Proceedings of the Edinburgh Mathematical Society (2008) **51**, 171–199 ©
DOI:10.1017/S001309150500177X Printed in the United Kingdom

AN ANALYTICAL APPROACH TO HEAT KERNEL ESTIMATES ON STRONGLY RECURRENT METRIC SPACES

JIAXIN HU

Department of Mathematical Sciences, Tsinghua University, Beijing 100084, People's Republic of China (hujiaxin@mail.tsinghua.edu.cn)

(Received 5 December 2005)

Abstract In this paper we prove that *sub-Gaussian* estimates of *heat kernels* of regular Dirichlet forms are equivalent to the regularity of measures, two-sided bounds of *effective resistances* and the locality of semigroups, on *strongly recurrent* compact metric spaces. Upper bounds of effective resistances imply the compact embedding theorem for domains of Dirichlet forms, and give rise to the existence of Green functions with zero Dirichlet boundary conditions. Green functions play an important role in our analysis. Our emphasis in this paper is on the *analytic* aspects of deriving two-sided sub-Gaussian bounds of heat kernels. We also give the *probabilistic* interpretation for each of the main analytic steps.

Keywords: heat kernel; effective resistance; α-set; Dirichlet form; Green function

2000 *Mathematics subject classification:* Primary 32W05
Secondary 35R20

1. Introduction

In recent years, people have studied *heat kernels* or *transition densities* on fractals, and have obtained elegant *sub-Gaussian* estimates of the form

$$\begin{aligned} a_1 t^{-\alpha/\beta} \exp(-b_1(t^{-1/\beta} d(x,y))^{\gamma_1}) &\leqslant p(t,x,y) \\ &\leqslant a_2 t^{-\alpha/\beta} \exp(-b_2(t^{-1/\beta} d(x,y))^{\gamma_2}) \end{aligned} \tag{1.1}$$

for heat kernels $p(t,x,y)$ on a certain class of fractals, where $a_i, b_i, \gamma_i > 0$ for $i = 1, 2$ and $\alpha > 0$, $\beta \geqslant 2$, and $d(x,y)$ is a metric on the fractal considered (see, for example, [**2**–**4**, **11**, **17**]). In order to obtain (1.1), the theory of Markov processes has been used extensively in the literature cited above. On the other hand, people are also interested in obtaining (1.1) in a purely analytic approach, without recourse to the theory of Markov processes. Recall that, for the classical case where $\beta = 2$, there have been analytic approaches to deriving (1.1) (see, for example, [**1**, **6**, **8**]). However, to our knowledge, no analytic approach is available for the fractal case or, more generally, for metric spaces where $\beta > 2$. Note that the existent analytic method for $\beta = 2$ is not applicable to the case where $\beta > 2$.

In this paper we will establish *analytically* sub-Gaussian estimates of the type (1.1) for heat kernels on compact metric spaces satisfying the chain condition. More precisely, we show that sub-Gaussian estimates for heat kernels of Dirichlet forms on strongly recurrent compact metric spaces are equivalent to the regularity of measures, polynomial growth of effective resistances and the locality of the semigroups (see Theorem 2.2). We explain here how to derive sub-Gaussian estimates of heat kernels *analytically*. Let X be a metric space, and let $B(x,r)$ be an open ball in X. The effective resistance $R(x, B(x,r)^{\mathrm{c}})$ between $x \in X$ and $B(x,r)^{\mathrm{c}}$ is well defined. This gives rise to the existence of the Green function $g^x_{B(x,r)^{\mathrm{c}}}(\cdot)$ vanishing on $B(x,r)^{\mathrm{c}}$. The Green function $g^x_{B(x,r)^{\mathrm{c}}}(y)$ has the same polynomial growth as the effective resistance $R(x, B(x,r)^{\mathrm{c}})$, for x and y sufficiently close. An important step is to bound from above the solution of the *linear heat equation* with initial values vanishing on the ball (see (5.55), below). This can be achieved by estimating the solution u_λ to a *linear elliptic equation* (see (5.52), below) for $\lambda > 0$ (see the crucial estimate (5.48)). We obtain (5.48) by applying the locality of the semigroup of the Dirichlet form, together with the estimate of the solution u_0 to a *Poisson equation* with zero boundary condition (see (5.20), (5.21), below). We estimate u_0 by using the regularity of the measure and the estimate of the Green function. Once we obtain the estimate (5.55), off-diagonal upper bounds of heat kernels easily follow by using on-diagonal upper bounds. Lower off-diagonal bounds of heat kernels are derived more easily by using the chain argument (see, for example, [**2**, **16**]).

We mention in passing here that there is a probabilistic interpretation for each of the main analytic steps for deriving (5.55), which will be pointed out in the remarks following it. (The reader may consult [**2**, **13**] for the probabilistic arguments.) A result similar to Theorem 2.2 was obtained on *strongly recurrent graphs* in [**5**] and on *resistance forms* in [**20**], by using the probability theory.

Finally, we remark that the method of this paper could be applicable to more general bounded or unbounded metric spaces. We will address this issue elsewhere.

Notation.

In the following, we keep c_i, ε_i, $i \geqslant 0$, fixed, and use c, c' and c'' to denote general constants. For two functions f and g, by $f \asymp g$ we mean that there is some $c > 0$ such that $c^{-1} f \leqslant g \leqslant cf$.

2. Preliminaries and main results

Let (X, d) be a compact metric space satisfying the *chain condition*, that is, there exists a constant $c_0 > 0$ such that, for any distinct points $x, y \in X$ and any integer $n \geqslant 1$, there exists a sequence of points $\{x_k\}_{k=0}^n$ in X such that $x_0 = x$, $x_n = y$ and

$$d(x_i, x_{i+1}) \leqslant c_0 n^{-1} d(x,y), \quad 0 \leqslant i \leqslant n-1. \tag{2.1}$$

Set $r_0 := \mathrm{diam}(X) < \infty$, the diameter of X. Let $B(x,r) = \{y \in X : d(y,x) < r\}$ be an open ball in X with centre x and radius r. Denote by $\partial B(x,r) = \{y \in X : d(y,x) = r\}$

the *boundary* of the ball $B(x,r)$. For *any* $x \in X$ and $r \in (0, r_0)$, the chain condition and the compactness of X imply that the boundary $\partial B(x,r)$ of *any* ball $B(x,r)$ is not empty.

Let μ be a Borel measure with $\operatorname{supp} \mu = X$. For simplicity, we assume that $\mu(X) = 1$. For $\alpha > 0$, we say that μ satisfies condition (A1) if μ is *α-regular*, that is, there exists some $c_1 > 0$ such that

$$c_1^{-1} r^\alpha \leqslant \mu(B(x,r)) \leqslant c_1 r^\alpha \tag{A1}$$

for all $x \in X$ and all $0 < r \leqslant r_0$. We call X an *α-set* if (A1) holds.

For $1 \leqslant p < \infty$, denote by $L^p(\mu) := L^p(X, d, \mu)$ the space of all p-integrable real-valued functions on X with the norm

$$\|u\|_p := \biggl(\int_X |u(x)|^p \,\mathrm{d}\mu(x) \biggr)^{1/p}.$$

Denote by $C(X)$ (or $C_0(X)$) the space of all continuous functions (or all continuous functions with compact support) on X with uniform norm. Let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(\mu)$. Recall that $(\mathcal{E}, \mathcal{F})$ is *regular* if $\mathcal{F} \cap C_0(X)$ is dense in $\mathcal{F}$ with norm

$$(\mathcal{E}(u) + \|u\|_2^2)^{1/2},$$

and dense in $C_0(X)$ with uniform norm. (Here we use the abbreviation $\mathcal{E}(u) := \mathcal{E}(u,u)$.) The form $(\mathcal{E}, \mathcal{F})$ is *local* if $\mathcal{E}(u,v) = 0$ for $u, v \in \mathcal{F}$ with disjoint supports, and *irreducible* if $\mathcal{E}(u) = 0$ when and only when u is constant [**12**]. Let $\{T_t\}_{t \geqslant 0}$ be the semigroup associated with $(\mathcal{E}, \mathcal{F})$. For any $t > 0$, the operator $\{T_t\}_{t \geqslant 0}$ may possess an integral kernel $p(t,x,y)$, termed the *heat kernel*, that is, for $t > 0$ and μ-almost all $x \in X$,

$$T_t u(x) = \int_X u(y) p(t,x,y) \,\mathrm{d}\mu(y) \tag{2.2}$$

for $u \in L^2(\mu)$. Since $\mathcal{E}(1) = 0$, we see that $T_t 1 = 1$ (see [**2**, Lemma 4.10, p. 50]). Therefore, if the heat kernel $p(t,x,y)$ exists, for $t > 0$ and μ-almost all $x \in X$, we have

$$\int_X p(t,x,y) \,\mathrm{d}\mu(y) = 1. \tag{2.3}$$

Note that $\{T_t\}_{t \geqslant 0}$ is *strongly continuous* in the $L^2(\mu)$-norm, that is

$$\lim_{t \to 0} \|T_t u - u\|_2 = 0, \quad u \in L^2(\mu). \tag{2.4}$$

Let H be the *infinitesimal generator* of $\{T_t\}_{t \geqslant 0}$ in the $L^2(\mu)$-norm, that is

$$\lim_{t \to 0} \|t^{-1}(T_t u - u) - Hu\|_2 = 0 \tag{2.5}$$

for $u \in \mathcal{D}(H)$, the *space* of all functions $u \in L^2(\mu)$ such that the above limit exists for some function $Hu \in L^2(\mu)$. Note that $\mathcal{D}(H)$ is dense in $L^2(\mu)$. We shall see below that the semigroup $\{T_t\}$ considered in this paper is actually the *Feller semigroup*, that is, $T_t f \geqslant 0$ for $f \geqslant 0$, $T_t f \leqslant 1$ for $f \leqslant 1$ and

(i) for any $t > 0$, the operator $T_t : C(X) \to C(X)$, and

(ii) $\|T_t u - u\|_{C(X)} \to 0$ as $t \to 0$ for $u \in C(X)$.

For a Feller semigroup $\{T_t\}$ and $u \in C(X)$, if there is a function $v \in C(X)$ such that

$$\lim_{t\to 0} \|t^{-1}(T_t u - u) - v\|_{C(X)} = 0, \tag{2.6}$$

then we define

$$\Delta u(x) = v(x), \quad x \in X. \tag{2.7}$$

Let $\mathcal{D}(\Delta)$ be the *domain* of Δ. Clearly, (2.6) implies that

$$\lim_{t\to 0} t^{-1}(T_t u(x) - u(x)) = v(x) \tag{2.8}$$

for *each* point $x \in X$. Note that H is the extension of the linear operator $\Delta : \mathcal{D}(\Delta) \to C(X)$. In the following, we will use the fact that

$$\mathcal{E}(u,v) = \lim_{t\to 0} \mathcal{E}_t(u,v) := \lim_{t\to 0} \left(\frac{u - T_t u}{t}, v\right), \quad u, v \in \mathcal{F}, \tag{2.9}$$

where $(\cdot\,,\cdot)$ is the inner product in $L^2(\mu)$ (see [**12**, Lemma 1.3.4, p. 22]).

For any non-empty open subset D of X, let

$$\mathcal{F}_D := \{u \in \mathcal{F} : u|_{D^c} = 0\},$$

where $D^c = X \setminus D$. If $(\mathcal{E}, \mathcal{F})$ is a regular Dirichlet form, then $(\mathcal{E}, \mathcal{F}_D)$ is also a regular Dirichlet form [**12**, Theorem 4.4.3, p. 154]. As in the case of the form $(\mathcal{E}, \mathcal{F})$, we denote by $\{T_t^D\}_{t\geqslant 0}$ and H_D the semigroup and generator for the form $(\mathcal{E}, \mathcal{F}_D)$, respectively. In particular, denote by $p_D(t,x,y)$ the heat kernel of $(\mathcal{E}, \mathcal{F}_D)$, if it exists. We extend $p_D(t,x,y)$ so that $p_D(t,x,y) = 0$ for $t > 0$ if $x \in D^c$ or $y \in D^c$. For any $u, v \in \mathcal{F}_D \subset \mathcal{F}$, observe that

$$\lim_{t\to 0} \left(\frac{u - T_t u}{t}, v\right) = \mathcal{E}(u,v) = \lim_{t\to 0} \left(\frac{u - T_t^D u}{t}, v\right) \tag{2.10}$$

for any non-empty open subset D. Finally, for any two non-empty open subsets $D_1 \subset D_2$ of X, it is known that

$$p_{D_1}(t,x,y) \leqslant p_{D_2}(t,x,y) \leqslant p(t,x,y) \tag{2.11}$$

for any $t > 0$ and μ-almost all $x, y \in X$, if all of them exist.

Remark 2.1. Let $(\{X_t\}_{t\geqslant 0}, \{\mathbb{P}_x\}_{x\in X})$ be the *Hunt process* corresponding to a regular Dirichlet form $(\mathcal{E}, \mathcal{F})$. Let $D_1 \subset D_2$ be two open subsets. Then, for any $t > 0$ and μ-almost all $x \in X$,

$$\begin{aligned} T_t^{D_1} u(x) = \mathbb{E}_x^{D_1}(u(X_t)) &:= \mathbb{E}_x(\mathbf{1}_{\{t<\tau_{D_1}\}} u(X_t)) \\ &\leqslant \mathbb{E}_x(\mathbf{1}_{\{t<\tau_{D_2}\}} u(X_t)) = T_t^{D_2} u(x) \\ &\leqslant \mathbb{E}_x(u(X_t)) = T_t u(x) \end{aligned} \tag{2.12}$$

for any non-negative bounded Borel function u, where

$$\tau_D = \inf\{t > 0 : X_t \notin D\}$$

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

is the first exit time of X_t from D [**12**, (4.1.2), p. 135]. Thus, (2.11) easily follows. The *analytic* proof of (2.11) for the classical case uses the maximum principle (see [**7**, Lemma 3.3]).

For the form $(\mathcal{E}, \mathcal{F})$, we define the *effective resistance* $R(x, y)$ for any two points $x, y \in X$ by

$$R(x,y)^{-1} = \inf\{\mathcal{E}(u) : u \in \mathcal{F},\ u(x) = 1 \text{ and } u(y) = 0\} \tag{2.13}$$

if $x \neq y$, and $R(x, y) = 0$ if $x = y$ (possibly $R(x, y) = \infty$ for some points $x, y \in X$). By (2.13) we see that

$$R(x,y) = \sup\left\{\frac{|u(y) - u(x)|^2}{\mathcal{E}(u)} : u \in \mathcal{F} \text{ and } \mathcal{E}(u) > 0\right\}, \tag{2.14}$$

which gives

$$|u(y) - u(x)|^2 \leqslant R(x,y)\mathcal{E}(u), \quad u \in \mathcal{F},\ x, y \in X. \tag{2.15}$$

We say that $R(x, y)$ satisfies condition (A2) if there exist a number $\gamma > 0$ and a constant $c_2 > 0$ such that

$$c_2^{-1} d(y,x)^\gamma \leqslant R(x,y) \leqslant c_2 d(y,x)^\gamma \tag{A2}$$

for all $x, y \in X$.

We say that $p(t, x, y)$ satisfies condition (A3) if there exist some constants $a_i, b_i > 0$, $i = 1, 2$, such that

$$a_1 t^{-\alpha/\beta} \exp(-b_1(t^{-1/\beta} d(x,y))^{\beta_0}) \leqslant p(t,x,y) \leqslant a_2 t^{-\alpha/\beta} \exp(-b_2(t^{-1/\beta} d(x,y))^{\beta_0}) \tag{A3}$$

for all $x, y \in X$ and all $0 < t \leqslant r_0^\beta$, where $\beta > 0$, $\beta_0 = \beta(\beta - 1)^{-1}$ and α is the same as in (A1).

We say that the semigroup $\{T_t\}_{t \geqslant 0}$ of the form $(\mathcal{E}, \mathcal{F})$ is of *local character* if, for any non-empty closed subset D of X,

$$\lim_{t \to 0} \frac{1}{t}(T_t \mathbf{1}_D)(x) = 0 \tag{2.16}$$

uniformly in $x \in X$ satisfying $d(x, D) \geqslant \delta > 0$ for any fixed δ (the function $\mathbf{1}_D$ is the indicator of D, that is $\mathbf{1}_D = 1$ on D, and $\mathbf{1}_D = 0$ elsewhere). Clearly, if (2.16) holds, then a regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ is local, by using the dominated convergence theorem. Moreover, if the heat kernel $p(t, x, y)$ of $\{T_t\}$ exists, it follows from (2.16) that $p(t, x, y)$ is of *local character* as well, that is, for any closed subset D and any fixed $\delta > 0$,

$$\lim_{t \to 0} \frac{1}{t} \int_D p(t,x,y)\, \mathrm{d}\mu(y) = 0 \tag{2.17}$$

uniformly in $x \in X$ with $d(x, D) \geqslant \delta$. Note that (2.17) was proved in [**21**] for the diffusion on $X = \mathbb{R}$ by using the probability method (see also [**10**]).

We now state the main result of this paper.

Theorem 2.2. *Let (X, d) be a compact metric space satisfying the chain condition. Assume that $(\mathcal{E}, \mathcal{F})$ is an irreducible regular Dirichlet form on $L^2(\mu)$. Then the following conditions are equivalent.*

(i) *The heat kernel $p(t, x, y)$ of $(\mathcal{E}, \mathcal{F})$ exists and satisfies (A3) with $\alpha < \beta$.*

(ii) *The measure μ satisfies (A1) and the effective resistance $R(x, y)$ satisfies (A2) with $\gamma = \beta - \alpha$, $\beta > 1$, and the semigroup $\{T_t\}$ of $(\mathcal{E}, \mathcal{F})$ is of local character.*

The proof of Theorem 2.2 will be given in §§ 3 and 5. It would be interesting to replace the locality of the semigroup $\{T_t\}$ in Theorem 2.2 (ii) by the locality of the form $(\mathcal{E}, \mathcal{F})$. We will explore this for the case of post-critically finite self-similar fractals introduced by Kigami [**18**].

Remarks 2.3. (1) Condition (A1) implies that α is the Hausdorff dimension of X [**9**]. The number β in (A3) is termed the *walk dimension*. The *spectral dimension* d_s of X is determined by the *Einstein relation* $d_\mathrm{s} = 2\alpha/\beta$. We say that X is *strongly recurrent* if $d_\mathrm{s} < 2$. Clearly, the X considered in this paper is strongly recurrent.

(2) The upper estimate in (A2) implies the Morrey–Sobolev inequality, that is

$$|u(y) - u(x)|^2 \leqslant cd(y, x)^\gamma \mathcal{E}(u) \tag{2.18}$$

for all $x, y \in X$ and all $u \in \mathcal{F}$, by virtue of (2.15). Thus, $\mathcal{F} \subset C(X)$. For a non-empty proper subset A of X, let $\mathcal{F}_A$ be equipped with norm $\mathcal{E}(u)^{1/2}$. If $\mathcal{F}_A$ is not empty, then $\mathcal{F}_A$ is *compactly embedded* in $C(X)$ by using the Ascoli–Arzelà theorem [**22**, p. 85], since X is compact, that is, any bounded sequence in $\mathcal{F}_A$ has a convergent subsequence in $C(X)$. The compact embedding will play an important role; in particular, it implies the existence of Green functions with zero boundary conditions (see § 4).

(3) If the upper bound of R in (A2) holds, then $R(x, y)$ is a metric on X (see § 4). Thus, the chain condition, together with (A2), implies that $\gamma \leqslant 1$. In fact, let $\{x_k\}_{k=0}^n$ be a chain connecting $x, y \in X (x \neq y)$ with $x_0 = x$ and $x_n = y$ for a large integer n. We see from (A2) and (2.1) that

$$\begin{aligned} c_2^{-1} d(y, x)^\gamma \leqslant R(x, y) &\leqslant \sum_{k=0}^{n-1} R(x_k, x_{k+1}) \leqslant \sum_{k=0}^{n-1} c_2 d(x_k, x_{k+1})^\gamma \\ &\leqslant \sum_{k=0}^{n-1} c_2 (c_0 n^{-1} d(x, y))^\gamma \leqslant cn^{1-\gamma} d(x, y)^\gamma, \end{aligned}$$

which implies that $n^{\gamma-1} \leqslant c$. Thus, $\gamma \leqslant 1$, and so $\beta = \alpha + \gamma \leqslant \alpha + 1$.

(4) If $R(x, y)$ defined as in (2.13) is shown to be a metric on X, we may take $d(y, x) = R(x, y)$ in (A2) with $\gamma = 1$. Then Theorem 2.2 says that, for an irreducible regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(\mu)$, the conditions (A1) and (2.16) are equivalent to (A3) with $\beta = \alpha + 1$ (see [**5**] for graphs, [**17**] for post-critically finite fractals with regular harmonic structure and [**20**] for effective forms).

3. Proof of (A3) $\Longrightarrow$ (A1) + (A2) + (2.16)

In this section we show that (i) $\Rightarrow$ (ii) in Theorem 2.2. The fact that (A3) implies (A1) and the upper bound of $R(x,y)$ in (A2) was actually obtained in [**16**]. We need only to prove the lower bound of $R(x,y)$ in (A2) and (2.16). For the reader's convenience, we outline the whole proof.

Proposition 3.1 (Grigor'yan *et al*. [16, Theorem 3.2]). *Assume that $p(t,x,y)$ satisfies (A3) (without the restriction that $\alpha < \beta$). Then μ satisfies (A1).*

Proof. In [**16**] X is assumed to be unbounded, and so $p(t,x,y)$ satisfies (A3) for all $0 < t < \infty$. One can slightly modify the proof in [**16**] to deal with the case in which X is bounded. Thus, Proposition 3.1 follows. □

For $\sigma > 0$, we define

$$W_\sigma(u) = \sup_{0<r<1} r^{-2\sigma} \int_X \left\{ \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |u(y) - u(x)|^2 \, \mathrm{d}\mu(y) \right\} \mathrm{d}\mu(x) \tag{3.1}$$

for $u \in L^2(\mu)$.

Proposition 3.2 (Grigor'yan *et al*. [16, Theorem 4.11 (iii)]). *Assume that μ satisfies (A1). If $\alpha < \beta$, then*

$$|u(y) - u(x)|^2 \leqslant cd(y,x)^{\beta-\alpha} W_{\beta/2}(u) \tag{3.2}$$

for all $x, y \in X$ and all $u \in C(X)$, for some $c > 0$.

Proof. In [**16**] the embedding (3.2) was obtained for $x, y \in X$ with $d(y,x) \leqslant \frac{1}{3}$. If $d(y,x) \geqslant \frac{1}{3}$, we let $\{x_k\}_{k=0}^{n}$ be a chain connecting x and y such that $d(x_k, x_{k+1}) \leqslant \frac{1}{3}$ for all $0 \leqslant k \leqslant n-1$; this can be done by taking $n = [3c_0 r_0] + 1$. For $u \in C(X)$, we use (3.2) for each pair (x_k, x_{k+1}), $0 \leqslant k \leqslant n-1$, and then sum over k to arrive at (3.2) for all $x, y \in X$ with $d(y,x) \geqslant \frac{1}{3}$. □

Proposition 3.3 (Grigor'yan *et al*. [16, Theorem 4.2]). *Assume that $p(t,x,y)$ satisfies (A3). Then*

$$c^{-1} W_{\beta/2}(u) \leqslant \mathcal{E}(u) \leqslant c W_{\beta/2}(u) \tag{3.3}$$

for all $u \in \mathcal{F}$, for some $c > 0$.

By (3.2) and (3.3), we see that if $p(t,x,y)$ satisfies (A3) with $\alpha < \beta$, then

$$|u(y) - u(x)|^2 \leqslant cd(y,x)^{\beta-\alpha} \mathcal{E}(u) \tag{3.4}$$

for all $x, y \in X$ and all $u \in \mathcal{F}$. This immediately gives

$$R(x,y) \leqslant cd(y,x)^{\beta-\alpha} \tag{3.5}$$

for all $x, y \in X$, by virtue of (2.14). So the upper bound of $R(x,y)$ in (A2) follows with $\gamma = \beta - \alpha > 0$. It remains to prove the lower bound of $R(x,y)$ in (A2).

Theorem 3.4. *Assume that $p(t,x,y)$ satisfies (A3) with $\alpha<\beta$. Then, for all $x,y\in X$,*

$$R(x,y)\geqslant c^{-1}d(y,x)^{\beta-\alpha}. \tag{3.6}$$

Proof. We first show that

$$\mathcal{E}(p(t,x,\cdot))\leqslant (et)^{-1}p(t,x,x) \tag{3.7}$$

for all $t>0$ and $x\in X$. Indeed, let $f\in L^2(\mu)$, and set

$$u(t,y)=T_tf(y),\quad y\in X.$$

By the spectral calculus, we have

$$\mathcal{E}(u(t,\cdot))=\int_0^\infty \lambda \mathrm{e}^{-2\lambda t}\,\mathrm{d}(E_\lambda f,f),$$

where $\{E_\lambda\}$ is the spectral representation of the generator of $\{T_t\}$. Noting that

$$\lambda \mathrm{e}^{-2\lambda t}\leqslant (2et)^{-1}$$

for $\lambda\geqslant 0$ and $t>0$, we see from above that

$$\mathcal{E}(u(t,\cdot))\leqslant (2et)^{-1}\|f\|_2^2.$$

Fix $x\in X$. Letting $f(y)=p(t,x,y)$, $y\in X$, we obtain $u(t,y)=p(2t,x,y)$. Thus, we have

$$\mathcal{E}(p(2t,x,\cdot))\leqslant (2et)^{-1}\|p(t,x,\cdot)\|_2^2=(2et)^{-1}p(2t,x,x),$$

proving (3.7) by replacing $2t$ by t. It follows from (3.7) that, using the upper diagonal bound of $p(t,x,y)$,

$$\mathcal{E}(p(t,x,\cdot))\leqslant (et)^{-1}p(t,x,x)\leqslant a_2\mathrm{e}^{-1}t^{-(1+\alpha/\beta)} \tag{3.8}$$

for all $0<t\leqslant r_0^\beta$ and all $x\in X$. Fix $y\in X$. By (A3), (2.15) and (3.8), we have

$$\begin{aligned} a_1t^{-\alpha/\beta}-a_2t^{-\alpha/\beta}\exp(-b_2(t^{-1/\beta}d(x,y))^{\beta_0}) &\leqslant p(t,x,x)-p(t,x,y)\\ &\leqslant R(x,y)^{1/2}\mathcal{E}(p(t,x,\cdot))^{1/2}\\ &\leqslant (a_2\mathrm{e}^{-1})^{1/2}R(x,y)^{1/2}t^{-(1+\alpha/\beta)/2} \end{aligned}$$

for all $0<t\leqslant r_0^\beta$. Let $t=(b^{-1}b_2)^{\beta/\beta_0}d(x,y)^\beta$, where b is so large that $a_2\exp(-b)\leqslant \frac12 a_1$ and $t\leqslant r_0^\beta$. It follows from above that

$$\tfrac12 a_1t^{-\alpha/\beta}\leqslant (a_2\mathrm{e}^{-1})^{1/2}R(x,y)^{1/2}t^{-(1+\alpha/\beta)/2},$$

and so

$$R(x,y)\geqslant ct^{(\beta-\alpha)/\beta}=cd(x,y)^{\beta-\alpha},$$

giving (3.6). □

Finally, we see that (2.17) (or equivalently (2.16)) easily follows from the upper bound of $p(t,x,y)$ in (A3) and the regularity of μ. Indeed, for a closed subset D of X and any point $x \in X$ with $\operatorname{dist}(x,D) \geqslant \delta > 0$, we have

$$\begin{aligned} t^{-1}\int_D p(t,x,y)\,\mathrm{d}\mu(y) &\leqslant t^{-1}\int_{B(x,\delta)^{\mathrm{c}}} p(t,x,y)\,\mathrm{d}\mu(y) \\ &\leqslant ct^{-1}\int_{\frac{1}{2}\delta t^{-1/\beta}}^{\infty} s^{\alpha-1}\exp(-c' s^{\beta_0})\,\mathrm{d}s \to 0 \end{aligned}$$

as $t \to 0$ (see, for example, [**16**, (3.7)]). If (A3) holds and (X,d) satisfies the chain condition, then $\beta \geqslant 2$ (see [**16**, (4.27), p. 2081]).

4. Green functions

In order to prove the other direction in Theorem 2.2, we need to investigate the existence of Green functions with zero boundary conditions. In this section, we assume only that $(\mathcal{E},\mathcal{F})$ is an irreducible regular Dirichlet form, and that the upper bound of R in (A2) holds. The variational problem (2.13) or (4.1) below possesses a unique solution, leading to the existence of Green functions with boundary conditions. This also generalizes the results on Green functions with boundary having finite points [**19**] to the case where the boundary may have infinite points.

For any non-empty subset A of X, we define the effective resistance $R(x,A)$ between any point $x \in X$ and A by $R(x,A) = 0$ if $x \in A$, and

$$R(x,A)^{-1} = \inf\{\mathcal{E}(u) : u \in \mathcal{F}^x_{A^{\mathrm{c}}}\} \tag{4.1}$$

if $x \notin A$, where

$$\mathcal{F}^x_{A^{\mathrm{c}}} := \{u \in \mathcal{F} : u(x) = 1 \text{ and } u|_A = 0\}, \quad x \notin A.$$

(Recall that $\mathcal{F} \subset C(X)$ by Remark 2.3 (2), and so $u \in \mathcal{F}$ is defined *pointwise* on X.) Note that $\mathcal{F}^x_{A^{\mathrm{c}}}$ may be empty for some subsets A of X. However, we have the following.

Proposition 4.1. *Assume that $(\mathcal{E},\mathcal{F})$ is an irreducible regular Dirichlet form on $L^2(\mu)$. Then, for any two disjoint non-empty closed subsets A and B of X,*

$$\{u \in \mathcal{F} : u|_A = 0,\ u|_B = 1\} \neq \emptyset$$

In particular, the set $\mathcal{F}^x_{A^{\mathrm{c}}}$ is not empty for any $x \notin A$ and any non-empty closed subset $A \subset X$.

Proof. Let A and B be two closed subsets of X with $A \cap B = \emptyset$. By the Urysohn theorem [**22**, p. 7], there exists a real-valued continuous function v on X such that $0 \leqslant v \leqslant 1$ on X, and $v|_A = 0$ and $v|_B = 1$. Since $\mathcal{E}$ is regular, there is a function $u_1 \in \mathcal{F}$ such that $u_1|_A \leqslant \frac{1}{3}$ and $u_1|_B \geqslant \frac{2}{3}$. Since $\mathcal{E}$ is irreducible, we see that $u_1 - \frac{1}{3} \in \mathcal{F}$. Define $u_+ = 0 \vee u$. Let

$$u = (3(u_1 - \tfrac{1}{3})_+) \wedge 1.$$

It is easily seen that u is the desired function by using the Markov property of $\mathcal{E}$. □

Proposition 4.2. *Assume that $(\mathcal{E},\mathcal{F})$ is an irreducible regular Dirichlet form on $L^2(\mu)$, and that the upper bound of R in (A2) holds. Then, for any non-empty closed subset A of X and any $x_0 \notin A$, the variational problem*

$$\lambda_0 = \inf\{\mathcal{E}(u) : u \in \mathcal{F}^{x_0}_{A^c}\} \tag{4.2}$$

possesses a unique solution in $[0,1]$, that is, there is a unique function $\psi^{x_0}_A \in \mathcal{F}^{x_0}_{A^c}$ with $0 \leqslant \psi^{x_0}_A \leqslant 1$ on X such that $\lambda_0 = \mathcal{E}(\psi^{x_0}_A)$.

Proof. The proof is standard, by using the compact embedding theorem. We outline the proof for the reader's convenience. Since $\mathcal{F}^{x_0}_{A^c}$ is not empty, we see that $\lambda_0 < \infty$. Let $\{u_n\}_{n\geqslant 1}$ be a minimizing sequence for (4.2), that is $\{u_n\}_{n\geqslant 1} \in \mathcal{F}^{x_0}_{A^c}$ and $\mathcal{E}(u_n) \to \lambda_0$ as $n \to \infty$. The compact embedding (see Remark 2.3 (2)) implies that there is a subsequence of $\{u_n\}$ (still denoted by $\{u_n\}$) and a function $\psi^{x_0}_A \in C(X)$ such that $\{u_n\}$ uniformly converges to $\psi^{x_0}_A$ as $n \to \infty$. Clearly, $\psi^{x_0}_A(x_0) = 1$ and $\psi^{x_0}_A|_A = 0$. We show that $\psi^{x_0}_A \in \mathcal{F}$. In fact, since $\mathcal{E}_t(u)$ increases to $\mathcal{E}(u)$ as $t \to 0$, from (2.9), using the dominated convergence theorem, we have

$$\mathcal{E}_t(\psi^{x_0}_A) = \lim_{n\to\infty} \mathcal{E}_t(u_n) \leqslant \lim_{n\to\infty} \mathcal{E}(u_n) = \lambda_0$$

for $t > 0$, which gives that $\mathcal{E}(\psi^{x_0}_A) \leqslant \lambda_0 < \infty$. Thus, $\psi^{x_0}_A \in \mathcal{F}$. Clearly, $\mathcal{E}(\psi^{x_0}_A) = \lambda_0$. To show that $\psi^{x_0}_A$ is unique, let $v \in \mathcal{F}^{x_0}_{A^c}$, satisfying $\mathcal{E}(v) = \lambda_0$. Note that

$$0 \leqslant \mathcal{E}\left(\frac{v-\psi^{x_0}_A}{2}\right) = \tfrac{1}{2}(\mathcal{E}(v) + \mathcal{E}(\psi^{x_0}_A)) - \mathcal{E}\left(\frac{v+\psi^{x_0}_A}{2}\right) \leqslant 0$$

since $\mathcal{E}(\frac{1}{2}(v+\psi^{x_0}_A)) \geqslant \lambda_0$, and so $v - \psi^{x_0}_A$ is constant by virtue of the irreducibility of $\mathcal{E}$. Thus, $v = \psi^{x_0}_A$ on X. It remains to show that $0 \leqslant \psi^{x_0}_A \leqslant 1$. But this easily follows from the Markov property of $\mathcal{E}$ and the uniqueness. Indeed, let $u = (0 \vee \psi^{x_0}_A) \wedge 1$. By the Markov property of $\mathcal{E}$, we see that $u \in \mathcal{F}^{x_0}_{A^c}$ and $\mathcal{E}(u) \leqslant \mathcal{E}(\psi^{x_0}_A)$. The uniqueness proved above implies that $\psi^{x_0}_A = u \in [0,1]$. □

For any two distinct points $x_0, y_0 \in X$, letting $A = \{y_0\}$, we see from Proposition 4.2 that there is a unique function $\psi^{x_0}_{y_0}$ with the property that $\psi^{x_0}_{y_0}(x_0) = 1$ and $\psi^{x_0}_{y_0}(y_0) = 0$ such that

$$R(x_0, y_0)^{-1} = \mathcal{E}(\psi^{x_0}_{y_0}). \tag{4.3}$$

Next we claim that, for any closed set $A \subset X$ and $x_0 \in X$, the function $\psi^{x_0}_A \in \mathcal{F}^{x_0}_{A^c}$ is the solution to (4.1) with $x = x_0$ if and only if

$$\mathcal{E}(\psi^{x_0}_A, v) = 0 \tag{4.4}$$

for all $v \in \mathcal{F}$ satisfying $v|_{A\cup\{x_0\}} = 0$. In fact, if $\psi^{x_0}_A$ is the solution to (4.1), we find that $\psi^{x_0}_A + tv \in \mathcal{F}^{x_0}_{A^c}$ for any $t \in \mathbb{R}$ and $v \in \mathcal{F}$ satisfying $v|_{A\cup\{x_0\}} = 0$. Thus,

$$\mathcal{E}(\psi^{x_0}_A) \leqslant \mathcal{E}(\psi^{x_0}_A + tv) = \mathcal{E}(\psi^{x_0}_A) + 2t\mathcal{E}(\psi^{x_0}_A, v) + t^2\mathcal{E}(v),$$

which means that $2t\mathcal{E}(\psi_A^{x_0}, v) + t^2\mathcal{E}(v) \geqslant 0$ for any $t \in \mathbb{R}$. Therefore, (4.4) follows. Conversely, if there is a function $\psi_A^{x_0} \in \mathcal{F}_{A^c}^{x_0}$ such that (4.4) holds, then we let $v = f - \psi_A^{x_0}$ for any $f \in \mathcal{F}_{A^c}^{x_0}$, and obtain

$$\mathcal{E}(f) = \mathcal{E}(v + \psi_A^{x_0}) = \mathcal{E}(v) + 2\mathcal{E}(\psi_A^{x_0}, v) + \mathcal{E}(\psi_A^{x_0}) = \mathcal{E}(v) + \mathcal{E}(\psi_A^{x_0}) \geqslant \mathcal{E}(\psi_A^{x_0}).$$

Thus, $\psi_A^{x_0}$ is the solution to (4.1), proving the claim.

For any non-empty closed set $A \subset X$ and any continuous function φ defined on A, we say that a function $f \in \mathcal{F}$ is *harmonic* on A^c with the boundary condition $f|_A = \varphi$ if

$$\mathcal{E}(f, v) = 0$$

for any $v \in \mathcal{F}_{A^c}$. Thus, by (4.4), the function $\psi_A^{x_0}$ is *harmonic* on $X \setminus A \cup \{x_0\}$ with boundary condition on $\psi_A^{x_0}|_A = 0$ and $\psi_A^{x_0}(x_0) = 1$. Note that a harmonic function is uniquely determined by its boundary condition, using the irreducibility of $\mathcal{E}$.

Definition 4.3. Let A be a closed subset of X and $x_0 \in X$. Define the *Green function* $g_A^{x_0}(\cdot) = R(x_0, A)\psi_A^{x_0}(\cdot)$ if $x_0 \notin A$, and $g_A^{x_0} \equiv 0$ if $x_0 \in A$.

Remark 4.4. The Green function defined above is the same as that introduced in [**19**] if A is a finite subset of X. If $A = \{y_0\}$ for $y_0 \in X$, we write the Green function $g_{y_0}^{x_0}(\cdot) = R(x_0, y_0)\psi_{y_0}^{x_0}(\cdot)$.

The Green function $g_A^{x_0}$ has the following properties (see [**19**] if A is a finite set).

(i) $g_A^{x_0} \geqslant 0$ on X, and $g_A^{x_0}(x) = 0$ if $x \in A$ or $x_0 \in A$.

(ii) $g_A^x(y) = g_A^y(x)$ for $x, y \in X$ (see (4.7), below).

(iii) $\mathcal{E}(g_A^{x_0}) = R(x_0, A) = g_A^{x_0}(x_0)$.

(iv) $g_A^{x_0}(x_0) \geqslant g_A^{x_0}(x)$ for all $x \in X$ (since $0 \leqslant \psi_{y_0}^{x_0} \leqslant 1$ on X).

Lemma 4.5. *Let $g_A^{x_0}$ be the Green function defined above for any closed subset A of X and any point $x_0 \in X$. Then, for any $u \in \mathcal{F}_{A^c}$,*

$$\mathcal{E}(g_A^{x_0}, u) = u(x_0). \tag{4.5}$$

Remark 4.6. It is easy to see that (4.5) fails if $u \notin \mathcal{F}_{A^c}$, for example, by letting $u \equiv 1$ on X.

Proof. If $x_0 \in A$, nothing can be proved. Now let $x_0 \notin A$. It suffices to show that

$$\mathcal{E}(\psi_A^{x_0}, u) = u(x_0)R(x_0, A)^{-1} \tag{4.6}$$

for $u \in \mathcal{F}_{A^c}$. We assume that $u(x_0) \neq 0$; otherwise, (4.5) follows from (4.4). Let $v(x) = \psi_A^{x_0}(x) - u(x_0)^{-1}u(x)$ for $x \in X$. Clearly, $v(x_0) = 0$ and $v|_A = 0$. Thus, it follows from (4.4) that

$$0 = \mathcal{E}(\psi_A^{x_0}, v) = \mathcal{E}(\psi_A^{x_0}) - u(x_0)^{-1}\mathcal{E}(\psi_A^{x_0}, u),$$

giving (4.6) by using the fact that $\mathcal{E}(\psi_A^{x_0}) = R(x_0, A)^{-1}$. □

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

Since $g_A^{y_0} \in \mathcal{F}_{A^c}$ for any $y_0 \in X$, it follows from (4.5) that

$$\mathcal{E}(g_A^{x_0}, g_A^{y_0}) = g_A^{x_0}(y_0) = g_A^{y_0}(x_0) \tag{4.7}$$

for any $x_0, y_0 \in X$. If $A = \{y_0\}$, from (4.5) we have

$$\mathcal{E}(g_{y_0}^{x_0}, u) = u(x_0) \tag{4.8}$$

for any $u \in \mathcal{F}$ with $u(y_0) = 0$. For general $u \in \mathcal{F}$, we let $\bar{u} = u - u(y_0)$, and (4.8) applied to $\bar{u}$ gives

$$\mathcal{E}(g_{y_0}^{x_0}, u) = u(x_0) - u(y_0) \tag{4.9}$$

for any $u \in \mathcal{F}$ and any $x_0, y_0 \in X$.

Lemma 4.7. *Let A be a non-empty closed subset of X and $x_0 \in X$. Then*

$$R(x_0, A) \leqslant R(x_0, y_0) + R(y_0, A) \tag{4.10}$$

for any $x_0, y_0 \in X$. In particular, if $A = \{z_0\}$, then R satisfies the triangle inequality

$$R(x_0, z_0) \leqslant R(x_0, y_0) + R(y_0, z_0) \tag{4.11}$$

for any $x_0, y_0, z_0 \in X$.

Proof. Let $x_0, y_0 \in X$, $x_0 \neq y_0$. Motivated by [**19**], we let $h(x) = g_A^{x_0}(x) - g_A^{y_0}(x)$ for $x \in X$. Note that $h(x_0) \geqslant 0$ and $h(y_0) \leqslant 0$, since $g_A^{x_0}(x_0) \geqslant g_A^{x_0}(x)$ for any $x \in X$. Since $h \in \mathcal{F}_{A^c}$, we see from (4.5) that

$$\mathcal{E}(h) = \mathcal{E}(g_A^{x_0} - g_A^{y_0}, h) = h(x_0) - h(y_0),$$

which combines with (2.15) to give

$$(h(x_0) - h(y_0))^2 \leqslant R(x_0, y_0)\mathcal{E}(h) = R(x_0, y_0)(h(x_0) - h(y_0)).$$

Therefore,

$$0 \leqslant h(x_0) \leqslant h(x_0) - h(y_0) \leqslant R(x_0, y_0),$$

and so

$$0 \leqslant g_A^{x_0}(x_0) - g_A^{y_0}(x_0) = h(x_0) \leqslant R(x_0, y_0). \tag{4.12}$$

Hence,

$$R(x_0, A) = g_A^{x_0}(x_0) \leqslant R(x_0, y_0) + g_A^{y_0}(x_0) \leqslant R(x_0, y_0) + R(y_0, A),$$

proving (4.10). □

Remark 4.8. The result in (4.10) was obtained by Kigami [**19**] for the case when A is a finite subset of X. Note that (4.11) implies that R is a metric on X (we assume that the upper bound in (A2) holds, so $R(x, y) < \infty$ for any $x, y \in X$; this is because, if $R(x_0, y_0) = \infty$ for some $x_0, y_0 \in X$, then $\mathcal{E}(\psi_{y_0}^{x_0}) = R(x_0, y_0)^{-1} = 0$, which would imply that $\psi_{y_0}^{x_0} \equiv$ const., which is a contradiction by Proposition 4.2). We call R the *effective resistance metric* on X.

For $y_0 \in A$, we observe that $R(y_0, A) = 0$, and (4.10) implies that

$$R(x_0, A) \leqslant \inf_{y \in A} R(x_0, y) \tag{4.13}$$

for any $x_0 \in X$ and any closed subset A of X. We next state that the Green function $g_A^{x_0}$ is uniformly Lipschitz in terms of R (see [**19**] for a finite subset A of X).

Lemma 4.9. *Let A be a closed subset of X and $x_0 \in X$. Then*

$$|g_A^{x_0}(x) - g_A^{x_0}(y)| \leqslant R(x, y), \tag{4.14}$$

for any $x, y \in X$.

Proof. The proof given here is motivated by [**19**, Lemma 4.9, p. 413]. Fix $x_0, x \in X$ temporally. Assume that $x_0, x \notin A$; otherwise (4.14) is clear. By (4.4), the function

$$u(y) = g_A^x(x) g_A^{x_0}(y) - g_A^{x_0}(x) g_A^x(y), \quad y \in X,$$

is harmonic on $X \setminus A \cup \{x_0, x\}$ with boundary conditions on $A \cup \{x_0, x\}$. Note that $u = 0$ on A, and $u(x) = 0$ and

$$u(x_0) = g_A^x(x) g_A^{x_0}(x_0) - g_A^{x_0}(x) g_A^x(x_0) \geqslant 0.$$

Therefore, we have $u \geqslant 0$ on X.* Thus,

$$g_A^x(x) g_A^{x_0}(y) \geqslant g_A^{x_0}(x) g_A^x(y)$$

for all $x_0, x, y \in X$. Therefore,

$$\begin{aligned} g_A^{x_0}(x) - g_A^{x_0}(y) &\leqslant g_A^{x_0}(x) - g_A^x(x)^{-1} g_A^{x_0}(x) g_A^x(y) \\ &= g_A^{x_0}(x) g_A^x(x)^{-1} (g_A^x(x) - g_A^x(y)) \\ &\leqslant g_A^x(x) - g_A^x(y) \\ &\leqslant R(x, y) \end{aligned}$$

for all $x_0, x, y \in X$, where the last inequality follows from (4.12). Exchanging x and y yields

$$g_A^{x_0}(y) - g_A^{x_0}(x) \leqslant R(y, x),$$

whence (4.14) follows. □

* In fact, if $u(x_0) = 0$, then $u \equiv 0$ on X by uniqueness, and if $u(x_0) > 0$, then the function $u/u(x_0)$ satisfies the variational problem

$$\inf\{\mathcal{E}(u) : u|_{A \cup \{x\}} = 0 \text{ and } u(x_0) = 1\},$$

and a similar argument to Proposition 4.2 in which A is replaced by $A \cup \{x\}$ shows that $0 \leqslant u/u(x_0) \leqslant 1$ on X.

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

5. Proof of (A1) + (A2) + (2.16) $\Longrightarrow$ (A3)

In this section we prove the other direction in Theorem 2.2, that is (A1), (A2) and (2.16) will imply (A3). We first obtain off-diagonal upper bounds of $p(t,x,y)$. The key is to estimate the solution of a linear elliptic equation (5.40) in the ball (see (5.48), below). The Green functions discussed above will be used. The locality of the semigroup plays an important role, which leads to a local maximum principle. We then derive lower bounds of $p(t,x,y)$ in a standard way by using the upper bound of $p(t,x,y)$, (2.15) and the chain condition.

5.1. On-diagonal upper bounds

Theorem 5.1. *Let (X,d,μ) be a measure metric space and let $(\mathcal{E},\mathcal{F})$ be a Dirichlet form (not necessarily local). If μ satisfies the lower bound in (A1) and R satisfies the upper bound in (A2), then the heat kernel $p(t,x,y)$ of $(\mathcal{E},\mathcal{F})$ exists, is continuous on $X\times X$ for each $t>0$ and satisfies*

$$p(t,x,y)\leqslant ct^{-\alpha/\beta} \tag{5.1}$$

for all $x,y\in X$ and $0<t\leqslant r_0^\beta$, where $\beta=\gamma+\alpha$.

Proof. Let f be a non-negative bounded function on X with $\|f\|_1\leqslant 1$. For $0<t\leqslant r_0^\beta$, we show that there exists a constant c independent of f and t such that

$$\|T_tf\|_\infty\leqslant ct^{-\alpha/\beta}. \tag{5.2}$$

The proof given here is motivated by [**5**] for graphs, but we do not assume *a priori* the existence of the heat kernel. To see this, note that $\|T_tf\|_1\leqslant\|f\|_1\leqslant 1$, and

$$\|T_{t/2}f\|_2^2=(T_{t/2}f,T_{t/2}f)=(f,T_tf)\leqslant\|T_tf\|_\infty\|f\|_1\leqslant\|T_tf\|_\infty,\quad t>0. \tag{5.3}$$

Fix $x_0\in X$, and define $B_0:=B(x_0,t^{1/\beta})$. Since

$$\int_{B_0}T_tf(x)\,\mathrm{d}\mu(x)\leqslant\|T_tf\|_1\leqslant 1,$$

using $(a+b)^2\leqslant 2(a^2+b^2)$ and Hölder's inequality, we have

$$\begin{aligned}T_tf(x_0)^2&=\left(\mu(B_0)^{-1}\int_{B_0}((T_tf(x_0)-T_tf(x))+T_tf(x))\,\mathrm{d}\mu(x)\right)^2\\&\leqslant 2\left(\mu(B_0)^{-1}\int_{B_0}(T_tf(x_0)-T_tf(x))\,\mathrm{d}\mu(x)\right)^2+2\left(\mu(B_0)^{-1}\int_{B_0}T_tf(x)\,\mathrm{d}\mu(x)\right)^2\\&\leqslant 2\mu(B_0)^{-1}\int_{B_0}(T_tf(x_0)-T_tf(x))^2\,\mathrm{d}\mu(x)+2\mu(B_0)^{-2}.\end{aligned} \tag{5.4}$$

Observing that

$$\mathcal{E}(T_tf)=-\frac{1}{2}\frac{\partial}{\partial t}\|T_tf\|_2^2,\quad t>0,$$

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

it follows that, using (2.15) and the upper bound of R in (A2),

$$\begin{aligned}(T_tf(x_0)-T_tf(x))^2 &\leqslant R(x_0,x)\mathcal{E}(T_tf)\leqslant c_2 d(x_0,x)^{\beta-\alpha}\mathcal{E}(T_tf)\\ &\leqslant -\frac{c_2}{2}t^{1-\alpha/\beta}\frac{\partial}{\partial t}\|T_tf\|_2^2\end{aligned}$$

for $x\in B_0$ and $t>0$. Therefore,

$$2\mu(B_0)^{-1}\int_{B_0}(T_tf(x_0)-T_tf(x))^2\,\mathrm{d}\mu(x)\leqslant -c_2t^{1-\alpha/\beta}\frac{\partial}{\partial t}\|T_tf\|_2^2,$$

which combines with (5.4) to yield that, using the lower bound of μ in (A1),

$$T_tf(x_0)^2\leqslant -c_2t^{1-\alpha/\beta}\frac{\partial}{\partial t}\|T_tf\|_2^2+ct^{-2\alpha/\beta} \tag{5.5}$$

for $0<t\leqslant r_0^\beta$, where c is independent of x_0, t and f. Set $\phi(t)=\|T_tf\|_\infty$. Note that ϕ is decreasing on $(0,\infty)$ because, for $s<t$,

$$\begin{aligned}\phi(t)=\|T_tf\|_\infty &= \sup_{\|g\|_1=1}(T_tf,g)=\sup_{\|g\|_1=1}(T_{t-s}(T_sf),g)\\ &=\sup_{\|g\|_1=1}(T_sf,T_{t-s}g)\leqslant \|T_sf\|_\infty\sup_{\|g\|_1=1}\|T_{t-s}g\|_1\leqslant\|T_sf\|_\infty=\phi(s).\end{aligned}$$

Since x_0 is an arbitrary point in X, we see from (5.5) that

$$\frac{\partial}{\partial t}\|T_tf\|_2^2\leqslant ct^{-1-\alpha/\beta}-c't^{-1+\alpha/\beta}\phi(t)^2. \tag{5.6}$$

Integrating (5.6) over $(\frac{1}{2}t,t)$, and then using (5.3) and the monotonicity of ϕ, we obtain

$$\begin{aligned}-\phi(t)&\leqslant\|T_tf\|_2^2-\|T_{t/2}f\|_2^2\\ &\leqslant c\int_{t/2}^t s^{-1-\alpha/\beta}\,\mathrm{d}s-c'\int_{t/2}^t s^{-1+\alpha/\beta}\phi(s)^2\,\mathrm{d}s\\ &\leqslant ct^{-\alpha/\beta}-c'\phi(t)^2\int_{t/2}^t s^{-1+\alpha/\beta}\,\mathrm{d}s\\ &=ct^{-\alpha/\beta}-c't^{\alpha/\beta}\phi(t)^2.\end{aligned}$$

Therefore,

$$c't^{\alpha/\beta}\phi(t)^2-\phi(t)-ct^{-\alpha/\beta}\leqslant 0,$$

which gives

$$0\leqslant\phi(t)\leqslant\frac{1+\sqrt{1+4cc'}}{2c't^{\alpha/\beta}}=ct^{-\alpha/\beta}, \tag{5.7}$$

proving (5.2) for $0<t\leqslant r_0^\beta$. The estimate (5.2) implies that the operator T_t is *ultracontractive* from $L^1(\mu)$ to $L^\infty(\mu)$ for $0<t\leqslant r_0^\beta$. Thus, the heat kernel $p(t,x,y)$ of $(\mathcal{E},\mathcal{F})$ exists [**13**], and (5.1) holds for all $0<t\leqslant r_0^\beta$ and almost all $x,y\in X$. Finally, using (3.7) and (5.1), we obtain the (Hölder) continuity of $p(t,x,y)$ by noting that

$$|p(t,x,y_1)-p(t,x,y_2)|^2\leqslant R(y_1,y_2)\mathcal{E}(p(t,x,\cdot))\leqslant ct^{-(1+\alpha/\beta)}d(y_1,y_2)^{\beta-\alpha}.$$

This completes the proof. □

The following proposition shows that the semigroup $\{T_t\}$ of the Dirichlet form $(\mathcal{E}, \mathcal{F})$ is a Feller semigroup if the hypotheses in Theorem 5.1 hold.

Proposition 5.2. *Assume that the hypotheses in Theorem 5.1 hold. Then $T_t u$ is continuous on X for any bounded function u on X and any $t > 0$. Moreover, $\{T_t\}$ is strongly continuous with uniform norm, that is*

$$\lim_{t\to 0} \|T_t u - u\|_{C(X)} = 0, \quad u \in C(X). \tag{5.8}$$

Proof. Let u be bounded on X. For a point $x_0 \in X$, let $\{x_k\}_{k\geqslant 1}$ be a sequence of points in X such that $x_k \to x_0$ as $k \to \infty$. Since the heat kernel $p(t, x, y)$ is continuous in x for $t > 0$ and $y \in X$, and $p(t, x_k, y) \leqslant ct^{-\alpha/\beta}$ by (5.1), using the dominated convergence theorem, we have

$$\lim_{k\to\infty} T_t u(x_k) = \lim_{k\to\infty} \int_X p(t, x_k, y)u(y)\,\mathrm{d}\mu(y) = \int_X p(t, x_0, y)u(y)\,\mathrm{d}\mu(y) = T_t u(x_0).$$

Thus, $T_t u$ is continuous on X for each $t > 0$. It remains to prove (5.8). To this end, note that, for any $u \in \mathcal{F}$,

$$\mathcal{E}(T_t u - u) = \int_0^\infty \lambda(\mathrm{e}^{-\lambda t} - 1)^2\,\mathrm{d}(E_\lambda u, u) \tag{5.9}$$

by using the spectral calculus, where $\{E_\lambda\}$ is the spectral family associated with the generator H of the form $(\mathcal{E}, \mathcal{F})$. Thus, by the dominated convergence theorem,

$$\lim_{t\to 0} \mathcal{E}(T_t u - u) = \lim_{t\to 0} \int_0^\infty \lambda(\mathrm{e}^{-\lambda t} - 1)^2\,\mathrm{d}(E_\lambda u, u) = 0. \tag{5.10}$$

On the other hand, from the upper bound in (A2) and (2.15) we have

$$\|u\|_{C(X)}^2 \leqslant c(\mathcal{E}(u) + \|u\|_2^2) \tag{5.11}$$

for all $u \in \mathcal{F}$. Replacing u by $T_t u - u$ in (5.11), we see that

$$\lim_{t\to 0} \|T_t u - u\|_{C(X)}^2 \leqslant c \lim_{t\to 0} (\mathcal{E}(T_t u - u) + \|T_t u - u\|_2^2) = 0 \tag{5.12}$$

for any $u \in \mathcal{F}$, by virtue of (5.10) and (2.4). Since $(\mathcal{E}, \mathcal{F})$ is regular and $\|T_t u\|_{C(X)} \leqslant \|u\|_{C(X)}$, we can easily see that (5.12) also holds for $u \in C(X)$. □

As in Theorem 5.1 and Proposition 5.2, we observe that, for any open subset D, the heat kernel $p_D(t, x, y)$ of the form $(\mathcal{E}, \mathcal{F}_D)$ also exists and is jointly continuous, and

$$\lim_{t\to 0} \|T_t^D u - u\|_{C(D)} = 0 \tag{5.13}$$

for any $u \in C_0(D)$.

We remark here that (5.8) was proved in [**18**, Lemma 5.2.7, p. 167] if X is a post-critically finite fractal with regular harmonic structure.

5.2. Off-diagonal upper bounds

We first give estimates of the effective resistance between any point $x_0 \in X$ and $B(x_0, r)^{\mathrm{c}}$ with $r \in (0, r_0]$. The locality of the form $(\mathcal{E}, \mathcal{F})$ will be used.

Proposition 5.3. *Let (X, d, μ) be a measure metric space satisfying the chain condition. Assume that $(\mathcal{E}, \mathcal{F})$ is a local regular Dirichlet form. If R satisfies (A2), then, for all $x_0 \in X$ and $0 < r \leqslant r_0$,*

$$R(x_0, B(x_0, r)^{\mathrm{c}}) \geqslant cr^{\gamma}. \tag{5.14}$$

Proof. The proof is essentially the same as in [**5**] for graphs or in [**20**] for effective forms. Let $D = \bar{B}(x_0, r) \setminus B(x_0, \frac{1}{2}r)$. For $y \in D$, let $\psi_y^{x_0}$ be the unique function such that $R(x_0, y)^{-1} = \mathcal{E}(\psi_y^{x_0})$ (recall that $\psi_y^{x_0}(y) = 0$ and $\psi_y^{x_0}(x_0) = 1$). It follows from (A2) that

$$\begin{aligned} \psi_y^{x_0}(x)^2 &= |\psi_y^{x_0}(x) - \psi_y^{x_0}(y)|^2 \\ &\leqslant R(x, y)\mathcal{E}(\psi_y^{x_0}) \\ &= R(x, y)R(x_0, y)^{-1} \\ &\leqslant c(d(y, x)r^{-1})^{\gamma} \leqslant \tfrac{1}{4} \end{aligned} \tag{5.15}$$

if $d(x, y) \leqslant \varepsilon_1 r$, where ε_1 is a small constant independent of x_0, x, y and r. Thus, $\psi_y^{x_0}(x) \leqslant \frac{1}{2}$ for $x \in B(y, \varepsilon_1 r)$. We cover D by N balls $\{B(y_i, \varepsilon_1 r)\}_{i=1}^N$. Since μ satisfies the doubling condition, the number N can be chosen independent of x_0, y_i and r. Define

$$v_0(x) = \min_{1 \leqslant i \leqslant N} \psi_{y_i}^{x_0}(x), \quad x \in X. \tag{5.16}$$

We see that $v_0(x_0) = 1$ since $\psi_{y_i}^{x_0}(x_0) = 1$, and $v_0 \leqslant \frac{1}{2}$ on D since $\psi_{y_i}^{x_0} \leqslant \frac{1}{2}$ on $B(y_i, \varepsilon_1 r)$, for $1 \leqslant i \leqslant N$. We claim that there exists some $c > 0$ independent of r and x_0 such that

$$\mathcal{E}(v_0) \leqslant cr^{-\gamma}. \tag{5.17}$$

In fact, from (5.16), for $x, z \in X$, we have

$$|v_0(x) - v_0(z)| \leqslant \max_{1 \leqslant i \leqslant N} |\psi_{y_i}^{x_0}(x) - \psi_{y_i}^{x_0}(z)| \leqslant \sum_{i=1}^{N} |\psi_{y_i}^{x_0}(x) - \psi_{y_i}^{x_0}(z)|.$$

Therefore, using (A2),

$$\begin{aligned} \mathcal{E}(v_0) &= \lim_{t \to 0} \frac{1}{2t} \int_X \int_X (v_0(x) - v_0(z))^2 p(t, x, z) \,\mathrm{d}\mu(z) \,\mathrm{d}\mu(x) \\ &\leqslant N \sum_{i=1}^{N} \lim_{t \to 0} \frac{1}{2t} \int_X \int_X |\psi_{y_i}^{x_0}(x) - \psi_{y_i}^{x_0}(z)|^2 p(t, x, z) \,\mathrm{d}\mu(z) \,\mathrm{d}\mu(x) \\ &= N \sum_{i=1}^{N} \mathcal{E}(\psi_{y_i}^{x_0}) \leqslant N^2 \max_{1 \leqslant i \leqslant N} R(x_0, y_i)^{-1} \leqslant cr^{-\gamma}, \end{aligned}$$

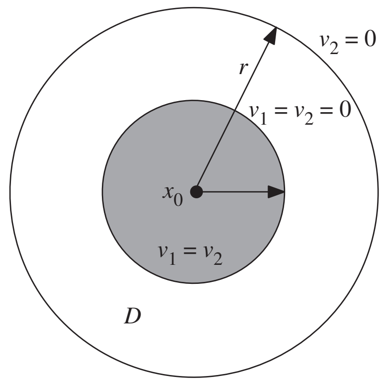

Figure 1. Functions v_1 and v_2.

proving (5.17). Let $v_1 = 2(v_0 - \frac{1}{2})_+$ and $v_2(\cdot) = v_1(\cdot)\mathbf{1}_{B(x_0,r)}(\cdot)$ (see Figure 1). By the locality of $\mathcal{E}$, we have $\mathcal{E}(v_2, v_1 - v_2) = 0$, and so

$$\mathcal{E}(v_2) = \mathcal{E}(v_1) - \mathcal{E}(v_1 - v_2) - 2\mathcal{E}(v_2, v_1 - v_2) \leqslant \mathcal{E}(v_1).$$

Noting that $v_2(x_0) = 1$, and $v_2 = 0$ on $B(x_0, r)^{\mathrm{c}}$, we see from (5.17) that

$$R(x_0, B(x_0, r)^{\mathrm{c}})^{-1} \leqslant \mathcal{E}(v_2) \leqslant \mathcal{E}(v_1) \leqslant 4\mathcal{E}(v_0) \leqslant cr^{-\gamma}, \tag{5.18}$$

proving the lemma. □

For $x_0 \in X$ and $0 < r < r_0$, define $B := B(x_0, r)$. Let φ_0 be a continuous function on X satisfying the condition that $0 \leqslant \varphi_0 \leqslant 1$, and

$$\varphi_0 = \begin{cases} 1 & \text{on } B(x_0, \frac{1}{2}r), \\ 0 & \text{on } B(x_0, \frac{2}{3}r)^{\mathrm{c}}. \end{cases} \tag{5.19}$$

Consider the Poisson equation with zero boundary condition

$$-H_B u_0 = \varphi_0 \quad \text{in } B, \tag{5.20}$$

$$u_0 = 0 \quad \text{in } B^{\mathrm{c}}, \tag{5.21}$$

where H_B is the generator of the Dirichlet form $(\mathcal{E}, \mathcal{F}_B)$. We say that a function u_0 defined on X is a *weak* solution to (5.20), (5.21) if $u_0 \in \mathcal{F}_B$ and

$$\mathcal{E}(u_0, v) = \int_B v(x)\varphi_0(x)\,\mathrm{d}\mu(x) \tag{5.22}$$

for any $v \in \mathcal{F}_B$. Note that a weak solution does not necessarily belong to the domain $\mathcal{D}(H_B)$ of H_B. If it does, and (5.20) holds pointwise in B, we call it a *strong* solution. Equation (5.20) with (5.21) has a unique weak solution,

$$u_0(x) = \int_X g^x_{B^{\mathrm{c}}}(y)\varphi_0(y)\,\mathrm{d}\mu(y) = \int_B g^x_{B^{\mathrm{c}}}(y)\varphi_0(y)\,\mathrm{d}\mu(y), \quad x \in X, \tag{5.23}$$

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

where $g^x_{B^{\mathrm{c}}}(\cdot) = R(x, B^{\mathrm{c}})\psi^x_{B^{\mathrm{c}}}(\cdot)$ is the Green function whose existence was proved in Proposition 4.2. In fact, from (5.23) and (4.5), for any $v \in \mathcal{F}_B$, we have

$$\begin{aligned}
\mathcal{E}(u_0, v) &= \lim_{t\to 0} \frac{1}{t} \int_X (u_0(x) - T_t u_0(x)) v(x)\,\mathrm{d}\mu(x) \\
&= \lim_{t\to 0} \frac{1}{t} \int_X \biggl(\int_X (g^y_{B^{\mathrm{c}}}(x) - (T_t g^y_{B^{\mathrm{c}}})(x)) v(x)\,\mathrm{d}\mu(x) \biggr) \varphi_0(y)\,\mathrm{d}\mu(y) \\
&= \lim_{t\to 0} \int_X \mathcal{E}_t(g^y_{B^{\mathrm{c}}}, v)\varphi_0(y)\,\mathrm{d}\mu(y) \\
&= \int_X \mathcal{E}(g^y_{B^{\mathrm{c}}}, v)\varphi_0(y)\,\mathrm{d}\mu(y) \\
&= \int_B v(y)\varphi_0(y)\,\mathrm{d}\mu(y), \qquad (5.24)
\end{aligned}$$

and so u_0 is a weak solution to (5.20), (5.21). The uniqueness easily follows from the irreducibility of $\mathcal{E}$.

Alternatively, the Green function $g^x_{B^{\mathrm{c}}}$ above may usefully be expressed as

$$g^x_{B^{\mathrm{c}}}(y) = \int_0^\infty p_B(t, x, y)\,\mathrm{d}t, \quad x, y \in B, \qquad (5.25)$$

where $p_B(t, x, y)$ is the heat kernel of $(\mathcal{E}, \mathcal{F}_B)$. (Note that $p_B(t, x, y)$ exists, and is continuous on $(0, \infty) \times B \times B$. The finiteness of the integral in (5.25) may be seen below.) In other words, the solution u_0 may also be written as

$$u_0(x) = \int_B \biggl(\int_0^\infty p_B(t, x, y)\,\mathrm{d}t \biggr) \varphi_0(y)\,\mathrm{d}\mu(y) = \int_0^\infty T^B_t \varphi_0(x)\,\mathrm{d}t, \quad x \in X. \qquad (5.26)$$

In fact, from (5.26) we have

$$u_0(x) - T^B_t u_0(x) = \int_0^t T^B_s \varphi_0(x)\,\mathrm{d}s,$$

and so, for $v \in \mathcal{F}_B$,

$$\mathcal{E}(u_0, v) = \lim_{t\to 0} \frac{1}{t} \int_B (u_0(x) - T^B_t u_0(x)) v(x)\,\mathrm{d}\mu(x) = \int_B v(x)\varphi_0(x)\,\mathrm{d}\mu(x).$$

Remark 5.4. Let X_t be the Hunt process associated with the regular Dirichlet form $(\mathcal{E}, \mathcal{F})$. Then the solution to (5.20), (5.21) can be written as

$$u_0(x) = \mathbb{E}_x \biggl(\int_0^{\tau_B} \varphi_0(X_t)\,\mathrm{d}t \biggr). \qquad (5.27)$$

In fact, using the Fubini theorem, we see from (5.26) that

$$\begin{aligned}
u_0(x) &= \int_0^\infty \mathbb{E}^B_x(\varphi_0(X_t))\,\mathrm{d}t \\
&= \int_0^\infty \mathbb{E}_x(\mathbf{1}_{\{t<\tau_B\}} \varphi_0(X_t))\,\mathrm{d}t
\end{aligned}$$

$$= \mathbb{E}_x\biggl(\int_0^\infty \mathbf{1}_{\{t<\tau_B\}}\varphi_0(X_t)\,\mathrm{d}t\biggr)$$
$$= \mathbb{E}_x\biggl(\int_0^{\tau_B} \varphi_0(X_t)\,\mathrm{d}t\biggr). \tag{5.28}$$

We now estimate u_0.

Proposition 5.5. *Assume that all the hypotheses in Proposition 5.3 hold, and that μ satisfies (A1). Let u_0 be the solution to (5.20), (5.21). There then exist constants $c_3, c_4, \varepsilon_2 > 0$ independent of x_0 and r such that*

$$c_3 r^\beta \leqslant \min_{x\in B(x_0,\varepsilon_2 r)} u_0(x) \leqslant \max_{x\in X} u_0(x) \leqslant c_4 r^\beta. \tag{5.29}$$

Proof. The key is to estimate the Green function $g^x_{B^{\mathrm{c}}}$. This can be done by using (A1) and (A2). Using (4.13) and the upper bound of R in (A2), we have

$$g^x_{B^{\mathrm{c}}}(y) \leqslant g^x_{B^{\mathrm{c}}}(x) = R(x, B^{\mathrm{c}}) \leqslant \inf_{z\in B^{\mathrm{c}}} R(x,z) \leqslant c_2 r^\gamma \tag{5.30}$$

for all $x, y \in X$. Thus, it follows from (5.23) and (A1) that, using $\gamma = \beta - \alpha$,

$$u_0(x) \leqslant c_2 r^\gamma \mu(B(x_0, r)) \leqslant c_4 r^{\gamma+\alpha} = c_4 r^\beta$$

for all $x \in X$, proving the third inequality in (5.29). On the other hand, let $\psi^x_{B^{\mathrm{c}}}$ be such that $R(x, B^{\mathrm{c}})^{-1} = \mathcal{E}(\psi^x_{B^{\mathrm{c}}})$, for $x \in B(x_0, \frac{1}{2}r)$. It follows from (5.14) and (A2) that, for $x \in B(x_0, \frac{1}{2}r)$,

$$\begin{aligned}(1 - \psi^x_{B^{\mathrm{c}}}(y))^2 &= (\psi^x_{B^{\mathrm{c}}}(x) - \psi^x_{B^{\mathrm{c}}}(y))^2 \\ &\leqslant R(x,y)\mathcal{E}(\psi^x_{B^{\mathrm{c}}}) \\ &= R(x,y)R(x,B^{\mathrm{c}})^{-1} \\ &\leqslant R(x,y)R(x,B(x,\tfrac{1}{4}r)^{\mathrm{c}})^{-1} \\ &\leqslant c(d(x,y)r^{-1})^\gamma \\ &\leqslant \tfrac{1}{4},\end{aligned}$$

if $y \in B(x, 2\varepsilon_2 r)$ for a small $\varepsilon_2 \in (0, \frac{1}{2})$ independent of x_0 and r. Thus, $\psi^x_{B^{\mathrm{c}}}(y) \geqslant \frac{1}{2}$. Therefore, for all $x, y \in B(x_0, \varepsilon_2 r)$,

$$g^x_{B^{\mathrm{c}}}(y) = R(x, B^{\mathrm{c}})\psi^x_{B^{\mathrm{c}}}(y) \geqslant \tfrac{1}{2}R(x, B(x, \tfrac{1}{4}r)^{\mathrm{c}}) \geqslant c r^\gamma, \tag{5.31}$$

where c is independent of x_0 and r. Hence, from (5.23) and (A1) we have

$$u_0(x) = \int_B g^x_{B^{\mathrm{c}}}(y)\varphi_0(y)\,\mathrm{d}\mu(y) \geqslant \int_{B(x_0,\varepsilon_2 r)} g^x_{B^{\mathrm{c}}}(y)\,\mathrm{d}\mu(y) \geqslant c_3 r^\beta,$$

for all $x \in B(x_0, \varepsilon_2 r)$. This proves the first inequality in (5.29). □

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

Remark 5.6. By Remark 5.4, the estimates in (5.29) may be obtained by estimating $\mathbb{E}_x(\tau_B)$ and $\mathbb{E}_x(\tau_{B(x_0,r/2)})$ by using the fact that

$$\mathbb{E}_x(\tau_{B(x_0,r/2)}) \leqslant u_0(x) \leqslant \mathbb{E}_x(\tau_B), \quad x \in B. \tag{5.32}$$

For $x_0 \in X$ and $0 < r < r_0$, let $B = B(x_0, r)$ as before. The estimates in (5.29) give rise to a lower integral estimate for the heat kernel $p_B(t, x, y)$, that is, there exist two constants $c_5 \in (0, 1)$ and $c_6 > 0$ independent of x_0, r and t such that, for all $x \in B(x_0, \varepsilon_2 r)$ and $t > 0$,

$$\int_B p_B(t, x, y) \, \mathrm{d}\mu(y) \geqslant c_5 - c_6 r^{-\beta} t. \tag{5.33}$$

Indeed, let u_0 be the solution to (5.20), (5.21). Using the semigroup property, we see from (5.26) that, for $t > 0$,

$$\begin{aligned}
\int_t^\infty &\left(\int_B p_B(s, x, y)\varphi_0(y) \, \mathrm{d}\mu(y) \right) \mathrm{d}s \\
&= \int_0^\infty \left(\int_B p_B(s + t, x, y)\varphi_0(y) \, \mathrm{d}\mu(y) \right) \mathrm{d}s \\
&= \int_0^\infty \left(\int_B \left\{ \int_B p_B(t, x, z) p_B(s, z, y) \, \mathrm{d}\mu(z) \right\} \varphi_0(y) \, \mathrm{d}\mu(y) \right) \mathrm{d}s \\
&= \int_B u_0(z) p_B(t, x, z) \, \mathrm{d}\mu(z) \\
&\leqslant \max_{z \in B} u_0(z) \int_B p_B(t, x, z) \, \mathrm{d}\mu(z).
\end{aligned}$$

Therefore, noting that

$$\int_B p_B(s, x, y) \, \mathrm{d}\mu(y) \leqslant 1 \quad \text{and} \quad 0 \leqslant \varphi_0 \leqslant 1,$$

we have

$$\begin{aligned}
u_0(x) &= \int_0^\infty \left(\int_B p_B(s, x, y)\varphi_0(y) \, \mathrm{d}\mu(y) \right) \mathrm{d}s \\
&= \int_0^t \left(\int_B p_B(s, x, y)\varphi_0(y) \, \mathrm{d}\mu(y) \right) \mathrm{d}s + \int_t^\infty \left(\int_B p_B(s, x, y)\varphi_0(y) \, \mathrm{d}\mu(y) \right) \mathrm{d}s \\
&\leqslant t + \max_{z \in B} u_0(z) \int_B p_B(t, x, z) \, \mathrm{d}\mu(z),
\end{aligned}$$

which yields

$$\int_B p_B(t, x, z) \, \mathrm{d}\mu(z) \geqslant \frac{u_0(x) - t}{\max u_0}, \quad x \in X, \ t > 0. \tag{5.34}$$

Thus, (5.33) follows by using (5.29). Combining (5.33) and (2.11), we see that

$$\int_{B(x_0,r)^{\mathrm{c}}} p(t,x,y)\,\mathrm{d}\mu(y) = 1 - \int_B p(t,x,y)\,\mathrm{d}\mu(y) \leqslant 1 - \int_B p_B(t,x,y)\,\mathrm{d}\mu(y) \leqslant \varepsilon + c_6 r^{-\beta} t \tag{5.35}$$

for all $t > 0$ and $x \in B(x_0, \varepsilon_2 r)$, where $\varepsilon = 1 - c_5 \in (0,1)$.

Remark 5.7. From the viewpoint of probability theory, we have

$$\int_{B(x_0,r)^{\mathrm{c}}} p(t,x,y)\,\mathrm{d}\mu(y) = T_t \mathbf{1}_{B^{\mathrm{c}}}(x) = \mathbb{P}_x(X_t \in B^{\mathrm{c}}) \leqslant \mathbb{P}_x(\tau_B \leqslant t). \tag{5.36}$$

Thus, (5.35) can be obtained by estimating $\mathbb{P}_x(\tau_B \leqslant t)$. The reader may consult [**2**, Lemma 3.16, p. 33] for the estimates on $\mathbb{P}_x(\tau_B \leqslant t)$ for the special case $x = x_0$ when assuming that $\mathbb{E}_y(\tau_{B(y,r)}) \asymp r^\beta$ for any $y \in X$.

Remark 5.8. The estimate (5.33) is interesting. To see this, note that, for any $\lambda > 0$ and $x \in X$,

$$\begin{aligned}\lambda \int_0^\infty \mathrm{e}^{-\lambda t} T_t^B 1(x)\,\mathrm{d}t &= \lambda \int_0^\infty \mathrm{e}^{-\lambda t} \mathbb{E}_x(\mathbf{1}_{\{t<\tau_B\}})\,\mathrm{d}t \\ &= \mathbb{E}_x\biggl(\int_0^{\tau_B} \lambda \mathrm{e}^{-\lambda t}\,\mathrm{d}t\biggr) \\ &= 1 - \mathbb{E}_x(\mathrm{e}^{-\lambda \tau_B}).\end{aligned} \tag{5.37}$$

On the other hand, we see from (5.33) that, for all $x \in B(x_0, \varepsilon_2 r)$,

$$\begin{aligned}\lambda \int_0^\infty \mathrm{e}^{-\lambda t} T_t^B 1(x)\,\mathrm{d}t &= \lambda \int_0^\infty \mathrm{e}^{-\lambda t}\biggl(\int_B p_B(t,x,y)\,\mathrm{d}\mu(y)\biggr)\mathrm{d}t \\ &\geqslant \lambda \int_0^\infty \mathrm{e}^{-\lambda t}(c_5 - c_6 r^{-\beta} t)\,\mathrm{d}t \\ &= c_5 - c(\lambda r^\beta)^{-1}.\end{aligned} \tag{5.38}$$

Combining (5.37) and (5.38), we obtain

$$\mathbb{E}_x(\mathrm{e}^{-\lambda \tau_B}) = 1 - \lambda \int_0^\infty \mathrm{e}^{-\lambda t} T_t^B 1(x)\,\mathrm{d}t \leqslant (1 - c_5) + c(\lambda r^\beta)^{-1} \tag{5.39}$$

for all $x \in B(x_0, \varepsilon_2 r)$ and $\lambda > 0$.

Estimate (5.35) is useful, but it is not good enough to obtain optimal off-diagonal upper bounds of $p(t,x,y)$. We need a more delicate estimate than (5.35) (see (5.55), below). In order to do this, we first estimate the solution to an elliptic equation in the ball $B = B(x_0, r)$. Let Δ be defined as in (2.7). For $\lambda > 0$, consider a function $u_1 \in \mathcal{D}(\Delta)$ satisfying the equation

$$\Delta u_1 = \lambda u_1 \quad \text{in } B. \tag{5.40}$$

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

Here u_1 is a *strong* solution, that is $u_1 \in \mathcal{D}(\Delta)$ satisfies (5.40) pointwise in B. The solution u_1 of (5.40) exists, which coincides in $B(x_0, \frac{1}{2}r)$ with u_λ determined by (5.52) or (5.53), below.

Lemma 5.9. *Assume that all the hypotheses (A1), (A2) and (2.16) hold. Let u_1 satisfy (5.40). If $0 \leqslant u_1 \leqslant 1$ on $\bar{B}$, then there exists some $\varepsilon_3 \in (0,1)$ independent of x_0, r and λ such that*

$$u_1(x) \leqslant \varepsilon_3 \tag{5.41}$$

for all $x \in B(x_0, \varepsilon_2 r)$, provided that λr^β is sufficiently large.

Proof. Set $u = 1 - u_1$. Then u satisfies

$$(\lambda - \Delta)u = \lambda \quad \text{in } B. \tag{5.42}$$

Let φ_0 be as before; see (5.19). Let $v \in \mathcal{F}_B$ be the solution to the equation

$$(\lambda - \Delta_B)v = \lambda\varphi_0 \quad \text{in } B, \tag{5.43}$$

where Δ_B is defined in the same way as in (2.7) for the semigroup $\{T_t^B\}$ of the form $(\mathcal{E}, \mathcal{F}_B)$. It is easy to see that

$$v(x) = \lambda \int_0^\infty \mathrm{e}^{-\lambda t} T_t^B \varphi_0(x)\,\mathrm{d}t, \quad x \in B. \tag{5.44}$$

Note that v is a *strong* solution to (5.43), that is, the function $v \in \mathcal{D}(\Delta_B)$ satisfies (5.43) pointwise in B. Indeed, one can easily verify from (5.44) that

$$\lim_{t\to 0} \|t^{-1}(T_t^B v - v) - \lambda(v - \varphi_0)\|_{C(B)} = 0$$

by using (5.13), and so $\Delta_B v = \lambda(v - \varphi_0)$ in B. Now we claim that

$$v(x) \leqslant u(x), \quad x \in \bar{B}. \tag{5.45}$$

Indeed, letting $h = u - v$, we see that h is continuous on $\bar{B}$ since both u and v are continuous on $\bar{B}$. Since $h|_{\partial B} = u|_{\partial B} \geqslant 0$, the inequality (5.45) holds if we can show that $h \geqslant 0$ in B. Assume that this is not true. Then there would be a point $y_0 \in B$ such that $h(y_0) = \min_{\bar{B}} h < 0$. Let $\delta > 0$ be so small that $B(y_0, \delta) \subset B$. Thus, it follows from (2.17) that

$$\begin{aligned}
\lim_{t\to 0} \frac{1}{t}(T_t h(y_0) - h(y_0)) &= \lim_{t\to 0} \frac{1}{t}\biggl\{ \int_{B(y_0,\delta)} (h(y) - h(y_0))p(t, y_0, y)\,\mathrm{d}\mu(y) \\
&\qquad + \int_{B(y_0,\delta)^{\mathrm{c}}} (h(y) - h(y_0))p(t, y_0, y)\,\mathrm{d}\mu(y)\biggr\} \\
&\geqslant \lim_{t\to 0} \frac{1}{t} \int_{B(y_0,\delta)^{\mathrm{c}}} (h(y) - h(y_0))p(t, y_0, y)\,\mathrm{d}\mu(y) \\
&= 0.
\end{aligned}$$

Therefore, by (5.42) and (5.43) and using (2.11),

$$\begin{aligned} 0 > \lambda h(y_0) &= \lambda(1-\varphi_0(y_0)) + \Delta u(y_0) - \Delta_B v(y_0) \\ &\geqslant \lim_{t\to 0} \frac{T_t u(y_0) - u(y_0) - (T_t^B v(y_0) - v(y_0))}{t} \\ &\geqslant \lim_{t\to 0} \frac{T_t h(y_0) - h(y_0)}{t} \geqslant 0, \end{aligned} \tag{5.46}$$

which is a contradiction. This proves (5.45). Finally, by (5.33) with B replaced by $B(x_0, \frac{1}{2}r)$,

$$\begin{aligned} T_t^B \varphi_0(x) &= \int_B p_B(t,x,y)\varphi_0(y)\,\mathrm{d}\mu(y) \\ &\geqslant \int_{B(x_0,r/2)} p_{B(x_0,r/2)}(t,x,y)\,\mathrm{d}\mu(y) \geqslant c_5 - 2^\beta c_6 r^{-\beta} t, \end{aligned}$$

and so

$$v(x) = \lambda \int_0^\infty \mathrm{e}^{-\lambda t} T_t^B \varphi_0(x)\,\mathrm{d}t \geqslant c_5 - c(r^\beta t)^{-1}.$$

This combines with (5.45), for all $x \in B(x_0, \varepsilon_2 r)$, to give

$$u_1(x) = 1 - u(x) \leqslant 1 - v(x) \leqslant (1-c_5) + c(\lambda r^\beta)^{-1} \leqslant \varepsilon_3, \tag{5.47}$$

if λr^β is sufficiently large. □

The estimate (5.41) will give a more accurate bound of u_1 at the point x_0.

Lemma 5.10. *Assume that the hypotheses in Lemma 5.9 hold. There then exist* $c, c_7 > 0$ *independent of* x_0, r *and* λ *such that*

$$u_1(x_0) \leqslant c \exp(-c_7 \lambda^{1/\beta} r). \tag{5.48}$$

Proof. We consider only the case where λr^β is large; otherwise, (5.48) is trivial since $u_1 \leqslant 1$ on $\bar{B}$. Let $r' = r/n$, where $n \geqslant 2$ is an integer to be determined below. Denote by x_i the maximum point of u_1 on the closed ball $\bar{B}(x_0, ir')$ for $1 \leqslant i \leqslant n$ (since u_1 is continuous, each of such points x_i exists). Set $m_i = u_1(x_i)$ for $1 \leqslant i \leqslant n$.

Consider the ball $B_i := B(x_i, \frac{1}{2}r')$. For $1 \leqslant i \leqslant n-1$, define

$$v_i(x) = \frac{u_1(x)}{m_{i+1}} \quad \text{for } x \in X.$$

Then v_i satisfies the condition that $\Delta v_i = \lambda v_i$ in B_i (see Figure 2).

By Lemma 5.9 with x_0 replaced by x_i and r by $\frac{1}{2}r'$, and noting that $0 \leqslant v_i \leqslant 1$ on $\bar{B}_i$, we find that $v_i(x_i) \leqslant \varepsilon_3$ if $\lambda(\frac{1}{2}r')^\beta \geqslant c$ for some large c. Therefore,

$$m_i \leqslant \varepsilon_3 m_{i+1}, \quad 1 \leqslant i \leqslant n-1,$$

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

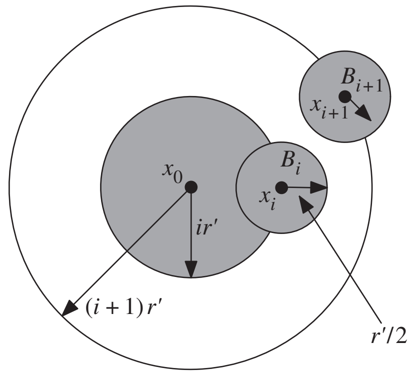

Figure 2. $\Delta v_i = \lambda v_i$ in B_i, and $0 \leqslant v_i \leqslant 1$ on $\bar{B}_i$.

which gives

$$m_1 \leqslant \varepsilon_3^{n-1} m_n \leqslant \varepsilon_3^{n-1}$$

by noting that $m_n \leqslant 1$. Setting $c' = -\log \varepsilon_3 > 0$, we see that

$$u_1(x_0) \leqslant m_1 = u_1(x_1) \leqslant \exp(-c'(n-1)) \leqslant \exp(-c_7 \lambda^{1/\beta} r),$$

by choosing the largest integer n so that $\lambda(\frac{1}{2}r')^\beta \geqslant c$. Therefore, (5.48) follows. □

The proof of Lemma 5.10 given here is motivated by [**15**, Lemma 5.4] (see also [**14**, Lemma 7.3]) on infinite graphs.

Remark 5.11. It follows from (5.39) that

$$\mathbb{E}_{x_0}(\mathrm{e}^{-\lambda \tau_B}) \leqslant \varepsilon_3 \tag{5.49}$$

if λr^β is sufficiently large. By the locality of $\mathcal{E}$ and using the chain argument (see [**2**, pp. 34, 35] or [**13**, Theorem 9.1, (iv) ⇒ (v)]), one can find from (5.49) that

$$\mathbb{E}_{x_0}(\mathrm{e}^{-\lambda \tau_B}) \leqslant c \exp(-c_7 \lambda^{1/\beta} r) \tag{5.50}$$

by applying the strong Markov property of the diffusion $(\{X_t\}_{t \geqslant 0}, \{\mathbb{P}_x\}_{x \in X})$ of the form $(\mathcal{E}, \mathcal{F})$. On the other hand, letting $B_1 = B(x_0, \frac{1}{2}r)$, we see from (5.37) that

$$v(x_0) = \lambda \int_0^\infty \mathrm{e}^{-\lambda t} T_t^B \varphi_0(x_0)\,\mathrm{d}t \geqslant \lambda \int_0^\infty \mathrm{e}^{-\lambda t} T_t^{B_1} 1(x_0)\,\mathrm{d}t = 1 - \mathbb{E}_{x_0}(\mathrm{e}^{-\lambda \tau_{B_1}}). \tag{5.51}$$

Thus, combining (5.47) and (5.50) with r replaced by $\frac{1}{2}r$, it follows that

$$u_1(x_0) = 1 - u(x_0) \leqslant 1 - v(x_0) \leqslant \mathbb{E}_{x_0}(\mathrm{e}^{-\lambda \tau_{B_1}}) \leqslant c \exp(-c' \lambda^{1/\beta} r).$$

Hence, (5.48) can be also obtained by using probability theory.

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

To obtain the key estimate (5.55) below, we introduce a function u_λ on X determined by the equation

$$(\lambda - \Delta)u_\lambda = \lambda\varphi_1 \quad \text{on } X \tag{5.52}$$

for $\lambda > 0$, where $0 \leqslant \varphi_1 \leqslant 1$ is a continuous function on X satisfying $\varphi_1(x) = 1$ for $x \in B(x_0, \frac{2}{3}r)^{\mathrm{c}}$ and $\varphi_1(x) = 0$ for $x \in B(x_0, \frac{1}{2}r)$. Observe that (5.52) has a unique strong solution

$$u_\lambda(x) = \lambda \int_0^\infty \exp(-\lambda t) T_t\varphi_1(x)\,\mathrm{d}t, \quad x \in X, \tag{5.53}$$

for any $\lambda > 0$, by using (5.8). By (5.53), we have $0 \leqslant u_\lambda \leqslant 1$ on X. Note that, for $x \in B(x_0, r)^{\mathrm{c}}$, using (5.33) with x_0 replaced by x, we have

$$\begin{aligned} T_t\varphi_1(x) &\geqslant \int_{B(x_0,2r/3)^{\mathrm{c}}} p(t,x,y)\,\mathrm{d}\mu(y) \\ &\geqslant \int_{B(x,r/3)} p(t,x,y)\,\mathrm{d}\mu(y) \\ &\geqslant \int_{B(x,r/3)} p_{B(x,r/3)}(t,x,y)\,\mathrm{d}\mu(y) \\ &\geqslant c_5 - c_6(\tfrac{1}{3}r)^{-\beta}t \geqslant \tfrac{1}{2}c_5 \end{aligned}$$

if $r^\beta t^{-1} \geqslant c$ for a large c. Thus, from (5.53) we have

$$u_\lambda(x) \geqslant \lambda \int_0^{cr^\beta} \mathrm{e}^{-\lambda t} T_t\varphi_1(x)\,\mathrm{d}t \geqslant \tfrac{1}{2}c_5(1 - \mathrm{e}^{-c\lambda r^\beta}) \geqslant \tfrac{1}{4}c_5$$

if λr^β is large. Therefore,

$$u_\lambda(x) \geqslant \tfrac{1}{4}c_5 \mathbf{1}_{B(x_0,r)^{\mathrm{c}}}(x) \tag{5.54}$$

for all $x \in X$ if λr^β is large.

Theorem 5.12. *Let (X, d, μ) be a metric measure space satisfying the chain condition, and let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form. Assume that the hypotheses (A1), (A2) and (2.16) all hold. Then, for $x_0 \in X$ and $r \in (0, r_0)$,*

$$\int_{B(x_0,r)^{\mathrm{c}}} p(t, x_0, y)\,\mathrm{d}\mu(y) \leqslant c\exp(-c_8(r^\beta t^{-1})^{1/(\beta-1)}), \tag{5.55}$$

where c, c_8 are independent of x_0, r and t.

Proof. Let u_λ be as in (5.52). Observe that $u_\lambda \in \mathcal{D}(\Delta)$ satisfies

$$\Delta u_\lambda = \lambda u_\lambda \quad \text{in } B(x_0, \tfrac{1}{2}r).$$

Thus, it follows from (5.48) that

$$u_\lambda(x_0) \leqslant c\exp(-c_7\lambda^{1/\beta}r). \tag{5.56}$$

Define

$$w_\lambda(t,x) = \mathrm{e}^{\lambda t} u_\lambda(x), \quad t>0,\ x\in X. \tag{5.57}$$

Clearly, for $x \in X$ and $t > 0$,

$$\frac{\partial w_\lambda}{\partial t}(t,x) = \lambda \mathrm{e}^{\lambda t} u_\lambda(x) = \Delta w_\lambda(x) + \lambda \mathrm{e}^{\lambda t}\varphi_1(x)$$

with initial value $w_\lambda(0,x) = u_\lambda(x)(x \in X)$. Thus, we see that, using (5.54),

$$\begin{aligned} w_\lambda(t,x) &= \int_X u_\lambda(y)p(t,x,y)\,\mathrm{d}\mu(y) + \lambda\int_0^t \mathrm{e}^{\lambda s}\,\mathrm{d}s \int_X \varphi_1(y)p(t-s,x,y)\,\mathrm{d}\mu(y) \\ &\geqslant \tfrac{1}{4}c_5 \int_{B(x_0,r)^{\mathrm{c}}} p(t,x,y)\,\mathrm{d}\mu(y) \end{aligned} \tag{5.58}$$

for $t > 0$, $x \in X$ and any $\lambda > 0$, if $\lambda r^\beta \geqslant c$ for a large c. For a suitable λ, we need to bound $w_\lambda(t,x_0)$ for $t > 0$. Indeed, from (5.57) and (5.56) we obtain

$$w_\lambda(t,x_0) \leqslant c\exp(\lambda t - c_7\lambda^{1/\beta} r) \leqslant c\exp(-c_7(t^{-1}r^\beta)^{1/(\beta-1)}),$$

by letting $2\lambda t = c_7\lambda^{1/\beta} r$, if $t^{-1}r^\beta \geqslant c$. Thus, it follows from (5.58) that

$$\int_{B(x_0,r)^{\mathrm{c}}} p(t,x_0,y)\,\mathrm{d}\mu(y) \leqslant \frac{4}{c_5} w_\lambda(t,x_0) \leqslant c\exp(-c_7(t^{-1}r^\beta)^{1/(\beta-1)}), \tag{5.59}$$

giving (5.55), if $t^{-1}r^\beta \geqslant c$ for a large c. However, if $t^{-1}r^\beta \leqslant c$, then (5.55) is obvious. □

Remark 5.13. Note that from (5.50) one can easily obtain

$$\begin{aligned} \mathbb{P}_{x_0}(\tau_B \leqslant t) &\leqslant \mathrm{e}^{\lambda t}\mathbb{E}_{x_0}(\mathrm{e}^{-\lambda\tau_B}) \\ &\leqslant c\exp(\lambda t - c_7\lambda^{1/\beta} r) \\ &\leqslant c\exp(-c_7(t^{-1}r^\beta)^{1/(\beta-1)}), \end{aligned} \tag{5.60}$$

by choosing a suitable λ, as above. Thus, (5.55) can alternatively be obtained in this way, by noting the fact that

$$\int_{B^{\mathrm{c}}} p(t,x_0,y)\,\mathrm{d}\mu(y) = \mathbb{P}_{x_0}(X_t \in B^{\mathrm{c}}) \leqslant \mathbb{P}_{x_0}(\tau_B \leqslant t).$$

We are now in a position to derive off-diagonal upper bounds of $p(t,x,y)$ by using (5.1) and (5.55). Fix $x_0, y_0 \in X$, $x_0 \neq y_0$, and let $r = \frac{1}{2}d(x_0,y_0)$. By (5.1) and (5.55), we have

$$\begin{aligned} \int_{B(x_0,r)^{\mathrm{c}}} p(t,x_0,z)p(t,z,y_0)\,\mathrm{d}\mu(z) &\leqslant ct^{-\alpha/\beta}\int_{B(x_0,r)^{\mathrm{c}}} p(t,x_0,z)\,\mathrm{d}\mu(z) \\ &\leqslant ct^{-\alpha/\beta}\exp(-c'(t^{-1}d(y_0,x_0)^\beta)^{1/(\beta-1)}). \end{aligned}$$

https://doi.org/10.1017/S001309150500177X Published online by Cambridge University Press

The above estimate is true if we exchange x_0 and y_0. Therefore, by the semigroup property of $p(t,x,y)$, it follows that

$$\begin{aligned}p(2t,x_0,y_0) &= \int_X p(t,x_0,z)p(t,z,y_0)\,\mathrm{d}\mu(z)\\ &\leqslant \int_{B(x_0,r)^c} p(t,x_0,z)p(t,z,y_0)\,\mathrm{d}\mu(z) + \int_{B(y_0,r)^c} p(t,x_0,z)p(t,z,y_0)\,\mathrm{d}\mu(z)\\ &\leqslant ct^{-\alpha/\beta}\exp(-c'(t^{-1}d(y_0,x_0)^\beta)^{1/(\beta-1)}),\end{aligned}$$

giving the upper bound of $p(t,x,y)$ in (A3) for $0<t\leqslant r_0$.

5.3. Off-diagonal lower bounds

Lower bounds of $p(t,x,y)$ can be obtained in a standard way. We sketch the proof, as follows. We first deduce on-diagonal lower bounds. Indeed, letting $B_1 := B(x,(\lambda_1 t)^{1/\beta})$ for some $\lambda_1>0$ to be specified later on, we see from (5.33) that

$$\int_{B_1} p(t,x,y)\,\mathrm{d}\mu(y) \geqslant \int_{B_1} p_{B_1}(t,x,y)\,\mathrm{d}\mu(y) \geqslant c_5 - c_6\lambda_1^{-1} = \tfrac{1}{2}c_5$$

if $2\lambda_1^{-1}c_6 = c_5$. Thus, for all $x\in X$ and $0<t\leqslant r_0$, using Hölder's inequality,

$$\begin{aligned}p(2t,x,x) &= \int_X p(t,x,y)^2\,\mathrm{d}\mu(y)\\ &\geqslant \int_{B_1} p(t,x,y)^2\,\mathrm{d}\mu(y)\\ &\geqslant \mu(B_1)^{-1}\biggl(\int_{B_1} p(t,x,y)\,\mathrm{d}\mu(y)\biggr)^2\\ &\geqslant ct^{-\alpha/\beta},\end{aligned}\tag{5.61}$$

giving an on-diagonal lower bound of $p(t,x,y)$. We next derive *near-diagonal* lower bounds of $p(t,x,y)$. Observe that, by (3.7) and (5.1),

$$\mathcal{E}(p(t,x,\cdot)) \leqslant ct^{-(1+\alpha/\beta)}$$

for all $x\in X$ and $0<t\leqslant r_0^\beta$. Therefore, by (5.61),

$$\begin{aligned}p(t,x,y) &\geqslant p(t,x,x) - |p(t,x,x)-p(t,x,y)|\\ &\geqslant ct^{-\alpha/\beta} - R(x,y)^{1/2}\mathcal{E}(p(t,x,\cdot))^{1/2}\\ &\geqslant ct^{-\alpha/\beta} - c_9(d(x,y)^{\beta-\alpha}t^{-(1+\alpha/\beta)})^{1/2}\\ &\geqslant 2^{-1}ct^{-\alpha/\beta}\end{aligned}\tag{5.62}$$

if $d(x,y)\leqslant \varepsilon_4 t^{1/\beta}$ for a small ε_4 independent of t, x and y. Finally, off-diagonal lower bounds of $p(t,x,y)$ follow from (5.61) and (5.62) by using the chain argument (see [**2**, pp. 36, 37] or [**16**, Corollary 3.5]). We omit the details.

Acknowledgements. The author thanks K. J. Falconer for his comments, and Alexander Grigor'yan for pointing out the simpler proof of (3.7). The author also thanks the referee for his/her suggestions. This work was partly supported by NSFC Grant no. 10371062 and the Alexander von Humboldt Foundation.

References

1. D. G. Aronson, Bounds for the fundamental solution of a parabolic equation, *Bull. Am. Math. Soc.* **73** (1967), 890–896.
2. M. T. Barlow, Diffusions on fractals, in *Lectures on Probability Theory and Statistics*, Lecture Notes on Mathematics, Volume 1690, pp. 1–121 (Springer, 1998).
3. M. T. Barlow and R. F. Bass, Brownian motion and harmonic analysis on Sierpiński carpets, *Can. J. Math.* **51** (1999), 673–744.
4. M. T. Barlow and E. A. Perkins, Brownian motion on the Sierpiński gasket, *Prob. Theory Relat. Fields* **79** (1988), 543–623.
5. M. Barlow, T. Coulhon and T. Kumagai, Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs, *Commun. Pure Appl. Math.* **58** (2005), 1642–1677.
6. E. B. Davies, *Heat kernels and spectral theory* (Cambridge University Press, 1989).
7. J. Dodziuk, Maximum principle for parabolic inequalities and the heat flow on open manifolds, *Indiana Univ. Math. J.* **32** (1983), 703–716.
8. E. B. Fabes and D. W. Stroock, A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash, *Arch. Ration. Mech. Analysis* **96** (1986), 327–338.
9. K. J. Falconer, *Fractal geometry: mathematical foundation and applications* (Wiley, 1992).
10. W. Feller, The general diffusion operator and positivity preserving semi-groups in one dimension, *Ann. Math. (2)* **60** (1954), 417–436.
11. P. J. Fitzsimmons, B. M. Hambly and T. Kumagai, Transition density estimates for Brownian motion on affine nested fractals, *Commun. Math. Phys.* **165** (1994), 595–620.
12. M. Fukushima, Y. Oshima and M. Takeda, *Dirichlet forms and symmetric Markov processes* (de Gruyter, Berlin, 1994).
13. A. Grigor'yan, Heat kernel upper bounds on fractal spaces, preprint (2005).
14. A. Grigor'yan and A. Telcs, Sub-Gaussian estimates of heat kernels on infinite graphs, *Duke Math. J.* **109** (2001), 451–510.
15. A. Grigor'yan and A. Telcs, Harnack inequalities and sub-Gaussian estimates for random walks, *Math. Annln* **324** (2002), 521–556.
16. A. Grigor'yan, J. Hu and K.-S. Lau, Heat kernels on metric-measure spaces and an application to semilinear elliptic equations, *Trans. Am. Math. Soc.* **355** (2003), 2065–2095.
17. B. M. Hambly and T. Kumagai, Transition density estimates for diffusion processes on post critically finite self-similar fractals, *Proc. Lond. Math. Soc.* **78** (1999), 431–458.
18. J. Kigami, *Analysis on fractals* (Cambridge University Press, 2001).
19. J. Kigami, Harmonic analysis for resistance forms, *J. Funct. Analysis* **204** (2003), 399–444.
20. T. Kumagai, Heat kernel estimates and parabolic Harnack inequalities on graphs and resistance forms, *Publ. RIMS Kyoto* **40** (2004), 793–818.
21. D. Ray, Stationary Markov processes with continuous paths, *Trans. Am. Math. Soc.* **82** (1956), 452–493.
22. K. Yosida, *Functional analysis* (Springer, 1980).