RANDOM WALKS AND INDUCED DIRICHLET FORMS ON SELF-SIMILAR SETS

SHI-LEI KONG, KA-SING LAU and TING-KAM LEONARD WONG

Abstract

Let K be a self-similar set satisfying the open set condition. Following Kaimanovich's elegant idea [25], it has been proved that on the symbolic space X of K a natural augmented tree structure $\mathfrak{E}$ exists; it is hyperbolic, and the hyperbolic boundary $\partial_H X$ with the Gromov metric is Hölder equivalent to K. In this paper we consider certain reversible random walks with return ratio $0 < \lambda < 1$ on $(X, \mathfrak{E})$. We show that the Martin boundary $\mathcal{M}$ can be identified with $\partial_H X$ and K. With this setup and a device of Silverstein [41], we obtain precise estimates of the Martin kernel and the Naïm kernel in terms of the Gromov product. Moreover, the Naïm kernel turns out to be a jump kernel satisfying the estimate $\Theta(\xi, \eta) \asymp |\xi - \eta|^{-(\alpha+\beta)}$, where α is the Hausdorff dimension of K and β depends on λ. For suitable β, the kernel defines a regular non-local Dirichlet form on K. This extends the results of Kigami [27] concerning random walks on certain trees with Cantor-type sets as boundaries (see also [5]).

Contents

1 Introduction

Let $\mathbb{D}$ be the open unit disk, and let $\mathbb{T}$ be the boundary circle parameterized by $\{\theta : 0 \le \theta < 2\pi\}$. Let

$$\mathcal{E}_{\mathbb{D}}(u, v) = \int_{\mathbb{D}} \nabla u(x) \nabla v(x) dx \tag{1.1}$$

be the standard Dirichlet form on $\mathbb{D}$. In classical analysis, it is well-known that a function $\varphi \in L^1(\mathbb{T})$ can be extended to a harmonic functions on $\mathbb{D}$ via the Poisson integral

$$(H\varphi)(x) = \int_{\mathbb{T}} \varphi(\theta) K(x, \theta) d\theta, \qquad x \in \mathbb{D},$$

where $K(x, \theta)$ is the Poisson kernel. Furthermore, there is an induced Dirichlet form on $\mathbb{T}$ defined by

$$\mathcal{E}_{\mathbb{T}}(\varphi, \psi) = \mathcal{E}_{\mathbb{D}}(H\varphi, H\psi).$$

Indeed, it can be shown that

$$\mathcal{E}_{\mathbb{T}}(\varphi, \psi) = \frac{1}{16\pi} \int_{\mathbb{T}} \int_{\mathbb{T}} (\varphi(\theta) - \varphi(\theta'))(\psi(\theta) - \psi(\theta')) \, \frac{1}{\sin^2(\frac{\theta - \theta'}{2})} \, d\theta d\theta'. \tag{1.2}$$

This integral is called the *Douglas integral* (see [14, Section 1.2]). From the probabilistic point of view, the Dirichlet form in (1.1) is associated with a Brownian motion on $\mathbb{D}$. The hitting distribution of the Brownian motion at the boundary $\mathbb{T}$ (starting from 0) is the uniform distribution $\frac{d\theta}{2\pi}$; the induced Dirichlet form in (1.2) corresponds to the reflecting Brownian motion on $\overline{\mathbb{D}}$ time-changed by its local time on $\mathbb{T}$, and defines a jump process on $\mathbb{T}$ which is a Cauchy process [6].

The above consideration has a counterpart in Markov chain theory. Let $\{Z_n\}_{n=0}^{\infty}$ be a transient Markov chain on an infinite discrete set X with transition probability P. According to the discrete potential theory of Markov chains [12, 44, 45], a chain starting at the reference point ϑ will converge to the *Martin boundary* $\mathcal{M}$ at infinity, and defines a hitting distribution $\nu = \nu_\vartheta$ on $\mathcal{M}$. Also, there is a Martin kernel $K(x, \xi)$, $x \in X, \xi \in \mathcal{M}$, which plays the same role as the Poisson kernel: if we define

$$(H\varphi)(x) = \int_{\mathcal{M}} \varphi(\xi) K(x, \xi) d\nu(\xi), \qquad x \in X,$$

then $u = H\varphi$ is harmonic on X, i.e., $u = Pu$. We call a Markov chain *reversible* if the transition probability is of the form $P(x, y) = c(x, y)/m(x)$, where $c(x, y) = c(y, x) \ge 0$, and $m(x) = \sum_{y \in X} c(x, y)$. We define a graph energy $\mathcal{E}_X$ on X by

$$\mathcal{E}_X[u] = \frac{1}{2} \sum_{x, y \in X} c(x, y) |u(x) - u(y)|^2. \tag{1.3}$$

In [41], Silverstein showed that for such Markov chain, there is an energy form $\mathcal{E}_{\mathcal{M}}$ on $\mathcal{M}$ that satisfies

$$\mathcal{E}_{\mathcal{M}}[\varphi] = \mathcal{E}_X[H\varphi],$$

and $\mathcal{E}_{\mathcal{M}}$ has the expression

$$\mathcal{E}_{\mathcal{M}}[\varphi] = \int_{\mathcal{M}} \int_{\mathcal{M}} |\varphi(\xi) - \varphi(\eta)|^2 \, \Theta(\xi, \eta) \, d\nu(\xi) d\nu(\eta). \tag{1.4}$$

2

Here $\Theta(\cdot, \cdot)$ is called the *Naïm's Θ-kernel* (or simply, the *Naïm kernel*). It was first introduced in classical potential theory by Naïm [36], and a general Douglas integral formula (corresponding to (1.2)) on Euclidean domain was proved by Doob [11]. Recently, Georgakopoulos introduced a class of "group-walk random graphs", and he outlined a study of the Poisson boundary and the Naïm kernel using electrical network theory in [15].

The domain of $\mathcal{E}_{\mathcal{M}}$ in (1.4) consists of square integrable functions φ such that $\mathcal{E}_{\mathcal{M}}[\varphi] < \infty$. If the domain is dense in $L^2(\mathcal{M}, \nu)$, then $\mathcal{E}_{\mathcal{M}}$ defines a non-local Dirichlet form. In the analysis of fractals, there is a large literature on the study of local and non-local Dirichlet forms as well as their associated heat kernels on self-similar sets, d-sets and more general metric measure spaces [6,7,16–18,22,26,27,38,39,42,43]. In many cases, a non-local form can be obtained by subordination of a local form, and has a jump kernel with order $|\xi - \eta|^{-(\alpha+\beta)}$, where α is the Hausdorff dimension of the underlying set and β is the walk dimension of the corresponding stable-like process [7, 42].

For a self-similar set K generated by an iterated function system (IFS), there is a symbolic space (coding space) Σ^* which gives a convenient representation of any $\xi \in K$ (analogous to the dyadic expansion of a real number). If the IFS satisfies the *open set condition* (OSC), then the representation is unique for generic points of K. Recently, there are studies of random walks on Σ^* such that K can be identified with the Martin boundary $\mathcal{M}$ under the canonical homeomorphism [8–10, 24, 25, 31, 32, 34]. In particular, Kaimanovich [25] used the Sierpiński gasket to introduce a natural *augmented tree* (Sierpiński graph) by adding horizontal edges on Σ^* according to the intersections of the cells from the IFS. This work brings into play the hyperbolic structure and hyperbolic boundary which are powerful tool for studying random walks on the graph. Kaimanovich's augmented tree was extended to general self-similar sets in [33, 35].

In this paper we will study random walks on the augmented trees $(\Sigma^*, \mathfrak{E})$ and their induced Dirichlet forms on K. We investigate a class of reversible random walks on $(\Sigma^*, \mathfrak{E})$ so that

(i) the self-similar set K can be identified with the Martin boundary, and

(ii) the hitting distribution $\nu = \nu_\vartheta$ is the normalized Hausdorff measure on K, and the Naïm kernel is of order $|\xi - \eta|^{-(\alpha+\beta)}$.

As a special case, Kigami [27] studied reversible random walks on trees where the Martin boundaries are Cantor-type sets. He used the resistance metric to give explicit expressions of the hitting distribution ν and the Martin kernel $K(x, y)$. Also, under the volume doubling property of ν with respect to the resistance metric, he studied the associated jump process and estimates of the heat kernel. These results were extended to the non-compact case in [28]. Recently, a duality between reversible random walks on trees and certain jump processes on the boundaries was studied in detail in [5].

Let $\{S_i\}_{i=1}^N$ denote an IFS of contractive similitudes, and let K denote the self-similar set generated. For simplicity, here we only state the case where the IFS is *homogenous* (all the maps S_i have the same contraction ratio r). Following standard notations, we let $\Sigma = \{1, 2, \ldots, N\}$, $\Sigma^n = \{\mathbf{x} = i_1 \cdots i_n : i_1, \cdots, i_n \in \Sigma\}$, $n \geq 1$ (by convention $\Sigma^0 = \{\vartheta\}$)

and use $|\mathbf{x}|$ to denote the length of $\mathbf{x}$; also we let $\Sigma^* := \bigcup_{n=0}^{\infty} \Sigma^n$, the set of finite words, and $\Sigma^{\infty} := \{i_1 i_2 \cdots : i_1, i_2, \cdots \in \Sigma\}$, the set of infinite words. There is a natural surjection $\kappa : \Sigma^{\infty} \to K$ defined by

$$\{\kappa(\omega)\} = \bigcap_{m \geq 0} S_{i_1 i_2 \cdots i_m}(K), \qquad \omega = i_1 i_2 \cdots \in \Sigma^{\infty},$$

where $S_{i_1 i_2 \cdots i_m} = S_{i_1} \circ S_{i_2} \circ \cdots \circ S_{i_m}$. Hence each $\xi \in K$ admits a symbolic representation (coding) $\omega \in \Sigma^{\infty}$.

The symbolic space Σ^* has a natural tree structure $\mathfrak{E}_v$ which is referred to as the set of "vertical edges". To consider (i) above, we enrich this simple graph by adding more "horizontal edges", denoted by $\mathfrak{E}_h$, to reflect the behaviors of the neighboring cells of K at each level:

$$(\mathbf{x}, \mathbf{y}) \in \mathfrak{E}_h \text{ if } \mathbf{x} \neq \mathbf{y}, \ |\mathbf{x}| = |\mathbf{y}| \text{ and } \inf_{\xi, \eta \in K} |S_{\mathbf{x}}(\xi) - S_{\mathbf{y}}(\eta)| \leq \gamma \cdot r^{|\mathbf{x}|}, \tag{1.5}$$

where $\gamma > 0$ is any fixed constant (see Definition (3.4)). We write $X_n = \bigcup_{j=1}^{n} \Sigma^j$, and $X = \Sigma^*$ for brevity, and let $\mathfrak{E} = \mathfrak{E}_v \cup \mathfrak{E}_h$. We call $(X, \mathfrak{E})$ the *augmented tree* of K. This notion was invented by Kaimanovich in [25], where the last condition in (1.5) was replaced by $S_{\mathbf{x}}(K) \cap S_{\mathbf{y}}(K) \neq \emptyset$. For general IFS (even with the OSC), there are certain technical difficulties to identify K with the graph at infinity [LW1]; the modification in (1.5) is from [LW3] which avoids the superfluous conditions in [LW1].

Recall that for a hyperbolic graph X, there is a *Gromov product* $(x|y)$ for $x, y \in X$, a *Gromov metric* $\varrho_a(x, y) = e^{-a(x|y)}$ on X, and a *hyperbolic boundary* $\partial_H X$ with $\varrho_a(\cdot, \cdot)$ extended to $\partial_H X$ (see Section 2 below for more details).

Theorem 1.1. [25, 33, 35] *For any IFS the augmented tree $(X, \mathfrak{E})$ is hyperbolic. Moreover, there is a canonical identification $\iota : \partial_H X \to K$ (independent of $\gamma > 0$), such that K and $\partial_H X$ are Hölder equivalent, i.e., $\varrho_a(\xi, \eta) \asymp |\iota(\xi) - \iota(\eta)|^{-a/\log r}$.*

For a homogenous IFS $\{S_i\}_{i=1}^{N}$ satisfying the open set condition (OSC), the self-similar set has Hausdorff dimension $\alpha = \log N / |\log r|$. For $0 < \lambda < 1$, we define a class of reversible random walks on $(X, \mathfrak{E})$ with conductance

$$c(\mathbf{x}, \mathbf{x}^-) = (\lambda^{-1} r^{\alpha})^{|\mathbf{x}|}, \qquad \mathbf{x} \in X.$$

and $c(\mathbf{x}, \mathbf{y}) \asymp c(\mathbf{x}, \mathbf{x}^-)$ for $(\mathbf{x}, \mathbf{y}) \in \mathfrak{E}_h$ where $\mathbf{x}^-$ is the parent of $\mathbf{x}$ (here $\asymp$ means the two terms dominate each other by two constants C_1, C_2 independent of the variables), and call it a λ-*natural random walk* (λ-NRW) (see Definition 4.4). Here r^{α} corresponds to the natural weight $p_i = r^{\alpha}$ of the IFS $\{S_i\}_{i=1}^{N}$ which generates, as a self-similar measure, the normalized Hausdorff measure $\mathcal{H}^{\alpha}$ on K; the parameter λ is the ratio of the probabilities for the walk to go upward or downward at each $\mathbf{x}$ (see (4.1) and Section 4). Note that a simple random walk (SRW) has transition probability

$$P(\mathbf{x}, \mathbf{y}) = \frac{1}{\deg(\mathbf{x})} \qquad \text{for} \quad (x, y) \in \mathfrak{E},$$

4

where $\deg(\mathbf{x})$ is the number of edges joining $\mathbf{x}$. It is easy to see that a SRW is a $1/N$-NRW with $c(\mathbf{x}, \mathbf{x}^-) = c(\mathbf{x}, \mathbf{y}) = 1$ for $(\mathbf{x}, \mathbf{y}) \in \mathfrak{E}_h$. Using Theorem 1.1 and a well-known result of Ancona [1, 2, 44] on uniformly irreducible random walks on hyperbolic graphs, we prove the following theorem (Theorem 5.1) which extends [25, Theorem 4.7] for the Sierpiński graph with the simple random walk.

Theorem 1.2. *Suppose in addition, the IFS satisfies the OSC, let $\{Z_n\}_{n=0}^{\infty}$ be a λ-NRW on $(X, \mathfrak{E})$. Then the hyperbolic boundary $\partial_H X$, the Martin boundary $\mathcal{M}$ and the self-similar set K are all homeomorphic under the canonical mapping.*

The above theorem allows us to identify K with the two boundaries. More importantly, we are able to use the hyperbolic structure to give sharp estimates of the Martin and Naïm kernels for the λ-NRW on $(X, \mathfrak{E})$.

Theorem 1.3. *The hitting distribution $\nu = \nu_\vartheta$ of the λ-NRW on $(X, \mathfrak{E})$ is the normalized Hausdorff measure $\mathcal{H}^\alpha$ on K, and the Martin kernel satisfies the estimate*

$$K(\mathbf{x}, \xi) \asymp \lambda^{|\mathbf{x}| - (\mathbf{x}|\xi)} r^{-\alpha(\mathbf{x}|\xi)}, \qquad \mathbf{x} \in X, \ \ \xi \in K.$$

To derive the hitting distribution, a main part is to show that $F_n(\vartheta, \mathbf{x}) := \mathbb{P}_\vartheta(Z_{\tau_n} = \mathbf{x})$ equals $r^{\alpha n}$ for $|\mathbf{x}| = n$, where τ_n is the first hitting time at level n (Theorem 4.6). The result then follows from a limiting argument (Theorem 5.6). Recall that the Martin kernel is given by

$$K(\mathbf{x}, \mathbf{y}) = \frac{G(\mathbf{x}, \mathbf{y})}{G(\vartheta, \mathbf{y})} = \frac{F(\mathbf{x}, \mathbf{y})}{F(\vartheta, \mathbf{y})}, \qquad \mathbf{x}, \mathbf{y} \in X,$$

where $G(\mathbf{x}, \mathbf{y})$ is the Green function, and $F(\mathbf{x}, \mathbf{y})$ is the probability of the chain ever visiting $\mathbf{y}$ from $\mathbf{x}$. The estimates of $F(\mathbf{x}, \vartheta)$ (Proposition 4.1), $F(\vartheta, \mathbf{x})$ (Theorem 4.6) and $F(\mathbf{x}, \mathbf{y})$ (Theorem 5.3) are the core of the proof, they need substantial use of the reversibility of the chain as well as the hyperbolicity of $(X, \mathfrak{E})$.

Following Silverstein [41], the Naïm kernel is defined by

$$\Theta(\mathbf{x}, \mathbf{y}) = \frac{K(\mathbf{x}, \mathbf{y})}{G(\mathbf{x}, \vartheta)} = \frac{K(\mathbf{x}, \mathbf{y})}{F(\mathbf{x}, \theta) G(\vartheta, \vartheta)}.$$

It is the discrete analogue of the kernel studied by [36] in classical potential theory. It is easy to extend $\Theta(\mathbf{x}, \mathbf{y})$ to $\Theta(\mathbf{x}, \eta)$ for $\eta \in K$, but the extension to $\Theta(\xi, \eta)$ for $\xi \in K$ is much more involved. In [36], the extension involves Cartan's fine topology; in Markov chain theory, Silverstein [41] proved the identity

$$\Theta(\xi, \eta) = \lim_{k \to \infty} \sum_{\mathbf{z} \in X} \ell_k^\xi(\mathbf{z}) \Theta(\mathbf{z}, \eta), \quad \xi, \eta \in K,$$

where $\ell_k^\xi(\mathbf{z})$ is the probability for the ξ-process to last exit the kth-level at $\mathbf{z}$ (see (6.1), (6.2)). Analyzing the above limit, we have (Theorem 6.3)

Theorem 1.4. *For the λ-NRW on $(X, \mathfrak{E})$ with the energy form as in (1.3), the Naïm kernel is asymptotically given by*

$$\Theta(\xi, \eta) \asymp (\lambda r^\alpha)^{-(\xi|\eta)}, \quad \xi, \eta \in K.$$

5

Consequently by Theorem 1.1 and (1.4)*, the induced energy form satisfies*

$$\mathcal{E}_K[\varphi] \asymp \int_K \int_K |\varphi(\xi) - \varphi(\eta)|^2 \, |\xi - \eta|^{-(\alpha+\beta)} d\nu(\xi) d\nu(\eta)$$

with $\beta = \log \lambda / \log r$.

The domain $\mathcal{D}_K$ of $\mathcal{E}_K$ consists of all functions in $L^2(K, \nu)$ that possess finite $\mathcal{E}_K$-energies, and equals the Besov space $\Lambda_{2,2}^{\alpha,\beta/2}$ on K [42]. Let β^* be the critical exponent such that $\Lambda_{2,2}^{\alpha,\beta/2} \cap C(K)$ is dense in $C(K)$ for all $0 < \beta < \beta^*$, then the pair $(\mathcal{E}_K, \mathcal{D}_K)$ with $\beta \in (0, \beta^*)$ is a non-local Dirichlet form. There are certain values of β for which the Dirichlet forms are known to be regular [7, 16, 17, 42], and there are some standard examples where the values of β^* can be determined explicitly [22, 30]. These issues will be discussed in Section 7, and more detail in the forthcoming paper [29]. For a more recent development, Grigor'yan and Yang [19] reconstructed the corresponding reflective process of the λ-NRW on the Sierpiński graph and its boundary analogous to the Brownian motion on the closed unit disk, and use this device and Theorem 1.4 to study the critical exponent β^*.

In the text the theorems are presented in greater generality. We will define a *pre-augmented tree* on an N-ary tree that has the hyperbolic property (Definition 3.2). Many results for the random walks will hold on such graphs. For an IFS $\{S_i\}_{i=1}^N$ that is not homogeneous, we need to re-adjust the levels Σ^n as in (3.2) and restrict to the natural weights $p_i = r_i^\alpha$, where r_i is the contraction ratio of S_i, and α satisfies $\sum_{i=1}^N r_i^\alpha = 1$. Furthermore, we can set up a more general λ-*quasi-natural random walk* (λ-QNRW) (Definition 4.4), which corresponds to weights $\{p_i\}_{i=1}^N$, that gives a doubling self-similar measure, and Theorems 1.2 and 1.3 are still valid (the hitting distribution becomes the self-similar measure defined by the weights $\{p_i\}_{i=1}^N$).

For the organization of the paper, we first recall some basic facts about hyperbolic graphs and the potential theory of Markov chains in Section 2. In Section 3 we introduce the augmented tree $(X, \mathfrak{E})$ on the symbolic space and study its hyperbolic structure. In Section 4, we study the classes of λ-QNRW and λ-NRW and estimate the associated probabilities. Theorems 1.2 and 1.3 are proved in Section 5. After a brief review of the Naïm kernel, we prove Theorem 1.4 in Section 6. Finally, the induced Dirichlet forms will be discussed in Section 7.

2   Preliminaries

This section presents a quick summary of hyperbolic graphs and the potential theory of Markov chains we need throughout the paper. For more details, the reader can refer to the excellent monograph [44] by Woess.

Let X be a countable set. An *(undirected simple) graph* is a pair $(X, \mathfrak{E})$ where $\mathfrak{E}$ is a symmetric subset of $X \times X \setminus \{(x, x) : x \in X\}$. We call $x \in X$ a vertex, and $(x, y) \in \mathfrak{E}$ an edge (denoted by $x \sim y$). For $x \in X$, the *degree* of x, denoted by $\deg(x)$, is the number of edges joining x. We say that $(X, \mathfrak{E})$ is of *bounded degree* if $\sup_{x \in X} \deg(x) < \infty$. For $x, y \in X$, a *path* from x to y with *length* n is a finite sequence $[x_0, x_1, \ldots, x_n]$ such that $x_0 = x, x_n = y$,

and $(x_i, x_{i+1}) \in \mathfrak{E}$ for all $0 \le i \le n-1$. We always assume that $(X, \mathfrak{E})$ is *connected*, i.e., any two vertices can be connected by a path. A graph is called a *tree* if any two points can be connected by a unique non-self-intersecting path.

We use $\pi(x, y)$ to denote a path connecting two vertices x and y with minimum length, and call it a *geodesic path*. The length of $\pi(x, y)$ defines a metric $d(x, y)$ on X. Let ϑ be any fixed vertex in X regarded as the *root*. We define the *Gromov product* by

$$(x|y) = \frac{1}{2}(|x| + |y| - d(x, y)), \qquad x, y \in X, \tag{2.1}$$

where $|x| = d(\vartheta, x)$.

Definition 2.1. *We say that $(X, \mathfrak{E})$ is hyperbolic if there exists $\delta \ge 0$ such that*

$$(x|y) \ge \min\{(x|z), (z|y)\} - \delta, \qquad \forall\, x, y, z \in X. \tag{2.2}$$

Clearly, every tree is hyperbolic with $\delta = 0$. For a hyperbolic graph $(X, \mathfrak{E})$, we choose $a > 0$ satisfying $e^{\delta a} < \sqrt{2}$, and define

$$\varrho_a(x, y) = e^{-a(x|y)}, \ x \ne y, \quad \text{and} \quad = 0 \ \text{if } x = y. \tag{2.3}$$

This $\varrho_a(\cdot, \cdot)$ on $X \times X$ is equivalent to a metric, and is called *Gromov metric* for convenience.

Definition 2.2. *Let $\widehat{X}_H$ denote the completion of (X, ϱ_a). The hyperbolic boundary of $(X, \mathfrak{E})$ is defined as $\partial_H X = \widehat{X}_H \setminus X$.*

The hyperbolic boundary $\partial_H X$ is compact and can be regarded as the collection of all (infinite) geodesic rays starting at ϑ, where two rays $(x_n)_n$ and $(y_n)_n$ are identified if $\lim_{n \to \infty}(x_n|y_n) = \infty$.

Next we consider random walks on a graph $(X, \mathfrak{E})$. Let $\{Z_n\}_{n=0}^{\infty}$ be a Markov chain on X with transition probability P. We write $\mathbb{P}(\cdot \mid Z_0 = x) = \mathbb{P}_x(\cdot)$ for short. A function u on X is said to be *harmonic* (with respect to P) if $Pu = u$ where $Pu(x) = \sum_{y \in X} P(x, y)u(y) = \mathbb{E}_x u(Z_1)$. The *Green function* with respect to P is $G(x, y) := \sum_{k=0}^{\infty} P^k(x, y)$. We always assume that the Markov chain is irreducible and transient, i.e., $0 < G(x, y) < \infty$ for any $x, y \in X$. We define $F(x, y)$ as the probability that the chain starting from x ever visits y, i.e.,

$$F(x, y) := \mathbb{P}_x(\exists\, n \ge 0 \text{ such that } Z_n = y). \tag{2.4}$$

It is clear that $G(x, y) = F(x, y)G(y, y)$.

Fix a reference point $\vartheta \in X$. We define the *Martin kernel* by

$$K(x, y) := \frac{G(x, y)}{G(\vartheta, y)} = \frac{F(x, y)}{F(\vartheta, y)}, \quad x, y \in X.$$

The following definition is taken from [44].

Definition 2.3. *The Martin compactification of (X, P) is the minimal compactification $\widehat{X}$ of X such that the Martin kernel $K(x, \cdot)$ extends continuously to $\widehat{X}$ for all $x \in X$. The set $\mathcal{M} = \widehat{X} \setminus X$ is called the Martin boundary of (X, P).*

It is known that under the Martin topology, $\{Z_n\}_{n=0}^{\infty}$ converges almost surely to a $\mathcal{M}$-valued random variable Z_{∞}. For $x \in X$, we define the hitting distribution ν_x to be the distribution of Z_{∞} on $\mathcal{M}$ under $\mathbb{P}_x$. The measure ν_x is absolutely continuous with respect to ν_{ϑ}, and the Radon-Nikodym density is $d\nu_x/d\nu_{\vartheta} = K(x, \cdot)$. For $u \in L^1(\mathcal{M}, \nu_{\vartheta})$, we define its *Poisson integral* by

$$Hu(x) = \mathbb{E}_x u(Z_{\infty}) = \int_{\mathcal{M}} K(x, \xi) u(\xi) d\nu_{\vartheta}(\xi), \quad x \in X.$$

Clearly Hu is harmonic on X since $K(\cdot, \xi)$ is harmonic for all $\xi \in \mathcal{M}$. A nonnegative harmonic function h on X is called *minimal* if $h(\vartheta) = 1$ and if $h' \leq h$ such that h' is nonnegative harmonic functions on X, then h'/h is constant. We define the *minimal Martin boundary* as $\mathcal{M}_{\min} = \{\xi \in \mathcal{M} : K(\cdot, \xi) \text{ is minimal}\}$. Then $Z_{\infty} \in \mathcal{M}_{\min}$ with probability 1, and ν_{ϑ} is supported on $\mathcal{M}_{\min}$. We will make use of $\mathcal{M}_{\min}$ in Section 6 (Lemma 6.1).

We say that P has *bounded range* if $\sup\{d(x, y) : P(x, y) > 0 \text{ for some } x, y \in X\} < \infty$, and is *uniformly irreducible* if there exist $\epsilon > 0$ and k_0 such that for any $x \sim y$, there exist $k \leq k_0$ with $P^k(x, y) \geq \epsilon$. The *spectral radius* of P is

$$r(P) = \limsup_{n \to \infty} (P^n(x, y))^{1/n} \in (0, 1], \quad x, y \in X.$$

(Note that the limsup is independent of x and y [44].) The following important result is due to Ancona [1, 2], and the specific version we use is taken from [44, Theorem 27.1].

Theorem 2.4. (Ancona) *Suppose $(X, \mathfrak{E})$ is a hyperbolic graph, P is uniformly irreducible with bounded range, and $r(P) < 1$. Then for any $\delta \geq 0$, there is a constant $C_{\delta} \geq 1$ such that for any $x, y \in X$ and u within distance δ from a geodesic segment between x and y,*

$$F(x, u)F(u, y) \leq F(x, y) \leq C_{\delta} F(x, u)F(u, y). \tag{2.5}$$

Moreover, the Martin boundary equals the minimal Martin boundary, and is homeomorphic to the hyperbolic boundary.

We focus on the class of (transient) *reversible random walks* on $(X, \mathfrak{E})$ where the transition function P satisfies

$$P(x, y) = \frac{c(x, y)}{m(x)}, \quad x \sim y, \tag{2.6}$$

and is 0 elsewhere. Here $c(x, y) = c(y, x) > 0$ for each pair $(x, y) \in \mathfrak{E}$. We call $c(x, y)$ the *conductance* of the edge (x, y), and $m(x) = \sum_{y \in X, y \sim x} c(x, y) > 0$ the *total conductance* at x. To apply Theorems 2.4, we need to check the condition $r(P) < 1$. Here we cite a geometric characterization of this condition in [44, Theorem 10.3] which will be used in Section 5. For $A \subset X$, let

$$\partial A = \{(x, y) \in \mathfrak{E} : x \in A, \ y \notin A\}$$

be the *boundary* of A and let $c(\partial A) = \sum_{(x,y) \in \partial A} c(x, y)$ (analogous to the surface area of A). We also view $m(A) = \sum_{x \in A} m(x)$ as the "volume" of A.

Proposition 2.5. *Suppose P is a reversible random walk on $(X, \mathfrak{E})$. Then $r(P) < 1$ if and only if (X, P) satisfies the strong isoperimetric inequality: there exists $\eta > 0$ such that*

$$m(A) \leq \eta c(\partial A)$$

for all finite subsets $A \subset X$.

3 Self-similar sets and augmented trees

Let $\{S_i\}_{i=1}^N$, $N \geq 2$, be an *iterated function system (IFS)* of contractive similitudes on $\mathbb{R}^d$ where S_i has contraction ratios $r_i \in (0,1)$. The *self-similar set K* of the IFS is the unique non-empty compact set in $\mathbb{R}^d$ satisfying

$$K = \bigcup_{i=1}^N S_i(K).$$

For a set of positive probability weights $\{p_i\}_{i=1}^N$, there is a unique self-similar measure μ satisfying the identity

$$\mu(\cdot) = \sum_{i=1}^N p_i \mu(S_i^{-1}(\cdot)). \tag{3.1}$$

The IFS is said to satisfy the *open set condition* (OSC) if there exists a nonempty bounded open set U such that $S_i(U) \subset U$, and $S_i(U) \cap S_j(U) = \emptyset$ for $i \neq j$. The OSC is the most basic condition imposed on the IFS. It is known that under this condition, the Hausdorff dimension of K, denoted by $\dim_H(K)$, equals α where $\sum_{i=1}^N r_i^\alpha = 1$. Furthermore, if we take $p_i = r_i^\alpha, i = 1, \cdots, N$, then the self-similar measure is the normalized α-Hausdorff measure on K. We call such weights $\{p_i\}_{i=1}^N$ the *natural weight* of the IFS.

The IFS gives rise to a symbolic space. Let Σ^* and Σ^∞ be respectively the sets of finite indices (words) and infinite indices as in Section 1. For $\mathbf{x} = i_1 \cdots i_n \in \Sigma^*$, we let $S_{\mathbf{x}}$ be the composition $S_{\mathbf{x}} = S_{i_1} \circ \cdots \circ S_{i_n}$, and the contraction ratio of $S_{\mathbf{x}}$ is denoted by $r_{\mathbf{x}} = r_{i_1} \cdots r_{i_n}$. There exists a natural surjection $\kappa : \Sigma^\infty \to K$ defined by

$$\{\kappa(\omega)\} = \bigcap_{m \geq 0} S_{i_1 i_2 \cdots i_m}(K), \qquad \omega = i_1 i_2 \cdots \in \Sigma^\infty.$$

Hence each $\xi \in K$ admits a symbolic representation (coding) $\omega \in \Sigma^\infty$. Under the OSC, the representation is unique, except a μ-null set, for any self-similar measure μ.

The finite word space Σ^* has a natural tree structure where $\mathbf{x} \sim \mathbf{y}$ if $\mathbf{y} = \mathbf{x}i$ or $\mathbf{x} = \mathbf{y}i$ for some $i \in \Sigma$. Here the empty word ϑ is the root. The canonical metric on Σ^* is $\rho_r(\mathbf{x}, \mathbf{y}) = r^{\mathbf{x} \wedge \mathbf{y}}$ $(0 < r < 1)$, where $\mathbf{x} \wedge \mathbf{y} = \min\{k : i_{k+1} \neq j_{k+1}\}$. Note that $\mathbf{x} \wedge \mathbf{y}$ coincides with the Gromov product, and ρ_r is a visual metric as in (2.3); the (hyperbolic) boundary is Σ^∞, and is a Cantor-type set. The symbolic space is a convenient tool to study the self-similar set K, but obviously it misses many properties of K. Following Kaimanovich's idea [25], we will define an "augmented tree" by adding more edges to Σ^*.

First, to deal with an IFS with different contraction ratios, we modify the symbolic space by grouping together words that have approximately the same contraction ratios. Let $r = \min\{r_i : 1 \leq i \leq N\}$. Define, for $n \geq 1$,

$$\mathcal{J}_n = \{\mathbf{x} = i_1 \cdots i_k \in \Sigma^* : r_{\mathbf{x}} \leq r^n < r_{i_1 \cdots i_{k-1}}\}, \tag{3.2}$$

and $\mathcal{J}_0 = \{\vartheta\}$ by convention. Clearly, $\mathcal{J}_n \cap \mathcal{J}_m = \emptyset$ for $n \neq m$. We define the *modified symbolic space* as $X = \bigcup_{k=0}^{\infty} \mathcal{J}_k$. We write $|\mathbf{x}| = n$ if $\mathbf{x} \in \mathcal{J}_n$, and use $\mathbf{x}^-$ to denote the *parent* of $\mathbf{x} \in X$, i.e., the unique word in X such that $|\mathbf{x}^-| = |\mathbf{x}| - 1$. We also define $\mathbf{x}^{-k}, k \geq 2$ for the unique word in X such that $|\mathbf{x}^{-k}| = |\mathbf{x}| - k$ and $\mathbf{x} = \mathbf{x}^{-k}\mathbf{z}$ for some $\mathbf{z} \in \Sigma^*$. In particular, if the IFS $\{S_i\}_{i=1}^N$ has equal contraction ratio (*homogeneous* IFS), then $\mathcal{J}_n = \Sigma^n$ and $\mathbf{x}^{-k}$ is just $\mathbf{x}$ with the last k alphabets deleted. The following basic lemma is known.

Lemma 3.1. *Suppose the OSC holds, and let α be the Hausdorff dimension of K. Then*

(i) *if μ is the self-similar measure with natural weights $p_i = r_i^\alpha$, then for $\mathbf{x} \in \mathcal{J}_n$, $p_{\mathbf{x}} = \mu(S_{\mathbf{x}}(K)) \asymp r^{\alpha n}$;*

(ii) $r^{-\alpha n} \leq \#\mathcal{J}_n < r^{-\alpha(n+1)}$.

Proof. Note that $p_{\mathbf{x}} = r_{\mathbf{x}}^\alpha$, and by definition $r_{\mathbf{x}} \in (r^{n+1}, r^n]$ for $\mathbf{x} \in \mathcal{J}_n$. Also the OSC implies that $p_{\mathbf{x}} = \mu(S_{\mathbf{x}}(K))$, and hence (i) follows. Observe that

$$1 = \mu(K) = \sum\nolimits_{\mathbf{x} \in \mathcal{J}_n} \mu(S_{\mathbf{x}}(K)) = \sum\nolimits_{\mathbf{x} \in \mathcal{J}_n} r_{\mathbf{x}}^\alpha.$$

As $r_{\mathbf{x}} \in (r^{n+1}, r^n]$, we have $(\#\mathcal{J}_n)r^{\alpha(n+1)} < 1 \leq (\#\mathcal{J}_n)r^{\alpha n}$. This implies (ii). $\square$

To introduce a graph structure on X, we let

$$\mathfrak{E}_v = \{(\mathbf{x}, \mathbf{y}) \in X \times X : \ \mathbf{x} = \mathbf{y}^- \text{ or } \mathbf{y} = \mathbf{x}^-\},$$

be the set of edges of the original tree structure on X, and use the notation $\mathbf{x} \sim_v \mathbf{y}$ for $(\mathbf{x}, \mathbf{y}) \in \mathfrak{E}_v$. We call $\mathfrak{E}_v$ the set of *vertical edges* of X. Now we augment the tree by adding more *horizontal edges*.

Definition 3.2. *Let X be the modified symbolic space associated with $\{S_i\}_{i=1}^N$. We call $(X, \mathfrak{E})$ a pre-augmented tree if $\mathfrak{E} = \mathfrak{E}_v \cup \mathfrak{E}_h$, where $\mathfrak{E}_h$ satisfies the condition*

$$(\mathbf{x}, \mathbf{y}) \in \mathfrak{E}_h \Rightarrow |\mathbf{x}| = |\mathbf{y}|, \ \text{with } \mathbf{x}^- = \mathbf{y}^- \text{ or } (\mathbf{x}^-, \mathbf{y}^-) \in \mathfrak{E}_h. \tag{3.3}$$

We write $\mathbf{x} \sim_h \mathbf{y}$ if $(\mathbf{x}, \mathbf{y}) \in \mathfrak{E}_h$.

For any vertices $\mathbf{x}, \mathbf{y}$ in $(X, \mathfrak{E})$, we say that the geodesic $\pi(\mathbf{x}, \mathbf{y})$ is a *horizontal geodesic* if it consists of horizontal edges only, and define *vertical geodesics* analogously. The geodesic between two vertices $\mathbf{x}, \mathbf{y}$ may not be unique (see Figure 1), but there is always a *canonical geodesic* such that there exist $\mathbf{u}, \mathbf{v} \in \pi(\mathbf{x}, \mathbf{y})$ with

(i) $\pi(\mathbf{x}, \mathbf{y}) = \pi(\mathbf{x}, \mathbf{u}) \cup \pi(\mathbf{u}, \mathbf{v}) \cup \pi(\mathbf{v}, \mathbf{y})$, where $\pi(\mathbf{u}, \mathbf{v})$ is a horizontal geodesic and $\pi(\mathbf{x}, \mathbf{u})$, $\pi(\mathbf{v}, \mathbf{y})$ are vertical geodesics.

(ii) for any geodesic $\pi'(\mathbf{x}, \mathbf{y})$, $\text{dist}(\vartheta, \pi(\mathbf{x}, \mathbf{y})) \leq \text{dist}(\vartheta, \pi'(\mathbf{x}, \mathbf{y}))$.

If we let $\ell = \text{dist}(\vartheta, \pi(\mathbf{u}, \mathbf{v}))$, and $h = d(\mathbf{u}, \mathbf{v})$, then the canonical geodesic has a simple geometric interpretation for the Gromov product (2.1),

$$(\mathbf{x}|\mathbf{y}) = \ell - \frac{1}{2}h. \tag{3.4}$$

We will use this expression frequently. As a consequence we have (see also [25, Theorem 3.13] for the "no big squares" condition)
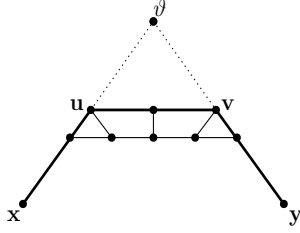
10

Figure 1: Canonical geodesic

Proposition 3.3. [33, Theorem 2.3] *A pre-augmented tree* $(X, \mathfrak{E})$ *is hyperbolic if and only if there exists* $M > 0$ *such that the lengths of all horizontal geodesics are bounded by* M.

The concept of pre-augmented tree is rather flexible (see Remark 2 in the following). For our purpose, we will use the following specific horizontal edge set.

Definition 3.4. [35] *For* $\gamma > 0$, *we define a horizontal edge set* $\mathfrak{E}_h^\gamma$ *on* X *by*

$$\mathfrak{E}_h^\gamma = \{(\mathbf{x}, \mathbf{y}) \in X \times X : \ \mathbf{x} \neq \mathbf{y}, \ |\mathbf{x}| = |\mathbf{y}| \ and \ \inf_{\xi, \eta \in K} |S_\mathbf{x}(\xi) - S_\mathbf{y}(\eta)| \leq \gamma \cdot r^{|\mathbf{x}|}\}. \qquad (3.5)$$

The augmented tree $(X, \mathfrak{E}^\gamma)$ *is the graph with edge set* $\mathfrak{E}^\gamma = \mathfrak{E}_v \cup \mathfrak{E}_h^\gamma$.

It is direct to check that $(X, \mathfrak{E}^\gamma)$ is a pre-augmented tree. As the constant $\gamma > 0$ has no significance on the edges in the levels $\mathcal{J}_n$ when n is large, we will omit the superscript γ in the notations when there is no confusion. The following theorem is the main reason for introducing the augmented tree.

Theorem 3.5. [25, 33, 35] *Let* $\{S_i\}_{i=1}^N$ *be an IFS of contractive similitudes and let* K *be the self-similar set. Let* $(X, \mathfrak{E})$ *be the augmented tree. Then*

(i) $(X, \mathfrak{E})$ *is hyperbolic;*

(ii) *there is a canonical identification* $\iota : \partial_H X \to K$ *(independent of* $\gamma > 0$*) defined by* $\{\iota(\xi)\} = \bigcap_n S_{\mathbf{x}_n}(K)$ *and* $(\mathbf{x}_n)_n$ *is a geodesic ray converging to* ξ. *Under this map,* K *is Hölder equivalent to* $\partial_H X$ *in the sense that* $\varrho_a(\xi, \eta) \asymp |\iota(\xi) - \iota(\eta)|^{-a/\log r}$.

Remark 1. Let $e_0 = 0$ and e_i, $1 \leq i \leq d$, be the standard basis vectors in $\mathbb{R}^d$. Let $\{S_i\}_{i=0}^d$ be an IFS on $\mathbb{R}^d$ defined by

$$S_i(\xi) = e_i + \frac{1}{2}(\xi - e_i), \quad \xi \in \mathbb{R}^d.$$

Then the self-similar set K is the *d-dimensional Sierpiński gasket.* When $d = 1$, K is simply the unit interval; when $d = 2$, K is the standard Sierpiński gasket. Theorem 3.5 was first proved in [25, Section 3] for the Sierpiński gaskets with a slightly different horizontal edge set $\mathfrak{E}_h^0$ (see Section 1; call this augmented tree the Sierpiński graph). The extension to general self-similar sets was in [33], where we need to assume, in addition, the OSC and some geometric condition on K. For many standard cases, $\mathfrak{E}_h^\gamma = \mathfrak{E}_h^0$ for small γ, but there

are examples where they are different for any $\gamma > 0$. The setup in Definition 3.4 with $\gamma > 0$ allows us to avoid the delicate behavior on $S_{\mathbf{x}}(K) \cap S_{\mathbf{y}}(K)$ without changing the boundary. Consequently, in [35], the technical assumptions in [33] were eliminated, and Theorem 3.5 was stated in the most general setting.

Remark 2. A pre-augmented tree can be constructed easily, and can be trivial; for example, we can connect all the vertices in each level by edges and still form a hyperbolic graph. On the other hand, if we augment the tree by suitable edges, we can obtain other interesting graphs. For example, starting with the binary augmented tree of the interval ($d = 1$ in Remark 1), we can add one more horizontal edge on each level to connect the two end vertices. This new graph is a pre-augmented tree and is hyperbolic (see Figure 2 or [37]). It is easy to modify the proof of Theorem 3.5 and show that the hyperbolic boundary is Hölder equivalent to the unit circle.
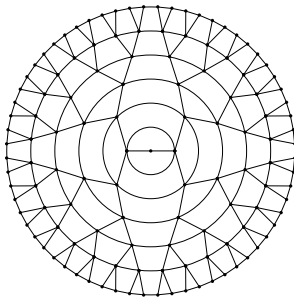

Figure 2: The circle as a hyperbolic boundary

That an augmented tree has bounded degree is important when we consider random walks on it. This property is verified when the IFS has the OSC.

Proposition 3.6. [35, Theorem 1.4] *The IFS $\{S_i\}_{i=1}^N$ satisfies the OSC if and only if the augmented tree $(X, \mathfrak{E})$ has bounded degree.*

To conclude this section, we will use the special form of the canonical geodesic to provide more information of the Gromov product in the augmented tree, which will be needed in Section 6. Note that $\mathbf{x}_n \in \mathcal{J}_n$ for any geodesic ray $(\mathbf{x}_n)_{n=0}^\infty$ in X (starting from ϑ).

Lemma 3.7. *Let $(X, \mathfrak{E})$ be a pre-augmented tree as in Definition 3.2. Let $\mathbf{z} \in X$, and let $(\mathbf{x}_n)_{n=0}^\infty$ and $(\mathbf{y}_n)_{n=0}^\infty$ be two distinct geodesic rays from ϑ. Then we have*

(i) for $n \geq |\mathbf{z}| := k$, $d(\mathbf{x}_n, \mathbf{z}) - |\mathbf{x}_n| = d(\mathbf{x}_k, \mathbf{z}) - k$;

(ii) $\{(\mathbf{x}_n|\mathbf{z})\}_n$ is increasing, and $(\mathbf{x}_n|\mathbf{z}) = (\mathbf{x}_k|\mathbf{z})$ for $n \geq k$;

(iii) $\{(\mathbf{x}_n|\mathbf{y}_n)\}_n$ is increasing, and there exists ℓ such that $(\mathbf{x}_n|\mathbf{y}_n) = (\mathbf{x}_\ell|\mathbf{y}_\ell)$ for $n \geq \ell$.

Proof. (i) For $n \geq |\mathbf{z}| := k$, consider the canonical geodesic $\pi(\mathbf{x}_n, \mathbf{z}) = \pi(\mathbf{x}_n, \mathbf{u}) \cup \pi(\mathbf{u}, \mathbf{v}) \cup \pi(\mathbf{v}, \mathbf{z})$, where $\pi(\mathbf{u}, \mathbf{v})$ is a horizontal geodesic, and $\pi(\mathbf{x}_n, \mathbf{u})$, $\pi(\mathbf{v}, \mathbf{z})$ are vertical geodesics.

Since $|\mathbf{x}_n| = n \geq |k| \geq |\mathbf{u}| = |\mathbf{v}|$, $\mathbf{x}_k$ lies on the vertical geodesic $\pi(\mathbf{x}_n, \mathbf{u})$. Hence (i) follows from

$$d(\mathbf{x}_n, \mathbf{z}) - |\mathbf{x}_n| = d(\mathbf{u}, \mathbf{z}) - |\mathbf{u}| = d(\mathbf{u}, \mathbf{z}) + (d(\mathbf{x}_k, \mathbf{u}) - |\mathbf{x}_k|) = d(\mathbf{x}_k, \mathbf{z}) - k.$$

(ii) That $\{(\mathbf{x}_n|\mathbf{z})\}$ is increasing in n follows by checking the definition of the Gromov product. Then by (i), we have

$$(\mathbf{x}_n|\mathbf{z}) = \frac{1}{2}(|\mathbf{z}| + |\mathbf{x}_n| - d(\mathbf{x}_n, \mathbf{z})) = \frac{1}{2}(2k - d(\mathbf{x}_k, \mathbf{z})) = (\mathbf{x}_k|\mathbf{z}), \quad n \geq k.$$

(iii) By the same proof as the above, we see that $\{(\mathbf{x}_n|\mathbf{y}_n)\}_n$ is increasing. Let M be the maximal length of the horizontal geodesics in $(X, \mathfrak{E})$ (by Proposition 3.3). Since the two rays are distinct, we have $\lim_{n\to\infty}(\mathbf{x}_n|\mathbf{y}_n) < \infty$. Then there exists a positive integer ℓ such that $d(\mathbf{x}_n, \mathbf{y}_n) > M$ for any $n > \ell$. That canonical geodesic $\pi(\mathbf{x}_n, \mathbf{y}_n)$ is not horizontal means both $\mathbf{x}_{n-1}$ and $\mathbf{y}_{n-1}$ lie in the two vertical segments of $\pi(\mathbf{x}_n, \mathbf{y}_n)$ respectively. Now (iii) follows since

$$(\mathbf{x}_n|\mathbf{y}_n) = 2n - d(\mathbf{x}_n, \mathbf{y}_n) = 2(n-1) - d(\mathbf{x}_{n-1}, \mathbf{y}_{n-1}) = (\mathbf{x}_{n-1}|\mathbf{y}_{n-1}), \quad n > \ell.$$

$\square$

Corollary 3.8. *Let $(X, \mathfrak{E})$ be an augmented tree as in Definition 3.4. Let $\xi, \eta \in \partial_H X$, $\xi \neq \eta$, and let $(\mathbf{x}_n), (\mathbf{x}'_n), (\mathbf{y}_n), (\mathbf{y}'_n)$ be geodesic rays, where $(\mathbf{x}_n), (\mathbf{x}'_n)$ converge to ξ, and $(\mathbf{y}_n), (\mathbf{y}'_n)$ converge to η. Then for any $\mathbf{z} \in X$,*

$$\Big| \lim_{n\to\infty}(\mathbf{x}_n|\mathbf{z}) - \lim_{n\to\infty}(\mathbf{x}'_n|\mathbf{z})\Big| \leq \frac{1}{2}, \quad and \quad \Big| \lim_{n\to\infty}(\mathbf{x}_n|\mathbf{y}_n) - \lim_{n\to\infty}(\mathbf{x}'_n|\mathbf{y}'_n)\Big| \leq 1.$$

Proof. We first observe from Theorem 3.5 that $\xi \in S_{\mathbf{x}_n}(K) \cap S_{\mathbf{x}'_n}(K)$. Thus $S_{\mathbf{x}_n}(K) \cap S_{\mathbf{x}'_n}(K) \neq \emptyset$, and by (3.5) either $\mathbf{x}_n = \mathbf{x}'_n$ or $\mathbf{x}_n \sim_h \mathbf{x}'_n$. It follows that $d(\mathbf{x}_n, \mathbf{x}'_n) \leq 1$ for all n. Then the triangle inequality implies

$$|(\mathbf{x}_n|\mathbf{z}) - (\mathbf{x}'_n|\mathbf{z})| = \frac{1}{2}|d(\mathbf{x}'_n, \mathbf{z}) - d(\mathbf{x}_n, \mathbf{z})| \leq \frac{1}{2}d(\mathbf{x}'_n, \mathbf{x}_n) \leq \frac{1}{2}. \tag{3.6}$$

For the second part, we need only to replace the above $\mathbf{z}$ by $\mathbf{y}_n$ and $\mathbf{y}'_n$ respectively and apply the triangle inequality. $\square$

From the corollary, we can extend the Gromov product to $\partial_H X$ by

$$(\xi|\mathbf{z}) = \sup\Big\{ \lim_{n\to\infty}(\mathbf{x}_n|\mathbf{z})\Big\}, \qquad \xi \in \partial_H X, \quad \mathbf{z} \in X, \tag{3.7}$$

and

$$(\xi|\eta) = \sup\Big\{ \lim_{n\to\infty}(\mathbf{x}_n|\mathbf{y}_n)\Big\}, \qquad \xi, \eta \in \partial_H X, \tag{3.8}$$

where the supremum is taken over all geodesic rays $(\mathbf{x}_n)_n$ and $(\mathbf{y}_n)_n$ that converge to ξ and η respectively.

13

4 Constant return ratio and quasi-natural RW

Let $\{Z_n\}_{n=0}^{\infty}$ be a reversible random walk on the pre-augmented tree $(X, \mathfrak{E})$. For $\mathbf{x} \in X \setminus \{\vartheta\}$, we define the *return ratio* at $\mathbf{x}$ by

$$\lambda(\mathbf{x}) = \frac{P(\mathbf{x}, \mathbf{x}^-)}{\displaystyle\sum_{\mathbf{y}:\mathbf{y}^-=\mathbf{x}} P(\mathbf{x}, \mathbf{y})} = \frac{c(\mathbf{x}, \mathbf{x}^-)}{\displaystyle\sum_{\mathbf{y}:\mathbf{y}^-=\mathbf{x}} c(\mathbf{x}, \mathbf{y})}. \tag{4.1}$$

For $\lambda > 0$, we introduce the following condition on the transition probability P:

(R_λ) (*Constant return ratio*) For any $\mathbf{x} \in X \setminus \{\vartheta\}$, $\lambda(\mathbf{x}) \equiv \lambda$ is a constant.

For example, the SRW on the augmented tree of a homogeneous IFS $\{S_i\}_{i=1}^{N}$ satisfies condition (R_λ) with $\lambda = N^{-1}$.

For a fixed level $m \geq 1$, let $X_m := \bigcup_{k=0}^{m} \mathcal{J}_k$ and $\mathfrak{E}_m := \mathfrak{E}|_{X_m \times X_m}$. Consider the following random walk $\{Z_n^{(m)}\}_{n=0}^{\infty}$ with transition probability P_m on the graph $(X_m, \mathfrak{E}_m)$: for $\mathbf{x}$, $\mathbf{y} \in X_m$,

$$P_m(\mathbf{x}, \mathbf{y}) = \begin{cases} P(\mathbf{x}, \mathbf{y}), & \text{if } \mathbf{x} \sim \mathbf{y} \text{ and } |\mathbf{x}| < m, \\ 1, & \text{if } \mathbf{x} = \mathbf{y} \text{ and } |\mathbf{x}| = m, \\ 0, & \text{otherwise.} \end{cases}$$

This is the restriction of P on $X_m \setminus \mathcal{J}_m$, where the vertices in $\mathcal{J}_m$ are absorbing states of P_m. Let $F_m(\mathbf{x}, \mathbf{y})$ be the probability of ever visiting $\mathbf{y}$ from $\mathbf{x}$ in X_m, i.e.,

$$F_m(\mathbf{x}, \mathbf{y}) := \mathbb{P}(\exists \, n \geq 0 \text{ such that } Z_n^{(m)} = \mathbf{y} \mid Z_0^{(m)} = \mathbf{x}).$$

We begin by finding the value of $F_m(\mathbf{x}, \vartheta)$ for a reversible random walk with condition (R_λ), which is the most basic identity to be used in this section and Theorem 5.3.

Proposition 4.1. *Let $\{Z_n\}_{n=0}^{\infty}$ be a reversible random walk on a pre-augmented tree $(X, \mathfrak{E})$ that satisfies (R_λ) for some $\lambda > 0$. Then for $m \geq 1$ and $\mathbf{x} \in X_m$,*

$$F_m(\mathbf{x}, \vartheta) = \begin{cases} \dfrac{\lambda^{|\mathbf{x}|} - \lambda^m}{1 - \lambda^m}, & \text{if } \lambda \neq 1, \\[2mm] \dfrac{m - |\mathbf{x}|}{m}, & \text{if } \lambda = 1. \end{cases} \tag{4.2}$$

Consequently, $F(\mathbf{x}, \vartheta) = \lambda^{|\mathbf{x}|}$ if $0 \leq \lambda < 1$, and $= 1$ if $\lambda \geq 1$.

Remark. We thank Professor J. Kigami for informing us for the following proof which shortened the original one.

Proof. For $n \geq 0$, let $t(n) = \inf\{k \geq n : |Z_k| \neq |Z_n|\}$ be the first time that the random walk jumps to a different level from time n. Define a sequence of stopping times $\{n_k\}_{k=0}^{\infty}$ by letting $n_0 = 0$ and $n_k = t(n_{k-1})$ for $k \geq 1$. For $k \geq 0$, let $L_k = |Z_{n_k}|$ denote the level of the chain. It can be checked directly that $\{L_k\}_{k=0}^{\infty}$ is a birth and death chain on the nonnegative integers with the following transition probabilities: $P_L(0, 1) = 1$, $P_L(\ell, \ell - 1) = \lambda/(1 + \lambda)$, and $P_L(\ell, \ell + 1) = 1/(1 + \lambda)$ for $\ell \geq 1$. Let T_ℓ be the first time that L_k visits level ℓ, then

14

the expression (4.2) of $F_m(\mathbf{x}, \vartheta) = \mathbb{P}(T_0 < T_m \mid L_0 = |\mathbf{x}|)$ for $\mathbf{x} \in X_m$ follows from standard calculations for birth and death chains.

Observe that for $|\mathbf{x}| < m$, $F_m(\mathbf{x}, \vartheta) \nearrow F(\mathbf{x}, \vartheta)$ as $m \to \infty$, the second part follows by taking limit of (4.2). $\qquad\square$

Since our main interest is on transient random walks, we will assume $\lambda \in (0, 1)$ throughout the paper. As a simple consequence of Proposition 4.1 we have the follow result on the Green function, which will be needed to consider $F_m(\vartheta, \mathbf{x})$ and $F(\vartheta, \mathbf{x})$.

Lemma 4.2. *Let $\{Z_n\}_{n=0}^\infty$ be a reversible random walk on a pre-augmented tree $(X, \mathfrak{E})$, and $G_m(\cdot, \cdot)$ be the Green function of $\{Z_n^{(m)}\}_{n=0}^\infty$. If $\{Z_n\}_{n=0}^\infty$ satisfies condition (R_λ) with $\lambda \in (0, 1)$, then $G_m(\vartheta, \vartheta) = \frac{1-\lambda^m}{1-\lambda}$ and $G(\vartheta, \vartheta) = \frac{1}{1-\lambda}$.*

Proof. Note that $G_m(\vartheta, \vartheta) = \frac{1}{1-U_m(\vartheta, \vartheta)}$, where $U_m(\vartheta, \vartheta) := \mathbb{P}(\exists n \geq 1$ such that $Z_n^{(m)} = \vartheta \mid Z_0^{(m)} = \vartheta)$, the probability that the random walk returns to ϑ after starting at the root ϑ). The lemma follows from this identity together with the one-step formula $U_m(\vartheta, \vartheta) = \sum_{\mathbf{x} \in \mathcal{J}_1} P_m(\vartheta, \mathbf{x}) F_m(\mathbf{x}, \vartheta)$ and Proposition 4.1. $\qquad\square$

Proposition 4.3. *Let $\{Z_n\}_{n=0}^\infty$ be a reversible random walk on a pre-augmented tree $(X, \mathfrak{E})$ that satisfies condition (R_λ). Then for $m \geq 1$ and $\mathbf{x} \in \mathcal{J}_m$,*

$$F_m(\vartheta, \mathbf{x}) = \frac{c(\mathbf{x}^-, \mathbf{x})\lambda^{m-1}}{m(\vartheta)} \ . \tag{4.3}$$

Proof. For $\mathbf{x} \in \mathcal{J}_m$, using the reversibility $m(\vartheta)G_m(\vartheta, \mathbf{x}^-) = m(\mathbf{x}^-)G_m(\mathbf{x}^-, \vartheta)$, together with Proposition 4.1 and Lemma 4.2, we can evaluate $F_m(\vartheta, \mathbf{x}) = G_m(\vartheta, \mathbf{x}^-)P(\mathbf{x}^-, \mathbf{x})$ directly to get (4.3). $\qquad\square$

In view of the expression of $F_m(\vartheta, \mathbf{x})$ in (4.3), we will introduce a class of conductance $c(x, y)$ so that the limits exist as $m \to \infty$. (We can not take the limit directly, as $\mathbf{x} \in \mathcal{J}_m$ depends on m.)

Definition 4.4. *Let $(X, \mathfrak{E})$ be a pre-augmented tree of an IFS $\{S_i\}_{j=1}^N$. A reversible random walk $\{Z_n\}_{n=0}^\infty$ on $(X, \mathfrak{E})$ with conductance $c : X \times X \to [0, \infty)$ is called quasi-natural with return ratio $\lambda \in (0, 1)$ (λ-QNRW) if for a set of probability weights $\{p_i\}_{i=1}^N$,*

(i) $c(\mathbf{x}, \mathbf{x}^-) = p_{\mathbf{x}}\lambda^{-m}$, for $\mathbf{x} \in \mathcal{J}_m$, $m \geq 1$, where $p_{\mathbf{x}} = p_{i_1} \cdots p_{i_k}$ if $\mathbf{x} = i_1 \cdots i_k$,

(ii) $c(\mathbf{x}, \mathbf{y}) \asymp c(\mathbf{x}, \mathbf{x}^-)$ for $\mathbf{y} \sim_h \mathbf{x} \in \mathcal{J}_m$, $m \geq 1$, and

(iii) $c(\mathbf{x}, \mathbf{y}) = 0$ if $\mathbf{y} \nsim \mathbf{x}$.

(Here the bounds of $\asymp$ are independent of $\mathcal{J}_m$, $m \geq 1$). Furthermore, if the IFS satisfies the OSC and $p_i = r_i^\alpha$, the natural weights of the IFS, we call $\{Z_n\}_{n=0}^\infty$ a natural random walk with return ratio λ (λ-NRW) on $(X, \mathfrak{E})$.

For a homogeneous IFS, using $\Sigma^n = \mathcal{J}_n$, it is direct to check from $P(\mathbf{x}, \mathbf{y}) = \frac{c(\mathbf{x}, \mathbf{y})}{m(\mathbf{x})}$, $\mathbf{x} \sim \mathbf{y}$ that $P(\mathbf{x}, \mathbf{x}^-) = C_{\mathbf{x}}\lambda$, and $P(\mathbf{x}, \mathbf{x}i) = C_{\mathbf{x}}p_i$ for some $C_{\mathbf{x}} > 0$. Clearly the walk has constant return ratio λ. More generally, we have

15

Proposition 4.5. *Let $\{Z_n\}_{n=0}^{\infty}$ be a λ-QNRW on the pre-augmented tree $(X, \mathfrak{E})$. Then $\{Z_n\}_{n=0}^{\infty}$ satisfies condition (R_λ). Also $m(\vartheta) = \lambda^{-1}$, and if $(X, \mathfrak{E})$ is of bounded degree, then*

$$m(\mathbf{x}) \asymp c(\mathbf{x}, \mathbf{x}^-), \qquad \mathbf{x} \in X. \tag{4.4}$$

Proof. For $\mathbf{x} \in X \setminus \{\vartheta\}$, by (4.1), we have $\lambda(\mathbf{x}) = \lambda p_\mathbf{x} / (\sum_{\mathbf{y}: \mathbf{y}^- = \mathbf{x}} p_\mathbf{y}) = \lambda$. Hence the walk satisfies condition (R_λ). Also, $m(\vartheta) = \sum_{\mathbf{y} \in \mathcal{J}_1} p_\mathbf{y} \lambda^{-1} = \lambda^{-1}$, and (4.4) is straightforward by the bounded degree assumption. $\square$

We now apply the previous results to make the following conclusion.

Theorem 4.6. *Let $\{Z_n\}_{n=0}^{\infty}$ be a λ-QNRW on a pre-augmented tree $(X, \mathfrak{E})$ with bounded degree. Then*

$$F_m(\vartheta, \mathbf{x}) = p_\mathbf{x}, \qquad \mathbf{x} \in \mathcal{J}_m, \tag{4.5}$$

Furthermore, $F(\vartheta, \mathbf{x}) \asymp p_\mathbf{x}$ (the bounds depend on λ) for any $\mathbf{x} \in X$.

Proof. For $m \geq 1$ and for $\mathbf{x} \in \mathcal{J}_m$, Propositions 4.3 and 4.5 imply that

$$F_m(\vartheta, \mathbf{x}) = \frac{c(\mathbf{x}, \mathbf{x}^-) \lambda^{m-1}}{m(\vartheta)} = \frac{p_\mathbf{x} \lambda^{-m} \cdot \lambda^{m-1}}{\lambda^{-1}} = p_\mathbf{x}.$$

Note that we cannot directly send $m \to \infty$ as $\mathbf{x} \in \mathcal{J}_m$ depends on m. Instead, we observe that

$$F(\vartheta, \mathbf{x}) \geq F_m(\vartheta, \mathbf{x}) = p_\mathbf{x}.$$

Also, note that P is reversible, and hence $m(\mathbf{x}) G(\mathbf{x}, \vartheta) = m(\vartheta) G(\vartheta, \mathbf{x})$ for any vertex $\mathbf{x} \in X$. Hence, for $\mathbf{x} \in X \setminus \{\vartheta\}$, by Proposition 4.1, Lemma 4.2 and Proposition 4.5, we have

$$\begin{aligned}
F(\vartheta, \mathbf{x}) = \frac{G(\vartheta, \mathbf{x})}{G(\mathbf{x}, \mathbf{x})} &\leq G(\vartheta, \mathbf{x}) && \text{(by } G(\mathbf{x}, \mathbf{x}) \geq 1) \\
&= \frac{m(\mathbf{x})}{m(\vartheta)} G(\mathbf{x}, \vartheta) \leq C_1 p_\mathbf{x} \lambda^{-|\mathbf{x}|} F(\mathbf{x}, \vartheta) G(\vartheta, \vartheta) \\
&= C_1 p_\mathbf{x} \lambda^{-|\mathbf{x}|} \cdot \left(\lambda^{|\mathbf{x}|} \cdot \frac{1}{1 - \lambda} \right) = \frac{C_1}{1 - \lambda} p_\mathbf{x}.
\end{aligned}$$

This shows that $F(\vartheta, \mathbf{x}) \asymp p_\mathbf{x}$. $\square$

Remark 1. In general, we cannot expect a λ-QNRW to have the *strict reversibility* property (i.e., $0 < M^{-1} \leq c(\mathbf{x}, \mathbf{y}) \leq M$ for any $\mathbf{x} \sim \mathbf{y}$). More precisely, we can show that if the IFS satisfies the OSC, then strict reversibility of λ-QNRW implies that $\lambda = r^\alpha$, where $r = \min_{1 \leq i \leq N} r_i$.

Indeed, suppose strict reversibility holds, and let $c(\mathcal{J}_m) = \sum_{\mathbf{x} \in \mathcal{J}_m} c(\mathbf{x}, \mathbf{x}^-)$, $m \geq 0$. Then by (4.1), we have $c(\mathcal{J}_{m+1}) = \lambda^{-m} c(\mathcal{J}_1) = \lambda^{-m} m(\vartheta)$. Recall that $r^{-\alpha m} \leq \#\mathcal{J}_m < r^{-\alpha(m+1)}$ where α satisfies $\sum_{i=1}^{N} r_i^\alpha = 1$ (Lemma 3.1). Hence

$$0 < M^{-1} \leq \frac{c(\mathcal{J}_{m+1})}{\#\mathcal{J}_{m+1}} \leq \frac{\lambda^{-m} m(\vartheta)}{r^{-\alpha(m+1)}}, \qquad m \geq 0.$$

16

This implies $\lambda \leq r^\alpha$. On the other hand,

$$M \geq \frac{c(\mathcal{J}_{m+1})}{\#\mathcal{J}_{m+1}} \geq \frac{\lambda^{-m} m(\vartheta)}{r^{-\alpha(m+2)}}, \qquad m \geq 0$$

yields $\lambda \geq r^\alpha$, and hence $\lambda = r^\alpha$. It follows that in Definition 4.4(i), we must have $p_i = r_i^\alpha$, the natural weight to have the strict reversibility property.

Remark 2. Conditions (i) and (ii) of the λ-QNRW imply that $c(\mathbf{x}, \mathbf{x}^-) \asymp c(\mathbf{x}, \mathbf{y}) \asymp c(\mathbf{y}, \mathbf{y}^-)$ for $\mathbf{x} \sim_h \mathbf{y}$. It follows that $p_\mathbf{x} \asymp p_\mathbf{y}$ for all $\mathbf{x} \sim_h \mathbf{y}$, which is a strong restriction on the choice of possible probability weights $\{p_i\}_{i=1}^N$. For the λ-NRW this restriction is fulfilled, as the natural weights satisfy $p_\mathbf{x} = r_\mathbf{x}^\alpha \asymp r^{\alpha|\mathbf{x}|}$ for all $\mathbf{x} \in X$. In particular, for a homogeneous IFS that satisfies the OSC and has contraction ratio r, the SRW on a pre-augmented tree $(X, \mathfrak{E})$ is a $\frac{1}{N}$-NRW, as $p_i = 1/N = r^\alpha$, $\lambda = N^{-1}$ and $c(\mathbf{x}, \mathbf{x}^-) = 1$.

However, for the λ-QNRW, we have to choose special probability weights so that $p_\mathbf{x} \asymp p_\mathbf{y}$, $\mathbf{x} \sim_h \mathbf{y}$. For instance, if K is the d-dimensional Sierpiński graph, one can derive from $p_\mathbf{x} \asymp p_\mathbf{y}$, $\mathbf{x} \sim_h \mathbf{y}$ that all p_i's are equal. Hence the λ-QNRW on the corresponding augmented tree must be the λ-NRW.

Example 4.7. If K is the Sierpiński carpet, and the IFS is $\{S_i\}_{i=1}^8$ on $\mathbb{R}^2$, where $S_i(z) = (z + q_i)/2$, $i = 1, 2, \ldots, 8$, and the q_i's are the four vertices and four mid-points of the edges of a square, and is labeled clockwise starting from the top left corner. Consider the pre-augmented tree $(X, \mathfrak{E})$ with $\mathfrak{E}_h$ defined by $\mathbf{x} \sim_h \mathbf{y}$ if $|\mathbf{x}| = |\mathbf{y}|$ and $\dim_H\big(K_\mathbf{x} \cap K_\mathbf{y}\big) = 1$. Then it is direct to check that one obtains a λ-QNRW with probability weight $\{p_i\}_{i=1}^8$ on $(X, \mathfrak{E})$ if and only if $p_1 = p_3 = p_5 = p_7$, $p_2 = p_6$, and $p_4 = p_8$. Note that with this weight, the self-similar measure is a *doubling measure* [46], i.e, there exists a constant $C > 0$ such that for any $\xi \in K$ and any $\delta > 0$, we have $\mu(B(\xi; 2\delta)) \leq C\mu(B(\xi; \delta))$.

In general, we can make use of a result in [46] to characterize the probability weights $\{p_i\}_{i=1}^N$ that admit a λ-QNRW.

Theorem 4.8. *Let $\{S_i\}_{i=1}^N$ be an IFS with OSC, K be the self-similar set, and μ be the self-similar measure generated by $\{p_i\}_{i=1}^N$. Then $\{p_i\}_{i=1}^N$ admits a λ-QNRW on the augmented tree $(X, \mathfrak{E})$ if and only if μ is a doubling measure on K.*

Proof. Let $K_\mathbf{x} = S_\mathbf{x}(K)$. For $F \subset \mathbb{R}^d$, let $|F| = \text{diam}(F)$, and $B(F; \delta) := \{\xi \in \mathbb{R}^d : \text{dist}(\xi, F) \leq \delta\}$ be the closed δ-neighborhood of F. Consider the following two conditions:

(i) if there exist constants C_1, $C_2 > 0$ such that for any $\mathbf{w}$, $\mathbf{v} \in \Sigma^* \setminus \{\vartheta\}$ that satisfy $K_\mathbf{w} \subset B(K_\mathbf{v}; C_1 r_\mathbf{v})$, then $p_\mathbf{w} \leq C_2 p_\mathbf{v}$;

(ii) Replace the quantifiers of (i) to "if for any $C_1 > 0$, there exists $C_2 > 0$ such that $\ldots$".

In [46, Theorem 2.3], it was proved that (i) $\Rightarrow \mu$ is a doubling measure $\Rightarrow$ (ii). We will prove (ii) $\Rightarrow \{p_i\}_{i=1}^N$ admits a QNRW $\Rightarrow$ (i). Hence all four conditions are equivalent, proving the theorem.

17

Assume (ii) and let $C_1 = (\gamma + |K|)r^{-1}$, where $r = \min_{1 \le i \le N} r_i$, and γ is as in (3.5). Then by (3.5), $\mathbf{x} \sim_h \mathbf{y}$ implies

$$K_{\mathbf{x}} \subset B(K_{\mathbf{y}}; \ \gamma r^{|\mathbf{x}|} + r_{\mathbf{x}}|K|)$$
$$\subset B(K_{\mathbf{y}}; \ (\gamma + |K|)r^{|\mathbf{x}|}) \subset B(K_{\mathbf{y}}; \ C_1 r_{\mathbf{y}}).$$

Then by (ii), $p_{\mathbf{x}} \le C_2 p_{\mathbf{y}}$ for some $C_2 > 0$; by the same reason, $p_{\mathbf{y}} \le C_2 p_{\mathbf{x}}$. Hence for $0 < \lambda < 1$, $c(\mathbf{x}, \mathbf{x}^-) = p_{\mathbf{x}} \lambda^{-|\mathbf{x}|} \asymp p_{\mathbf{y}} \lambda^{-|\mathbf{y}|} = c(\mathbf{y}, \mathbf{y}^-)$ for all $\mathbf{x} \sim_h \mathbf{y}$. We can choose, for instance, $c(\mathbf{x}, \mathbf{y}) = \sqrt{c(\mathbf{x}, \mathbf{x}^-)c(\mathbf{y}, \mathbf{y}^-)}$ to get a λ-QNRW on $(X, \mathfrak{E})$.

Next we assume $\{p_i\}_{i=1}^N$ admits a λ-QNRW on $(X, \mathfrak{E})$. We first define an integer $N_0 = \lceil \log r / \log r' \rceil$ where $r' = \max_{1 \le i \le N} r_i$. Then for $p = \min_{1 \le i \le N} p_i$, and for any $\mathbf{z} \in X$, we have $p_{\mathbf{z}} \ge p^{N_0} p_{\mathbf{z}^-}$ since the difference of the length of the words $\mathbf{z}$ and $\mathbf{z}^-$ is at most N_0.

The λ-QNRW implies that $p_{\mathbf{x}} \lambda^{-|\mathbf{x}|} = c(\mathbf{x}, \mathbf{x}^-) \asymp c(\mathbf{x}, \mathbf{y}) \asymp c(\mathbf{y}, \mathbf{y}^-) = p_{\mathbf{y}} \lambda^{-|\mathbf{y}|}$ for any $\mathbf{x} \sim_h \mathbf{y}$. Therefore there exists $C > 0$ such that $p_{\mathbf{x}} \le C p_{\mathbf{y}}$ for any $\mathbf{x} \sim_h \mathbf{y}$. Let $C_1 = \min\{\gamma, (1-r)(2r)^{-1}|K|\} > 0$. It follows that for any $\mathbf{y} \in \mathcal{J}_m$,

$$|B(K_{\mathbf{y}}; \ C_1 r_{\mathbf{y}})| \le |K_{\mathbf{y}}| + 2C_1 r_{\mathbf{y}}$$
$$\le (1 + 2(1-r)(2r)^{-1})r^m |K| = r^{m-1}|K|.$$

Hence for $\mathbf{x} \in X$ such that $K_{\mathbf{x}} \subset B(K_{\mathbf{y}}; \ C_1 r_{\mathbf{y}})$, $|\mathbf{x}| \ge m - 1$ (by (3.2)). We claim that $p_{\mathbf{x}} \le C p^{-N_0} p_{\mathbf{y}}$. Indeed, let $\mathbf{x}_m \in \mathcal{J}_m$ such that it is on the same vertical path of $\mathbf{x}$. Then either $K_{\mathbf{x}} \subset K_{\mathbf{x}_m}$ or $K_{\mathbf{x}_m} \subset K_{\mathbf{x}}$. We have three distinct cases:

(a) If $\mathbf{x}_m = \mathbf{y}$, then $p_{\mathbf{x}} \le p_{\mathbf{y}^-} \le p^{-N_0} p_{\mathbf{y}}$.

(b) If $\mathbf{x}_m \ne \mathbf{y}$ and $|\mathbf{x}| = m - 1$, then $\mathbf{x}_m^- = \mathbf{x}$ and $K_{\mathbf{x}_m} \subset K_{\mathbf{x}} \subset B(K_{\mathbf{y}}; \ C_1 r_{\mathbf{y}})$, which implies $\mathrm{dist}(K_{\mathbf{x}_m}, K_{\mathbf{y}}) \le C_1 r_{\mathbf{y}} \le \gamma r^m$, hence $\mathbf{x}_m \sim_h \mathbf{y}$. This shows that $p_{\mathbf{x}} \le p^{-N_0} p_{\mathbf{x}_m} \le C p^{-N_0} p_{\mathbf{y}}$.

(c) If $\mathbf{x}_m \ne \mathbf{y}$ and $|\mathbf{x}| \ge m$, then $K_{\mathbf{x}} \subset K_{\mathbf{x}_m}$ and $\mathrm{dist}(K_{\mathbf{x}_m}, K_{\mathbf{y}}) \le \mathrm{dist}(K_{\mathbf{x}}, K_{\mathbf{y}}) \le C_1 r_{\mathbf{y}} \le \gamma r^m$, hence $\mathbf{x}_m \sim_h \mathbf{y}$. This shows that $p_{\mathbf{x}} \le p_{\mathbf{x}_m} \le C p_{\mathbf{y}}$.

To conclude the proof, we let $\mathbf{w}, \mathbf{v} \in \Sigma^* \setminus \{\vartheta\}$. We can choose $\mathbf{u}, \mathbf{t} \in X$ such that $S_{\mathbf{u}}(K) \subset K_{\mathbf{w}} \subset K_{\mathbf{u}^-}$, and $K_{\mathbf{t}} \subset K_{\mathbf{v}} \subset K_{\mathbf{t}^-}$. If $K_{\mathbf{w}} \subset B(K_{\mathbf{v}}; \ C_1 r_{\mathbf{v}})$ as in the assumption in condition (i), then $K_{\mathbf{u}} \subset B(K_{\mathbf{t}^-}; \ C_1 r_{\mathbf{t}^-})$, hence

$$p_{\mathbf{w}} \le p_{\mathbf{u}^-} \le p^{-N_0} p_{\mathbf{u}} \le C p^{-2N_0} p_{\mathbf{t}^-} \le C p^{-3N_0} p_{\mathbf{t}} \le C p^{-3N_0} p_{\mathbf{v}}.$$

Let $C_2 = C p^{-3N_0}$. Then (i) follows. $\qquad\square$

5   Martin boundary and hitting distribution

In this section we focus on the boundary behavior of natural random walks on augmented trees. We first show that Ancona's Theorem (Theorem 2.4) can be applied to identify the Martin and hyperbolic boundaries with the self-similar sets. This extends [25, Theorem 4.7] for the SRW on the Sierpiński graph.

Theorem 5.1. *Let $(X, \mathfrak{E})$ be a pre-augmented tree which is hyperbolic and has bounded degree, and let $\{Z_n\}_{n=0}^{\infty}$ be a λ-QNRW on $(X, \mathfrak{E})$. Then the transition probability P satisfies the hypotheses in Theorem 2.4. Hence $\mathcal{M} = \mathcal{M}_{\min}$, and the hyperbolic boundary $\partial_H X$ and $\mathcal{M}$ are homeomorphic under the canonical mapping.*

In particular, if the IFS satisfies the OSC, and $(X, \mathfrak{E})$ is the augmented tree as in (3.5), then $\partial_H X$, $\mathcal{M}$ and the self-similar set K are all homeomorphic under the canonical mapping.

Proof. We need to check that the conditions in Ancona's theorem (Theorem 2.4) are satisfied. Clearly P is of bounded range. That

$$P(\mathbf{x}, \mathbf{y}) = \frac{c(\mathbf{x}, \mathbf{y})}{m(\mathbf{x})} \geq \eta > 0, \qquad \mathbf{x} \sim \mathbf{y} \tag{5.1}$$

follows from Definition 4.4 and (4.4). Hence P is uniformly irreducible. It remains to show that $r(P) < 1$. By Proposition 2.5, it suffices to show that (X, P) satisfies the strong isoperimetric inequality.

Consider the subtree $T = (X, \mathfrak{E}_v)$, and restrict the random walk to T. The transition probability is given by $P_T(\mathbf{x}, \mathbf{x}^-) = \frac{c(\mathbf{x}, \mathbf{x}^-)}{m_T(\mathbf{x})}$, and $= 0$ otherwise, where $m_T(\mathbf{x}) = c(\mathbf{x}, \mathbf{x}^-) + \sum_{\mathbf{y}: \mathbf{y}^- = \mathbf{x}} c(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \in X \setminus \{\vartheta\}$. Note that $m_T(\vartheta) = m(\vartheta)$. We claim that the transition probability P_T satisfies the following strong isoperimetric inequality:

$$m_T(A) \leq \frac{1 + \lambda}{1 - \lambda} c_T(\partial A) \tag{5.2}$$

for all finite subsets $A \subset X$. To prove the claim, we can first assume that A is nonempty and connected (otherwise check (5.2) for each connected component and then sum them up). We use induction on $n := \min\{m : A \subset X_m\}$. It is clear that if $n = 0$, then $A = \{\vartheta\}$ and $m_T(A) = m(\vartheta) = c_T(\partial A)$ implies (5.2). Suppose that for some $n \geq 1$, (5.2) holds for any finite connected subset in X_{n-1}. Consider a connected subset $A \subset X$ such that $\min\{m : A \subset X_m\} = n$. If $\#A = 1$, then $m_T(A) = c_T(\partial A)$ implies (5.2). Now we suppose $\#A \geq 2$. Then for any $\mathbf{x} \in A \cap \mathcal{J}_n$, $\mathbf{x}^- \in A$ since T is a tree and A is connected. Hence

$$
\begin{aligned}
m_T(A) &= m_T(A \setminus \mathcal{J}_n) + \sum_{\mathbf{x} \in A \cap \mathcal{J}_n} \left(c(\mathbf{x}, \mathbf{x}^-) + \sum_{\mathbf{y}: \mathbf{y}^- = \mathbf{x}} c(\mathbf{x}, \mathbf{y}) \right) \\
&= m_T(A \setminus \mathcal{J}_n) + \sum_{\mathbf{x} \in A \cap \mathcal{J}_n} (1 + \lambda^{-1}) c(\mathbf{x}, \mathbf{x}^-) \qquad \text{(by (4.1))} \\
&\leq \frac{1 + \lambda}{1 - \lambda} \left(c_T(\partial(A \setminus \mathcal{J}_n)) + \sum_{\mathbf{x} \in A \cap \mathcal{J}_n} (\lambda^{-1} - 1) c(\mathbf{x}, \mathbf{x}^-) \right) \quad \text{(by induction)} \\
&= \frac{1 + \lambda}{1 - \lambda} \left(c_T(\partial(A \setminus \mathcal{J}_n)) + \sum_{\mathbf{x} \in A \cap \mathcal{J}_n} \left(\sum_{\mathbf{y}: \mathbf{y}^- = \mathbf{x}} c(\mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{x}^-) \right) \right) \quad \text{(by (4.1))} \\
&= \frac{1 + \lambda}{1 - \lambda} c_T(\partial A) \qquad \text{(use tree property)}
\end{aligned}
$$

This completes the proof of the claim. Next observe that for any $\mathbf{x} \in X \setminus \{\vartheta\}$, there exists

19

$C_1 > 0$ such that

$$m(\mathbf{x}) \leq C_1 p_\mathbf{x} \lambda^{-|\mathbf{x}|} = C_1 c(\mathbf{x}, \mathbf{x}^-) \qquad \text{(by bounded degree and (4.4))}$$
$$= \frac{C_1 \lambda}{1 + \lambda} \Big(\sum_{\mathbf{y}:\mathbf{y}^- = \mathbf{x}} c(\mathbf{x}, \mathbf{y}) + c(\mathbf{x}, \mathbf{x}^-) \Big) \quad \text{(by (4.1))}$$
$$= \frac{C_1 \lambda}{1 + \lambda} m_T(\mathbf{x}).$$

This together with the above claim imply the following strong isoperimetric inequality on $(X, \mathfrak{E})$:

$$m(A) \leq \frac{C_1 \lambda}{1 + \lambda} m_T(A) \leq \frac{C_1 \lambda}{1 - \lambda} c_T(\partial A) \leq \frac{C_1 \lambda}{1 - \lambda} c(\partial A), \quad A \subset X \text{ finite.}$$

For the last part, we observe the OSC implies that the the augmented tree $(X, \mathcal{E})$ is of bounded degree (Proposition 3.6), and Theorem 3.5 and the above give the homeomorphism among K, $\partial_H X$ and $\mathcal{M}$. $\qquad\qquad\square$

Since $r(P) < 1$ as shown in the above proof, we can apply the first part of Theorem 2.4 to obtain the same inequality for $F(\mathbf{x}, \mathbf{y})$ of the λ-QNRW.

Corollary 5.2. *Let $(X, \mathfrak{E})$ be a hyperbolic pre-augmented tree with bounded degree, and let $\{Z_n\}_{n=0}^\infty$ be a λ-QNRW on $(X, \mathfrak{E})$. Then for $\delta \geq 0$, there is a constant $C_\delta \geq 1$ such that for any $\mathbf{x}$, $\mathbf{y} \in X$ and $\mathbf{u}$ within distance δ from a geodesic segment between $\mathbf{x}$ and $\mathbf{y}$,*

$$F(\mathbf{x}, \mathbf{u}) F(\mathbf{u}, \mathbf{y}) \leq F(\mathbf{x}, \mathbf{y}) \leq C_\delta F(\mathbf{x}, \mathbf{u}) F(\mathbf{u}, \mathbf{y}). \qquad (5.3)$$

The above corollary allows us to give a useful estimate for $F(\mathbf{x}, \mathbf{y})$ in terms of the Gromov product, which will be essential in the estimates of the Martin kernel and the Naïm kernel. For this, we need to restrict our consideration to the IFS satisfying the OSC, and to the class of natural random walks with natural weights $p_i = r_i^\alpha, 1 \leq i \leq N$. In this case, we have $p_\mathbf{x} = \mu(S_\mathbf{x}(K)) = r_\mathbf{x}^\alpha \asymp r^{\alpha m}$ for $\mathbf{x} \in \mathcal{J}_m$, where $r = \min_{1 \leq i \leq N}\{r_i\}$, α is the Hausdorff dimension of K, and μ is the self-similar measure corresponding to $\{p_i\}_{i=1}^N$ (Lemma 3.1).

Theorem 5.3. *Let $\{S_i\}_{i=1}^N$ be an IFS satisfying the OSC, and let $\{Z_n\}_{n=0}^\infty$ be a λ-NRW on the augmented tree $(X, \mathfrak{E})$. Then*

$$F(\mathbf{x}, \mathbf{y}) \asymp \lambda^{|\mathbf{x}| - (\mathbf{x}|\mathbf{y})} r^{\alpha(|\mathbf{y}| - (\mathbf{x}|\mathbf{y}))}, \qquad \forall \, \mathbf{x}, \mathbf{y} \in X$$

where $(\mathbf{x}|\mathbf{y})$ is the Gromov product. Moreover, the Martin kernel satisfies

$$K(\mathbf{x}, \mathbf{y}) \asymp \lambda^{|\mathbf{x}| - (\mathbf{x}|\mathbf{y})} r^{-\alpha(\mathbf{x}|\mathbf{y})}, \qquad \mathbf{x}, \mathbf{y} \in X,$$

and also for $K(\mathbf{x}, \xi)$ by replacing $\mathbf{y}$ with $\xi \in \partial_H X (\approx \mathcal{M} \approx K)$ in the above estimate.

Proof. Consider the canonical geodesic $\pi(\mathbf{x}, \mathbf{y}) = \pi(\mathbf{x}, \mathbf{u}) \cup \pi(\mathbf{u}, \mathbf{v}) \cup \pi(\mathbf{v}, \mathbf{y})$ as in Section 3. It follows from Corollary 5.2 that

$$F(\mathbf{x}, \mathbf{y}) \asymp F(\mathbf{x}, \mathbf{u})F(\mathbf{u}, \mathbf{v})F(\mathbf{v}, \mathbf{y}) \asymp \frac{F(\mathbf{x}, \vartheta)}{F(\mathbf{u}, \vartheta)} \cdot F(\mathbf{u}, \mathbf{v}) \cdot \frac{F(\vartheta, \mathbf{y})}{F(\vartheta, \mathbf{v})}.$$

Since $d(\mathbf{u}, \mathbf{v}) \leq M$ (Proposition 3.3), (5.1) yields $\eta^M \leq F(\mathbf{u}, \mathbf{v}) \leq 1$. Applying Theorem 4.1 $(F(\mathbf{w}, \vartheta) = \lambda^{|\mathbf{w}|})$, Theorem 4.6 $(F(\vartheta, \mathbf{w}) \asymp r^{\alpha|\mathbf{w}|})$ and $d(\mathbf{u}, \mathbf{v}) \leq M$ to the other two factors of the above expression, we have

$$F(\mathbf{x}, \mathbf{y}) \asymp \lambda^{d(\mathbf{x}, \mathbf{u})} r^{\alpha d(\mathbf{v}, \mathbf{y})} \asymp \lambda^{d(\mathbf{x}, \mathbf{u})} r^{\alpha d(\mathbf{v}, \mathbf{y})} (\lambda r^{\alpha})^{d(\mathbf{u}, \mathbf{v})/2} = \lambda^{|\mathbf{x}| - (\mathbf{x}|\mathbf{y})} r^{\alpha(|\mathbf{y}| - (\mathbf{x}|\mathbf{y}))}.$$

For the estimate of the Martin kernel, we recall that $K(\mathbf{x}, \mathbf{y}) = \frac{G(\mathbf{x}, \mathbf{y})}{G(\vartheta, \mathbf{y})} = \frac{F(\mathbf{x}, \mathbf{y})}{F(\vartheta, \mathbf{y})}$, and the Gromov product $(\mathbf{x}|\xi)$ is defined for $\xi \in \partial_H X$ (see (3.7)). Using the above estimate of $F(\mathbf{x}, \mathbf{y})$, the second part follows. $\qquad\square$

In the following, we will assume the IFS satisfies the OSC, so the augmented tree $(X, \mathfrak{E})$ has bounded degree (Proposition 3.6). By Theorem 5.1, we identify K with the Martin boundary $\mathcal{M}$ of a λ-QNRW $\{Z_n\}_{n=0}^{\infty}$. We will study the hitting distribution on $\mathcal{M} \approx K$ starting from the root ϑ.

First we construct a projection $\{Z_n\}_{n=0}^{\infty}$ on X onto K. For $\{S_i\}_{i=1}^{N}$ satisfying the OSC, we let O be an open set in the OSC such that $O \cap K \neq \emptyset$ [40]. Define a projection $\iota : X \to K$ by selecting arbitrarily

$$\iota(\mathbf{x}) \in S_{\mathbf{x}}(O \cap K), \quad \mathbf{x} \in X.$$

We extend ι to $\widehat{X} = X \cup \mathcal{M}$ by defining $\iota(\xi) = \xi$ for $\xi \in \mathcal{M}(\approx K)$. Then a sequence $\{\mathbf{x}_n\}_n \subset X$ with $|\mathbf{x}_n| \to \infty$ converges to $\xi \in \widehat{X}$ if and only if $\iota(\mathbf{x}_n) \to \iota(\xi) \in K$ in the Euclidean topology. Hence ι is continuous on $\widehat{X}$, and $\iota|_{\mathcal{M}}$ is the canonical homeomorphism of $\mathcal{M}$ onto K. Also we have

$$\iota(Z_{\infty}) = \lim_{n \to \infty} \iota(Z_n) \quad \mathbb{P}_{\vartheta} - a.e. \tag{5.4}$$

Let ν_{ϑ} be the distribution of Z_{∞}, and ν the distribution of $\iota(Z_{\infty})$. It is direct to check that $\nu = \nu_{\vartheta} \circ \iota^{-1}$ on K. For later use, we shall write $\mathbb{P} = \mathbb{P}_{\vartheta}$ and $\mathbb{E} = \mathbb{E}_{\vartheta}$ if there is no confusion.

For $\ell \geq 1$, let $\tau_{\ell} := \inf\{n \geq 0 : Z_n \in \mathcal{J}_{\ell}\}$ be the first hitting time of $\mathcal{J}_{\ell}$. Then by Theorem 4.6, the distribution of $Z_{\tau_{\ell}}$ is given by

$$\mathbb{P}(Z_{\tau_{\ell}} = \mathbf{x}) = p_{\mathbf{x}} = \mu(S_{\mathbf{x}}(K)), \quad \mathbf{x} \in \mathcal{J}_{\ell}, \tag{5.5}$$

where μ is the self-similar measure associated with the probability weight $\{p_i\}_{i=1}^{N}$. Note that under the OSC, $\mu(S_{\mathbf{x}}(K) \cap S_{\mathbf{y}}(K)) = 0$ for $\mathbf{x} \in \mathcal{J}_m$, $\mathbf{x} \neq \mathbf{y}$.

Proposition 5.4. *For $\mathbf{x} \in \mathcal{J}_m$ and $\ell \geq m$,*

$$\mathbb{P}(\iota(Z_{\tau_{\ell}}) \in S_{\mathbf{x}}(K)) = \mu(S_{\mathbf{x}}(K)). \tag{5.6}$$

Also $\mathbb{P}(\iota(Z_{\tau_{\ell}}) \in S_{\mathbf{x}}(K) \cap S_{\mathbf{y}}(K)) = 0$ for any $\mathbf{x}, \mathbf{y} \in \mathcal{J}_m$, $\mathbf{x} \neq \mathbf{y}$.

Proof. Let $\mathbf{x}, \mathbf{z} \in \mathcal{J}_m$, $\mathbf{z} \neq \mathbf{x}$. Since $\iota(\mathbf{z}) \in S_{\mathbf{z}}(O \cap K)$, the OSC implies that $S_{\mathbf{z}}(O) \cap S_{\mathbf{x}}(O) = \emptyset$. Hence $\iota(\mathbf{z}) \notin \overline{S_{\mathbf{x}}(O)} \supset S_{\mathbf{x}}(K)$. It follows that for $\mathbf{x} \in \mathcal{J}_m$ and $\iota(Z_{\tau_\ell}(\omega)) \in S_{\mathbf{x}}(K)$, $\ell \geq m$, $Z_{\tau_\ell}(\omega)$ must be of the form $\mathbf{xu}$ for some $\mathbf{u} \in \Sigma^*$ and $\mathbf{xu} \in \mathcal{J}_\ell$. By (5.5),

$$\mathbb{P}(\iota(Z_{\tau_\ell}) \in S_{\mathbf{x}}(K)) = \sum_{\mathbf{u} \in \Sigma^* : \mathbf{xu} \in \mathcal{J}_\ell} \mathbb{P}(Z_{\tau_\ell} = \mathbf{xu})$$
$$= \sum_{\mathbf{u} \in \Sigma^* : \mathbf{xu} \in \mathcal{J}_\ell} \mu(S_{\mathbf{xu}}(K)) = \mu(S_{\mathbf{x}}(K)).$$

Using the fact that $\mathbb{P}(Z_{\tau_\ell} = \mathbf{xz}, \ Z_{\tau_\ell} = \mathbf{yz'}) = 0$ and adopting the same argument, the second part follows. $\qquad\square$

To obtain the hitting distribution ν_ϑ, a simple minded approach is to take limit of ℓ in (5.6) to obtain $\mathbb{P}(\iota(Z_\infty) \in S_{\mathbf{x}}(K)) = \mu(S_{\mathbf{x}}(K))$. However this requires that $\chi_{S_{\mathbf{x}}(K)}(\iota(Z_{\tau_\ell})) \to \chi_{S_{\mathbf{x}}(K)}(\iota(Z_\infty))$ a.e., but we cannot make such conclusion directly, as the convergence for the composition $f(\iota(Z_{\tau_\ell}))$ generally requires the continuity of f [4]. We need to go through a more detailed analysis.

Lemma 5.5. *For any two distinct* $\mathbf{x}, \mathbf{x}' \in \mathcal{J}_m$, *we have*

$$\nu(S_{\mathbf{x}}(K) \cap S_{\mathbf{x}'}(K)) = 0.$$

Proof. For $\mathbf{x} \in \mathcal{J}_m$, and $k > m$, we let

$$\mathcal{F}_{k,\mathbf{x}} = \{\mathbf{yu} \in \mathcal{J}_k : \ \mathbf{y} \in \mathcal{J}_m \setminus \{\mathbf{x}\}, \ \mathbf{u} \in \Sigma^*, \ S_{\mathbf{yu}}(K) \cap S_{\mathbf{x}}(K) \neq \emptyset\}$$

Let $F_{k,\mathbf{x}} = \bigcup_{\mathbf{w} \in \mathcal{F}_{k,\mathbf{x}}} S_{\mathbf{w}}(K)$. Then $F_{k,\mathbf{x}}$ is a decreasing sequence of sets in k and the limit set is $\bigcup_{\mathbf{y} \in \mathcal{J}_m \setminus \{\mathbf{x}\}}(S_{\mathbf{y}}(K) \cap S_{\mathbf{x}}(K))$, which is a μ-null set by the OSC. This implies that $\lim_{k \to \infty} \mu(F_{k,\mathbf{x}}) = 0$.

Let $T_{k,\mathbf{x}} = S_{\mathbf{x}}(K) \bigcup F_{k,\mathbf{x}}$. By a similar argument as in Proposition 5.4, we can show that

$$\mathbb{P}(\iota(Z_{\tau_\ell}) \in T_{k,\mathbf{x}} \cap T_{k,\mathbf{y}}) = \mu(T_{k,\mathbf{x}} \cap T_{k,\mathbf{y}}) \qquad \text{for } \ell \geq m. \tag{5.7}$$

Observe that $K \setminus T_{k,\mathbf{x}}$ has positive Euclidian distance to $S_{\mathbf{x}}(K)$. Hence $\iota(Z_\infty) \in S_{\mathbf{x}}(K)$ implies that $\iota(Z_n)$ is eventually in $T_{k,\mathbf{x}}$, which means $\lim_{n \to \infty} \chi_{T_{k,\mathbf{x}}}(\iota(Z_n)) = 1$.

Now for two distinct $\mathbf{x}, \mathbf{x}' \in \mathcal{J}_m$, if $S_{\mathbf{x}}(K) \cap S_{\mathbf{x}'}(K) = \emptyset$, then the lemma is obviously true. Hence assume that $S_{\mathbf{x}}(K) \cap S_{\mathbf{x}'}(K) \neq \emptyset$. Let $T_{k,\mathbf{x}}$ and $T_{k,\mathbf{x}'}$ be defined as the above. Then

$$\nu(S_{\mathbf{x}}(K) \cap S_{\mathbf{x}'}(K)) = \mathbb{P}(\iota(Z_\infty) \in S_{\mathbf{x}}(K) \cap S_{\mathbf{x}'}(K))$$
$$\leq \mathbb{E}\left(\liminf_{n \to \infty} \chi_{T_{k,\mathbf{x}} \cap T_{k,\mathbf{x}'}}(\iota(Z_n))\right)$$
$$\leq \liminf_{n \to \infty} \mathbb{E}(\chi_{T_{k,\mathbf{x}} \cap T_{k,\mathbf{x}'}}(\iota(Z_n))) \qquad \text{(by Fatou's lemma)}$$
$$\leq \liminf_{\ell \to \infty} \mathbb{P}(\iota(Z_{\tau_\ell}) \in T_{k,\mathbf{x}} \cap T_{k,\mathbf{x}'}) \tag{5.8}$$
$$= \mu(T_{k,\mathbf{x}} \cap T_{k,\mathbf{x}'}) \qquad \text{(by (5.7))}.$$

By the limit in the first paragraph, we have $\nu(S_{\mathbf{x}}(K) \cap S_{\mathbf{x}'}(K)) \leq \lim_{k \to \infty} \mu(T_{k,\mathbf{x}} \cap T_{k,\mathbf{x}'}) = \mu((S_{\mathbf{x}}(K) \cap S_{\mathbf{x}'}(K)) = 0$, and this completes the proof. $\qquad\square$

Theorem 5.6. *Let $\{S_i\}_{i=1}^N$ be an IFS with the OSC, and let K be the self-similar set. Suppose $\{Z_n\}_{n=0}^\infty$ is a λ-QNRW on the augmented tree $(X, \mathfrak{E})$ associated with a set of probability weights $\{p_i\}_{i=1}^N$. Then the hitting distribution ν_ϑ is the self-similar measure μ with weights $\{p_i\}_{i=1}^N$ on K.*

Proof. To prove the theorem, it suffices to show that $\nu(S_\mathbf{x}(K)) = p_\mathbf{x}$ for $m \geq 1$ and $\mathbf{x} \in \mathcal{J}_m$. Then ν is the unique self-similar measure with weights $\{p_i\}_{i=1}^N$. To this end, we fix $m \geq 1$ and $\mathbf{x} \in \mathcal{J}_m$. Let

$$U_\mathbf{x} = S_\mathbf{x}(K) \setminus \bigcup_{\mathbf{y} \in \mathcal{J}_m : \mathbf{y} \neq \mathbf{x}} S_\mathbf{y}(K)$$

be the 'interior' of the m-cell $S_\mathbf{x}(K)$. Then any sequence of points in K that has a limit in $U_\mathbf{x}$ must be in $U_\mathbf{x}$ eventually. Therefore $\iota(Z_\infty) \in U_\mathbf{x}$ implies $\lim_{n \to \infty} \chi_{U_\mathbf{x}}(\iota(Z_n)) = 1$. Hence by using Fatou's lemma as in Lemma 5.4, we have

$$
\begin{aligned}
\nu(S_\mathbf{x}(K)) &= \nu(U_\mathbf{x}) && \text{(by Lemma 5.5)} \\
&\leq \liminf_{\ell \to \infty} \ \mathbb{P}(\iota(Z_{\tau_\ell}) \in U_\mathbf{x}) && \text{(as in (5.8))} \\
&\leq \liminf_{\ell \to \infty} \ \mathbb{P}(\iota(Z_{\tau_\ell}) \in S_\mathbf{x}(K)) \\
&= \mu(S_\mathbf{x}(K)) = p_\mathbf{x} && \text{(by Proposition 5.4).}
\end{aligned}
$$

Summing the above inequality over $\mathbf{x} \in \mathcal{J}_m$, we have $1 = \nu(K) \leq \sum_{\mathbf{x} \in \mathcal{J}_m} \nu(S_\mathbf{x}(K)) \leq \sum_{\mathbf{x} \in \mathcal{J}_m} p_\mathbf{x} = 1$. Thus $\nu(S_\mathbf{x}(K)) = p_\mathbf{x}$ for $\mathbf{x} \in \mathcal{J}_m$. This completes the proof of the theorem. $\square$

6   Estimation of the Naïm kernel

Let P be a transient reversible random walk on a countable set X. For $\mathbf{x}, \mathbf{y} \in X$, we define the Θ-kernel by

$$\Theta(\mathbf{x}, \mathbf{y}) = \frac{K(\mathbf{x}, \mathbf{y})}{G(\mathbf{x}, \vartheta)} = \frac{F(\mathbf{x}, \mathbf{y})}{F(\mathbf{x}, \vartheta) G(\vartheta, \vartheta) F(\vartheta, \mathbf{y})}.$$

This kernel is clearly positive and symmetric (by $m(\mathbf{x})G(\mathbf{x}, \mathbf{y}) = m(\mathbf{y})G(\mathbf{y}, \mathbf{x})$) on $X \times X$, and can be extended to $\Theta(\mathbf{x}, \eta)$ on $X \times \widehat{X}$ continuously (as $K(\mathbf{x}, \mathbf{y})$ does), but it is more difficult to extend to $\widehat{X} \times \widehat{X}$. We first recall the extension of $\Theta(\xi, \eta)$ on $\widehat{X} \times \widehat{X}$ in [41]. For $\xi \in \mathcal{M}$, we define the ξ-*process* by setting the transition probability

$$P^\xi(\mathbf{x}, \mathbf{y}) = \frac{P(\mathbf{x}, \mathbf{y})K(\mathbf{y}, \xi)}{K(\mathbf{x}, \xi)}. \tag{6.1}$$

(It is also known as the *h-transform* of P [12], by taking the harmonic function $h = K(\cdot, \xi)$ on X.) Clearly P^ξ is a transition probability since $K(\cdot, \xi)$ is harmonic on X. For the ξ-process, we denote the corresponding ever-visiting probability and Green function by $F^\xi(\cdot, \cdot)$ and $G^\xi(\cdot, \cdot)$ respectively. The following lemma can be found in [45, Corollary 7.51].

Lemma 6.1. *For $\xi \in \mathcal{M}$, let ν_ϑ^ξ be the hitting distribution of the ξ-process. Then $\xi \in \mathcal{M}_{\min}$ if and only if $\nu_\vartheta^\xi = \delta_\xi$, the point mass at ξ.*

For $m \geq 0$, let τ_m^* be the last visit time of $X_m = \bigcup_{k=0}^m \mathcal{J}_k$ by $\{Z_n\}_{n=0}^\infty$. Note that τ_m^* is not a stopping time. For $\xi \in K$ and $\mathbf{z} \in X$, let

$$\ell_m^\xi(\mathbf{z}) = \mathbb{P}_\vartheta^\xi(Z_{\tau_m^*} = \mathbf{z}),$$

where $\mathbb{P}_\vartheta^\xi$ is the probability corresponding to the ξ-process. The following lemma is from [41, Lemma 1.9].

Lemma 6.2. *Let $\xi, \eta \in \mathcal{M}$. Then the sum $\sum_{\mathbf{z} \in X} \ell_m^\xi(\mathbf{z}) \Theta(\mathbf{z}, \eta)$ is increasing in m.*

In view of Lemmas 6.1 and 6.2, Silverstein [41] extended the Naïm kernel on $\mathcal{M} \times \mathcal{M}$ by $\Theta(\xi, \eta) = \lim_{k \to \infty} \sum_{\mathbf{z} \in X} \ell_k^\xi(\mathbf{z}) \Theta(\mathbf{z}, \eta)$ for $\xi, \eta \in \mathcal{M}$. We will apply this to the λ-NRW on the augmented tree. Since the walk satisfies $|Z_{n+1}| - |Z_n| \in \{0, \pm 1\}$, the identity reduces to

$$\Theta(\xi, \eta) = \lim_{k \to \infty} \sum_{\mathbf{z} \in \mathcal{J}_k} \ell_k^\xi(\mathbf{z}) \Theta(\mathbf{z}, \eta). \tag{6.2}$$

Here is the main theorem of this section.

Theorem 6.3. *Let $\{S_i\}_{i=1}^N$ be an IFS that satisfies the OSC, and let K be the self-similar set. Let $\{Z_n\}_{n=0}^\infty$ be a λ-NRW on the augmented tree $(X, \mathfrak{E})$. Then*

$$\Theta(\xi, \eta) \asymp (\lambda r^\alpha)^{-(\xi|\eta)}, \quad \xi, \eta \in \partial_H X, \ \xi \neq \eta,$$

where $\alpha = \dim_H K$. Consequently, by Theorem 3.5, we have

$$\Theta(\xi, \eta) \asymp |\xi - \eta|^{-(\alpha+\beta)}, \quad \xi, \eta \in K, \ \xi \neq \eta,$$

with $\beta = \frac{\log \lambda}{\log r}$.

Proof. Let $\xi, \eta \in \partial_H X$, $\xi \neq \eta$. For $\mathbf{x} \in X$, Proposition 4.1 and Theorem 5.3 imply

$$\Theta(\mathbf{x}, \eta) = \frac{K(\mathbf{x}, \eta)}{F(\mathbf{x}, \vartheta) G(\vartheta, \vartheta)} \asymp (\lambda r^\alpha)^{-(\mathbf{x}|\eta)}. \tag{6.3}$$

We fix two geodesic rays $(\mathbf{x}_n)_{n=0}^\infty$ and $(\mathbf{y}_n)_{n=0}^\infty$ starting from ϑ such that $(\mathbf{x}_n)$ tends to ξ, $(\mathbf{y}_n)$ tends to η, and $(\xi|\eta) = \lim_{n \to \infty} (\mathbf{x}_n|\mathbf{y}_n)$.

We first estimate $\Theta(\xi, \eta)$ from below. Following Lemma 3.7(iii), there exists a positive integer ℓ such that $(\xi|\eta) = (\mathbf{x}_n|\mathbf{y}_n)$ for all $n \geq \ell$. Let $T_\xi = S_{\mathbf{x}_\ell}(K) \cup (\bigcup_{\mathbf{y} \sim_h \mathbf{x}_\ell} S_\mathbf{y}(K))$. For $\mathbf{z} \in \iota^{-1}(T_\xi)$ with $|\mathbf{z}| \geq \ell$, let $\mathbf{z}_\ell \in \mathcal{J}_\ell \cap \pi(\vartheta, \mathbf{z})$. Then $d(\mathbf{z}_\ell, \mathbf{x}_\ell) = 0$ or 1. By Lemma 3.7(ii), (iii) and the triangle inequality (see (3.6)), we have

$$(\mathbf{z}|\eta) \geq (\mathbf{z}_\ell|\mathbf{y}_\ell) \geq (\mathbf{x}_\ell|\mathbf{y}_\ell) - \frac{1}{2} = (\xi|\eta) - \frac{1}{2}. \tag{6.4}$$

As K is identified with $\mathcal{M}$ (Theorem 5.1), Lemma 6.1 implies that $\iota(Z_n)$ converges to ξ $\mathbb{P}_\vartheta^\xi$-almost surely. Let $t_\xi = \sup\{n \geq \ell : \iota(Z_n) \notin T_\xi\} + 1$, i.e., the first time that Z_n enters

24

$\iota^{-1}(T_\xi)$ and then stays inside. Clearly $\mathbb{P}_\vartheta^\xi(t_\xi < \infty) = 1$. Therefore we can choose an integer m_1 such that $\mathbb{P}_\vartheta^\xi(t_\xi \leq \tau_{m_1}^*) \geq 1/2$. By Lemma 6.2, (6.3) and (6.4),

$$
\begin{aligned}
\Theta(\xi, \eta) &\geq \sum_{\mathbf{z} \in \mathcal{J}_{m_1} \cap \iota^{-1}(T_\xi)} \ell_{m_1}^\xi(\mathbf{z}) \Theta(\mathbf{z}, \eta) \\
&\geq c_1 (\lambda r^\alpha)^{-(\xi|\eta)} \sum_{\mathbf{z} \in \mathcal{J}_{m_1} \cap \iota^{-1}(T_\xi)} \ell_{m_1}^\xi(\mathbf{z}) \\
&\geq c_1 (\lambda r^\alpha)^{-(\xi|\eta)} \cdot \mathbb{P}_\vartheta^\xi(t_\xi \leq \tau_{m_1}^*) \geq c_2 (\lambda r^\alpha)^{-(\xi|\eta)}.
\end{aligned}
$$

To obtain the upper bound, we first observe that $\ell_k^\xi(\mathbf{z}) \leq G^\xi(\vartheta, \mathbf{z}) = G(\vartheta, \mathbf{z})K(\mathbf{z}, \xi)$. This together with the reversibility $G(\vartheta, \mathbf{z}) = (m(\mathbf{z})/m(\vartheta))G(\mathbf{z}, \vartheta)$ imply

$$
\begin{aligned}
\Theta(\xi, \eta) &\leq (m(\vartheta))^{-1} \lim_{k \to \infty} \sum_{\mathbf{z} \in \mathcal{J}_k} m(\mathbf{z}) K(\mathbf{z}, \xi) K(\mathbf{z}, \eta) \\
&\asymp \lim_{k \to \infty} \sum_{\mathbf{z} \in \mathcal{J}_k} (\lambda r^\alpha)^{|\mathbf{z}| - (\mathbf{z}|\mathbf{x}_k) - (\mathbf{z}|\mathbf{y}_k)}
\end{aligned}
$$

(the $\asymp$ is a consequence of Proposition 4.5, Theorem 5.3 and Lemma 3.7(ii)). Let $\Phi(\mathbf{z})$ denote the summand in the above sum. Using Lemma 3.7(iii) and the triangle inequality, we have

$$
\Phi(\mathbf{z}) \leq (\lambda r^\alpha)^{d(\mathbf{x}_k, \mathbf{y}_k)/2 - k} = (\lambda r^\alpha)^{-(\mathbf{x}_k|\mathbf{y}_k)} \leq (\lambda r^\alpha)^{-(\xi|\eta)}, \quad \mathbf{z} \in \mathcal{J}_k. \tag{6.5}
$$

For any $\mathbf{w} \in \mathcal{J}_k$, $k \geq 1$, let $\mathcal{F}_k(\mathbf{w}) = \{\mathbf{z} \in \mathcal{J}_k : d(\mathbf{z}, \mathbf{w}) \leq M\}$, where M is the maximal length of the horizontal geodesics. As $(X, \mathfrak{E})$ is of bounded degree (Lemma 3.6), it is direct to check that $\#\mathcal{F}_k(\mathbf{w}) \leq C_1$ for some $C_1 > 0$ independent of $\mathbf{w}$ and k. Our aim is to show that for sufficient large n, the sum $\sum_{\mathbf{z} \in \mathcal{J}_{k+n}} \Phi(\mathbf{z})$ is concentrated on the n-th descendants of $\mathcal{F}_k(\mathbf{x}_k) \cup \mathcal{F}_k(\mathbf{y}_k)$. This enables us to obtain the desired upper bound.

To this end, for $\mathbf{z} \in \mathcal{J}_k$, we let $\mathcal{J}_n(\mathbf{z})$ denote the set of n-th descendants of $\mathbf{z}$ in $\mathcal{J}_{k+n}$. For $\mathbf{z} \notin \mathcal{F}_k(\mathbf{x}_k)$ and $\mathbf{z}^{(1)} \in \mathcal{J}_1(\mathbf{z})$, it is clear that $d(\mathbf{z}^{(1)}, \mathbf{x}_{k+1}) \geq d(\mathbf{z}, \mathbf{x}_k) > M$, which yields $\mathbf{z}^{(1)} \notin \mathcal{F}_{k+1}(\mathbf{x}_{k+1})$. Moreover, by the same argument as in Lemma 3.7(iii), we have $(\mathbf{z}^{(1)}|\mathbf{x}_{k+1}) = (\mathbf{z}|\mathbf{x}_k)$. Let $\mathcal{F}_k^c = \mathcal{J}_k \setminus (\mathcal{F}_k(\mathbf{x}_k) \cup \mathcal{F}_k(\mathbf{y}_k))$. It follows that $\Phi(\mathbf{z}^{(1)}) = \lambda r^\alpha \Phi(\mathbf{z})$ for any $\mathbf{z} \in \mathcal{F}_k^c$, and hence inductively,

$$
\Phi(\mathbf{z}^{(n)}) = (\lambda r^\alpha)^n \Phi(\mathbf{z}), \qquad \mathbf{z} \in \mathcal{F}_k^c, \ \ \mathbf{z}^{(n)} \in \mathcal{J}_n(\mathbf{z}). \tag{6.6}
$$

Also by the same proof as Lemma 3.1, we have

$$
r^{-\alpha n} \leq \#\mathcal{J}_n(\mathbf{z}) \leq r^{-\alpha(n+1)}. \tag{6.7}
$$

Now choose n_0 such that $\delta = r^{-\alpha} \lambda^{n_0} < 1$. It follows from (6.5)-(6.7) that for any $k \geq 1$,

$$
\begin{aligned}
\sum_{\mathbf{w} \in \mathcal{J}_{k+n_0}} \Phi(\mathbf{w}) &= \left(\sum_{\mathbf{z} \in \mathcal{F}_k^c} \sum_{\mathbf{w} \in \mathcal{J}_{n_0}(\mathbf{z})} + \sum_{\mathbf{z} \in \mathcal{F}_k(\mathbf{x}_k)} \sum_{\mathbf{w} \in \mathcal{J}_{n_0}(\mathbf{z})} + \sum_{\mathbf{z} \in \mathcal{F}_k(\mathbf{y}_k)} \sum_{\mathbf{w} \in \mathcal{J}_{n_0}(\mathbf{z})} \right) \Phi(\mathbf{w}) \\
&\leq r^{-\alpha} \lambda^{n_0} \sum_{\mathbf{z} \in \mathcal{F}_k^c} \Phi(\mathbf{z}) + r^{-\alpha(n_0+1)} \left(\sum_{\mathbf{z} \in \mathcal{F}_k(\mathbf{x}_k)} + \sum_{\mathbf{z} \in \mathcal{F}_k(\mathbf{y}_k)} \right) (\lambda r^\alpha)^{-(\xi|\eta)} \\
&\leq \delta \sum_{\mathbf{z} \in \mathcal{J}_k} \Phi(\mathbf{z}) + C_2 (\lambda r^\alpha)^{-(\xi|\eta)}.
\end{aligned}
$$

By iteration, we have for any integer $q \geq 1$, and any $k \in \{1, 2, \cdots, n_0\}$,

$$\sum_{\mathbf{w} \in \mathcal{J}_{k+n_0 q}} \Phi(\mathbf{w}) \leq \delta^q \sum_{\mathbf{z} \in \mathcal{J}_k} \Phi(\mathbf{z}) + \frac{C_2}{1-\delta}(\lambda r^\alpha)^{-(\xi|\eta)}$$

$$\leq \left(\#\mathcal{J}_k + \frac{C_2}{1-\delta}\right) \cdot (\lambda r^\alpha)^{-(\xi|\eta)} \qquad \text{(by (6.5))}$$

$$\leq \left(\#\mathcal{J}_{n_0} + \frac{C_2}{1-\delta}\right) \cdot (\lambda r^\alpha)^{-(\xi|\eta)}.$$

This completes the upper estimation of $\Theta(\xi, \eta)$. $\qquad\square$

We remark that for homogeneous IFS, the SRW is a special example of λ-NRW with $\lambda = 1/N = r^\alpha$. Hence by Theorem 6.3, $\Theta(\xi, \eta) \asymp |\xi - \eta|^{-2\alpha}$. Also for the example in Section 3, Remark 2 (see Figure 2), the above theorem (with slight adjustment) shows that for the SRW there, the Naïm kernel $\Theta(\xi, \eta) \asymp |\xi - \eta|^{-2}$, which resembles the classical case of the kernel in the Douglas integral (1.2) on the unit circle.

We also remark that there is an analogous result for the more general random walks on N-ary trees investigated by Kigami [27]. For the λ-QNRW on the N-homogeneous tree T defined by a set of probability weights $\{p_i\}_{i=1}^N$, the effective resistance between $\mathbf{x} \in T$ and $\Sigma_{\mathbf{x}}$ with respect to the sub-tree $T_{\mathbf{x}}$ is $R_{\mathbf{x}} = \lambda^{|\mathbf{x}|+1}/((1-\lambda)p_{\mathbf{x}})$, which can be derived from the combined resistance by iteration. Since the hitting distribution ν is the self-similar measure of $\{p_i\}_{i=1}^N$ on the Cantor set Σ^∞ as Martin boundary, we have $\nu(\Sigma_{\mathbf{x}}^\infty) = p_{\mathbf{x}}$. In the notations of [27], $D_{\mathbf{x}} := \nu(\Sigma_{\mathbf{x}}^\infty)R_{\mathbf{x}} = \lambda^{|\mathbf{x}|+1}/(1-\lambda)$, $\lambda_{\mathbf{x}} = 1/D_{\mathbf{x}} = (1-\lambda)/\lambda^{|\mathbf{x}|+1}$, and $N(\xi, \eta) = (\xi|\eta)$. Following [27, Theorem 5.6], we can rewrite his jump kernel $J(\xi, \eta)$ for such λ-QNRW as

$$J(\xi, \eta) = \frac{1-\lambda}{2\lambda}\left(1 + \sum_{m=1}^{(\xi|\eta)} \frac{1-\lambda}{\lambda^m p_{[\xi,\eta]_m}}\right) \asymp \lambda^{-(\xi|\eta)}p_{[\xi,\eta]}^{-1}, \quad \xi, \eta \in \Sigma^\infty, \qquad (6.8)$$

where $[\xi, \eta]_m$ is the unique word in Σ^m such that both ξ and η belong to $\Sigma_{[\xi,\eta]_m}$, and $[\xi, \eta] := [\xi, \eta]_{(\xi|\eta)}$. In particular, for the λ-NRW on the N-homogeneous tree, $p_{\mathbf{x}} = N^{-|\mathbf{x}|}$, and we have the following estimate:

$$J(\xi, \eta) = \frac{(1-\lambda)(N-1)}{2(N-\lambda)} + \frac{N(1-\lambda)^2(N/\lambda)^{(\xi|\eta)}}{2\lambda(N-\lambda)} \asymp (N/\lambda)^{(\xi|\eta)} = (\lambda r^\alpha)^{-(\xi|\eta)}, \qquad (6.9)$$

where $\alpha = |\log N/\log r|$ is the Hausdorff dimension of the Cantor set Σ^∞. It coincides with Theorem 6.3.

7 Induced Dirichlet forms

In this section we use the λ-NRW and the Naïm kernel to induce an energy form on K, and make some remarks about the known results for it to be a Dirichlet form. First we summarize some general results that hold for all transient reversible random walks.

Let $(X, \mathfrak{E})$ be a countably infinite, connected, locally finite graph, and let $\{Z_n\}_{n=0}^\infty$ be a transient reversible random walk on (X, E) with conductance c and total conductance m. The *graph energy* of a (real) function f on X is defined by

$$\mathcal{E}_X[f] = \frac{1}{2} \sum_{x,y \in X, x \sim y} c(x,y)(f(x) - f(y))^2.$$

We let $\mathcal{D}_X$, the domain of $\mathcal{E}_X$, be the set of $f : X \to \mathbb{R}$ with $\mathcal{E}_X[f] < \infty$. Since $P(x,y) = c(x,y)/m(x)$, we have

$$\mathcal{E}_X[f] = \frac{1}{2} \sum_{x \in X} m(x) \sum_{y \in X, y \sim x} P(x,y)(f(y) - f(x))^2$$

$$= \frac{1}{2} \sum_{x \in X} m(x)\mathbb{E}_x[(f(Z_1) - f(Z_0))^2]. \tag{7.1}$$

The graph energy defines a non-negative definite symmetric form $\mathcal{E}_X(f,g)$ on $\mathcal{D}_X$ by polarization. Moreover, if we fix any $x_0 \in X$, then $\mathcal{D}_X$ is a Hilbert space under the inner product

$$\langle f, g \rangle_0 := f(x_0)g(x_0) + \mathcal{E}_X(f,g), \quad f, g \in \mathcal{D}_X,$$

and the convergence in $(\mathcal{D}_X, \mathcal{E}_X)$ implies pointwise convergence [44, Lemma 2.4].

Proposition 7.1. *Let $\mathcal{M}$ be the Martin boundary of $\{Z_n\}_{n=1}^\infty$, and let $\nu = \nu_\vartheta$ be the hitting distribution. Suppose $f \in \mathcal{D}_X$ and is harmonic. Then $\{f(Z_n)\}_{n=0}^\infty$ converges almost surely and in L^2, and there exists $u \in L^2(\mathcal{M}, \nu)$ such that $\lim_{n \to \infty} f(Z_n) = u(Z_\infty)$. Moreover, u is uniquely determined ν-a.e., and $f = Hu$ where*

$$Hu(x) = \int_\mathcal{M} u(\xi)K(x,\xi)\, d\nu(\xi), \quad x \in X.$$

Proof. The first statement was actually proved in [3, Theorem 1.1] without the assumption that f is harmonic. We only need this result when f is harmonic, and the proof is easy for this case. Fix $x \in X$, we claim that $\mathbb{E}_x[f(Z_n)^2]$ is bounded. In fact, Since f is harmonic, $\{f(Z_n)\}_{n=0}^\infty$ is a martingale under $\mathbb{P}_x$. It follows that

$$\mathbb{E}_x[f(Z_n)^2] \leq f(x)^2 + \sum_{k=1}^\infty \mathbb{E}_x\big[(f(Z_k) - f(Z_{k-1}))^2\big]$$

$$= f(x)^2 + \sum_{k=1}^\infty \sum_{y \in X} P^k(x,y)\mathbb{E}_y\big[(f(Z_1) - f(Z_0))^2\big]$$

$$= f(x)^2 + \sum_{y \in X} G(x,y)\mathbb{E}_y\big[(f(Z_1) - f(Z_0))^2\big].$$

Using the reversibility, we have $G(x,y) = \frac{m(y)}{m(x)}G(y,x) \leq \frac{m(y)}{m(x)}G(x,x)$. Therefore, by (7.1),

$$\mathbb{E}_x[f(Z_n)^2] \leq f(x)^2 + 2\frac{G(x,x)}{m(x)}\mathcal{E}_X[f],$$

and the claim follows. Consequently, $\{f(Z_n)\}_{n=0}^\infty$ is an L^2-bounded martingale under $\mathbb{P}_x$. The martingale convergence theorem implies the almost surely convergence and L^2-convergence.

Let $Y = \lim_{n\to\infty} f(Z_n)$. Note that Y is a *final* random variable, i.e. $Y \circ \theta = Y$ a.s. where θ is the shift operator on the path space. Hence there exists a measurable function u on $\mathcal{M}$ such that $Y = u(Z_\infty)$ a.s. (under any $\mathbb{P}_x$) (see [12]). It is clear that u is unique ν-a.e. Now since $Y \in L^2(\mathbb{P}_\vartheta)$, we have

$$\int_{\mathcal{M}} u(\xi)^2 d\nu(\xi) = \mathbb{E}_\vartheta(u(Z_\infty)^2) = \mathbb{E}_\vartheta(Y^2) < \infty.$$

That $f = Hu$ follows from

$$f(Z_n) = \mathbb{E}[u(Z_\infty)|\mathcal{F}_n] = \mathbb{E}[u(Z_\infty)|Z_n] = \mathbb{E}_{Z_n}[u(Z_\infty)] = Hu(Z_n),$$

and the irreducibility of the chain. $\square$

For a reversible random walk on $(X, \mathfrak{E})$, the energy form induces a bilinear form $(\mathcal{E}_{\mathcal{M}}, \mathcal{D}_{\mathcal{M}})$ on $L^2(\mathcal{M}, \nu)$ defined by

$$\mathcal{E}_{\mathcal{M}}(u, v) = \mathcal{E}_X(Hu, Hv), \quad u, v \in \mathcal{D}_{\mathcal{M}}, \tag{7.2}$$

where $\mathcal{D}_{\mathcal{M}} = \{u \in L^2(\mathcal{M}, \nu) : Hu \in \mathcal{D}_X\}$ is the domain of $\mathcal{E}_{\mathcal{M}}$. It follows from Proposition 7.1 that

$$\mathcal{D}_{\mathcal{M}} = \{u \in L^2(\mathcal{M}, \nu) : \ \exists \text{ harmonic } f \in \mathcal{D}_X \text{ such that } u(Z_\infty) = \lim_{n\to\infty} f(Z_n) \text{ a.s.}\}. \tag{7.3}$$

Theorem 7.2. [41, Theorem 3.5] *The induced bilinear form $(\mathcal{E}_{\mathcal{M}}, \mathcal{D}_{\mathcal{M}})$ has the expression*

$$\mathcal{E}_{\mathcal{M}}[u] = \mathcal{E}_X[Hu] = \frac{1}{2} m(\vartheta) \int_{\mathcal{M}} \int_{\mathcal{M}} (u(\xi) - u(\eta))^2 \Theta(\xi, \eta) d\nu(\xi) d\nu(\eta), \quad u \in \mathcal{D}_{\mathcal{M}},$$

where $\Theta(\xi, \eta)$ is the Naïm kernel defined in (6.2). Moreover, $\mathcal{D}_{\mathcal{M}} = \{u \in L^2(\mathcal{M}, \nu) : \mathcal{E}_{\mathcal{M}}[u] < \infty\}$.

As a direct consequence of Theorem 6.3 and the above, we have

Theorem 7.3. *Let $\{S_i\}_{i=1}^N$ be an IFS on $\mathbb{R}^d$ that satisfies the OSC, and let K be the self-similar set. Let $\{Z_n\}_{n=0}^\infty$ be a λ-NRW on the augmented tree $(X, \mathfrak{E})$. Then*

$$\mathcal{E}_K[u] \asymp \int_K \int_K (u(\xi) - u(\eta))^2 \, |\xi - \eta|^{-(\alpha+\beta)} d\nu(\xi) d\nu(\eta), \quad u \in \mathcal{D}_K,$$

where $\alpha = \dim_H K$, and $\beta = \frac{\log \lambda}{\log r}$.

Next we recall the definition of Dirichlet form (see [6, 14]). Let $\mathfrak{X}$ be a locally compact separable metric space together with a positive Radon measure μ such that $\text{supp}(\mu) = \mathfrak{X}$; also let $C_0(\mathfrak{X})$ denote the space of continuous function on $\mathfrak{X}$ with compact support.

Definition 7.4. *A Dirichlet form $(\mathcal{E}, \mathcal{D})$ on $L^2(\mathfrak{X}, \mu)$ is a bilinear form which is symmetric, non-negative definite, closed, Markovian and densely defined on $L^2(\mathfrak{X}, \mu)$. It is called regular if the subspace $\mathcal{D} \cap C_0(\mathfrak{X})$ is dense in $\mathcal{D}$ with the $\mathcal{E}_1$-norm, and is dense in $C_0(\mathfrak{X})$ with the supremum norm. It is called local if for two functions $u, v \in \mathcal{D}$ having disjoint compact supports, $\mathcal{E}(u, v) = 0$.*

It is easy to see that the bilinear form $\mathcal{E}_K(\cdot,\cdot)$ in Theorem 7.3 is symmetric, non-negative definite, closed and Markovian. For it to be a Dirichlet form, we need to show that its domain $\mathcal{D}_K$ is dense in $L^2(K,\nu)$. In view of the fact that the Naïm kernel in Theorem 7.3 satisfies the estimate $\Theta(\xi,\eta) \asymp |\xi - \eta|^{-(\alpha+\beta)}$, we introduce the following Besov space on $L^2(\mathfrak{X},\mu)$ (see [16, 17, 21, 22, 42]).

For convenience, we assume that $\mathfrak{X} \subset \mathbb{R}^d$ and is equipped with the Euclidean distance; we also assume that $\mathfrak{X}$ has Hausdorff dimension α, and $\mu(B(x;r)) \asymp r^\alpha$ for all $x \in \mathfrak{X}$ and $0 < r < 1$. We call such $\mathfrak{X}$ an α-set [23]. For $\sigma > 0$, and for $u \in L^2(\mathfrak{X},\mu)$, we define

$$\mathcal{N}_{2,2}^{\alpha;\sigma}(u) = \int_0^\infty \frac{dr}{r}\frac{1}{r^{\alpha+2\sigma}} \iint\limits_{\{\xi,\eta\in\mathfrak{X}:|\xi-\eta|<r\}} (u(\xi) - u(\eta))^2 d\mu(\xi)d\mu(\eta), \tag{7.4}$$

and the *Besov spaces* $\Lambda_{2,2}^{\alpha;\sigma} = \{u \in L^2(\mathfrak{X},\mu) : \mathcal{N}_{2,2}^{\alpha;\sigma}(u) < \infty\}$ with the associated norms $\|u\|_{\Lambda_{2,2}^{\alpha,\sigma}}^2 = \|u\|_2^2 + \mathcal{N}_{2,2}^{\alpha;\sigma}(u)$. The space can be trivial when σ is a large value. (For example, in Euclidean space $\mathbb{R}^d$, $\Lambda_{2,2}^{d,1} = \{0\}$.) We introduce an important quantity $\beta^* \in [0,+\infty]$ which is intrinsic to the underlying space $\mathfrak{X}$ [16, 42]:

$$\beta^* := \sup\{\beta > 0 : \Lambda_{2,2}^{\alpha,\beta/2} \cap C_0(\mathfrak{X}) \text{ is dense in } C_0(\mathfrak{X})\}. \tag{7.5}$$

It is called the *critical exponent* of the family $\left\{\Lambda_{2,2}^{\alpha,\beta/2}\right\}_{\beta>0}$. The value of β^* is already known for some standard cases: for the Euclidean space $\mathbb{R}^d$, we have $\alpha = d, \beta^* = 2$. For d-dimensional Sierpiński gasket, then $\alpha = \log(d+1)/\log 2$, and $\beta^* = \log(d+3)/\log 2$ [22]. There are also extensions to nested fractals and related Besov spaces [38, 39], and evaluation of some other specific cases [30]. For Cantor-type set as the boundary of an infinite binary tree, it follows from [27, Theorem 5.6] that $\beta^* = \infty$. In general, we know that if a metric measure space $(\mathfrak{X},\mu)$ satisfies $\mu(B(x;r)) \asymp r^\alpha$ for all $x \in \mathfrak{X}$ and $0 < r < 1$, then $\beta^* \geq 2$. If in addition $\mathfrak{X}$ satisfies *the chain condition* [16], then $\beta^* \leq \alpha + 1$.

Continuing the statement in Theorem 7.3, we have the following conclusion.

Theorem 7.5. *With the same assumption and notations as in Theorem 7.3, we have*

$$\mathcal{E}_K[u] \asymp \mathcal{N}_{2,2}^{\alpha,\beta/2}(u), \quad u \in \mathcal{D}_K, \tag{7.6}$$

and $\mathcal{D}_K$ is the Besov space $\Lambda_{2,2}^{\alpha,\beta/2}$. Therefore, if $\beta < \beta^*$, then $(\mathcal{E}_K,\mathcal{D}_K)$ is a non-local Dirichlet form on $L^2(K,\nu)$.

The proof of (7.6) and that $\mathcal{D}_K$ is the Besov space $\Lambda_{2,2}^{\alpha,\beta/2}$ are in [42]. The following proposition deals with the regularity of the induced Dirichlet form.

Proposition 7.6. *Let $\{S_i\}_{i=1}^N$ be an IFS of contractive similitudes with the OSC, and let K be the self-similar set. For a λ-NRW on the augmented tree $(X,\mathfrak{E})$, if either (i) $\lambda \in (r^2,1)$, or (ii) $\alpha < \beta^*$ and $\lambda \in (r^{\beta^*}, r^\alpha)$, then the induced form $(\mathcal{E}_K,\mathcal{D}_K)$ is a regular non-local Dirichlet form on $L^2(K,\nu)$.*

Moreover, if $2 \leq \alpha < \beta^*$ and $\lambda \in [r^\alpha, r^2]$, let $\mathcal{D}_K^* = \overline{C(K) \cap \mathcal{D}_K}$, where the closure is taken under the norm $\|\cdot\|_{\Lambda_{2,2}^{\alpha,\beta/2}}$. Then $(\mathcal{E}_K,\mathcal{D}_K^*)$ is a regular non-local Dirichlet form on $L^2(K,\nu)$.

Proof. It is well-known that K equipped with Euclidean metric and α-Hausdorff measure is an α-set. As $\beta^* \geq 2$, the assumption $\lambda \in (r^2, 1)$ in (i) implies $0 < \beta < 2(\leq \beta^*)$. It follows from [42, Theorem 3] that the Dirichlet form $(\mathcal{E}_K, \mathcal{D}_K)$ is regular.

For (ii), we have $\alpha < \beta < \beta^*$, and the domain $\mathcal{D}_K$ is embedded into the Lipschitz space of Hölder exponent $(\beta - \alpha)/2$ (see [16, Theorem 4.11]). Hence $(\mathcal{E}_K, \mathcal{D}_K)$ is regular.

For the last part, we have $2 \leq \beta \leq \alpha < \beta^*$. Let $\beta_0 = (\alpha + \beta^*)/2$. Then $\alpha < \beta_0 < \beta^*$. Hence by the above paragraph, $\mathcal{D}_K^{(\beta_0)}$ consists of certain Lipschitz functions of order $(\beta_0 - \alpha)/2$, and $\mathcal{D}_K^{(\beta_0)} = C(K) \cap \mathcal{D}_K^{(\beta_0)} \subset C(K) \cap \mathcal{D}_K \subset \mathcal{D}_K^*$. This implies that $C(K) \cap \mathcal{D}_K$ is dense in $L^2(K, \nu)$ and in $C(K)$, and $(\mathcal{E}_K, \mathcal{D}_K^*)$ is regular. $\square$

Remark. In the proof of the last part, we cannot prove $C(K) \cap \mathcal{D}_K$ is dense in $\mathcal{D}_K$ with the $\mathcal{E}_1$-norm. Hence we use $\mathcal{D}_K^* = \overline{C(K) \cap \mathcal{D}_K}$ instead of the original $\mathcal{D}_K$. Also we do not know the regularity of the Dirichlet form for $2 \leq \beta < \beta^* \leq \alpha$.

By [14, Theorems 7.2.1, 7.2.2], we know that a regular Dirichlet form $(\mathcal{E}_K, \mathcal{D}_K)$ on $L^2(K, \nu)$ generates an associated Hunt jump process with transition density function (heat kernel) $p(t, \xi, \eta)$. For $0 < \beta < 2$, i.e., $\lambda \in (r^2, 1)$, it has been shown in [7] that the heat kernel satisfies the following estimate:

$$p(t, \xi, \eta) \asymp \min \left\{ t^{-\alpha/\beta}, \frac{t}{|\xi - \eta|^{\alpha+\beta}} \right\}, \quad \xi, \eta \in K, \ 0 < t \leq 1.$$

We do not have estimates of $p(t, \xi, \eta)$ for $\alpha < \beta < \beta^*$ or for other cases in general.

The Besov spaces at the critical exponent β^* are particularly important, and are not completely understood. In fact, there is another class of Besov spaces $\Lambda_{2,\infty}^{\alpha,\sigma}$ involved (see [16, 17]). For example on $\mathbb{R}^d$, $\beta^* = 2$, and we have $\Lambda_{2,2}^{d,1}(\mathbb{R}^d) = \{0\}$, but $\Lambda_{2,\infty}^{d,1}$ equals the Sobolev space $W_2^1(\mathbb{R}^d)$, the domain of classical Dirichlet form that generates the Gaussian heat kernel. Similar situations hold for self-similar sets (and more general metric measure spaces) that admit local Dirichlet forms and subgaussian kernels [16, 17]. In forthcoming paper [29], we will give a more detailed discussion, and provide a criterion to determine the exponent β^*.

References

[1] A. Ancona, Negatively curved manifolds, elliptic operators, and the Martin boundary. Ann. Math. **125** (1987), 495–536.

[2] A. Ancona, Positive harmonic functions and hyperbolicity, Potential Theory: Surveys and Problems, Lecture Notes in Math. vol. 1344, pp. 1–23. Springer, Heidelberg (1988).

[3] A. Ancona, R. Lyons, Y. Peres, Crossing estimates and convergence of Dirichlet functions along random walk and diffusion paths, Ann. Probab. **27** (1999), 970–989.

[4] R. Bartle, J. Joichi, The preservation of convergence of measurable functions under composition, Proc. Amer. Math. Soc. **12** (1961), 122–126.

[5] A. Bendikov, A. Grigor'yan, C. Pittet, W. Woess, Isotropic Markov semigroups on ultra-metric spaces, Russian Math. Survey **69** (2014), 589–680.

[6] Z.Q. Chen, M. Fukushima, Symmetric Markov Processes, time changes and boundary theory, London Math. Soc. Mono. vol. 35. Princeton University Press, Princeton (2011).

[7] Z.Q. Chen, T. Kumagai, Heat kernel estimates for stable-like processes on d-sets. Stochastic Processes and their Applications **108** (2003), 27–62.

[8] M. Denker, H. Sato, Sierpiński gasket as a Martin boundary I: Martin kernel, Potential Anal. **14** (2001), 211–232.

[9] M. Denker, H. Sato, Sierpiński gasket as a Martin boundary II: The intrinsic metric, Publ. RIMS, Kyoto Univ. **35** (1999), 769–794.

[10] M. Denker, H. Sato, Reflections on harmonic analysis of the Sierpiński gasket, Math. Nachr. **241**, 32–55 (2002).

[11] J. Doob, Boundary properties of functions with finite Dirichlet integrals, Ann. Inst. Fourier **12** (1962), 573–621.

[12] E. Dynkin, Boundary theory of Markov processes (the discrete case), Russian Math. Surveys **24** (1969), 1–42.

[13] K. Falconer, Fractal Geometry. Mathematical Foundations and Applications, John Wiley and Sons, Ltd., Chichester (1990).

[14] M. Fukushima, Y. Oshima, M. Takeda, Dirichlet forms and symmetric Markov processes, De Gruyter Studies in Mathematics, vol. 19. Walter de Gruyter & Co., Berlin (1994).

[15] A. Georgakopoulos, Group-walk random graphs, Groups, Graphs and Random walks, London Math. Soc. Lecture Note Series, vol. 436, pp. 190–204. Cambridge University Press, Cambridge (2017).

[16] A. Grigor'yan, J.X. Hu, K.S. Lau, Heat kernels on metric-measure spaces and an application to semilinear elliptic equations, Trans. Amer. Math. Soc. **355** (2003), 2065–2095.

[17] A. Grigor'yan, J.X. Hu, K.S. Lau, Heat kernels on metric spaces, Geometry and Analysis of Fractals. Springer Proc. Math. Stat. vol. 88, pp. 147–207. Springer, Heidelberg (2014).

[18] A. Grigor'yan, J.X. Hu, K.S. Lau, Estimates of heat kernels for non-local regular Dirichlet forms. Tran. Amer. Math. Soc. **366** (2014), 6397–6441.

[19] A. Grigor'yan, M. Yang, Determination of walk dimension of Sierpiński gasket without using diffusion, J. Fractal Geom. (to appear).

[20] M. Gromov, Hyperbolic groups, Essays in group theory, Math. Sci. Res. Inst. Publ., vol. 8, pp. 75–263. Springer, New York (1987).

[21] J.X. Hu, T. Kumagai, Nash-type inequalities and heat kernels for non-local Dirichlet forms, Kyushu J. Math. **60** (2006), 245–265.

[22] A. Jonsson, Brownian motion on fractals and function spaces, Math. Zeit. **222** (1996), 495–504.

[23] A. Jonsson, H. Wallin, Function spaces on subsets of $\mathbb{R}^n$, Math. Reports 2, Part 1. Harwood Academic Publ. (1984).

[24] H. Ju, K.S. Lau, X.Y. Wang, Post-critically finite fractal and Martin boundary, Tran. Amer. Math. Soc. **364** (2012), 103–118.

[25] V. Kaimanovich, Random walks on Sierpiński graphs: hyperbolicity and stochastic homogenization, Fractals in Graz 2001, Trends Math., pp. 145–183. Birkhuser, Basel (2003).

[26] J. Kigami, Analysis on Fractals. Cambridge Tracts in Mathematics vol. 143. Cambridge University Press, Cambridge (2001).

[27] J. Kigami, Dirichlet forms and associated heat kernels on the Cantor set induced by random walks on trees, Adv. Math. **225** (2010), 2674–2730.

[28] J. Kigami, Transitions on a noncompact Cantor set and random walks on its defining tree, Ann. Inst. Henri Poincaré Probab. Stat. **49** (2013), 1090–1129.

[29] S.L. Kong, K.S. Lau, Critical exponents of induced Dirichlet forms on self-similar sets, arXiv:1612.01708.

[30] T. Kumagai, Estimates of transition densities for Brownian motion on nested fractals, Prob. Th. Rel. Fields **96** (1993), 205-224.

[31] K.S. Lau, S.M. Ngai, Martin boundary and exit space on the Sierpiński gasket. Sci. China Math. **55** (2012), 475–494.

[32] K.S. Lau, S.M. Ngai, Boundary theory on Hata tree, Nonlinear Analysis **95** (2014), 292–307.

[33] K.S. Lau, X.Y. Wang, Self-similar sets as hyperbolic boundaries, Indiana Univ. Math. J. **58** (2009), 1777–1795.

[34] K.S. Lau, X.Y. Wang, Denker-Sato type Markov chains on self-similar sets. Math. Zeit. **280** (2015), 401–420.

[35] K.S. Lau, X.Y. Wang, On hyperbolic graphs induced by iterated function systems, Adv. Math. **313** (2017), 357–378.

[36] L. Naïm, Sur le role de la frontiere de R. S. Martin dans la theorie du potentiel, Ann. Inst. Fourier **7** (1957), 183–281.

[37] V. Nekrashevych, Self-similar groups, Math. Surveys & Monog. vol. 117. Amer. Math. Soc., Providence (2005).

[38] K. Pietruska-Pałuba, Some function spaces related to the brownian motion on simple nested fractals, Stochastics and Stochastics Reports **67** (1999), 267–285.

[39] K. Pietruska-Pałuba, On function spaces related to fractional diffusions on d-sets, Stochastics and Stochastics Reports **70** (2000), 153–164.

[40] A. Schief, Separation properties for self-similar sets, Proc. Amer. Math. Soc. **122** (1994), 111–115.

[41] M. Silverstein, Classification of stable symmetric Markov chains, Indiana Univ. Math. J. **24** (1974), 29–77.

[42] A. Stós, Symmetric α-stable processes on d-sets, Bull. Polish Acad. Sci. Math. **48** (2000), 237–245.

[43] R.S. Strichartz, Differential equations on fractals: a tutorial, Princeton University Press, Princeton (2006).

[44] W. Woess, Random Walks on Infinite Graphs and Groups. Cambridge University Press, Cambridge (2000).

[45] W. Woess, Denumerable Markov Chains. European Math. Soc., Switzerland (2009).

[46] P.L. Yung, Doubling properties of self-similar measures, Indiana Univ. Math. J. **56** (2007), 965–990.

SHI-LEI KONG, Department of Mathematics, The Chinese University of Hong Kong, Hong Kong
slkong@math.cuhk.edu.hk

KA-SING LAU, Department of Mathematics, The Chinese University of Hong Kong, Hong Kong
kslau@math.cuhk.edu.hk

TING-KAM LEONARD WONG, Department of Mathematics, University of Southern California, Los Angeles, CA 90089 USA
tkleonardwong@gmail.com