aligned}$$

and hence $\mathcal{E}$ is Markov.

As $\mathcal{B}$ is regular, it has a core $\mathcal{C} \subset \mathcal{F}$. Let $u \in \mathcal{F}$. As $\mathcal{C}$ is a core for $\mathcal{B}$, there exist $u_n \in \mathcal{C}$ such that $\mathcal{B}_1(u - u_n, u - u_n) \to 0$. Since $\mathcal{A} \leq \mathcal{B}$, $\mathcal{A}_1(u_n - u, u_n - u) \to 0$ also, and so $\mathcal{E}_1(u_n - u, u_n - u) \to 0$. Thus $\mathcal{C}$ is dense in $\mathcal{F}$ in the $\mathcal{E}_1$ norm (and it is dense in $C(F)$ in the supremum norm since it is a core for $\mathcal{B}$), so $\mathcal{E}$ is regular.

Let $A \subset F$ be invariant for the semigroup corresponding to $\mathcal{E}$. By [17, Theorem 1.6.1], this is equivalent to the following: $1_A u \in \mathcal{F}$ for all $u \in \mathcal{F}$ and

$$\mathcal{E}(u,v) = \mathcal{E}(1_A u, 1_A v) + \mathcal{E}(1_{F-A} u, 1_{F-A} v) \qquad \forall u,v \in \mathcal{F}. \tag{2.4}$$

Once we have $1_A u \in \mathcal{F}$, since $(1_A u)(1_{F-A} u) = 0$ we have $\mathcal{A}(1_A u, 1_{F-A} u) = 0$, and we obtain (2.4) for $\mathcal{A}$ also. Using [17, Theorem 1.6.1] again, we see that A is invariant for the semigroup corresponding to $\mathcal{A}$. Since $\mathcal{A}$ is irreducible, we conclude that either $m(A) = 0$ or $m(X - A) = 0$ holds and hence that $(\mathcal{E}, \mathcal{F})$ is irreducible. □

Remark 2.2 This should be compared with the situation for Dirichlet forms on finite sets, which is the context of the uniqueness results in [33, 41]. In that case the Dirichlet forms are not local, and given $\mathcal{A}$, $\mathcal{B}$ satisfying (2.2) there may exist $\delta_0 > 0$ such that $(1+\delta)\mathcal{B} - \mathcal{A}$ fails to be a Dirichlet form for $\delta \in (0, \delta_0)$.

For the remainder of this section we assume that $(\mathcal{E}, \mathcal{F})$ is a local regular Dirichlet form on $L^2(F, m)$, that $1 \in \mathcal{F}$ and $\mathcal{E}(1,1) = 0$. We write T_t for the semigroup associated with $\mathcal{E}$, and X for the associated diffusion.

Lemma 2.3 *T_t is recurrent and conservative.*

Proof. T_t is recurrent by [17, Theorem 1.6.3]. Hence by [17, Lemma 1.6.5] T_t is conservative. □

Let D be a Borel subset of F. We write T_D for the hitting time of D, and τ_D for the exit time of D:

$$T_D = T_D^X = \inf\{t \geq 0 : X_t \in D\}, \qquad \tau_D = \tau_D^X = \inf\{t \geq 0 : X_t \notin D\}. \tag{2.5}$$

Let $\overline{T}_t$ be the semigroup of X killed on exiting D, and $\overline{X}$ be the killed process. Set

$$q(x) = \mathbb{P}^x(\tau_D = \infty),$$

and

$$E_D = \{x : q(x) = 0\}, \qquad Z_D = \{x : q(x) = 1\}. \tag{2.6}$$

Lemma 2.4 *Let D be a Borel subset of F. Then $m(D - (E_D \cup Z_D)) = 0$. Further, E_D and Z_D are invariant sets for the killed process $\overline{X}$, and Z_D is invariant for X.*

Proof. If $f \geq 0$,

$$\langle \overline{T}_t(f1_{E_D}), 1_{D-E_D}q\rangle = \langle f1_{E_D}, \overline{T}_t(1_{D-E_D}q)\rangle \leq \langle f1_{E_D}, \overline{T}_t q\rangle = 0.$$

So $\overline{T}_t(f1_{E_D}) = 0$ on $D - E_D$ and hence (see [17, Lemma 1.6.1(ii)]) E_D is invariant for $\overline{X}$.

Let $A = \{x : P^x(\tau_D < \infty) > 0\} = Z_D^c$. The set A is an invariant set of the process X by [17, Lemma 4.6.4]. Using the fact that $\overline{X} = X$, $\mathbb{P}^x$-a.s. for $x \in Z_D$ and [17, Lemma 1.6.1(ii)], we see that A is an invariant set of the process $\overline{X}$ as well. So we see that Z_D is invariant both for X and $\overline{X}$. In order to prove $m(D-(E_D \cup Z_D)) = 0$, it suffices to show that $\mathbb{E}^x[\tau_D] < \infty$ for a.e. $x \in A \cap D$. Let U_D be the resolvent of the killed process $\overline{X}$. Since $A \cap D$ is of finite measure, the proof of Lemma 1.6.5 or Lemma 1.6.6 of [17] give $U_D 1(x) < \infty$ for a.e. $x \in A \cap D$, so we obtain $\mathbb{E}^x[\tau_D] < \infty$. □

Note that in the above proof we do not use the boundedness of D, but only the fact that $m(D) < \infty$.

Next, we give some general facts on harmonic and caloric functions. Let D be a Borel subset in F and let $h : F \to \mathbb{R}$. There are two possible definitions of h being harmonic in D. The probabilistic one is that h is harmonic in D if $h(X_{t\wedge\tau_{D'}})$ is a uniformly integrable martingale under $\mathbb{P}^x$ for q.e. x whenever D' is a relatively open subset of D. The Dirichlet form definition is that h is harmonic with respect to $\mathcal{E}$ in D if $h \in \mathcal{F}$ and $\mathcal{E}(h, u) = 0$ whenever $u \in \mathcal{F}$ is continuous and the support of u is contained in D.

The following is well known to experts. We will use it in the proofs of Lemma 4.9 and Lemma 4.24. (See [15] for the equivalence of the two notions of harmonicity in a very general framework.) Recall that $\mathbb{P}^x(\tau_D < \infty) = 1$ for $x \in E_D$.

Proposition 2.5 *(a) Let $(\mathcal{E}, \mathcal{F})$ and D satisfy the above conditions, and let $h \in \mathcal{F}$ be bounded. Then h is harmonic in a domain D in the probabilistic sense if and only if it is harmonic in the Dirichlet form sense.*
(b) If h is a bounded Borel measurable function in D and D' is a relatively open subset of D, then $h(X_{t\wedge\tau_{D'}})$ is a martingale under $\mathbb{P}^x$ for q.e. $x \in E_D$ if and only if $h(x) = \mathbb{E}^x[h(X_{\tau_{D'}})]$ for q.e. $x \in E_D$.

Proof. (a) By [17, Theorem 5.2.2], we have the Fukushima decomposition $h(X_t) - h(X_0) = M_t^{[h]} + N_t^{[h]}$, where $M^{[h]}$ is a square integrable martingale additive functional of finite energy and $N^{[h]}$ is a continuous additive functional having zero energy (see [17, Section 5.2]). We need to consider the Dirichlet form $(\mathcal{E}, \mathcal{F}_D)$ where $\mathcal{F}_D = \{f \in \mathcal{F} : \mathrm{supp}(f) \subset D\}$, and denote the corresponding semigroup as P_t^D.

If h is harmonic in the Dirichlet form sense, then by the discussion in [17, p. 218] and [17, Theorem 5.4.1], we have $\mathbb{P}^x(N_t^{[h]} = 0,\ \forall t < \tau_D) = 1$ q.e. $x \in F$. Thus, h is harmonic in the probabilistic sense. Here the notion of the spectrum from [17, Sect. 2.3] and especially [17, Theorem 2.3.3] are used.

To show that being harmonic in the probabilistic sense implies being harmonic in the Dirichlet form sense is the delicate part of this proposition. Since Z_D is

P_t^D-invariant (by Lemma 2.4) and $h(X_t)$ is a bounded martingale under $\mathbb{P}^x$ for $x \in Z_D$, we have

$$P_t^D(h1_{Z_D})(x) = 1_{Z_D}(x)P_t^D h(x) = 1_{Z_D}(x)E^x[h(X_t)] = h1_{Z_D}(x).$$

Thus by [17, Lemma 1.3.4], we have $h1_{Z_D} \in \mathcal{F}$ and $\mathcal{E}(h1_{Z_D}, v) = 0$ for all $v \in \mathcal{F}$. Next, note that on Z_D^c we have $H_B h = h$, according to the definition of H_B on page 150 of [17] and Lemma 2.4, which implies $H_B\left(h1_{Z_D^c}\right) = h1_{Z_D^c}$. Then from [17, Theorem 4.6.5], applied with $\widetilde{u} = h1_{Z_D^c} = h - h1_{Z_D} \in \mathcal{F}$ and $B^c = D$, we conclude that $h1_{Z_D^c}$ is harmonic in the Dirichlet form sense. Thus $h = h1_{Z_D^c} + h1_{Z_D}$ is harmonic in the Dirichlet form sense in D.

(b) If $h(X_{t\wedge\tau_{D'}})$ is a martingale under $\mathbb{P}^x$ for q.e. $x \in E_D$, then $\mathbb{E}^x[h(X_{s\wedge\tau_{D'}})] = \mathbb{E}^x[h(X_{t\wedge\tau_{D'}})]$ for q.e. $x \in E_D$ and for all $s, t \geq 0$, where we can take $s \downarrow 0$ and $t \uparrow \infty$ and interchange the limit and the expectation since h is bounded. Conversely, if $h(x) = E^x[h(X_{\tau_{D'}})]$ for q.e. $x \in E_D$, then by the strong Markov property, $h(X_{t\wedge\tau_{D'}}) = E^x[h(X_{\tau_{D'}})|\mathcal{F}_{t\wedge\tau_{D'}}]$ under $\mathbb{P}^x$ for q.e. $x \in E_D$, so $h(X_{t\wedge\tau_{D'}})$ is a martingale under $\mathbb{P}^x$ for q.e. $x \in E_D$. □

We call a function $u : \mathbb{R}_+ \times F \to \mathbb{R}$ caloric in D in the probabilistic sense if $u(t, x) = \mathbb{E}^x[f(X_{t\wedge\tau_D})]$ for some bounded Borel $f : F \to \mathbb{R}$. It is natural to view $u(t, x)$ as the solution to the heat equation with boundary data defined by $f(x)$ outside of D and the initial data defined by $f(x)$ inside of D. We call a function $u : \mathbb{R}_+ \times F \to \mathbb{R}$ caloric in D in the Dirichlet form sense if there is a function h which is harmonic in D and a bounded Borel $f_D : F \to \mathbb{R}$ which vanishes outside of D such that $u(t, x) = h(x) + \overline{T}_t f_D$. Note that $\overline{T}_t$ is the semigroup of X killed on exiting D, which can be either defined probabilistically as above or, equivalently, in the Dirichlet form sense by Theorems 4.4.3 and A.2.10 in [17].

Proposition 2.6 *Let $(\mathcal{E}, \mathcal{F})$ and D satisfy the above conditions, and let $f \in \mathcal{F}$ be bounded and $t \geq 0$. Then*

$$\mathbb{E}^x[f(X_{t\wedge\tau_D})] = h(x) + \overline{T}_t f_D$$

q.e., where $h(x) = \mathbb{E}^x[f(X_{\tau_D})]$ is the harmonic function that coincides with f on D^c, and $f_D(x) = f(x) - h(x)$.

Proof. By Proposition 2.5, h is uniquely defined in the probabilistic and Dirichlet form senses, and $h(x) = \mathbb{E}^x[h(X_{t\wedge\tau_D})]$. Note that $f_D(x)$ vanishes q.e. outside of D. Then we have $\mathbb{E}^x[f_D(X_{t\wedge\tau_D})] = \overline{T}_t f_D$ by Theorems 4.4.3 and A.2.10 in [17]. □

Note that the condition $f \in \mathcal{F}$ can be relaxed (see the proof of Lemma 4.9).

We show a general property of local Dirichlet forms which will be used in the proof of Proposition 2.21. Note that it is not assumed that $\mathcal{E}$ admits a *carré du champ*. Since $\mathcal{E}$ is regular, $\mathcal{E}(f, f)$ can be written in terms of a measure $\Gamma(f, f)$, the energy measure of f, as follows. Let $\mathcal{F}_b$ be the elements of $\mathcal{F}$ that are essentially bounded. If $f \in \mathcal{F}_b$, then $\Gamma(f,f)$ is defined to be the unique smooth Borel measure on F satisfying

$$\int_F g d\Gamma(f,f) = 2\mathcal{E}(f, fg) - \mathcal{E}(f^2, g), \qquad g \in \mathcal{F}_b.$$

Lemma 2.7 *If $\mathcal{E}$ is a local regular Dirichlet form with domain $\mathcal{F}$, then for any $f \in \mathcal{F} \cap L^\infty(F)$ we have $\Gamma(f,f)(A) = 0$, where $A = \{x \in F : f(x) = 0\}$.*

Proof. Let σ^f be the measure on $\mathbb{R}$ which is the image of the measure $\Gamma(f,f)$ on F under the function $f : F \to \mathbb{R}$. By [13, Theorem 5.2.1, Theorem 5.2.3] and the chain rule, σ^f is absolutely continuous with respect to one-dimensional Lebesgue measure on $\mathbb{R}$. Hence $\Gamma(f,f)(A) = \sigma^f(\{0\}) = 0$. □

Lemma 2.8 *Given a m-symmetric Feller process on F, the corresponding Dirichlet form $(\mathcal{E}, \mathcal{F})$ is regular.*

Proof. First, we note the following: if H is dense in $L^2(F,m)$, then $U^1(H)$ is dense in $\mathcal{F}$, where U^1 is the 1-resolvent operator. This is because $U^1 : L^2 \to \mathcal{D}(\mathcal{L})$ is an isometry where the norm of $g \in \mathcal{D}(\mathcal{L})$ is given by $\|g\|_{\mathcal{D}(\mathcal{L})} := \|(I - \mathcal{L})g\|_2$, and $\mathcal{D}(\mathcal{L}) \subset \mathcal{F}$ is a continuous dense embedding (see, for example [17, Lemma 1.3.3(iii)]). Here $\mathcal{L}$ is the generator corresponding to $\mathcal{E}$. Since $C(F)$ is dense in L^2 and $U^1(C(F)) \subset \mathcal{F} \cap C(F)$ as the process is Feller, we see that $\mathcal{F} \cap C(F)$ is dense in $\mathcal{F}$ in the $\mathcal{E}_1$-norm.

Next we need to show that $u \in C(F)$ can be approximated with respect to the supremum norm by functions in $\mathcal{F} \cap C(F)$. This is easy, since $T_t u \in \mathcal{F}$ for each t, is continuous since we have a Feller process, and $T_t u \to u$ uniformly by [39, Lemma III.6.7]. □

Remark 2.9 The proof above uses the fact that F is compact. However, it can be easily generalized to a Feller process on a locally compact separable metric space by a standard truncation argument – for example by using [17, Lemma 1.4.2(i)].

2.2 Generalized Sierpinski carpets

Let $d \geq 2$, $F_0 = [0,1]^d$, and let $L_F \in \mathbb{N}$, $L_F \geq 3$, be fixed. For $n \in \mathbb{Z}$ let $\mathcal{Q}_n$ be the collection of closed cubes of side L_F^{-n} with vertices in $L_F^{-n}\mathbb{Z}^d$. For $A \subseteq \mathbb{R}^d$, set

$$\mathcal{Q}_n(A) = \{Q \in \mathcal{Q}_n : \text{int}(Q) \cap A \neq \emptyset\}.$$

For $Q \in \mathcal{Q}_n$, let Ψ_Q be the orientation preserving affine map (i.e. similitude with no rotation part) which maps F_0 onto Q. We now define a decreasing sequence (F_n) of closed subsets of F_0. Let $1 \leq m_F \leq L_F^d$ be an integer, and let F_1 be the union of m_F distinct elements of $\mathcal{Q}_1(F_0)$. We impose the following conditions on F_1.

(H1) (Symmetry) F_1 is preserved by all the isometries of the unit cube F_0.

(H2) (Connectedness) $\mathrm{Int}(F_1)$ is connected.

(H3) (Non-diagonality) Let $m \geq 1$ and $B \subset F_0$ be a cube of side length $2L_F^{-m}$, which is the union of 2^d distinct elements of $\mathcal{Q}_m$. Then if $\mathrm{int}(F_1 \cap B)$ is non-empty, it is connected.

(H4) (Borders included) F_1 contains the line segment $\{x : 0 \leq x_1 \leq 1, x_2 = \cdots = x_d = 0\}$.

We may think of F_1 as being derived from F_0 by removing the interiors of $L_F^d - m_F$ cubes in $\mathcal{Q}_1(F_0)$. Given F_1, F_2 is obtained by removing the same pattern from each of the cubes in $\mathcal{Q}_1(F_1)$. Iterating, we obtain a sequence $\{F_n\}$, where F_n is the union of m_F^n cubes in $\mathcal{Q}_n(F_0)$. Formally, we define

$$F_{n+1} = \bigcup_{Q \in \mathcal{Q}_n(F_n)} \Psi_Q(F_1) = \bigcup_{Q \in \mathcal{Q}_1(F_1)} \Psi_Q(F_n), \quad n \geq 1.$$

We call the set $F = \cap_{n=0}^{\infty} F_n$ a generalized Sierpinski carpet (GSC). The Hausdorff dimension of F is $d_f = d_f(F) = \log m_F / \log L_F$. Later on we will also discuss the unbounded GSC $\widetilde{F} = \cup_{k=0}^{\infty} L_F^k F$, where $rA = \{rx : x \in A\}$.

Let

$$\mu_n(dx) = (L_F^d / m_F)^n 1_{F_n}(x) dx,$$

and let μ be the weak limit of the μ_n; μ is a constant multiple of the Hausdorff x^{d_f} - measure on F. For $x, y \in F$ we write $d(x, y)$ for the length of the shortest path in F connecting x and y. Using (H1)–(H4) we have that $d(x, y)$ is comparable with the Euclidean distance $|x - y|$.

Remark 2.10 1. There is an error in [5], where it was only assumed that (H3) above holds when $m = 1$. However, that assumption is not strong enough to imply the connectedness of the set J_k in [5, Theorem 3.19]. To correct this error, we replace the (H3) in [5] by the (H3) in the current paper.
2. The *standard SC* in dimension d is the GSC with $L_F = 3$, $m_F = 3^d - 1$, and with F_1 obtained from F_0 by removing the middle cube. We have allowed $m_F = L_F^d$, so that our GSCs do include the 'trivial' case $F = [0, 1]^d$. The 'Menger sponge' (see the picture on [32], p. 145) is one example of a GSC, and has $d = 3$, $L_F = 3$, $m_F = 20$.

Definition 2.11 Define:

$$\mathcal{S}_n = \mathcal{S}_n(F) = \{Q \cap F : Q \in \mathcal{Q}_n(F)\}.$$

We will need to consider two different types of interior and boundary for subsets of F which consist of unions of elements of $\mathcal{S}_n$. First, for any $A \subset F$ we write $\mathrm{int}_F(A)$ for the interior of A with respect to the metric space (F, d), and $\partial_F(A) = \overline{A} - \mathrm{int}_F(A)$. Given any $U \subset \mathbb{R}^d$ we write U^o for the interior of U in with respect to the usual topology on $\mathbb{R}^d$, and $\partial U = \overline{U} - U^o$ for the usual boundary of U. Let A be a finite union of elements of $\mathcal{S}_n$, so that $A = \cup_{i=1}^k S_i$, where $S_i = F \cap Q_i$ and $Q_i \in \mathcal{Q}_n(F)$. Then we define $\mathrm{int}_r(A) = F \cap ((\cup_{i=1}^k Q_i)^o)$, and $\partial_r(A) = A - \mathrm{int}_r(A)$. We have $\mathrm{int}_r(A) = A - \partial(\cup_{i=1}^k Q_i)$. (See Figure 3).

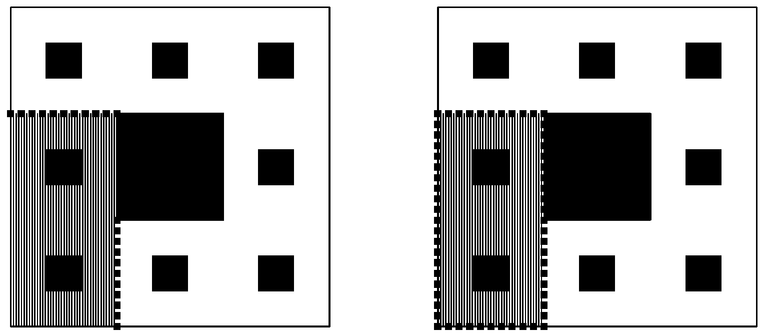

Figure 3: Illustration for Definition 2.11 in the case of the standard Sierpinski carpet and $n = 1$. Let A be the shaded set. The thick dotted lines give $\mathrm{int}_F A$ on the left, and $\mathrm{int}_r A$ on the right.

Definition 2.12 We define the folding map $\varphi_S : F \to S$ for $S \in \mathcal{S}_n(F)$ as follows. Let $\overline{\varphi}_0 : [-1, 1] \to \mathbb{R}$ be defined by $\overline{\varphi}_0(x) = |x|$ for $|x| \le 1$, and then extend the domain of $\overline{\varphi}_0$ to all of $\mathbb{R}$ by periodicity, so that $\overline{\varphi}_0(x + 2n) = \overline{\varphi}_0(x)$ for all $x \in \mathbb{R}$, $n \in \mathbb{Z}$. If y is the point of S closest to the origin, define $\varphi_S(x)$ for $x \in F$ to be the point whose i^{th} coordinate is $y_i + L_F^{-n}\overline{\varphi}_0(L_F^n(x_i - y_i))$.

It is straightforward to check the following

Lemma 2.13 *(a) φ_S is the identity on S and for each $S' \in \mathcal{S}_n$, $\varphi_S : S' \to S$ is an isometry.*
(b) If $S_1, S_2 \in \mathcal{S}_n$ then

$$\varphi_{S_1} \circ \varphi_{S_2} = \varphi_{S_1}. \tag{2.7}$$

(c) Let $x, y \in F$. If there exists $S_1 \in \mathcal{S}_n$ such that $\varphi_{S_1}(x) = \varphi_{S_1}(y)$, then $\varphi_S(x) = \varphi_S(y)$ for every $S \in \mathcal{S}_n$.
(d) Let $S \in \mathcal{S}_n$ and $S' \in \mathcal{S}_{n+1}$. If $x, y \in F$ and $\varphi_S(x) = \varphi_S(y)$ then $\varphi_{S'}(x) = \varphi_{S'}(y)$.

Given $S \in \mathcal{S}_n$, $f : S \to \mathbb{R}$ and $g : F \to \mathbb{R}$ we define the unfolding and restriction operators by

$$U_S f = f \circ \varphi_S, \quad R_S g = g|_S.$$

Using (2.7), we have that if $S_1, S_2 \in \mathcal{S}_n$ then

$$U_{S_2} R_{S_2} U_{S_1} R_{S_1} = U_{S_1} R_{S_1}. \tag{2.8}$$

Definition 2.14 We define the *length* and *mass* scale factors of F to be L_F and m_F respectively.

Let D_n be the network of diagonal crosswires obtained by joining each vertex of a cube $Q \in \mathcal{Q}_n$ to the vertex at the center of the cube by a wire of unit resistance – see [4, 34]. Write R_n^D for the resistance across two opposite faces of D_n. Then it is proved in [4, 34] that there exists ρ_F such that there exist constants C_i, depending only on the dimension d, such that

$$C_1 \rho_F^n \le R_n^D \le C_2 \rho_F^n. \tag{2.9}$$

We remark that $\rho_F \le L_F^2/m_F$ – see [5, Proposition 5.1].

2.3 F-invariant Dirichlet forms

Let $(\mathcal{E}, \mathcal{F})$ be a local regular Dirichlet form on $L^2(F, \mu)$. Let $S \in \mathcal{S}_n$. We set

$$\mathcal{E}^S(g, g) = \frac{1}{m_F^n} \mathcal{E}(U_S g, U_S g). \tag{2.10}$$

and define the domain of $\mathcal{E}^S$ to be $\mathcal{F}^S = \{g : g \text{ maps } S \text{ to } \mathbb{R}, U_S g \in \mathcal{F}\}$. We write $\mu_S = \mu|_S$.

Definition 2.15 Let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(F, \mu)$. We say that $\mathcal{E}$ is an *F-invariant Dirichlet form* or that $\mathcal{E}$ *is invariant with respect to all the local symmetries of F* if the following items (1)–(3) hold:

(1) If $S \in \mathcal{S}_n(F)$, then $U_S R_S f \in \mathcal{F}$ (i.e. $R_S f \in \mathcal{F}^S$) for any $f \in \mathcal{F}$.

(2) Let $n \ge 0$ and S_1, S_2 be any two elements of $\mathcal{S}_n$, and let Φ be any isometry of $\mathbb{R}^d$ which maps S_1 onto S_2. (We allow $S_1 = S_2$.) If $f \in \mathcal{F}^{S_2}$, then $f \circ \Phi \in \mathcal{F}^{S_1}$ and

$$\mathcal{E}^{S_1}(f \circ \Phi, f \circ \Phi) = \mathcal{E}^{S_2}(f, f). \tag{2.11}$$

(3) For all $f \in \mathcal{F}$

$$\mathcal{E}(f, f) = \sum_{S \in \mathcal{S}_n(F)} \mathcal{E}^S(R_S f, R_S f). \tag{2.12}$$

We write $\mathfrak{E}$ for the set of F-invariant, non-zero, local, regular, conservative Dirichlet forms.

Remark 2.16 We cannot exclude at this point the possibility that the energy measure of $\mathcal{E} \in \mathfrak{E}$ may charge the boundaries of cubes in $\mathcal{S}_n$. See Remark 5.3.

We will not need the following definition of scale invariance until we come to the proof of Corollary 1.3 in Section 5.

Definition 2.17 Recall that Ψ_Q, $Q \in \mathcal{Q}_1(F_1)$ are the similitudes which define F_1. Let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(F, \mu)$ and suppose that

$$f \circ \Psi_Q \in \mathcal{F} \text{ for all } Q \in \mathcal{Q}_1(F_1),\ f \in \mathcal{F}. \tag{2.13}$$

Then we can define the *replication* of $\mathcal{E}$ by

$$\mathcal{R}\mathcal{E}(f, f) = \sum_{Q \in \mathcal{Q}_1(F_1)} \mathcal{E}(f \circ \Psi_Q, f \circ \Psi_Q). \tag{2.14}$$

We say that $(\mathcal{E}, \mathcal{F})$ is *scale invariant* if (2.13) holds, and there exists $\lambda > 0$ such that $\mathcal{R}\mathcal{E} = \lambda \mathcal{E}$.

Remark 2.18 We do not have any direct proof that if $\mathcal{E} \in \mathfrak{E}$ then (2.13) holds. Ultimately, however, this will follow from Theorem 1.2.

Lemma 2.19 *Let $(\mathcal{A}, \mathcal{F}_1)$, $(\mathcal{B}, \mathcal{F}_2) \in \mathfrak{E}$ with $\mathcal{F}_1 = \mathcal{F}_2$ and $\mathcal{A} \geq \mathcal{B}$. Then $\mathcal{C} = (1+\delta)\mathcal{A} - \mathcal{B} \in \mathfrak{E}$ for any $\delta > 0$.*

Proof. It is easy to see that Definition 2.15 holds. This and Theorem 2.1 proves the lemma. □

Proposition 2.20 *If $\mathcal{E} \in \mathfrak{E}$ and $S \in \mathcal{S}_n(F)$, then $(\mathcal{E}^S, \mathcal{F}^S)$ is a local regular Dirichlet form on $L^2(S, \mu_S)$.*

Proof. (Local): If u, v are in $\mathcal{F}^S$ with compact support and v is constant in a neighborhood of the support of u, then $U_S u, U_S v$ will be in $\mathcal{F}$, and by the local property of $\mathcal{E}$, we have $\mathcal{E}(U_S u, U_S v) = 0$. Then by (2.10) we have $\mathcal{E}^S(u, v) = 0$.

(Markov): Given that $\mathcal{E}^S$ is local, we have the Markov property by the same proof as that in Theorem 2.1.

(Conservative): Since $1 \in \mathcal{F}$, $\mathcal{E}^S(1, 1) = 0$ by (2.10).

(Regular): If $h \in \mathcal{F}$ then by (2.12) $\mathcal{E}^S(R_S h, R_S h) \leq \mathcal{E}(h, h)$. Let $f \in \mathcal{F}^S$, so that $U_S f \in \mathcal{F}$. As $\mathcal{E}$ is regular, given $\varepsilon > 0$ there exists a continuous $g \in \mathcal{F}$ such that $\mathcal{E}_1(U_S f - g, U_S f - g) < \varepsilon$. Then $R_S U_S f - R_S g = f - R_S g$ on S, so

$$\begin{aligned} \mathcal{E}_1^S(f - R_S g, f - R_S g) &= \mathcal{E}_1^S(R_S U_S f - R_S g, R_S U_S f - R_S g) \\ &\leq \mathcal{E}_1(U_S f - g, U_S f - g) < \varepsilon. \end{aligned}$$

As $R_S g$ is continuous, we see that $\mathcal{F}^S \cap C(S)$ is dense in $\mathcal{F}^S$ in the $\mathcal{E}_1^S$ norm. One can similarly prove that $\mathcal{F}^S \cap C(S)$ is dense in $C(S)$ in the supremum norm, so the regularity of $\mathcal{E}^S$ is proved.

(Closed): If f_m is Cauchy with respect to $\mathcal{E}_1^S$, then $U_S f_m$ will be Cauchy with respect to $\mathcal{E}_1$. Hence $U_S f_m$ converges with respect to $\mathcal{E}_1$, and it follows that $R_S(U_S f_m) = f_m$ converges with respect to $\mathcal{E}_1^S$. □

Fix n and define for functions f on F

$$\Theta f = \frac{1}{m_F^n} \sum_{S \in \mathcal{S}_n(F)} U_S R_S f. \tag{2.15}$$

Using (2.8) we have $\Theta^2 = \Theta$, and so Θ is a projection operator. It is bounded on $C(F)$ and $L^2(F, \mu)$, and moreover by [40, Theorem 12.14] is an orthogonal projection on $L^2(F, \mu)$. Definition 2.15(1) implies that $\Theta : \mathcal{F} \to \mathcal{F}$.

Proposition 2.21 *Assume that $\mathcal{E}$ is a local regular Dirichlet form on F, T_t is its semigroup, and $U_S R_S f \in \mathcal{F}$ whenever $S \in \mathcal{S}_n(F)$ and $f \in \mathcal{F}$. Then the following are equivalent:*

(a) *For all* $f \in \mathcal{F}$, *we have* $\mathcal{E}(f,f) = \sum_{S\in\mathcal{S}_n(F)} \mathcal{E}^S(R_S f, R_S f)$;

(b) *for all* $f, g \in \mathcal{F}$

$$\mathcal{E}(\Theta f, g) = \mathcal{E}(f, \Theta g); \tag{2.16}$$

(c) $T_t \Theta f = \Theta T_t f$ *a.e for any* $f \in L^2(F,\mu)$ *and* $t \geq 0$.

Remark 2.22 Note that this proposition and the following corollary do not use all the symmetries that are assumed in Definition 2.15(2). Although these symmetries are not needed here, they will be essential later in the paper.

Proof. To prove that $(a) \Rightarrow (b)$, note that (a) implies that

$$\mathcal{E}(f,g) = \sum_{T\in\mathcal{S}_n(F)} \mathcal{E}^T(R_T f, R_T g) = \frac{1}{m_F^n} \sum_{T\in\mathcal{S}_n(F)} \mathcal{E}(U_T R_T f, U_T R_T g). \tag{2.17}$$

Then using (2.15), (2.17) and (2.8),

$$\begin{aligned}
\mathcal{E}(\Theta f, g) &= \frac{1}{m_F^n} \sum_{S\in\mathcal{S}_n(F)} \mathcal{E}(U_S R_S f, g) \\
&= \frac{1}{m_F^{2n}} \sum_{S\in\mathcal{S}_n(F)} \sum_{T\in\mathcal{S}_n(F)} \mathcal{E}(U_T R_T U_S R_S f, U_T R_T g) \\
&= \frac{1}{m_F^{2n}} \sum_{S\in\mathcal{S}_n(F)} \sum_{T\in\mathcal{S}_n(F)} \mathcal{E}(U_S R_S f, U_T R_T g).
\end{aligned}$$

Essentially the same calculation shows that $\mathcal{E}(f, \Theta g)$ is equal to the last line of the above with the summations reversed.

Next we show that $(b) \Rightarrow (c)$. If $\mathcal{L}$ is the generator corresponding to $\mathcal{E}$, $f \in \mathcal{D}(\mathcal{L})$ and $g \in \mathcal{F}$ then, writing $\langle f, g\rangle$ for $\int_F fg\,d\mu$, we have

$$\langle \Theta\mathcal{L} f, g\rangle = \langle \mathcal{L} f, \Theta g\rangle = -\mathcal{E}(f, \Theta g) = -\mathcal{E}(\Theta f, g)$$

by (2.16) and the fact that Θ is self-adjoint in the L^2 sense. By the definition of the generator corresponding to a Dirichlet form, this is equivalent to

$$\Theta f \in \mathcal{D}(\mathcal{L}) \text{ and } \Theta\mathcal{L} f = \mathcal{L}\Theta f.$$

By [40, Theorem 13.33], this implies that any bounded Borel function of $\mathcal{L}$ commutes with Θ. (Another good source on the spectral theory of unbounded self-adjoint operators is [38, Section VIII.5].) In particular, the L^2-semigroup T_t of $\mathcal{L}$ commutes with Θ in the L^2-sense. This implies (c).

In order to see that $(c) \Rightarrow (b)$, note that if $f, g \in \mathcal{F}$,

$$\begin{aligned}
\mathcal{E}(\Theta f, g) &= \lim_{t\to 0} t^{-1}\langle (I - T_t)\Theta f, g\rangle = \lim t^{-1}\langle \Theta(I - T_t) f, g\rangle \\
&= \lim t^{-1}\langle (I - T_t) f, \Theta g\rangle = \lim t^{-1}\langle f, (I - T_t)\Theta g\rangle \\
&= \mathcal{E}(f, \Theta g).
\end{aligned}$$

It remains to prove that $(b) \Rightarrow (a)$. This is the only implication that uses the assumption that $\mathcal{E}$ is local. It suffices to assume f and g are bounded.

First, note the obvious relation

$$\sum_{S \in \mathcal{S}_n(F)} \frac{1_S(x)}{N_n(x)} = 1 \tag{2.18}$$

for any $x \in F$, where

$$N_n(x) = \sum_{S \in \mathcal{S}_n(F)} 1_S(x) \tag{2.19}$$

is the number of cubes $\mathcal{S}_n$ whose interiors intersect F and which contain the point x. We break the remainder of the proof into a number of steps.

Step 1: We show that if $\Theta f = f$, then $\Theta(hf) = f(\Theta h)$. To show this, we start with the relationship $U_T R_T U_S R_S f = U_S R_S f$. Summing over $S \in \mathcal{S}_n(F)$ and dividing by m_F^n yields

$$U_T R_T f = U_T R_T \Theta(f) = \Theta f = f.$$

Since $R_S(f_1 f_2) = R_S(f_1) R_S(f_2)$ and $U_S(g_1 g_2) = U_S(g_1) U_S(g_2)$, we have

$$\Theta(hf) = \frac{1}{m_F^n} \sum_{S \in \mathcal{S}_n} (U_S R_S f)(U_S R_S h) = \frac{1}{m_F^n} \sum_{S \in \mathcal{S}_n} f(U_S R_S h) = f(\Theta h).$$

In particular, $\Theta(f^2) = f \Theta f = f^2$.

Step 2: We compute the adjoints of R_S and U_S. R_S maps $C(F)$, the continuous functions on F, to $C(S)$, the continuous functions on S. So R_S^* maps finite measures on S to finite measures on F. We have

$$\int f \, d(R_S^* \nu) = \int R_S f \, d\nu = \int 1_S(x) f(x) \, \nu(dx),$$

and hence

$$R_S^* \nu(dx) = 1_S(x) \, \nu(dx). \tag{2.20}$$

U_S maps $C(S)$ to $C(F)$, so U_S^* maps finite measures on F to finite measures on S. If ν is a finite measure on F, then using (2.18)

$$\begin{aligned} \int_S f \, d(U_S^* \nu) &= \int_F U_S f \, d\nu = \int_F f \circ \varphi_S(x) \, \nu(dx) \\ &= \int_F \Big(\sum_{T \in \mathcal{S}_n} \frac{1_T(x)}{N_n(x)} \Big) f \circ \varphi_S(x) \, \nu(dx) \\ &= \sum_T \int_T \frac{f \circ \varphi_S(x)}{N_n(x)} \, \nu(dx). \end{aligned} \tag{2.21}$$

Let $\varphi_{T,S} : T \to S$ be defined to be the restriction of φ_S to T; this is one-to-one and onto. If κ is a measure on T, define its pull-back $\varphi_{T,S}^* \kappa$ to be the measure on S given by

$$\int_S f \, d(\varphi_{T,S}^* \kappa) = \int_T (f \circ \varphi_{T,S}) \, d\kappa.$$

Write

$$\nu_T(dx) = \frac{1_T(x)}{N_n(x)}\,\nu(dx).$$

Then (2.21) translates to

$$\int_S f\, d(U_S^*\nu) = \sum_T \int_T f\, \varphi_{T,S}^*(\nu_T)(dx),$$

and thus

$$U_S^*\nu = \sum_{T\in\mathcal{S}_n} \varphi_{T,S}^*(\nu_T). \tag{2.22}$$

Step 3: We prove that if ν is a finite measure on F such that $\Theta^*\nu = \nu$ and $S \in \mathcal{S}_n$, then

$$\nu(F) = m_F^n \int_S \frac{1}{N_n(x)}\,\nu(dx). \tag{2.23}$$

To see this, recall that $\varphi_{T,V}^*(\nu_T)$ is a measure on V, and then by (2.20) and (2.22)

$$\begin{aligned}
\Theta^*\nu &= \frac{1}{m_F^n} \sum_{V\in\mathcal{S}_n} R_V^* U_V^*\nu \\
&= \frac{1}{m_F^n} \sum_{V\in\mathcal{S}_n} \sum_{T\in\mathcal{S}_n} \int 1_V(x)\, \varphi_{T,V}^*(\nu_T)(dx) \\
&= \frac{1}{m_F^n} \sum_V \sum_T \int \varphi_{T,V}^*(\nu_T)(dx).
\end{aligned}$$

On the other hand, using (2.18)

$$\nu(dx) = \sum_V \frac{1_V(x)}{N_n(x)}\,\nu(dx) = \sum_V \nu_V(dx).$$

Note that ν_V and $m_F^{-n}\sum_T \varphi_{T,V}^*(\nu_T)$ are both supported on V, and the only way $\Theta^*\nu$ can equal ν is if

$$\nu_V = m_F^{-n} \sum_{T\in\mathcal{S}_n} \varphi_{T,V}^*(\nu_T) \tag{2.24}$$

for each V. Therefore

$$\begin{aligned}
\int_S \frac{1}{N_n(x)}\,\nu(dx) = \nu_S(F) &= m_F^{-n} \sum_T \int 1_F(x)\, \varphi_{T,S}^*(\nu_T)(dx) \\
&= m_F^{-n} \sum_T \int 1_F \circ \varphi_{T,S}(x)\, \nu_T(dx) = m_F^{-n} \sum_T \int \nu_T(dx) \\
&= m_F^{-n} \sum_T \int \frac{1_T(x)}{N_n(x)}\,\nu(dx) = m_F^{-n} \int \nu(dx) = m_F^{-n}\nu(F).
\end{aligned}$$

Multiplying both sides by m_F^n gives (2.23).

Step 4: We show that if $\Theta f = f$, then

$$\Theta^*(\Gamma(f,f)) = \Gamma(f,f). \tag{2.25}$$

Using Step 1, we have for $h \in C(F) \cap \mathcal{F}$

$$\begin{aligned}\int_F h\,\Theta^*(\Gamma(f,f))(dx) &= \int_F \Theta h(x)\,\Gamma(f,f)(dx) = 2\mathcal{E}(f, f\Theta h) - \mathcal{E}(f^2, \Theta h)\\ &= 2\mathcal{E}(f, \Theta(fh)) - \mathcal{E}(\Theta f^2, h) = 2\mathcal{E}(\Theta f, fh) - \mathcal{E}(f^2, h)\\ &= 2\mathcal{E}(f, fh) - \mathcal{E}(f^2, h) = \int_F h\,\Gamma(f,f)(dx).\end{aligned}$$

This is the step where we used (b).

Step 5: We now prove (a). Note that if $g \in \mathcal{F} \cap L^\infty(F)$ and $A = \{x \in F : g(x) = 0\}$, then $\Gamma(g,g)(A) = 0$ by Lemma 2.7. By applying this to the function $g = f - U_S R_S f$, which vanishes on S, and using the inequality

$$\begin{aligned}\left|\Gamma(f,f)(B)^{1/2} - \Gamma(U_S R_S f, U_S R_S f)(B)^{1/2}\right| &\le \Gamma(g,g)(B)^{1/2}\\ &\le \Gamma(g,g)(S)^{1/2} = 0, \quad \forall B \subset S,\end{aligned}$$

(see page 111 in [17]), we see that

$$1_S(x)\Gamma(f,f)(dx) = 1_S(x)\Gamma(U_S R_S f, U_S R_S f)(dx) \tag{2.26}$$

for any $f \in \mathcal{F}$ and $S \in \mathcal{S}_n(F)$.

Starting from $U_T R_T U_S R_S f = U_S R_S f$, summing over $T \in \mathcal{S}_n$ and dividing by m_F^n shows that $\Theta(U_S R_S f) = U_S R_S f$. Applying Step 4 with f replaced by $U_S R_S f$,

$$\Theta^*(\Gamma(U_S R_S f, U_S R_S f))(dx) = \Gamma(U_S R_S f, U_S R_S f)(dx).$$

Applying Step 3 with $\nu = \Gamma(U_S R_S f, U_S R_S f)$, we see

$$\begin{aligned}\mathcal{E}(U_S R_S f, U_S R_S f) &= \Gamma(U_S R_S f, U_S R_S f)(F)\\ &= m_F^n \int_S \frac{1}{N_n(x)}\,\Gamma(U_S R_S f, U_S R_S f)(dx).\end{aligned}$$

Dividing both sides by m_F^n, using the definition of $\mathcal{E}^S$, and (2.26),

$$\mathcal{E}^S(R_S f, R_S f) = \int_S \frac{1}{N_n(x)}\,\Gamma(f,f)(dx). \tag{2.27}$$

Summing over $S \in \mathcal{S}_n$ and using (2.18) we obtain

$$\sum_S \mathcal{E}^S(R_S f, R_S f) = \int \Gamma(f,f)(dx) = \mathcal{E}(f,f),$$

which is (a). □

Corollary 2.23 *If $\mathcal{E} \in \mathfrak{E}$, $f \in \mathcal{F}$, $S \in \mathcal{S}_n(F)$, and $\Gamma_S(R_S f, R_S f)$ is the energy measure of $\mathcal{E}^S$, then*

$$\Gamma_S(R_S f, R_S f)(dx) = \frac{1}{N_n(x)} \Gamma(f,f)(dx), \quad x \in S.$$

We finish this section with properties of sets of capacity zero for F-invariant Dirichlet forms. Let $A \subset F$ and $S \in \mathcal{S}_n$. We define

$$\Theta(A) = \varphi_S^{-1}(\varphi_S(A)). \tag{2.28}$$

Thus $\Theta(A)$ is the union of all the sets that can be obtained from A by local reflections. We can check that $\Theta(A)$ does not depend on S, and that

$$\Theta(A) = \{x : \Theta(1_A)(x) > 0\}.$$

Lemma 2.24 *If $\mathcal{E} \in \mathfrak{E}$ then*

$$\mathrm{Cap}(A) \leq \mathrm{Cap}(\Theta(A)) \leq m_F^{2n}\,\mathrm{Cap}(A)$$

for all Borel sets $A \subset F$.

Proof. The first inequality holds because we always have $A \subset \Theta(A)$. To prove the second inequality it is enough to assume that A is open since the definition of the capacity uses an infimum over open covers of A, and Θ transforms an open cover of A into an open cover of $\Theta(A)$. If $u \in \mathcal{F}$ and $u \geq 1$ on A, then $m_F^n \Theta u \geq 1$ on $\Theta(A)$. This implies the second inequality because $\mathcal{E}(\Theta u, \Theta u) \leq \mathcal{E}(u,u)$, using that Θ is an orthogonal projection with respect to $\mathcal{E}$, that is, $\mathcal{E}(\Theta f, g) = \mathcal{E}(f, \Theta g)$. □

Corollary 2.25 *If $\mathcal{E} \in \mathfrak{E}$, then $Cap(A) = 0$ if and only if $Cap(\Theta(A)) = 0$. Moreover, if f is quasi-continuous, then Θf is quasi-continuous.*

Proof. The first fact follows from Lemma 2.24. Then the second fact holds because Θ preserves continuity of functions on Θ-invariant sets. □

3 The Barlow-Bass and Kusuoka-Zhou Dirichlet forms

In this section we prove that the Dirichlet forms associated with the diffusions on F constructed in [3, 5, 30] are F-invariant; in particular this shows that $\mathfrak{E}$ is non-empty and proves Proposition 1.1. A reader who is only interested in the uniqueness statement in Theorem 1.2 can skip this section.

3.1 The Barlow-Bass processes

The constructions in [3, 5] were probabilistic and almost no mention was made of Dirichlet forms. Further, in [5] the diffusion was constructed on the unbounded fractal $\widetilde{F}$. So before we can assert that the Dirichlet forms are F-invariant, we need to discuss the corresponding forms on F. Recall the way the processes in [3, 5] were constructed was to let W_t^n be normally reflecting Brownian motion on F_n, and to let $X_t^n = W_{a_n t}^n$ for a suitable sequence (a_n). This sequence satisfied

$$c_1(m_F \rho_F / L_F^2)^n \le a_n \le c_2(m_F \rho_F / L_F^2)^n, \tag{3.1}$$

where ρ_F is the resistance scale factor for F. It was then shown that the laws of the X^n were tight and that resolvent tightness held. Let U_n^λ be the λ-resolvent operator for X^n on F_n. The two types of tightness were used to show there exist subsequences n_j such that $U_{n_j}^\lambda f$ converges uniformly on F if f is continuous on F_0 and that the $\mathbb{P}^x$ law of X^{n_j} converges weakly for each x. Any such a subsequential limit point was then called a Brownian motion on the GSC. The Dirichlet form for W^n is $\int_{F_n} |\nabla f|^2 \, d\mu_n$ and that for X^n is

$$\mathcal{E}_n(f,f) = a_n \int_{F_n} |\nabla f(x)|^2 \, \mu_n(dx),$$

both on $L^2(F, \mu_n)$.

Fix any subsequence n_j such that the laws of the X^{n_j}'s converge, and the resolvents converge. If X is the limit process and T_t the semigroup for X, define

$$\mathcal{E}_{BB}(f,f) = \sup_{t>0} \frac{1}{t} \langle f - T_t f, f \rangle$$

with the domain $\mathcal{F}_{BB}$ being those $f \in L^2(F, \mu)$ for which the supremum is finite.

We will need the fact that if U_n^λ is the λ-resolvent operator for X^n and f is bounded on F_0, then $U_n^\lambda f$ is equicontinuous on F. This is already known for the Brownian motion constructed in [5] on the unbounded fractal $\widetilde{F}$, but now we need it for the process on F with reflection on the boundaries of F_0. However the proof is very similar to proofs in [3, 5], so we will be brief. Fix x_0 and suppose x, y are in $B(x_0, r) \cap F_n$. Then

$$\begin{aligned} U_n^\lambda f(x) &= \mathbb{E}^x \int_0^\infty e^{-\lambda t} f(X_t^n)\, dt \\ &= \mathbb{E}^x \int_0^{S_r^n} e^{-\lambda t} f(X_t^n)\, dt + \mathbb{E}^x (e^{-\lambda S_r^n} - 1) U_n^\lambda f(X_{S_r^n}^n) + \mathbb{E}^x U_n^\lambda f(X_{S_r^n}^n), \end{aligned} \tag{3.2}$$

where S_r^n is the time of first exit from $B(x_0, r) \cap F_n$. The first term in (3.2) is bounded by $\|f\|_\infty \mathbb{E}^x S_r^n$. The second term in (3.2) is bounded by

$$\lambda \|U_n^\lambda f\|_\infty \mathbb{E}^x S_r^n \le \|f\|_\infty \mathbb{E}^x S_r^n.$$

We have the same estimates in the case when x is replaced by y, so

$$|U_n^\lambda f(x) - U_n^\lambda f(y)| \le |\mathbb{E}^x U_n^\lambda f(X^n_{S^n_r}) - \mathbb{E}^y U_n^\lambda f(X^n_{S^n_r})| + \delta_n(r),$$

where $\delta_n(r) \to 0$ as $r \to 0$ uniformly in n by [5, Proposition 5.5]. But $z \to \mathbb{E}^z U_n^\lambda f(X^n_{S^n_r})$ is harmonic in the ball of radius $r/2$ about x_0. Using the uniform elliptic Harnack inequality for X^n_t and the corresponding uniform modulus of continuity for harmonic functions ([5, Section 4]), taking $r = |x-y|^{1/2}$, and using the estimate for $\delta_n(r)$ gives the equicontinuity.

It is easy to derive from this that the limiting resolvent U^λ satisfies the property that $U^\lambda f$ is continuous on f whenever f is bounded.

Theorem 3.1 *Each $\mathcal{E}_{BB}$ is in $\mathfrak{E}$.*

Proof. We suppose a suitable subsequence n_j is fixed, and we write $\mathcal{E}$ for the corresponding Dirichlet form $\mathcal{E}_{BB}$. First of all, each X^n is clearly conservative, so $T^n_t 1 = 1$. Since we have $T^{n_j}_t f \to T_t f$ uniformly for each f continuous, then $T_t 1 = 1$. This shows X is conservative, and $\mathcal{E}(1,1) = \sup_t \langle 1 - T_t 1, 1\rangle = 0$.

The regularity of $\mathcal{E}$ follows from Lemma 2.8 and the fact that the processes constructed in [5] are μ-symmetric Feller (see the above discussion, [5, Theorem 5.7] and [3, Section 6]). Since the process is a diffusion, the locality of $\mathcal{E}$ follows from [17, Theorem 4.5.1].

The construction in [3, 5] gives a nondegenerate process, so $\mathcal{E}$ is non-zero. Fix ℓ and let $S \in \mathcal{S}_\ell(F)$. It is easy to see from the above discussion that $U_S R_S f \in \mathcal{F}$ for any $f \in \mathcal{F}$. Before establishing the remaining properties of F-invariance, we show that Θ_ℓ and T_t commute, where Θ_ℓ is defined in (2.15), but with $\mathcal{S}_n(F)$ replaced by $\mathcal{S}_\ell(F)$. Let $\langle f, g\rangle_n$ denote $\int_{F_n} f(x)g(x)\,\mu_n(dx)$. The infinitesimal generator for X^n is a constant times the Laplacian, and it is clear that this commutes with Θ_ℓ. Hence U_n^λ commutes with Θ_ℓ, or

$$\langle \Theta_\ell U_n^\lambda f, g\rangle_n = \langle U_n^\lambda \Theta_\ell f, g\rangle_n. \tag{3.3}$$

Suppose f and g are continuous and f is nonnegative. The left hand side is $\langle U_n^\lambda f, \Theta_\ell g\rangle_n$, and if n converges to infinity along the subsequence n_j, this converges to

$$\langle U^\lambda f, \Theta_\ell g\rangle = \langle \Theta_\ell U^\lambda f, g\rangle.$$

The right hand side of (3.3) converges to $\langle U^\lambda \Theta_\ell f, g\rangle$ since $\Theta_\ell f$ is continuous if f is. Since X_t has continuous paths, $t \to T_t f$ is continuous, and so by the uniqueness of the Laplace transform, $\langle \Theta_\ell T_t f, g\rangle = \langle T_t \Theta_\ell f, g\rangle$. Linearity and a limit argument allows us to extend this equality to all $f \in L^2(F)$. The implication (c) $\Rightarrow$ (a) in Proposition 2.21 implies that $\mathcal{E} \in \mathfrak{E}$. □

3.2 The Kusuoka-Zhou Dirichlet form

Write $\mathcal{E}_{KZ}$ for the Dirichlet form constructed in [30]. Note that this form is self-similar.

Theorem 3.2 $\mathcal{E}_{KZ} \in \mathfrak{E}$.

Proof. One can see that $\mathcal{E}_{KZ}$ satisfies Definition 2.15 because of the self-similarity. The argument goes as follows. Initially we consider $n = 1$, and suppose $f \in \mathcal{F} = \mathcal{D}(\mathcal{E}_{KZ})$. Then [30, Theorem 5.4] implies $U_S R_S f \in \mathcal{F}$ for any $S \in \mathcal{S}_1(F)$. This gives us Definition 2.15(1).

Let $S \in \mathcal{S}_1(F)$ and $S = \Psi_i(F)$ where Ψ_i is one of the contractions that define the self-similar structure on F, as in [30]. Then we have

$$f \circ \Psi_i = (U_S R_S f) \circ \Psi_i = (U_S R_S f) \circ \Psi_j$$

for any i, j. Hence by [30, Theorem 6.9], we have

$$\begin{aligned}\mathcal{E}_{KZ}(U_S R_S f, U_S R_S f) &= \rho_F m_F^{-1} \sum_j \mathcal{E}_{KZ}((U_S R_S f) \circ \Psi_j, (U_S R_S f) \circ \Psi_j) \\ &= \rho_F \mathcal{E}_{KZ}(f \circ \Psi_i, f \circ \Psi_i).\end{aligned}$$

By [30, Theorem 6.9] this gives Definition 2.15(3), and moreover

$$\mathcal{E}^S(f, f) = \rho_F m_F^{-1} \mathcal{E}_{KZ}(f \circ \Psi_i, f \circ \Psi_i).$$

Definition 2.15(2) and the rest of the conditions for $\mathcal{E}_{KZ}$ to be in $\mathfrak{E}$ follow from (1), (3) and the results of [30]. The case $n > 1$ can be dealt with by using the self-similarity. □

Proof of Proposition 1.1 This is immediate from Theorems 3.1 and 3.2. □

4 Diffusions associated with F-invariant Dirichlet forms

In this section we extensively use notation and definitions introduced in Section 2, especially Subsections 2.2 and 2.3. We fix a Dirichlet form $\mathcal{E} \in \mathfrak{E}$. Let $X = X^{(\mathcal{E})}$ be the associated diffusion, $T_t = T_t^{(\mathcal{E})}$ be the semigroup of X and $\mathbb{P}^x = \mathbb{P}^{x,(\mathcal{E})}$, $x \in F - \mathcal{N}_0$, the associated probability laws. Here $\mathcal{N}_0$ is a properly exceptional set for X. Ultimately (see Corollary 1.4) we will be able to define $\mathbb{P}^x$ for all $x \in F$, so that $\mathcal{N}_0 = \emptyset$.

4.1 Reflected processes and the Markov property

Theorem 4.1 *Let $S \in \mathcal{S}_n(F)$ and $Z = \varphi_S(X)$. Then Z is a μ_S-symmetric Markov process with Dirichlet form $(\mathcal{E}^S, \mathcal{F}^S)$, and semigroup $T_t^Z f = R_S T_t U_S f$. Write $\widetilde{\mathbb{P}}^y$ for the laws of Z; these are defined for $y \in S - \mathcal{N}_2^Z$, where $\mathcal{N}_2^Z$ is a properly exceptional set for Z. There exists a properly exceptional set $\mathcal{N}_2$ for X such that for any Borel set $A \subset F$,*

$$\widetilde{\mathbb{P}}^{\varphi_S(x)}(Z_t \in A) = \mathbb{P}^x(X_t \in \varphi_S^{-1}(A)), \quad x \in F - \mathcal{N}_2. \tag{4.1}$$

Proof. Denote $\varphi = \varphi_S$. We begin by proving that there exists a properly exceptional set $\mathcal{N}_2$ for X such that

$$\mathbb{P}^x(X_t \in \varphi^{-1}(A)) = T_t 1_{\varphi^{-1}(A)}(x) = T_t 1_{\varphi^{-1}(A)}(y) = \mathbb{P}^y(X_t \in \varphi^{-1}(A)) \qquad (4.2)$$

whenever $A \subset S$ is Borel, $\varphi(x) = \varphi(y)$, and $x, y \in F - \mathcal{N}_2$. It is sufficient to prove (4.2) for a countable base (A_m) of the Borel σ-field on F. Let $f_m = 1_{A_m}$. Since $T_t 1_{\varphi^{-1}(A_m)} = T_t U_S f_m$, it is enough to prove that there exists a properly exceptional set $\mathcal{N}_2$ such that for $m \in \mathbb{N}$,

$$T_t U_S f_m(x) = T_t U_S f_m(y), \quad \text{ if } x, y \in F - \mathcal{N}_2 \text{ and } \varphi(x) = \varphi(y). \qquad (4.3)$$

By (2.8), $\Theta(U_S f) = U_S f$. Using Proposition 2.21,

$$\Theta T_t U_S f = T_t \Theta U_S f_m = T_t U_S f,$$

for $f \in L^2$, where the equality holds in the L^2 sense.

Recall that we always consider quasi-continuous modifications of functions in $\mathcal{F}$. By Corollary 2.25, $\Theta T_t U_S f_m$ is quasi-continuous. Since [17, Lemma 2.1.4] tells us that if two quasi-continuous functions coincide μ-a.e., then they coincide q.e., we have that $\Theta(T_t U_S f_m) = T_t U_S f_m$ q.e. The definition of Θ implies that $\Theta(T_t U_S f_m)(x) = \Theta(T_t U_S f_m)(y)$ whenever $\varphi(x) = \varphi(y)$, so there exists a properly exceptional set $\mathcal{N}_{2,m}$ such that (4.3) holds. Taking $\mathcal{N}_2 = \cup_m \mathcal{N}_{2,m}$ gives (4.2). Using Theorem 10.13 of [16], Z is Markov and has semigroup $T_t^Z f = R_S T_t(U_S f)$. We take $\mathcal{N}_2^Z = \varphi(\mathcal{N}_2)$.

Using (4.3), $U_S R_S T_t U_S f = T_t U_S f$, and then

$$\langle T_t^Z f, g\rangle_S = \langle R_S T_t U_S f, g\rangle_S = m_F^{-n} \langle U_S R_S T_t U_S f, U_S g\rangle = m_F^{-n}\langle T_t U_S f, U_S g\rangle.$$

This equals $m_F^{-n}\langle U_S f, T_t U_S g\rangle$, and reversing the above calculation, we deduce that $\langle f, T_t^Z g\rangle = m_F^{-n}\langle U_S f, T_t U_S g\rangle$, proving that Z is μ_S-symmetric.

To identify the Dirichlet form of Z we note that

$$t^{-1}\langle T_t^Z f - f, f\rangle_S = m_F^{-n} t^{-1} \langle T_t U_S f - U_S f, U_S f\rangle.$$

Taking the limit as $t \to 0$, and using [17, Lemma 1.3.4], it follows that Z has Dirichlet form

$$\mathcal{E}_Z(f,f) = m_F^{-n}\mathcal{E}(U_S f, U_S f) = \mathcal{E}^S(f,f).$$

□

Lemma 4.2 *Let $S, S' \in \mathcal{S}_n$, $Z = \varphi_S(X)$, and Φ be an isometry of S onto S'. Then if $x \in S - \mathcal{N}$,*

$$\mathbb{P}^x(\Phi(Z) \in \cdot) = \mathbb{P}^{\Phi(x)}(Z \in \cdot).$$

Proof. By Theorem 4.1 and Definition 2.15(2) Z and $\Phi(Z)$ have the same Dirichlet form. The result is then immediate from [17, Theorem 4.2.7], which states that two Hunt processes are equivalent if they have the same Dirichlet forms, provided we exclude an F-invariant set of capacity zero. □

We say $S, S' \in \mathcal{S}_n(F)$ are *adjacent* if $S = Q \cap F$, $S' = Q' \cap F$ for $Q, Q' \in \mathcal{Q}_n(F)$, and $Q \cap Q'$ is a $(d-1)$-dimensional set. In this situation, let H be the hyperplane separating S, S'. For any hyperplane $H \subset \mathbb{R}^d$, let $g_H : \mathbb{R}^d \to \mathbb{R}^d$ be reflection in H. Recall the definition of $\partial_r D$, where D is a finite union of elements of $\mathcal{S}_n$.

Lemma 4.3 *Let $S_1, S_2 \in \mathcal{S}_n(F)$ be adjacent, let $D = S_1 \cup S_2$, let $B = \partial_r(S_1 \cup S_2)$, and let H be the hyperplane separating S_1 and S_2. Then there exists a properly exceptional set $\mathcal{N}$ such that if $x \in H \cap D - \mathcal{N}$, the processes $(X_t, 0 \le t \le T_B)$ and $(g_H(X_t), 0 \le t \le T_B)$ have the same law under $\mathbb{P}^x$.*

Proof. Let $f \in \mathcal{F}$ with support in the interior of D. Then Definition 2.15(3) and Proposition 2.20 imply that $\mathcal{E}(f,f) = \mathcal{E}^{S_1}(R_{S_1} f, R_{S_1} f) + \mathcal{E}^{S_2}(R_{S_2} f, R_{S_2} f)$. Definition 2.15(2) implies that $\mathcal{E}(f,f) = \mathcal{E}(f \circ g_H, f \circ g_H)$. Hence $(g_H(X_t), 0 \le t \le T_B)$ has the same Dirichlet form as $(X_t, 0 \le t \le T_B)$, and so they have the same law by [17, Theorem 4.2.7] if we exclude an F-invariant set of capacity zero. □

4.2 Moves by Z and X

At this point we have proved that the Markov process X associated with the Dirichlet form $\mathcal{E} \in \mathfrak{E}$ has strong symmetry properties. We now use these to obtain various global properties of X. The key idea, as in [5], is to prove that certain 'moves' of the process in F have probabilities which can be bounded below by constants depending only on the dimension d.

We need a considerable amount of extra technical notation, based on that in [5], which will only be used in this subsection.

We begin by looking at the process $Z = \varphi_S(X)$ for some $S \in \mathcal{S}_n$, where $n \ge 0$. Since our initial arguments are scale invariant, we can simplify our notation by taking $n = 0$ and $S = F$ in the next definition.

Definition 4.4 Let $1 \le i, j \le d$, with $i \ne j$, and set

$$\begin{aligned} H_i(t) &= \{x = (x_1, \dots, x_d) : x_i = t\},\ t \in \mathbb{R}; \\ L_i &= H_i(0) \cap [0, 1/2]^d; \\ M_{ij} &= \{x \in [0,1]^d :\ x_i = 0,\ \tfrac{1}{2} \le x_j \le 1,\ \text{and } 0 \le x_k \le \tfrac{1}{2} \text{ for } k \ne j\}. \end{aligned}$$

Let

$$\partial_e S = S \cap (\cup_{i=1}^d H_i(1)), \qquad D = S - \partial_e S.$$

We now define, for the process Z, the sets E_D and Z_D as in (2.6). The next proposition says that the corners and slides of [5] hold for Z, provided that $Z_0 \in E_D$.

Proposition 4.5 *There exists a constant q_0, depending only on the dimension d, such that*

$$\widetilde{\mathbb{P}}^x(T^Z_{L_j} < \tau^Z_D) \ge q_0, \quad x \in L_i \cap E_D, \tag{4.4}$$

$$\widetilde{\mathbb{P}}^x(T^Z_{M_{ij}} < \tau^Z_D) \ge q_0, \quad x \in L_i \cap E_D. \tag{4.5}$$

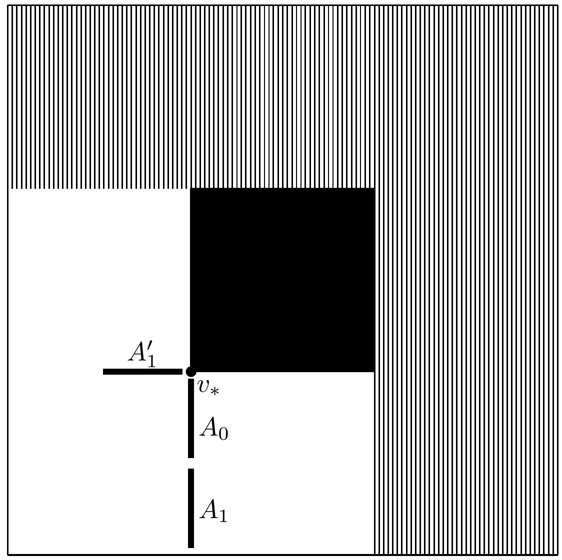

Figure 4: Illustration for Definition 4.6 in the case of the standard Sierpinski carpet and $n = 1$. The complement of D is shaded. The half-face A_1 corresponds to a slide move, and the half-face A_1' corresponds to a corner move. In this case Q_* is the lower left cube in $\mathcal{S}_1$.

These inequalities hold for any $n \geq 0$ provided we modify Definition 4.4 appropriately.

Proof. Using Lemma 4.2 this follows by the same reflection arguments as those used in the proofs of Proposition 3.5 – Lemma 3.10 of [5]. We remark that, inspecting these proofs, we can take $q_0 = 2^{-2d^2}$. □

We now fix $n \geq 0$. We call a set $A \subset \mathbb{R}^d$ a (level n) *half-face* if there exists $i \in \{1, \dots, d\}$, $a = (a_1, \dots, a_d) \in \frac{1}{2}\mathbb{Z}^d$ with $a_i \in \mathbb{Z}$ such that

$$A = \{x : x_i = a_i L_F^{-n}, \quad a_j L_F^{-n} \leq x_j \leq (a_j + 1/2) L_F^{-n} \quad \text{for } j \neq i\}.$$

(Note that a level n half-face need not be a subset of F.) For A as above set $\iota(A) = i$. Let $\mathcal{A}^{(n)}$ be the collection of level n half-faces, and

$$\mathcal{A}_F^{(n)} = \{A \in \mathcal{A}^{(n)} : A \subset F_n\}.$$

We define a graph structure on $\mathcal{A}_F^{(n)}$ by taking $\{A, B\}$ to be an edge if

$$\dim(A \cap B) = d - 2, \quad \text{and } A \cup B \subset Q \text{ for some } Q \in \mathcal{Q}_n.$$

Let $E(\mathcal{A}_F^{(n)})$ be the set of edges in $\mathcal{A}_F^{(n)}$. As in [5, Lemma 3.12] we have that the graph $\mathcal{A}_F^{(n)}$ is connected. We call an edge $\{A, B\}$ an *$i - j$ corner* if $\iota(A) = i$,

$\iota(B) = j$, and $i \neq j$ and call $\{A, B\}$ an $i-j$ *slide* if $\iota(A) = \iota(B) = i$, and the line joining the centers of A and B is parallel to the x_j axis. Any edge is either a corner or a slide; note that the move (L_i, L_j) is an $i-j$ corner, while (L_i, M_{ij}) is an $i-j$ slide.

For the next few results we need some further notation.

Definition 4.6 Let (A_0, A_1) be an edge in $E(\mathcal{A}_F^{(n)})$, and Q_* be a cube in $\mathcal{Q}_n(F)$ such that $A_0 \cup A_1 \subset Q_*$. Let v_* be the unique vertex of Q_* such that $v_* \in A_0$, and let R be the union of the 2^d cubes in $\mathcal{Q}_n$ containing v_*. Then there exist distinct $S_i \in \mathcal{S}_n$, $1 \leq i \leq m$ such that $F \cap R = \cup_{i=1}^m S_i$. Let $D = F \cap R^o$; thus

$$\overline{D} = F \cap R = \cup_{i=1}^m S_i.$$

Let S_* be any one of the S_i, and set $Z = \varphi_{S_*}(X)$. Write

$$\tau = \tau_D^X = \inf\{t \geq 0 : X_t \notin D\} = \inf\{t : Z_t \in \partial_r R\}. \tag{4.6}$$

Let

$$E_D = \{x \in D : \mathbb{P}^x(\tau < \infty) = 1\}. \tag{4.7}$$

We wish to obtain a lower bound for

$$\inf_{x \in A_0 \cap E_D} \mathbb{P}^x(T_{A_1}^X \leq \tau). \tag{4.8}$$

By Proposition 4.5 we have

$$\inf_{y \in A_0 \cap E_D} \widetilde{\mathbb{P}}^y(T_{A_1}^Z \leq \tau) \geq q_0. \tag{4.9}$$

Z hits A_1 if and only if X hits $\Theta(A_1)$, and one wishes to use symmetry to prove that, if $x \in A_0 \cap E_D$ then for some $q_1 > 0$

$$\mathbb{P}^x(T_{A_1}^X \leq \tau) \geq q_1 \widetilde{\mathbb{P}}^x(T_{A_1}^Z \leq \tau) \geq q_1 q_0. \tag{4.10}$$

This was proved in [5] in the context of reflecting Brownian motion on F_{n+k}, but the proof used the fact that sets of dimension $d-2$ were polar for this process. Here we need to handle the possibility that there may be times t such that X_t is in more than two of the S_i. We therefore need to consider the way that X leaves points y which are in several S_i.

Definition 4.7 Let $y \in E_D$ be in exactly k of the S_i, where $1 \leq k \leq m$. Let $S_1', \ldots, S_k'$ be the elements of $\mathcal{S}_n$ containing y. (We do not necessarily have that S_1 is one of the S_j'.) Let $D(y) = \mathrm{int}_r(\cup_{i=1}^k S_i')$; so that $\overline{D(y)} = \cup_{i=1}^k S_i'$. Let D_1, D_2 be open sets in F such that $y \in D_2 \subset \overline{D}_2 \subset D_1 \subset \overline{D}_1 \subset D(y)$. Assume further that $\Theta(D_i) \cap D(y) = D_i$ for $i = 1, 2$, and note that we always have $\Theta(D_i) \supset D_i$. For $f \in \mathcal{F}$ define

$$\Theta^{D_1} f = k^{-1} m_F^n 1_{D_1} \Theta f; \tag{4.11}$$

the normalization factor is chosen so that $\Theta^{D_1} 1_{D_1} = 1_{D_1}$.

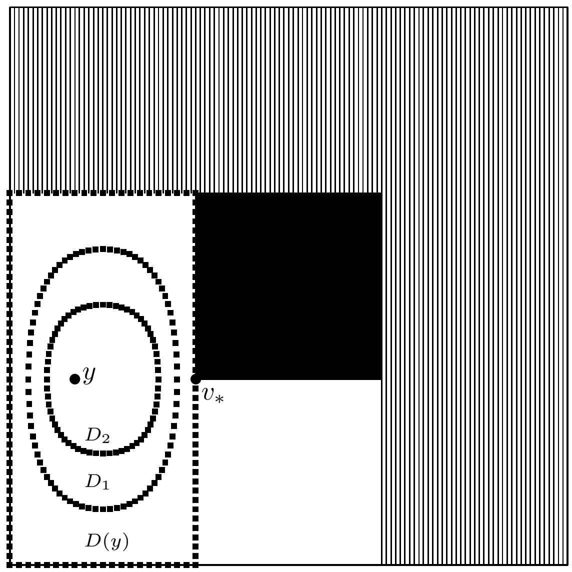

Figure 5: Illustration for Definition 4.7 in the case of the standard Sierpinski carpet and $n = 1$. The complement of D is shaded, and the dotted lines outline $D(y) \supset D_1 \supset D_2$.

As before we define $\mathcal{F}_{D_1} \subset \mathcal{F}$ as the closure of the set of functions $\{f \in \mathcal{F} : \operatorname{supp}(f) \subset D_1\}$. We denote by $\mathcal{E}_{D_1}$ the associated Dirichlet form and by $T_t^{D_1}$ the associated semigroup, which are the Dirichlet form and the semigroup of the process X killed on exiting D_1, by Theorems 4.4.3 and A.2.10 in [17]. For convenience, we state the next lemma in the situation of Definition 4.7, although it holds under somewhat more general conditions.

Lemma 4.8 *Let D_1, D_2 be as in Definition 4.7.*
(a) Let $f \in \mathcal{F}_{D_1}$. Then $\Theta^{D_1} f \in \mathcal{F}_{D_1}$. Moreover, for all $f, g \in \mathcal{F}_{D_1}$ we have

$$\mathcal{E}_{D_1}(\Theta^{D_1} f, g) = \mathcal{E}_{D_1}(f, \Theta^{D_1} g)$$

and $T_t^{D_1} \Theta^{D_1} f = \Theta^{D_1} T_t^{D_1} f$.
(b) If $h \in \mathcal{F}_{D_1}$ is harmonic (in the Dirichlet form sense) in D_2 then $\Theta^{D_1} h$ is harmonic (in the Dirichlet form sense) in D_2.
(c) If u is caloric in D_2, in the sense of Proposition 2.6, then $\Theta^{D_1} u$ is also caloric in D_2.

Proof. (a) By Definition 2.15, $\Theta f \in \mathcal{F}$. Let ψ be a function in $\mathcal{F}$ which has support in $D(y)$ and is 1 on D_1; such a function exists because $\mathcal{E}$ is regular and Markov. Then $\psi \Theta f \in \mathcal{F}$, and $\psi \Theta f = k m_F^{-n} \Theta^{D_1} f$. The rest of the proof follows from Proposition 2.21(b,c) because $\mathcal{E}(\Theta^{D_1} f, g) = k^{-1} m_F^n \mathcal{E}(\Theta f, g)$.

(b) Let $g \in \mathcal{F}$ with $\mathrm{supp}(g) \subset D_2$. Then

$$\mathcal{E}(\Theta^{D_1} h, g) = k^{-1} m_F^n \mathcal{E}(\Theta h, g) = k^{-1} m_F^n \mathcal{E}(h, \Theta g) = \mathcal{E}(h, \Theta^{D_1} g) = 0. \tag{4.12}$$

The final equality holds because h is harmonic on D_2 and $\Theta^{D_1} g$ has support in D_2. Relation (4.12) implies that $\Theta^{D_1} h$ is harmonic in D_2 by Proposition 2.5.
(c) We denote by $\overline{T}_t$ the semigroup of the process $\overline{X}_t$, which is X_t killed at exiting D_2. The same reasoning as in (a) implies that $\overline{T}_t \Theta^{D_1} = \Theta^{D_1} \overline{T}_t$. Hence (c) follows from (a), (b) and Proposition 2.6. □

Recall from (2.19) the definition of the "cube counting" function $N_n(z)$. Define the related "weight" function

$$r_S(z) = 1_S(z) N_n(z)^{-1}$$

for each $S \in \mathcal{S}_n(F)$. If no confusion can arise, we will denote $r_i(z) = r_{S_i'}(z)$.

Let $(\mathcal{F}_t^Z)$ be the filtration generated by Z. Since $\mathcal{F}_0^Z$ contains all $\mathbb{P}^x$ null sets, under the law $\mathbb{P}^x$ we have that $X_0 = x$ is $\mathcal{F}_0^Z$ measurable.

Lemma 4.9 *Let $y \in E_D$, D_1, D_2 be as in Definition 4.7. Write $V = \tau_{D_2}^X$.*
(a) If $U \subset \partial_F(D_2)$ satisfies $\Theta(U) \cap D(y) = U$, then

$$\mathbb{E}^y(r_i(X_V) 1_{(X_V \in U)}) = k^{-1} \widetilde{\mathbb{P}}^{\varphi_S(y)}(Z_V \in \varphi_S(U)), \quad \text{for } i = 1, \ldots, k = N_n(y). \tag{4.13}$$

(b) For any bounded Borel function $f : D_1 \to \mathbb{R}$ and all $0 \le t \le \infty$,

$$\mathbb{E}^y(f(X_{t\wedge V}) | \mathcal{F}_{t\wedge V}^Z) = \left(\Theta^{D_1} f\right)(Z_{t\wedge V}). \tag{4.14}$$

In particular

$$\mathbb{E}^y(r_i(X_{t\wedge V}) | \mathcal{F}_{t\wedge V}^Z) = k^{-1}. \tag{4.15}$$

Proof. Note that, by the symmetry of D_2, V is a $(\mathcal{F}_t^Z)$ stopping time.
(a) Let $f \in \mathcal{F}_{D_1}$ be bounded, and h be the function with support in D_1 which equals f in $D_1 - D_2$, and is harmonic (in the Dirichlet form sense) inside D_2. Then since $\varphi_{S_i'}(y) = y$ for $1 \le i \le k$,

$$\Theta^{D_1} h(y) = k^{-1} \sum_{i=1}^{k} h(\varphi_{S_i'}(y)) = h(y).$$

Since $\Theta^{D_1} h$ is harmonic (in the Dirichlet form sense) in D_2 and since $y \in E_D$, we have, using Proposition 2.5, that

$$h(y) = \Theta^{D_1} h(y) = \mathbb{E}^y (\Theta^{D_1} h)(X_V) = k^{-1} \mathbb{E}^y \sum_{i=1}^{k} h(\varphi_{S_i'}(X_V)).$$

Since $f = h$ on $\partial_F(D_2)$,

$$\mathbb{E}^y(f(X_V)) = h(y) = k^{-1} \mathbb{E}^y \sum_{i=1}^{k} f(\varphi_{S_i'}(X_V)).$$

Write δ_x for the unit measure at x, and define measures $\nu_i(\omega, dx)$ by

$$\nu_1(dx) = \delta_{X_V}(dx), \quad \nu_2(dx) = k^{-1} \sum_{i=1}^{k} \delta_{\varphi_{S'_i}(X_V)}(dx) = k^{-1} \sum_{i=1}^{k} \delta_{\varphi_{S'_i}(Z_V)}(dx).$$

Then we have

$$\mathbb{E}^y \int f(x)\nu_1(dx) = \mathbb{E}^y \int f(x)\nu_2(dx)$$

for $f \in \mathcal{F}_{D_1}$, and hence for all bounded Borel f defined on $\partial_F(D_2)$. Taking $f = r_i(x)1_U(x)$ then gives (4.13).
(b) We can take the cube S^* in Definition 4.6 to be S'_1. If g is defined on S^* then $U_S g$ is the unique extension of g to $\overline{D(y)}$ such that $\Theta^{D_1} U_S g = U_S g$ on $\overline{D(y)}$. Thus any function on S is the restriction of a function which is invariant with respect to Θ^{D_1}. We will repeatedly use the fact that if $\Theta^{D_1} g = g$ then $g(X_t) = g(Z_t)$, and so also $g(X_{t\wedge V}) = g(Z_{t\wedge V})$.

We break the proof into several steps.
Step 1. Let $T_t^{D_2}$ denote the semigroup of X stopped on exiting D_2, that is

$$T_t^{D_2} f(x) = \mathbb{E}^x f(X_{t\wedge V}).$$

If $f \in \mathcal{F}_{D_1}$ is bounded, then Proposition 2.6 and Lemma 4.8 imply that q.e. in D_2

$$T_t^{D_2} \Theta^{D_1} f = \Theta^{D_1} T_t^{D_2} f. \tag{4.16}$$

Note that by Proposition 2.6 and [17, Theorem 4.4.3(ii)], the notion "q.e." in D_2 coincides for the semigroups T, T^{D_2} and $\overline{T}$, where $\overline{T}$ is defined in Lemma 4.8.
Step 2. If $f, g \in \mathcal{F}_{D_1}$ are bounded and $\Theta^{D_1} g = g$, then we have $\Theta^{D_1}(gf) = g\Theta^{D_1} f$. Hence

$$T_t^{D_2}(g\Theta^{D_1} f) = T_t^{D_2}\Theta^{D_1}(gf) = \Theta^{D_1} T_t^{D_2}(gf). \tag{4.17}$$

Step 3. Let ν be a Borel probability measure on D_2. Set $\nu^* = (\Theta^{D_1})^*\nu$. Suppose that $\nu(\mathcal{N}_2) = 0$, where $\mathcal{N}_2$ is defined in Theorem 4.1. If f, g are as in the preceding paragraph, then we have

$$\begin{aligned}
\mathbb{E}^{\nu^*} g(Z_{t\wedge V}) f(X_{t\wedge V}) &= \int_{D_2} T_t^{D_2}\big(gf\big)(x) \left(\Theta^{D_1}\right)^* \nu(dx) \\
&= \int_{D_2} \Theta^{D_1}\big(T_t^{D_2}(gf)\big)(x)\nu(dx) \\
&= \int_{D_2} T_t^{D_2}\big(g\Theta^{D_1} f\big)(x)\nu(dx) \\
&= \mathbb{E}^{\nu} g(Z_{t\wedge V})\Theta^{D_1} f(X_{t\wedge V}) \\
&= \mathbb{E}^{\nu} g(Z_{t\wedge V})\Theta^{D_1} f(Z_{t\wedge V}),
\end{aligned} \tag{4.18}$$

where we use the definition of adjoint, (4.17) to interchange T^{D_2} and Θ^{D_1}, and that $g(X_{t\wedge V}) = g(Z_{t\wedge V})$.

Step 4. We prove by induction that if $\nu(\mathcal{N}_2)=0$, $m\ge 0$, $0<t_1<\cdots<t_m<t$, $g_1,\dots,g_m$ are bounded Borel functions satisfying $\Theta^{D_1}g_i=g_i$, and f is bounded and Borel, then

$$\mathbb{E}^{\nu^*}(\Pi_{i=1}^m g_i(Z_{t_i\wedge V}))f(X_{t\wedge V})=\mathbb{E}^{\nu}(\Pi_{i=1}^m g_i(Z_{t_i\wedge V}))\Theta^{D_1}f(Z_{t\wedge V}). \tag{4.19}$$

The case $m=0$ is (4.18). Suppose (4.19) holds for $m-1$. Then set

$$h(x)=\mathbb{E}^x(\Pi_{i=2}^m g_i(Z_{(t_i-t_1)\wedge V}))f(X_{(t-t_1)\wedge V}). \tag{4.20}$$

Write $\delta_x^*=(\delta_x)^*$. By (4.19) for $m-1$, provided x is such that $\delta_x^*(\mathcal{N}_2)=0$,

$$\Theta^{D_1}h(x)=\mathbb{E}^{\delta_x^*}(\Pi_{i=2}^m g_i(Z_{(t_i-t_1)\wedge V}))f(X_{(t-t_1)\wedge V}) \tag{4.21}$$

$$=\mathbb{E}^x(\Pi_{i=2}^m g_i(Z_{(t_i-t_1)\wedge V}))\Theta^{D_1}f(Z_{(t-t_1)\wedge V}). \tag{4.22}$$

So, using the Markov property, (4.18) and (4.21)

$$\begin{aligned}\mathbb{E}^{\nu^*}(\Pi_{i=1}^m g_i(Z_{t_i\wedge V}))f(X_{t\wedge V})&=\mathbb{E}^{\nu^*}g_1(Z_{t_1\wedge V})h(X_{t_1\wedge V})\\&=\mathbb{E}^{\nu}g_1(Z_{t_1\wedge V})\Theta^{D_1}h(X_{t_1\wedge V})\\&=\mathbb{E}^{\nu}g_1(Z_{t_1\wedge V})\mathbb{E}^{X_{t_1\wedge V}}\left((\Pi_{i=2}^m g_i(Z_{(t_i-t_1)\wedge V}))\Theta^{D_1}f(Z_{(t-t_1)\wedge V})\right)\\&=\mathbb{E}^{\nu}(\Pi_{i=1}^m g_i(Z_{t_i\wedge V}))\Theta^{D_1}f(Z_{t\wedge V}),\end{aligned}$$

which proves (4.19). Therefore since $(\delta_x^*)^*=\delta_x^*$,

$$\mathbb{E}^{\delta_x^*}(\Pi_{i=1}^m g_i(Z_{t_i\wedge V}))f(X_{t\wedge V})=\mathbb{E}^{\delta_x^*}(\Pi_{i=1}^m g_i(Z_{t_i\wedge V}))\Theta^{D_1}f(Z_{t\wedge V}),$$

and so

$$\mathbb{E}^{\delta_x^*}(f(X_{t\wedge V})|\mathcal{F}^Z_{t\wedge V})=\left(\Theta^{D_1}f\right)(Z_{t\wedge V}).$$

To obtain (4.14), observe that $\delta_y^*=\delta_y$. Equation (4.15) follows since $\Theta^{D_1}r_i(x)=k^{-1}$ for all $x\in D_1$. □

Corollary 4.10 *Let $f:D(y)\to\mathbb{R}$ be bounded Borel, and $t\ge 0$. Then*

$$\mathbb{E}^y(f(X_{t\wedge\tau})|\mathcal{F}^Z_{t\wedge\tau})=\left(\Theta^{D(y)}f\right)(Z_{t\wedge\tau}). \tag{4.23}$$

Proof. This follows from Lemma 4.9 by letting the regions D_i in Definition 4.7 increase to $D(y)$. □

Let (A_0,A_1), Z be as in Definition 4.6. We now look at X conditional on $\mathcal{F}^Z$. Write $W_i(t)=\varphi_{S_i}(Z_t)\in S_i$. For any t, we have that $X_{t\wedge\tau}$ is at one of the points $W_i(t\wedge\tau)$. Let

$$\begin{aligned}J_i(t)&=\{j:W_j(t\wedge\tau)=W_i(t\wedge\tau)\},\\M_i(t)&=\sum_{j=1}^m 1_{(W_j(t\wedge\tau)=W_i(t\wedge\tau))}=\#J_i(t),\\p_i(t)&=\mathbb{P}^x(X_{t\wedge\tau}=W_i(t\wedge\tau)|\mathcal{F}^Z_{t\wedge\tau})M_i(t)^{-1}=\mathbb{E}^x(r_i(X_{t\wedge\tau})|\mathcal{F}^Z_{t\wedge\tau}).\end{aligned}$$

Thus the conditional distribution of X_t given $\mathcal{F}^Z_{t\wedge\tau}$ is

$$\sum_{i=1}^{k} p_i(t)\delta_{W_i(t\wedge\tau)}. \tag{4.24}$$

Note that by the definitions given above, we have $M_i(t) = N_n(W_i(t))$ for $0 \le t < \tau$, which is the number of elements of $\mathcal{S}_n$ that contain $W_i(t)$.

To describe the intuitive picture, we call the W_i "particles." Each $W_i(t)$ is a single point, and for each t we consider the collection of points $\{W_i(t), 1 \le i \le m\}$. This is a finite set, but the number of distinct points depends on t. In fact, we have $\{W_i(t), 1 \le i \le m\} = \Theta\{X_t\} \cap D$. For each given t, X_t is equal to some of the $W_i(t)$. If X_t is in the r-interior of an element of $\mathcal{S}_n$, then all the $W_i(t)$ are distinct, and so there are m of them. In this case there is a single i such that $X_t = W_i(t)$. If Z_t is in a lower dimensional face, then there can be fewer than m distinct points $W_i(t)$, because some of them coincide and we can have $X_t = W_i(t) = W_j(t)$ for $i \ne j$. We call such a situation a "collision." There may be many kinds of collisions because there may be many different lower dimensional faces that can be hit.

Lemma 4.11 *The processes $p_i(t)$ satisfy the following:*
(a) If T is any $(\mathcal{F}^Z_t)$ stopping time satisfying $T \le \tau$ on $\{T < \infty\}$ then there exists $\delta(\omega) > 0$ such that

$$p_i(T + h) = p_i(T) \quad \text{for } 0 \le h < \delta.$$

(b) Let T be any $(\mathcal{F}^Z_t)$ stopping time satisfying $T \le \tau$ on $\{T < \infty\}$. Then for each $i = 1, \dots k$,

$$p_i(T) = \lim_{s\to T-} M_i(T)^{-1} \sum_{j\in J_i(T)} p_j(s).$$

Proof. (a) Let $D(y)$ be as defined as in Definition 4.7, and $D' = \varphi_S(D(X_T))$. Let

$$T_0 = \inf\{s \ge 0 : Z_s \notin D'\}, \quad T_1 = \inf\{s \ge T : Z_s \notin D'\};$$

note that $T_1 > T$ a.s. Let $s > 0$, ξ_0 be a bounded $\mathcal{F}^Z_T$ measurable r.v., and $\xi_1 = \prod_{j=1}^m f_j(Z_{(T+t_j)\wedge T_1})$, where f_j are bounded and measurable, and $0 \le t_1 < \dots < t_m \le s$. Write $\xi_1' = \prod_{j=1}^m f_j(Z_{(t_j)\wedge T_0})$. To prove that $p_i((T+s)\wedge T_1) = p_i(T)$ it is enough to prove that

$$\mathbb{E}^x \xi_0\xi_1 r_i(X_{(T+s)\wedge T_1}) = \mathbb{E}^x \xi_0\xi_1 p_i(T). \tag{4.25}$$

However,

$$\begin{aligned}
\mathbb{E}^x \xi_0\xi_1 r_i(X_{(T+s)\wedge T_1}) &= \mathbb{E}^x \Big(\xi_0 \mathbb{E}(\xi_1 r_i(X_{(T+s)\wedge T_1}) | \mathcal{F}^X_T)\Big) \\
&= \mathbb{E}^x \Big(\xi_0 \mathbb{E}^{X_T}(\xi_1' r_i(X_{s\wedge T_0}))\Big) \\
&= \mathbb{E}^x \Big(\xi_0 \sum_j p_j(T) \mathbb{E}^{W_j(T)}(\xi_1' r_i(X_{s\wedge T_0}))\Big). \qquad (4.26)
\end{aligned}$$

If $W_j(T) \notin S_i$ then

$$\mathbb{E}^{W_j(T)}(\xi_1' r_i(X_{s\wedge T_0})) = 0.$$

Otherwise, by (4.15) we have

$$\mathbb{E}^{W_j(T)}(\xi_1' r_i(X_{s\wedge T_0})) = M_i(T)^{-1}\widetilde{\mathbb{E}}^{Z_T}\xi_1'. \tag{4.27}$$

So,

$$\begin{aligned}\sum_j p_j(T)\mathbb{E}^{W_j(T)}(\xi_1' r_i(X_{s\wedge T_0})) &= \sum_j p_j(T) 1_{(j\in J_i(T))} M_i(T)^{-1}\widetilde{\mathbb{E}}^{Z_T}\xi_1' \\ &= p_i(T)\widetilde{\mathbb{E}}^{Z_T}\xi_1'. \end{aligned}\tag{4.28}$$

Here we used the fact that $p_j(T) = p_i(T)$ if $j \in J_i(T)$. Combining (4.26) and (4.28) we obtain (4.25).
(b) Note that $\sum_{j\in J_i(T)} r_j(x)$ is constant in a neighborhood of X_T. Hence

$$\lim_{s\to T-} \sum_{j\in J_i(T)} r_j(X_s) = \sum_{j\in J_i(T)} r_j(X_T),$$

and therefore

$$\lim_{s\to T-} \sum_{j\in J_i(T)} p_j(s) = \sum_{j\in J_i(T)} p_j(T) = M_i(T)p_i(T),$$

where the final equality holds since $p_i(T) = p_j(T)$ if $W_i(T) = W_j(T)$. □

Proposition 4.12 *Let* (A_0, A_1), Z *be as in Definition 4.6. There exists a constant* $q_1 > 0$, *depending only on* d, *such that if* $x \in A_0 \cap E_D$ *and* $T_0 \le \tau$ *is a finite* $(\mathcal{F}_t^Z)$ *stopping time, then*

$$\mathbb{P}^x(X_{T_0} \in S | \mathcal{F}_{T_0}^Z) \ge q_1. \tag{4.29}$$

Hence

$$\mathbb{P}^x(T_{A_1}^X \le \tau) \ge q_0 q_1. \tag{4.30}$$

Proof. In this proof we restrict t to $[0,\tau]$. Lemma 4.11 implies that each process $p_i(\cdot)$ is a 'pure jump' process, that is it is constant except at the jump times. (The lemma does not exclude the possibility that these jump times might accumulate.)

Let

$$\begin{aligned} K(t) &= \{i : p_i(t) > 0\}, \\ k(t) &= |K(t)|, \\ p_{\min}(t) &= \min\{p_i(t) : i \in K(t)\} = \min\{p_i(t) : p_i(t) > 0\}. \end{aligned}$$

Note that Lemma 4.11 implies that if $p_i(t) > 0$ then we have $p_i(s) > 0$ for all $s > t$. Thus K and k are non-decreasing processes. Choose $I(t)$ to be the smallest i such that $p_{I(t)}(t) = p_{\min}(t)$.

To prove (4.29) it is sufficient to prove that

$$p_{\min}(t) \geq 2^{-dk(t)} \geq 2^{-d2^d}, \quad 0 \leq t \leq \tau. \tag{4.31}$$

This clearly holds for $t = 0$, since $k(0) \geq 1$ and $p_i(0) = r_i(X_0)$, which is for each i either zero or at least 2^{-d}.

Now let

$$T = \inf\{t \leq \tau : p_{\min}(t) < 2^{-dk(t)}\}.$$

Since $p_i(T+h) = p_i(T)$ and $k(T+h) = k(T)$ for all sufficiently small $h > 0$, we must have

$$p_{\min}(T) < 2^{-dk(T)}, \quad \text{on } \{T < \infty\}. \tag{4.32}$$

Since Z is a diffusion, T is a predictable stopping time so there exists an increasing sequence of stopping times T_n with $T_n < T$ for all n, and $T = \lim_n T_n$. By the definition of T, (4.31) holds for each T_n. Let $A = \{\omega : k(T_n) < k(T) \text{ for all } n\}$. On A we have, writing $I = I(T)$, and using Lemma 4.11(b) and the fact that $k(T_n) \leq k(T) - 1$ for all n,

$$\begin{aligned} p_{\min}(T) = p_I(T) &= M_I(T)^{-1} \sum_{j \in J_I(T)} p_j(T) \\ &= \lim_{n\to\infty} M_I(T)^{-1} \sum_{j \in J_I(T)} p_j(T_n) \geq 2^{-d} \lim_{n\to\infty} p_{\min}(T_n) \\ &\geq 2^{-d} \lim_{n\to\infty} 2^{-dk(T_n)} \geq 2^{-d} 2^{-d(k(T)-1)} = 2^{-dk(T)}. \end{aligned}$$

On A^c we have

$$\begin{aligned} p_{\min}(T) &= \lim_{n\to\infty} M_I(T)^{-1} \sum_{j \in J_I(T)} p_j(T_n) \\ &\geq \lim_{n\to\infty} p_{\min}(T_n) \\ &\geq \lim_{n\to\infty} 2^{-dk(T_n)} = 2^{-dk(T)}. \end{aligned}$$

So in both case we deduce that $p_{\min}(T) \geq 2^{-dk(T)}$, contradicting (4.32). It follows that $\mathbb{P}(T < \infty) = 0$, and so (4.31) holds.

This gives (4.29), and using Proposition 4.5 we then obtain (4.30). □

4.3 Properties of X

Remark 4.13 μ is a doubling measure, so for each Borel subset H of F, almost every point of H is a point of density for H; see [44, Corollary IX.1.3].

Let I be a face of F_0 and let $F' = F - I$.

Proposition 4.14 *There exists a set $\mathcal{N}$ of capacity 0 such that if $x \notin \mathcal{N}$, then $\mathbb{P}^x(\tau_{F'} < \infty) = 1$.*

Proof. Let A be the set of x such that when the process starts at x, it never leaves x. Our first step is to show $F - A$ has positive measure. If not, for almost every x, $T_t f(x) = f(x)$, so

$$\frac{1}{t}\langle f - T_t f, f\rangle = 0.$$

Taking the supremum over $t > 0$, we have $\mathcal{E}(f,f) = 0$. This is true for every $f \in L^2$, which contradicts $\mathcal{E}$ being non-zero.

Recall the definition of E_S in (2.6). If $\mu(E_S \cap S) = 0$ for every $S \in \mathcal{S}_n(F)$ and $n \geq 1$ then $\mu(F - A) = 0$. Therefore there must exist n and $S \in \mathcal{S}_n(F)$ such that $\mu(E_S \cap S) > 0$. Let $\varepsilon > 0$. By Remark 4.13 we can find $k \geq 1$ so that there exists $S' \in \mathcal{S}_{n+k}(F)$ such that

$$\frac{\mu(E_S \cap S')}{\mu(S')} > 1 - \varepsilon.$$

Let $S'' \in \mathcal{S}_{n+k}$ be adjacent to S' and contained in S, and let g be the map that reflects $S' \cup S''$ across $S' \cap S''$. Define

$$J_i(S') = \cup\{T : T \in \mathcal{S}_{n+k+i},\, T \subset \mathrm{int}_r(S')\},$$

and define $J_i(S'')$ analogously. We can choose i large enough so that

$$\mu(E_S \cap J_i(S')) > (1 - 2\varepsilon)\mu(S'). \tag{4.33}$$

Let $x \in E_S \cap J_i(S')$. Since $x \in E_S$, the process started from x will leave S' with probability one. We can find a finite sequence of moves (that is, corners or slides) at level $n+k+i$ so that X started at x will exit S' by hitting $S' \cap S''$. By Proposition 4.12 the probability of X following this sequence of moves is strictly positive, so we have

$$\mathbb{P}^x(X(\tau_{S'}) \in S' \cap S'') > 0.$$

Starting from $x \in E_S$, the process can never leave E_S, so X will leave S' through $B = E_S \cap S' \cap S''$ with positive probability. By symmetry, X_t started from $g(x)$ will leave S'' in B with positive probability. So by the strong Markov property, starting from $g(x)$, the process will leave S with positive probability. We conclude $g(x) \in E_S$ as well. Thus $g(E_S \cap J_i(S')) \subset E_S \cap J_i(S'')$, and so by (4.33) we have

$$\mu(E_S \cap J_i(S'')) > (1 - 2\varepsilon)\mu(S'').$$

Iterating this argument, we have that for every $S_j \in \mathcal{S}_{n+k}(F)$ with $S_j \subset S$,

$$\mu(E_S \cap S_j) \geq \mu(E_S \cap J_i(S_j)) \geq (1 - 2\varepsilon)\mu(S_j).$$

Summing over the S_i's, we obtain

$$\mu(E_S \cap S) \geq (1 - 2\varepsilon)\mu(S).$$

Since ε was arbitrary, then $\mu(E_S \cap S) = \mu(S)$. In other words, starting from almost every point of S, the process will leave S.

By symmetry, this is also true for every element of $\mathcal{S}_n(F)$ isomorphic to S. Then using corners and slides (Proposition 4.12), starting at almost any $x \in F$, there is positive probability of exiting F'. We conclude that $E_{F'}$ has full measure.

The function $1_{E_{F'}}$ is invariant so $T_t 1_{E_{F'}} = 1$, a.e. By [17, Lemma 2.1.4], $T_t(1-1_{E_{F'}}) = 0$, q.e. Let $\mathcal{N}$ be the set of x where $T_t 1_{E_{F'}}(x) \neq 1$ for some rational t. If $x \notin \mathcal{N}$, then $\mathbb{P}^x(X_t \in E_{F'}) = 1$ if t is rational. By the Markov property, $x \in E_{F'}$. □

Lemma 4.15 *Let $U \subset F$ be open and non-empty. Then $\mathbb{P}^x(T_U < \infty) = 1$, q.e.*

Proof. This follows by Propositions 4.12 and 4.14. □

4.4 Coupling

Lemma 4.16 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let X and Z be random variables taking values in separable metric spaces E_1 and E_2, respectively, each furnished with the Borel σ-field. Then there exists $F : E_2 \times [0,1] \to E_1$ that is jointly measurable such that if U is a random variable whose distribution is uniform on $[0,1]$ which is independent of Z and $\widetilde{X} = F(Z,U)$, then (X,Z) and $(\widetilde{X}, Z)$ have the same law.*

Proof. First let us suppose $E_1 = E_2 = [0,1]$. We will extend to the general case later. Let $\mathbb{Q}$ denote the rationals. For each $r \in [0,1] \cap \mathbb{Q}$, $\mathbb{P}(X \le r \mid Z)$ is a $\sigma(Z)$-measurable random variable, hence there exists a Borel measurable function h_r such that $\mathbb{P}(X \le r \mid Z) = h_r(Z)$, a.s. For $r < s$ let $A_{rs} = \{z : h_r(z) > h_s(z)\}$. If $C = \cup_{r<s;\, r,s\in\mathbb{Q}} A_{rs}$, then $\mathbb{P}(Z \in C) = 0$. For $z \notin C$, $h_r(z)$ is nondecreasing in r for r rational. For $x \in [0,1]$, define $g_x(z)$ to be equal to x if $z \in C$ and equal to $\inf_{s>x, s\to x;\, s\in\mathbb{Q}} h_s(z)$ otherwise. For each z, let $f_x(z)$ be the right continuous inverse to $g_x(z)$. Finally let $F(z,x) = f_x(z)$.

We need to check that (X,Z) and $(\widetilde{X}, Z)$ have the same distributions. We have

$$\begin{aligned}\mathbb{P}(X \le x, Z \le z) &= \mathbb{E}[\mathbb{P}(X \le x \mid Z); Z \le z] = \lim_{s>x, s\in\mathbb{Q}, s\to x} \mathbb{E}[\mathbb{P}(X \le s \mid Z); Z \le z] \\ &= \lim \mathbb{E}[h_s(Z); Z \le z] = \mathbb{E}[g_x(Z); Z \le z].\end{aligned}$$

On the other hand,

$$\begin{aligned}\mathbb{P}(\widetilde{X} \le x, Z \le r) &= \mathbb{E}[\mathbb{P}(F(Z,U) \le x \mid Z); Z \le z] = \mathbb{E}[\mathbb{P}(f_U(Z) \le x \mid Z); Z \le z] \\ &= \mathbb{E}[\mathbb{P}(U \le g_x(Z) \mid Z); Z \le z] = \mathbb{E}[g_x(Z); Z \le z].\end{aligned}$$

For general E_1, E_2, let ψ_i be bimeasurable one-to-one maps from E_i to $[0,1]$, $i = 1,2$. Apply the above to $\overline{X} = \psi_1(X)$ and $\overline{Z} = \psi_2(Z)$ to obtain a function $\overline{F}$. Then $F(z,u) = \psi_1^{-1} \circ \overline{F}(\psi_2(z), u)$ will be the required function. □

We say that $x, y \in F$ are *m-associated*, and write $x \sim_m y$, if $\varphi_S(x) = \varphi_S(y)$ for some (and hence all) $S \in \mathcal{S}_m$. Note that by Lemma 2.13 if $x \sim_m y$ then also

$x \sim_{m+1} y$. One can verify that this is the same as the definition of $x{\sim_m}y$ given in [5].

The coupling result we want is:

Proposition 4.17 *(Cf. [5, Theorem 3.14].) Let $x_1, x_2 \in F$ with $x_1 \sim_n x_2$, where $x_1 \in S_1 \in \mathcal{S}_n(F)$, $x_2 \in S_2 \in \mathcal{S}_n(F)$, and let $\Phi = \varphi_{S_1}|_{S_2}$. Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ carrying processes X_k, $k = 1, 2$ and Z with the following properties.*
(a) Each X_k is an $\mathcal{E}$-diffusion started at x_k.
(b) $Z = \varphi_{S_2}(X_2) = \Phi \circ \varphi_{S_1}(X_1)$.
(c) X_1 and X_2 are conditionally independent given Z.

Proof. Let Y be the diffusion corresponding to the Dirichlet form $\mathcal{E}$ and let Y_1, Y_2 be processes such that Y_i is equal in law to Y started at x_i. Let $Z_1 = \Phi \circ \varphi_{S_1}(Y_1)$ and $Z_2 = \varphi_{S_2}(Y_2)$. Since the Dirichlet form for $\varphi_{S_i}(Y)$ is $\mathcal{E}^{S_i}$ and Z_1, Z_2 have the same starting point, then Z_1 and Z_2 are equal in law. Use Lemma 4.16 to find functions F_1 and F_2 such that $(F_i(Z_i, U), Z_i)$ is equal in law to (Y_i, Z_i), $i = 1, 2$, if U is an independent uniform random variable on $[0, 1]$.

Now take a probability space supporting a process Z with the same law as Z_i and two independent random variables U_1, U_2 independent of Z which are uniform on $[0, 1]$. Let $X_i = F_i(Z, U_i)$, $i = 1, 2$. We proceed to show that the X_i satisfy (a)-(c).

X_i is equal in law to $F_i(Z_i, U_i)$, which is equal in law to Y_i, $i = 1, 2$, which establishes (a). Similarly (X_i, Z) is equal in law to $(F(Z_i, U_i), Z_i)$, which is equal in law to (Y_i, Z_i). Since $Z_1 = \Phi \circ \varphi_{S_1}(Y_1)$ and $Z_2 = \varphi_{S_2}(Y_2)$, it follows from the equality in law that $Z = \Phi \circ \varphi_{S_1}(Y_1)$ and $Z = \varphi_{S_2}(Y_2)$. This establishes (b).

As $X_i = F_i(Z, U_i)$ for $i = 1, 2$, and Z, U_1, and U_2 are independent, (c) is immediate. □

Given a pair of $\mathcal{E}$-diffusions $X_1(t)$ and $X_2(t)$ we define the coupling time

$$T_C(X_1, X_2) = \inf\{t \geq 0 : X_1(t) = X_2(t)\}. \tag{4.34}$$

Given Propositions 4.12 and 4.17 we can now use the same arguments as in [5] to couple copies of X started at points $x, y \in F$, provided that $x{\sim_m}y$ for some $m \geq 1$.

Theorem 4.18 *Let $r > 0$, $\varepsilon > 0$ and $r' = r/L_F^2$. There exist constants q_3 and δ, depending only on the GSC F, such that the following hold:*
(a) Suppose x_1, $x_2 \in F$ with $||x_1 - x_2||_\infty < r'$ and $x_1{\sim_m}x_2$ for some $m \geq 1$. There exist $\mathcal{E}$-diffusions $X_i(t)$, $i = 1, 2$, with $X_i(0) = x_i$, such that, writing

$$\tau_i = \inf\{t \geq 0 : X_i(t) \notin B(x_1, r)\},$$

we have

$$\mathbb{P}\big(T_C(X_1, X_2) < \tau_1 \wedge \tau_2\big) > q_3. \tag{4.35}$$

(b) If in addition $||x_1 - x_2||_\infty < \delta r$ and $x_1{\sim_m}x_2$ for some $m \geq 1$ then

$$\mathbb{P}\big(T_C(X_1, X_2) < \tau_1 \wedge \tau_2\big) > 1 - \varepsilon. \tag{4.36}$$

Proof. Given Propositions 4.12 and 4.17, this follows by the same arguments as in [5], p. 694–701. □

4.5 Elliptic Harnack inequality

As mentioned in Section 2.1, there are two definitions of harmonic that we can give. We adopt the probabilistic one here. Recall that a function h is harmonic in a relatively open subset D of F if $h(X_{t\wedge\tau'_D})$ is a martingale under $\mathbb{P}^x$ for q.e. x whenever D' is a relatively open subset of D.

X satisfies the *elliptic Harnack inequality* if there exists a constant c_1 such that the following holds: for any ball $B(x,R)$, whenever u is a non-negative harmonic function on $B(x,R)$ then there is a quasi-continuous modification $\tilde{u}$ of u that satisfies

$$\sup_{B(x,R/2)} \tilde{u} \le c_1 \inf_{B(x,R/2)} \tilde{u}.$$

We abbreviate "elliptic Harnack inequality" by "EHI."

Lemma 4.19 *Let $\mathcal{E}$ be in $\mathfrak{E}$, $r\in(0,1)$, and h be bounded and harmonic in $B = B(x_0,r)$. Then there exists $\theta>0$ such that*

$$|h(x)-h(y)| \le C\Big(\frac{|x-y|}{r}\Big)^\theta (\sup_B |h|), \quad x,y\in B(x_0,r/2), \quad x\sim_m y. \tag{4.37}$$

Proof. As in [5, Proposition 4.1] this follows from the coupling in Theorem 4.18 by standard arguments. □

Proposition 4.20 *Let $\mathcal{E}$ be in $\mathfrak{E}$ and h be bounded and harmonic in $B(x_0,r)$. Then there exists a set $\mathcal{N}$ of $\mathcal{E}$-capacity 0 such that*

$$|h(x)-h(y)| \le C\Big(\frac{|x-y|}{r}\Big)^\theta (\sup_B |h|), \quad x,y\in B(x_0,r/2)-\mathcal{N}. \tag{4.38}$$

Proof. Write $B = B(x_0,r)$, $B' = B(x_0,r/2)$. By Lusin's theorem, there exist open sets $G_n\downarrow$ such that $\mu(G_n)\downarrow 0$, and h restricted to $G_n^c\cap B'$ is continuous. We will first show that h restricted to any G_n^c satisfies (4.37) except when one or both of x,y is in $\mathcal{N}_n$, a set of measure 0. If $G=\cap_n G_n$, then h on G^c is Hölder continuous outside of $\cup\mathcal{N}_n$, which is a set of measure 0. Thus h is Hölder continuous on all of B' outside of a set E of measure 0.

So fix n and let $H = G_n^c$. Let x,y be points of density for H; recall Remark 4.13. Let S_x and S_y be appropriate isometries of an element of $\mathcal{S}_k$ such that $x\in S_x$, $y\in S_y$, and $\mu(S_x\cap H)/\mu(S_x)\ge \frac{2}{3}$ and the same for S_y. Let Φ be the isometry taking S_x to S_y. Then the measure of $\Phi(S_x\cap H)$ must be at least two thirds the measure of S_y and we already know the measure of $S_y\cap H$ is at least two thirds that of S_y. Hence the measure of $(S_y\cap H)\cap(\Phi(S_x\cap H))$ is at least one third the measure of S_y. So there must exist points $x_k\in S_x\cap H$ and $y_k=\Phi(x_k)\in S_y\cap H$ that are m-associated for some m. The inequality (4.37) holds for each pair x_k, y_k.

We do this for each k sufficiently large and get sequences $x_k \in H$ tending to x and $y_k \in H$ tending to y. Since h restricted to H is continuous, (4.37) holds for our given x and y.

We therefore know that h is continuous a.e. on B'. We now need to show the continuity q.e., without modifying the function h. Let x, y be two points in B' for which $h(X_{t\wedge\tau_B})$ is a martingale under $\mathbb{P}^x$ and $\mathbb{P}^y$. The set of points $\mathcal{N}$ where this fails has $\mathcal{E}$-capacity zero. Let $R = |x-y| < r$ and let $\varepsilon > 0$. Since $\mu(E) = 0$, then by [17, Lemma 4.1.1], for each t, $T_t 1_E(x) = T_t(x, E) = 0$ for m-a.e. x. $T_t 1_E$ is in the domain of $\mathcal{E}$, so by [17, Lemma 2.1.4], $T_t 1_E = 0$, q.e. Enlarge $\mathcal{N}$ to include the null sets where $T_t 1_E \neq 0$ for some t rational. Hence if $x, y \notin \mathcal{N}$, then with probability one with respect to both $\mathbb{P}^x$ and $\mathbb{P}^y$, we have $X_t \notin E$ for t rational. Choose balls B_x, B_y with radii in $[R/4, R/3]$ and centered at x and y, resp., such that $\mathbb{P}^x(X_{\tau_{B_x}} \in \mathcal{N}) = \mathbb{P}^y(X_{\tau_{B_y}} \in \mathcal{N}) = 0$. By the continuity of paths, we can choose t rational and small enough that $\mathbb{P}^x(\sup_{s\le t} |X_s - X_0| > R/4) < \varepsilon$ and the same with x replaced by y. Then

$$
\begin{aligned}
|h(x) - h(y)| &= |\mathbb{E}^x h(X_{t\wedge\tau_{B_x}}) - \mathbb{E}^y h(X_{t\wedge\tau_{B_y}})| \\
&\le |\mathbb{E}^x [h(X_{t\wedge\tau_{B_x}}); t < \tau_{B_x}] - \mathbb{E}^y [h(X_{t\wedge\tau_{B_y}}); t < \tau_{B_y}]| + 2\varepsilon \|h\|_\infty \\
&\le C\Big(\frac{R}{r}\Big)^\theta \|h\|_\infty + 4\varepsilon\|h\|_\infty .
\end{aligned}
$$

The last inequality above holds because we have $\mathbb{P}^x(X_t \in \mathcal{N}) = 0$ and similarly for $\mathbb{P}^y$, points in B_x are at most $2R$ from points in B_y, and $X_{t\wedge\tau_{B_x}}$ and $X_{t\wedge\tau_{B_y}}$ are not in E almost surely. Since ε is arbitrary, this shows that except for x, y in a set of capacity 0, we have (4.37). □

Lemma 4.21 *Let $\mathcal{E} \in \mathfrak{E}$. Then there exist constants $\kappa > 0$, C_i, depending only on F, such that if $0 < r < 1$, $x_0 \in F$, $y, z \in B(x_0, C_1 r)$ then for all $0 < \delta < C_1$,*

$$
\mathbb{P}^y(T_{B(z,\delta r)} < \tau_{B(x_0,r)}) > \delta^\kappa. \tag{4.39}
$$

Proof. This follows by using corner and slide moves, as in [5, Corollary 3.24]. □

Proposition 4.22 *EHI holds for $\mathcal{E}$, with constants depending only on F.*

Proof. Given Proposition 4.20 and Lemma 4.21 this follows by the same argument as [5, Theorem 4.3]. □

Corollary 4.23 *(a) $\mathcal{E}$ is irreducible.*
(b) If $\mathcal{E}(f, f) = 0$ then f is a.e. constant.

Proof. (a) If A is an invariant set, then $T_t 1_A = 1_A$, or 1_A is harmonic on F. By EHI, either 1_A is never 0 except for a set of capacity 0 or else it is 0, q.e. Hence $\mu(A)$ is either 0 or 1. So $\mathcal{E}$ is irreducible.
(b) The equivalence of (a) and (b) in this setting is well known to experts. Suppose

that f is a function such that $\mathcal{E}(f,f)=0$, and that f is not a.e. constant. Then using the contraction property and scaling we can assume that $0 \le f \le 1$ and there exist $0<a<b<1$ such that the sets $A=\{x : f(x)<a\}$ and $B=\{x : f(x)>b\}$ both have positive measure. Let $g=b\wedge(a\vee f)$; then $\mathcal{E}(g,g)=0$ also. By Lemma 1.3.4 of [17], for any $t>0$,

$$\mathcal{E}^{(t)}(g,g)=t^{-1}\langle g-T_t g, g\rangle = 0.$$

So $\langle g, T_t g\rangle = \langle g,g\rangle$. By the semigroup property, $T_t^2=T_{2t}$, and hence $\langle T_t g, T_t g\rangle = \langle g, T_{2t} g\rangle = \langle g,g\rangle$, from which it follows that $\langle g-T_t g, g-T_t g\rangle = 0$. This implies that $g(x)=\mathbb{E}^x g(X_t)$ a.e. Hence the sets A and B are invariant for (T_t), which contradicts the irreducibility of $\mathcal{E}$. □

Given a Dirichlet form $(\mathcal{E},\mathcal{F})$ on F we define the *effective resistance* between subsets A_1 and A_2 of F by:

$$R_{\text{eff}}(A_1,A_2)^{-1}=\inf\{\mathcal{E}(f,f) : f\in\mathcal{F}, f\,|_{A_1}=0, f\,|_{A_2}=1\}. \tag{4.40}$$

Let

$$A(t)=\{x\in F : x_1=t\}, \qquad t\in[0,1]. \tag{4.41}$$

For $\mathcal{E}\in\mathfrak{E}$ we set

$$||\mathcal{E}|| = R_{\text{eff}}(A(0),A(1))^{-1}. \tag{4.42}$$

Let $\mathfrak{E}_1=\{\mathcal{E}\in\mathfrak{E} : ||\mathcal{E}||=1\}$.

Lemma 4.24 *If $\mathcal{E}\in\mathfrak{E}$ then $||\mathcal{E}||>0$.*

Proof. Write $\mathcal{H}$ for the set of functions u on F such that $u=i$ on $A(i)$, $i=0,1$. First, observe that $\mathcal{F}\cap\mathcal{H}$ is not empty. This is because, by the regularity of $\mathcal{E}$, there is a continuous function $u\in\mathcal{F}$ such that $u\le 0$ on the face $A(0)$ and $u\ge 1$ on the opposite face $A(1)$. Then the Markov property for Dirichlet forms says $0\vee(u\wedge 1)\in\mathcal{F}\cap\mathcal{H}$.

Second, observe that by Proposition 4.14 and the symmetry, $T_{A(0)}<\infty$ a.s., which implies that $(\mathcal{E},\mathcal{F}_{A(0)})$ is a transient Dirichlet form (see Lemma 1.6.5 and Theorem 1.6.2 in [17]). Here as usual we denote $\mathcal{F}_{A(0)}=\{f\in\mathcal{F} : f|_{A(0)}=0\}$. Hence $\mathcal{F}_{A(0)}$ is a Hilbert space with the norm $\mathcal{E}$. Let $u\in\mathcal{F}\cap\mathcal{H}$ and h be its orthogonal projection onto the orthogonal complement of $\mathcal{F}_{A(0)\cup A(1)}$ in this Hilbert space. It is easy to see that $\mathcal{E}(h,h)=||\mathcal{E}||$.

If we suppose that $||\mathcal{E}||=0$, then $h=0$ by Corollary 4.23. By our definition, h is harmonic in the complement of $A(0)\cup A(1)$ in the Dirichlet form sense, and so by Proposition 2.5 h is harmonic in the probabilistic sense and $h(x)=\mathbb{P}^x(X_{T_{A(0)\cup A(1)}}\in A(1))$. Thus, by the symmetries of F, the fact that $h=0$ contradicts the fact that $T_{A(1)}<\infty$ by Proposition 4.14.

An alternative proof of this lemma starts with defining h probabilistically and uses [14, Corollary 1.7] to show $h\in\mathcal{F}_{A(0)}$. □

4.6 Resistance estimates

Let now $\mathcal{E} \in \mathfrak{E}_1$. Let $S \in \mathcal{S}_n$ and let $\gamma_n = \gamma_n(\mathcal{E})$ be the conductance across S. That is, if $S = Q \cap F$ for $Q \in \mathcal{Q}_n(F)$ and $Q = \{a_i \le x_i \le b_i, i = 1, \dots, d\}$, then

$$\gamma_n = \inf\{\mathcal{E}^S(u,u) : u \in \mathcal{F}^S, u\,|_{\{x_1=a_1\}} = 0, u\,|_{\{x_1=b_1\}} = 1\}.$$

Note that γ_n does not depend on S, and that $\gamma_0 = 1$. Write $v_n = v_n^{\mathcal{E}}$ for the minimizing function. We remark that from the results in [4, 34] we have

$$C_1 \rho_F^n \le \gamma_n(\mathcal{E}_{BB}) \le C_2 \rho_F^n.$$

Proposition 4.25 *Let $\mathcal{E} \in \mathfrak{E}_1$. Then for $n, m \ge 0$*

$$\gamma_{n+m}(\mathcal{E}) \ge C_1 \gamma_m(\mathcal{E}) \rho_F^n. \tag{4.43}$$

Proof. We begin with the case $m = 0$. As in [4] we compare the energy of v_0 with that of a function constructed from v_n and the minimizing function on a network where each cube side L_F^{-n} is replaced by a diagonal crosswire.

Write D_n for the network of diagonal crosswires, as in [4, 34], obtained by joining each vertex of a cube $Q \in \mathcal{Q}_n$ to a vertex at the center of the cube by a wire of unit resistance. Let R_n^D be the resistance across two opposite faces of F in this network, and let f_n be the minimizing potential function.

Fix a cube $Q \in \mathcal{Q}_n$ and let $S = Q \cap F$. Let x_i, $i = 1, \dots 2^d$, be its vertices, and for each i let A_{ij}, $j = 1, \dots d$, be the faces containing x_i. Let A'_{ij} be the face opposite to A_{ij}. Let w_{ij} be the function, congruent to v_n, which is 1 on A_{ij} and zero on A'_{ij}. Set

$$u_i = \min\{w_{i1}, \dots w_{id}\}.$$

Note that $u_i(x_i) = 1$, and $u_i = 0$ on $\cup_j A'_{ij}$. Then

$$\mathcal{E}(u_i, u_i) \le \sum_j \mathcal{E}(w_{ij}, w_{ij}) = d\gamma_n.$$

Write $a_i = f(x_i)$, and $\overline{a} = 2^{-d} \sum_i a_i$. Then the energy of f_n in S is

$$\mathcal{E}_D^S(f_n, f_n) = \sum_i (a_i - \overline{a})^2.$$

Now define a function $g_S : S \to \mathbb{R}$ by

$$g_S(y) = \overline{a} + \sum_i (a_i - \overline{a}) u_i(y).$$

Then

$$\mathcal{E}^S(g_S, g_S) \le C\mathcal{E}(u_1, u_1) \sum_i (a_i - \overline{a})^2 \le C\gamma_n \mathcal{E}_D^S(f_n, f_n).$$

We can check from the definition of g_S that if two cubes Q_1, Q_2 have a common face A and $S_i = Q_i \cap F$, then $g_{S_1} = g_{S_2}$ on A. Now define $g : F \to \mathbb{R}$ by taking $g(x) = g_S(x)$ for $x \in S$. Summing over $Q \in \mathcal{Q}_n(F)$ we deduce that $\mathcal{E}(g,g) \le C\gamma_n (R_n^D)^{-1}$. However, the function g is zero on one face of F, and 1 on the opposite face. Therefore

$$1 = \gamma_0 = \mathcal{E}(v_0, v_0) \le \mathcal{E}(g,g) \le C\gamma_n (R_n^D)^{-1} \le C\gamma_n \rho_F^{-n},$$

which gives (4.43) in the case $m = 0$.

The proof when $m \ge 1$ is the same, except we work in a cube $S \in \mathcal{S}_m$ and use subcubes of side L_F^{-n-m}. □

Lemma 4.26 *We have*

$$C_1 \gamma_n \le \gamma_{n+1} \le C_2 \gamma_n. \tag{4.44}$$

Proof. The left-hand inequality is immediate from (4.43). To prove the right-hand one, let first $n = 0$. By Propositions 4.12 and 4.14, we deduce that $v_0 \ge C_3 > 0$ on $A(L_F^{-1})$; recall the definition in (4.41). Let $w = (v_0 \wedge C_3)/C_3$. Choose a cube $Q \in \mathcal{Q}_1(F_1)$ between the hyperplanes $A_1(0)$ and $A_1(L_F^{-1})$; $A_1(t)$ is defined in (4.41). Then

$$\begin{aligned}\gamma_1 = \mathcal{E}^{F_1}(v_1, v_1) &\le \mathcal{E}^{F_1}(w,w) \le \mathcal{E}(w,w) \\ &= C_3^{-2}\mathcal{E}(v_0 \wedge C_3, v_0 \wedge C_3) \le C_3^{-2}\mathcal{E}(v_0, v_0) = C_4 \gamma_0.\end{aligned}$$

Again the case $n \ge 0$ is similar, except we work in a cube $S \in \mathcal{S}_n$. □

Note that (4.43) and (4.44) only give a one-sided comparison between $\gamma_n(\mathcal{E})$ and $\gamma_n(\mathcal{E}_{BB})$; however this will turn out to be sufficient.

Set

$$\alpha = \log m_F / \log L_F, \qquad \beta_0 = \log(m_F \rho_F)/\log L_F.$$

By [5, Corollary 5.3] we have $\beta_0 \ge 2$, and so $\rho_F m_F \ge L_F^2$. Let

$$H_0(r) = r^{\beta_0}.$$

We now define a 'time scale function' H for $\mathcal{E}$. First note that by (4.43) we have, for $n \ge 0$, $k \ge 0$.

$$\frac{\gamma_n m_F^n}{\gamma_{n+k} m_F^{n+k}} \le C \rho_F^{-k} m_F^{-k}. \tag{4.45}$$

Since $\rho_F m_F \ge L_F^2 > 1$ there exists $k \ge 1$ such that

$$\gamma_n m_F^n < \gamma_{n+k} m_F^{n+k}, \quad n \ge 0. \tag{4.46}$$

Fix this k, let

$$H(L_F^{-nk}) = \gamma_{nk}^{-1} m_F^{-nk}, \quad n \ge 0, \tag{4.47}$$

and define H by linear interpolation on each interval $(L_F^{-(n+1)k}, L_F^{-nk})$. Set also $H(0) = 0$. We now summarize some properties of H.

Lemma 4.27 *There exist constants C_i and β', depending only on F such that the following hold.*
(a) H is strictly increasing and continuous on $[0,1]$.
(b) For any $n, m \geq 0$

$$H(L_F^{-nk-mk}) \leq C_1 H(L_F^{-nk}) H_0(L_F^{-mk}). \tag{4.48}$$

(c) For $n \geq 0$

$$H(L_F^{-(n+1)k}) \leq H(L_F^{-nk}) \leq C_2 H(L_F^{-(n+1)k}). \tag{4.49}$$

(d)

$$C_3 (t/s)^{\beta_0} \leq \frac{H(t)}{H(s)} \leq C_4 (t/s)^{\beta'} \text{ for } 0 < s \leq t \leq 1. \tag{4.50}$$

In particular H satisfies the 'fast time growth' condition of [20] and [10, Assumption 1.2].
(e) H satisfies 'time doubling':

$$H(2r) \leq C_5 H(r) \text{ for } 0 \leq r \leq 1/2. \tag{4.51}$$

(f) For $r \in [0,1]$,

$$H(r) \leq C_6 H_0(r).$$

Proof. (a), (b) and (c) are immediate from the definitions of H and H_0, (4.43) and (4.44). For (d), using (4.48) we have

$$\frac{H(L_F^{-kn})}{H(L_F^{-kn-km})} \geq C_7 \frac{H(L_F^{-kn})}{H(L_F^{-kn}) H_0(L_F^{-km})} = C_7 L_F^{km\beta_0} = C_7 \Big(\frac{L_F^{-kn}}{L_F^{-kn-km}} \Big)^{\beta_0},$$

and interpolating using (c) gives the lower bound in (4.50). For the upper bound, using (4.44),

$$\frac{H(L_F^{-kn})}{H(L_F^{-kn-km})} \leq C_8^{km} = L_F^{km\beta'} = \Big(\frac{L_F^{-kn}}{L_F^{-kn-km}} \Big)^{\beta'}, \tag{4.52}$$

where $\beta' = \log C_8 / \log L_F$, and again using (c) gives (4.50). (e) is immediate from (d). Taking $n = 0$ in (4.48) and using (c) gives (f). □

We say $\mathcal{E}$ satisfies the condition RES(H, c_1, c_2) if for all $x_0 \in F$, $r \in (0, L_F^{-1})$,

$$c_1 \frac{H(r)}{r^\alpha} \leq R_{\text{eff}}(B(x_0, r), B(x_0, 2r)^c) \leq c_2 \frac{H(r)}{r^\alpha}. \tag{$RES(H,c_1,c_2)$}$$

Proposition 4.28 *There exist constants C_1, C_2, depending only on F, such that $\mathcal{E}$ satisfies RES(H, C_1, C_2).*

Proof. Let k be the smallest integer so that $L_F^{-k} \leq \frac{1}{2} d^{-1/2} R$. Note that if $Q \in \mathcal{Q}_k$ and $x, y \in Q$, then $d(x,y) \leq d^{1/2} L_F^{-k} \leq \frac{1}{2} R$. Write $B_0 = B(x_0, R)$ and $B_1 = B(x_0, 2R)^c$.

We begin with the upper bound. Let S_0 be a cube in $\mathcal{Q}_k$ containing x_0: then $S_0 \cap F \subset B$. We can find a chain of cubes $S_0, S_1, \dots S_n$ such that $S_n \subset B_1$ and S_i is adjacent to S_{i+1} for $i = 0, \dots, n-1$. Let f be the harmonic function in $F - (S_0 \cup B_1)$ which is 1 on S_0 and 0 on B_1. Let $A_0 = S_0 \cap S_1$, and A_1 be the opposite face of S_1 to A_0. Then using the lower bounds for slides and corner moves, we have that there exists $C_1 \in (0,1)$ such that $f \geq C_1$ on A_1. So $g = (f - C_1)_+/(1 - C_1)$ satisfies $\mathcal{E}^{S_1}(g,g) \geq \gamma_k$. Hence

$$R_{\text{eff}}(S_0, B_1)^{-1} = \mathcal{E}(f,f) \geq \mathcal{E}^{S_1}(f,f) \geq (1 - C_1)^{-2}\gamma_k,$$

and by the monotonicity of resistance

$$R_{\text{eff}}(B_0, B_1) \leq R_{\text{eff}}(S_0, B_1) \leq C_2 \gamma_k^{-1},$$

which gives the upper bound in (RES(H, c_1, c_2)).

Now let $n = k + 1$ and let $S \in \mathcal{Q}_n$. Recall from Proposition 4.25 the definition of the functions v_n, w_{ij} and u_i. By the symmetry of v_n we have that $w_{ij} \geq \frac{1}{2}$ on the half of S which is closer to A_{ij}, and therefore $u_i(x) \geq \frac{1}{2}$ if $||x - x_i||_\infty \leq \frac{1}{2} L_F^{-n}$.

Now let $y \in L_F^{-n}\mathbb{Z}^d \cap F$, and let $V(y)$ be the union of the 2^d cubes in $\mathcal{Q}_n$ containing y. By looking at functions congruent to $2u_i \wedge 1$ in each of the cubes in $V(y)$, we can construct a function g_i such that $g_i = 0$ on $F - V(y)$, $g_i(z) = 1$ for $z \in F$ with $||z - y||_\infty \leq \frac{1}{2} L_F^{-n}$, and $\mathcal{E}(g_i, g_i) \leq C\gamma_n$. We now choose $y_1, \dots y_m$ so that $B_0 \subset \cup_i V(y_i)$: clearly we can take $m \leq C_5$. Then if $h = 1 \wedge (\sum_i g_i)$, we have $h = 1$ on B_0 and $h = 0$ on B_1. Thus

$$R_{\text{eff}}(B_0, B_1)^{-1} \leq \mathcal{E}(h,h) \leq \mathcal{E}\Big(\sum g_i, \sum g_i\Big) \leq C_6 \gamma_n,$$

proving the lower bound. □

4.7 Heat kernel estimates

We write h for the inverse of H, and $V(x,r) = \mu(B(x,r))$. We say that $p_t(x,y)$ satisfies HK($H; \eta_1, \eta_2, c_0$) if for $x, y \in F$, $0 < t \leq 1$,

$$p_t(x,y) \geq c_0^{-1} V(x, h(t))^{-1} \exp(-c_0 (H(d(x,y))/t)^{\eta_1}),$$
$$p_t(x,y) \leq c_0 V(x, h(t))^{-1} \exp(-c_0^{-1} (H(d(x,y))/t)^{\eta_2}).$$

The following equivalence is proved in [20]. (See also [10, Theorem 1.3, $(a) \Rightarrow (c)$] for a detailed proof of $(a) \Rightarrow (b)$, which is adjusted to our current setting.)

Theorem 4.29 *Let $H : [0,1] \to [0,\infty)$ be a strictly increasing function with $H(1) \in (0,\infty)$ that satisfies (4.51) and (4.50). Then the following are equivalent:*
(a) $(\mathcal{E}, \mathcal{F})$ satisfies (VD), (EHI) and $(RES(H, c_1, c_2))$ for some $c_1, c_2 > 0$.
(b) $(\mathcal{E}, \mathcal{F})$ satisfies $HK(H; \eta_1, \eta_2, c_0)$ for some $\alpha, \eta_1, \eta_2, c_0 > 0$.
Further the constants in each implication are effective.

By saying that the constants are 'effective' we mean that if, for example (a) holds, then the constants η_i, c_0 in (b) depend only on the constants c_i in (a), and the constants in (VD), (EHI) and (4.51) and (4.50).

Theorem 4.30 *X has a transition density $p_t(x,y)$ which satisfies $HK(H;\eta_1,\eta_2,C)$, where $\eta_1 = 1/(\beta_0 - 1)$, $\eta_2 = 1/(\beta' - 1)$, and the constant C depends only on F.*

Proof. This is immediate from Theorem 4.29, and Propositions 4.22 and 4.28. □

Let

$$\begin{aligned} J_r(f) &= r^{-\alpha}\int_F\int_{B(x,r)}|f(x)-f(y)|^2 d\mu(x)d\mu(y),\\ N_H^r(f) &= H(r)^{-1}J_r(f),\\ N_H(f) &= \sup_{0<r\le 1} N_H^r(f),\\ W_H &= \{f\in L^2(F,\mu) : N_H(f)<\infty\}. \end{aligned} \tag{4.53}$$

We now use Theorem 4.1 of [28], which we rewrite slightly for our context. (See also Theorem 1.4 of [10], which is adjusted to our current setting.) Let $r_j = L^{-kj}$, where k is as in the definition of H.

Theorem 4.31 *Suppose p_t satisfies $HK(H,\eta_1,\eta_2,C_0)$, and H satisfies (4.51) and (4.50). Then*

$$C_1\mathcal{E}(f,f)\le \limsup_{j\to\infty} N_H^{r_j}(f)\le N_H(f)\le C_2\mathcal{E}(f,f) \quad \text{for all } f\in W_H, \tag{4.54}$$

where the constants C_i depend only on the constants in (4.51) and (4.50), and in $HK(H;\eta_1,\eta_2,C_0)$. Further,

$$\mathcal{F} = W_H. \tag{4.55}$$

Theorem 4.32 *Let $(\mathcal{E},\mathcal{F})\in\mathfrak{E}_1$.*
(a) There exist constants $C_1, C_2 > 0$ such that for all $r\in[0,1]$,

$$C_1H_0(r)\le H(r)\le C_2H_0(r). \tag{4.56}$$

(b) $W_H = W_{H_0}$, and there exist constants C_3, C_4 such that

$$C_3N_{H_0}(f)\le\mathcal{E}(f,f)\le C_4N_{H_0}(f) \quad \text{for all } f\in W_H. \tag{4.57}$$

(c) $\mathcal{F} = W_{H_0}$.

Proof. (a) We have $H(r)\le C_2H_0(r)$ by Lemma 4.27, and so

$$N_H(f)\ge C_2^{-1}N_{H_0}(f). \tag{4.58}$$

Recall that $(\mathcal{E}_{BB},\mathcal{F}_{BB})$ is (one of) the Dirichlet forms constructed in [5]. By (4.58) and (4.55) we have $\mathcal{F}\subset\mathcal{F}_{BB}$. In particular, the function $v_0^{\mathcal{E}}\in\mathcal{F}_{BB}$ (see Subsection 4.6).

Now let

$$A = \limsup_{k\to\infty} \frac{H(r_k)}{H_0(r_k)};$$

we have $A \le C_2$.

Let $f \in \mathcal{F}$. Then by Theorem 4.31

$$\begin{aligned}
\mathcal{E}_{BB}(f,f) &\le C_3 \limsup_{j\to\infty} H_0(r_j)^{-1} J_{r_j}(f) \\
&= C_3 \limsup_{j\to\infty} \frac{H(r_j)}{H_0(r_j)} H(r_j)^{-1} J_{r_j}(f) \\
&\le C_3 \limsup_{j\to\infty} A N_H^{r_j}(f) \le C_4 A \mathcal{E}(f,f).
\end{aligned}$$

Taking $f = v_0^{\mathcal{E}}$,

$$1 \le \mathcal{E}_{BB}(v_0^{\mathcal{E}}, v_0^{\mathcal{E}}) \le C_4 A \mathcal{E}(v_0^{\mathcal{E}}, v_0^{\mathcal{E}}) = C_4 A. \tag{4.59}$$

Thus $A \ge C_5 = C_4^{-1}$. By Lemma 4.27(c) we have, for $n, m \ge 0$,

$$\frac{H(r_{n+m})}{H_0(r_{n+m})} \le C_6 \frac{H(r_n)}{H_0(r_n)}.$$

So, for any n

$$\frac{H(r_n)}{H_0(r_n)} \ge C_6^{-1} A \ge C_5/C_6,$$

and (a) follows.

(b) and (c) are then immediate by Theorem 4.31. □

Remark 4.33 (4.56) now implies that $p_t(x,y)$ satisfies $\mathrm{HK}(H_0, \eta_1, \eta_1, C)$ with $\eta_1 = 1/(\beta_0 - 1)$.

5 Uniqueness

Definition 5.1 Let $W = W_{H_0}$ be as defined in (4.53). Let $\mathcal{A}, \mathcal{B} \in \mathfrak{E}$. We say $\mathcal{A} \le \mathcal{B}$ if

$$\mathcal{B}(u,u) - \mathcal{A}(u,u) \ge 0 \text{ for all } u \in W.$$

For $\mathcal{A}, \mathcal{B} \in \mathfrak{E}$ define

$$\begin{aligned}
\sup(\mathcal{B}|\mathcal{A}) &= \sup\left\{ \frac{\mathcal{B}(f,f)}{\mathcal{A}(f,f)} : f \in W \right\}, \\
\inf(\mathcal{B}|\mathcal{A}) &= \inf\left\{ \frac{\mathcal{B}(f,f)}{\mathcal{A}(f,f)} : f \in W \right\}, \\
h(\mathcal{A},\mathcal{B}) &= \log\left(\frac{\sup(\mathcal{B}|\mathcal{A})}{\inf(\mathcal{B}|\mathcal{A})} \right);
\end{aligned}$$

h is Hilbert's projective metric and we have $h(\theta\mathcal{A}, \mathcal{B}) = h(\mathcal{A},\mathcal{B})$ for any $\theta \in (0,\infty)$. Note that $h(\mathcal{A},\mathcal{B}) = 0$ if and only if $\mathcal{A}$ is a nonzero constant multiple of $\mathcal{B}$.

Theorem 5.2 *There exists a constant C_F, depending only on the GSC F, such that if $\mathcal{A}, \mathcal{B} \in \mathfrak{E}$ then*

$$h(\mathcal{A}, \mathcal{B}) \leq C_F.$$

Proof. Let $\mathcal{A}' = \mathcal{A}/||\mathcal{A}||$, $\mathcal{B}' = \mathcal{B}/||\mathcal{B}||$. Then $h(\mathcal{A}, \mathcal{B}) = h(\mathcal{A}', \mathcal{B}')$. By Theorem 4.32 there exist C_i depending only on F such that (4.57) holds for both $\mathcal{A}'$ and $\mathcal{B}'$. Therefore

$$\frac{\mathcal{B}'(f,f)}{\mathcal{A}'(f,f)} \leq \frac{C_2}{C_1}, \qquad \text{for } f \in W,$$

and so $\sup(\mathcal{B}'|\mathcal{A}') \leq C_2/C_1$. Similarly, $\inf(\mathcal{B}'|\mathcal{A}') \geq C_1/C_2$, so $h(\mathcal{A}', \mathcal{B}') \leq 2\log(C_2/C_1)$. □

Proof of Theorem 1.2 By Proposition 1.1 we have that $\mathfrak{E}$ is non-empty.

Let $\mathcal{A}, \mathcal{B} \in \mathfrak{E}$, and $\lambda = \inf(\mathcal{B}|\mathcal{A})$. Let $\delta > 0$ and $\mathcal{C} = (1+\delta)\mathcal{B} - \lambda\mathcal{A}$. By Theorem 2.1, $\mathcal{C}$ is a local regular Dirichlet form on $L^2(F, \mu)$ and $\mathcal{C} \in \mathfrak{E}$. Since

$$\frac{\mathcal{C}(f,f)}{\mathcal{A}(f,f)} = (1+\delta)\frac{\mathcal{B}(f,f)}{\mathcal{A}(f,f)} - \lambda, \qquad f \in W,$$

we obtain

$$\sup(\mathcal{C}|\mathcal{A}) = (1+\delta)\sup(\mathcal{B}|\mathcal{A}) - \lambda,$$

and

$$\inf(\mathcal{C}|\mathcal{A}) = (1+\delta)\inf(\mathcal{B}|\mathcal{A}) - \lambda = \delta\lambda.$$

Hence for any $\delta > 0$,

$$e^{h(\mathcal{A},\mathcal{C})} = \frac{(1+\delta)\sup(\mathcal{B}|\mathcal{A}) - \lambda}{\delta\lambda} \geq \frac{1}{\delta}\left(e^{h(\mathcal{A},\mathcal{B})} - 1\right).$$

If $h(\mathcal{A}, \mathcal{B}) > 0$, this is not bounded as $\delta \to 0$, contradicting Theorem 5.2. We must therefore have $h(\mathcal{A}, \mathcal{B}) = 0$, which proves our theorem. □

Proof of Corollary 1.4 Note that Theorem 1.2 implies that the $\mathbb{P}^x$ law of X is uniquely defined, up to scalar multiples of the time parameter, for all $x \notin \mathcal{N}$, where $\mathcal{N}$ is a set of capacity 0. If f is continuous and X is a Feller process, the map $x \to \mathbb{E}^x f(X_t)$ is uniquely defined for all x by the continuity of $T_t f$. By a limit argument it is uniquely defined if f is bounded and measurable, and then by the Markov property, we see that the finite dimensional distributions of X under $\mathbb{P}^x$ are uniquely determined. Since X has continuous paths, the law of X under $\mathbb{P}^x$ is determined. (Recall that the the processes constructed in [5] are Feller processes.) □

Remark 5.3 In addition to (H1)-(H4), assume that the $(d-1)$-dimensional fractal $F \cap \{x_1 = 0\}$ also satisfies the conditions corresponding to (H1)-(H4). (This assumption is used in [22, Section 5.3].). Then one can show $\Gamma(f,f)(F \cap \partial F_0) = 0$ for all $f \in \mathcal{F}$ where $\Gamma(f,f)$ is the energy measure for $\mathcal{E} \in \mathfrak{E}$ and $f \in \mathcal{F}$. Indeed, by the uniqueness we know that $\mathcal{E}$ is self-similar, so the results in [22] can be applied.

For h given in [22, Proposition 3.8], we have $\Gamma(h,h)(F\cap\partial[0,1]^d)=0$ by taking $i\to\infty$ in the last inequality of [22, Proposition 3.8]. For general $f\in\mathcal{F}$, take an approximating sequence $\{g_m\}\subset\mathcal{F}$ as in the proof of Theorem 2.5 of [22]. Using the inequality

$$\begin{aligned}|\Gamma(g_m,g_m)(A)^{1/2}-\Gamma(f,f)(A)^{1/2}| &\le \Gamma(g_m-f,g_m-f)(A)^{1/2}\\ &\le 2\mathcal{E}(g_m-f,g_m-f)^{1/2},\end{aligned}$$

(see page 111 in [17]), we conclude that $\Gamma(f,f)(F\cap\partial[0,1]^d)=0$. Using the self-similarity, we can also prove that the energy measure does not charge the image of $F\cap\partial[0,1]^d$ by any of the contraction maps.

Remark 5.4 One question left over from [3, 5] is whether the sequence of approximating reflecting Brownian motions used to construct the Barlow-Bass processes converges. Let $\widetilde{X}^n_t=X^n_{c_nt}$, where X^n is defined in Subsection 3.1 and c_n is a normalizing constant. We choose c_n so that the expected time for $\widetilde{X}^n$ started at 0 to reach one of the faces not containing 0 is one. There will exist subsequences $\{n_j\}$ such that there is resolvent convergence for $\{\widetilde{X}^{n_j}\}$ and also weak convergence, starting at every point in F. Any of the subsequential limit points will have a Dirichlet form that is a constant multiple of one of the $\mathcal{E}_{BB}$. By virtue of the normalization and our uniqueness result, all the limit points are the same, and therefore the whole sequence $\{\widetilde{X}^n\}$ converges, both in the sense of resolvent convergence and in the sense of weak convergence for each starting point.

Acknowledgment. The authors thank Z.-Q. Chen, M. Fukushima, M. Hino, V. Metz, and M. Takeda for valuable discussions, and D. Croydon for correcting some typos.

References

[1] Alexander, S. and Orbach, R.: Density of states on fractals: “fractons”. *J. Physique (Paris) Lett.* **43** (1982), 625-631.

[2] Barlow, M.T.: *Diffusions on fractals.* Lectures on probability theory and statistics (Saint-Flour, 1995), 1–121, Lecture Notes in Math. **1690**, Springer, Berlin, 1998.

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

[4] Barlow, M.T., Bass, R.F.: On the resistance of the Sierpinski carpet. *Proc. R. Soc. London A.* **431** (1990) 345–360.

[5] Barlow, M.T., Bass, R.F.: Brownian motion and harmonic analysis on Sierpinski carpets. *Canad. J. Math.* **54** (1999), 673–744.

[6] Barlow, M.T., Bass, R.F., Kumagai, T.: Stability of parabolic Harnack inequalities on metric measure spaces. *J. Math. Soc. Japan* (2) **58** (2006), 485–519.

[7] Barlow, M.T., Kumagai, T.: Random walk on the incipient infinite cluster on trees. *Illinois J. Math.*, **50** (2006), 33–65.

[8] Barlow, M.T., Járai, A.A., Kumagai, T., Slade, G.: Random walk on the incipient infinite cluster for oriented percolation in high dimensions. *Commun. Math. Phys.*, **278** (2008), 385–431.

[9] Barlow, M.T., Bass, R.F., Kumagai, T.: Note on the equivalence of parabolic Harnack inequalities and heat kernel estimates.
`http://www.math.uconn.edu/~bass/papers/phidfapp.pdf`

[10] Barlow, M.T., Bass, R.F., Kumagai, T., Teplyaev, A.: Supplementary notes for "Uniqueness of Brownian motion on Sierpinski carpets".
`http://www.math.uconn.edu/~bass/papers/scuapp.pdf`

[11] Barlow, M.T., Perkins, E.A.: Brownian Motion on the Sierpinski Gasket. *Probab. Theory Rel. Fields* **79** (1988), 543–623.

[12] Ben-Avraham, D., Havlin, S.: *Diffusion and reactions in fractals and disordered systems.* Cambridge University Press, Cambridge, 2000.

[13] Bouleau, N., Hirsch, F.: *Dirichlet forms and analysis on Wiener space.* de Gruyter Studies in Mathematics, 14. Walter de Gruyter and Co., Berlin, 1991.

[14] Chen, Z.-Q.: On reflected Dirichlet spaces. *Probab. Theory Rel. Fields* **94** (1992), 135–162.

[15] Chen, Z.-Q.: On notions of harmonicity. *Proc. Amer. Math. Soc.*, to appear.

[16] Dynkin, E.B.: *Markov Processes - I.* Springer, Berlin, 1965.

[17] Fukushima, M., Oshima, Y., Takeda, M.: *Dirichlet Forms and Symmetric Markov Processes.* de Gruyter, Berlin, 1994.

[18] Fukushima, M., Shima, T.: On a spectral analysis for the Sierpiński gasket. *Potential Anal.* **1** (1992), 1–35.

[19] Goldstein, S.: Random walks and diffusions on fractals. Percolation theory and ergodic theory of infinite particle systems (Minneapolis, Minn., 1984–1985), 121–129, IMA Vol. Math. Appl. **8**, Springer, New York, 1987.

[20] Grigor'yan, A., Telcs, A.: Two-sided estimates of heat kernels in metric measure spaces. In preparation.

[21] Hambly, B.M., Metz, V., Teplyaev, A.: Admissible refinements of energy on finitely ramified fractals. *J. London Math. Soc.* **74** (2006), 93–112.

[22] Hino, M., Kumagai, T.: A trace theorem for Dirichlet forms on fractals. *J. Func. Anal.* **238** (2006), 578–611.

[23] Kesten, H.: Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincare Probab. Statist.*, **22** (1986), 425487.

[24] Kigami, J.: A harmonic calculus on the Sierpinski space. *Japan J. Appl. Math.* **6** (1989), 259–290.

[25] Kigami, J.: A harmonic calculus for p.c.f. self-similar sets. *Trans. Amer. Math. Soc.* **335** (1993), 721–755.

[26] Kigami, J.: *Analysis on fractals*. Cambridge Univ. Press, Cambridge, 2001.

[27] Kozma, G., Nachmias, A.: The Alexander-Orbach conjecture holds in high dimensions. Preprint 2008.

[28] Kumagai, T., Sturm, K.-T.: Construction of diffusion processes on fractals, d-sets, and general metric measure spaces. *J. Math. Kyoto Univ.* **45** (2005), no. 2, 307–327.

[29] Kusuoka, S.: A diffusion process on a fractal. Probabilistic methods in mathematical physics (Katata/Kyoto, 1985), 251–274, Academic Press, Boston, MA, 1987.

[30] Kusuoka, S., Zhou, X.Y.: Dirichlet forms on fractals: Poincaré constant and resistance. *Probab. Theory Related Fields* **93** (1992), no. 2, 169–196.

[31] Lindstrøm, T.: Brownian motion on nested fractals. *Mem. Amer. Math. Soc.* **420** (1990), 1–128.

[32] Mandelbrot, B.: *The Fractal Geometry of Nature*. W.H. Freeman, San Francisco, 1982.

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

[34] McGillivray, I.: Resistance in higher-dimensional Sierpiński carpets. *Potential Anal.* **16** (2002), no. 3, 289–303.

[35] Osada, H.: Singular time changes of diffusions on Sierpinski carpets. *Stochastic Process. Appl.* **116** (2006), 675–689.

[36] Peirone, R.: Convergence and uniqueness problems for Dirichlet forms on fractals. *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8)* **3** (2000), 431–460.

[37] Rammal, R., Toulouse G.: Random walks on fractal structures and percolation clusters. *J. Physique Lettres* **44** (1983) L13–L22.

[38] Reed, M., Simon, B.: *Methods of modern mathematical physics. I. Functional analysis*. Academic Press, 1980.

[39] Rogers, L.C.G., Williams, D.: *Diffusions, Markov Processes, and Martingales. Volume one: Foundations*, 2nd ed. Wiley, 1994.

[40] Rudin, W.: *Functional analysis.* McGraw-Hill, 1991.

[41] Sabot, C.: Existence and uniqueness of diffusions on finitely ramified self-similar fractals. *Ann. Sci. École Norm. Sup. (4)* **30** (1997), 605–673.

[42] Schmuland, B.: On the local property for positivity preserving coercive forms. Dirichlet forms and stochastic processes (Beijing, 1993), 345–354, de Gruyter, Berlin, 1995.

[43] Strichartz, R.S.: *Differential Equations on Fractals: a Tutorial.* Princeton University Press, Princeton, NJ, 2006.

[44] Torchinsky, A.: *Real-Variable Methods in Harmonic Analysis.* Academic Press, Orlando FL, 1986.