NASH-TYPE INEQUALITIES AND HEAT KERNELS FOR NON-LOCAL DIRICHLET FORMS

JIAXIN HU AND TAKASHI KUMAGAI

ABSTRACT. We use an elementary method to obtain *Nash-type inequalities* for *non-local Dirichlet forms* on *d-sets*. We obtain two-side estimates for the corresponding *heat kernels* if the walk dimensions of heat kernels are less than 2; these estimates are obtained by combining probabilistic and analytic methods. Our arguments partly simplify those in Bass and Levin (2002) and Chen and Kumagai (2003).
2000 MSC: Primary 60J75; Secondary 60J35; 46E35
Key Words and Phrases: Dirichlet forms; fractals; heat kernels; jump processes; Lévy systems; Nash's inequality; parabolic Harnack inequality
Running Head: Non-local Dirichlet Forms

1. INTRODUCTION

In recent years, there has been growing interest in Brownian motions on fractals, see [2, 3, 14] etc. The Brownian motion on a fractal is a *diffusion*, and gives rise to a *local regular Dirichlet form* on the fractal. The corresponding heat kernel or *transition density* generally satisfies *sub-Gaussian* upper and lower estimates. In the estimates, there appears a new parameter, termed the *walk dimension* α of the diffusion. For a classical Euclidean domain, we always have that $\alpha = 2$ if the domain is equipped with the Lebesgue measure. However, for a fractal or other porous media, we usually have $\alpha > 2$. Roughly speaking, this means that a random walker needs more time to go across a fractal domain than a regular one.

Given a Brownian motion whose heat kernel satisfies sub-Gaussian estimates, one may use the *subordination* technique to obtain another type of processes−jump processes, see for example [6, 16, 17, 21]. A jump process corresponds to a *non-local* Dirichlet form whose *domain* is a Besov space.

There is an alternative direction. That is, one may obtain heat kernel bounds for jump processes without assuming a priori the existence of a Brownian motion on a d-set. (Many basic fractals like Sierpinski gaskets or carpets are d-sets). In [8], Chen and Kumagai took this direction and obtained detailed heat kernel estimates. In that paper, embedding theorems of Besov spaces in Euclidean spaces are used in several places. Besides, the proofs of the heat kernel estimates are quite complicated. In this paper, we consider d-sets which are not necessarily embedded into Euclidean spaces (but they are required to satisfy (4.1) instead) and obtain detailed heat kernel bounds similar to those obtained in [8]. Our aim is twofold;
1) To give a proof of the Nash-type inequality which does not rely on embedding theorems.
2) To simplify several parts of the proof in [8].

Date: August 27, 2005. JH was partially supported by NSFC, Grant No. 10371062 and the Alexander von Humboldt Foundation. TK was partially supported by Ministry of Education, Japan, Grant-in-Aid for Scientific Research for Young Scientists (B) 16740052.

The outline of the paper is as follows. Given a d-set F, we first introduce the functional dimension d_w of (F,μ) by means of function spaces, which coincides with the walk dimension. We then define Besov spaces with parameter $\alpha < d_w$, which give rise to non-local Dirichlet forms. In Section 3, we use an elementary approach initiated by [20, 1] to derive Nash-type inequalities for these Besov spaces. Note that we do not need embedding theorems. This gives the existence and the on-diagonal upper bound of a heat kernel $p(t,x,y)$ with any walk dimension $\alpha \in (0, d_w)$, by the well-known result in [7]. In Section 4, we give a proof of two-sided heat kernel estimates when $\alpha < 2$. Several estimates such as a hitting probability estimate (Theorem 4.6) and tightness (Proposition 4.7) can be deduced in a simpler way than those in [8].

We will use c, with or without subscripts, to denote strictly positive finite constants whose values are insignificant and may change from line to line.

2. Non-local Dirichlet forms

Let (F,ρ) be a connected locally compact separable metric space. Let μ be a *d-measure* for some $d>0$, that is, there is a positive constant $c_0 \ge 1$ such that

$$c_0^{-1}\, r^d \le \mu(B(x,r)) \le c_0\, r^d \tag{2.1}$$

for all $x \in F$ and all $0 < r \le r_0 := \operatorname{diam}(F) \le \infty$, where $B(x,r) = \{z \in F : \rho(z,x) < r\}$ is a ball in F with radius r and center x. It turns out that d is the *Hausdorff dimension* of F. Let $L^p(\mu) := L^p(F,\mu), 1 \le p \le \infty$ be the usual space of real-valued p-integrable functions on F with respect to μ, with the norm

$$\|f\|_p := \left(\int_F |f(x)|^p d\mu(x)\right)^{1/p} \quad (1 \le p \le \infty)$$

(with the obvious modification when $p=\infty$). For $\alpha > 0$, let

$$\mathcal{E}_p[u] := \mathcal{E}_p^{(\alpha)}[u] = \int_F\int_F \frac{|u(x)-u(y)|^p}{\rho(x,y)^{d+p\alpha/2}} d\mu(y)d\mu(x),\ 1<p<\infty, \tag{2.2}$$

for $u \in \mathcal{D}(\mathcal{E}_p)$, where $\mathcal{D}(\mathcal{E}_p)$ is the *domain* of $\mathcal{E}_p$

$$\mathcal{D}(\mathcal{E}_p) := \{u \in L^p(\mu) : \mathcal{E}_p[u] < \infty\}. \tag{2.3}$$

By Fatou's lemma, it is easy to see that $\mathcal{D}(\mathcal{E}_p)$ is a Banach space under the norm $(\|u\|_p^p + \mathcal{E}_p[u])^{\frac{1}{p}}$. Define a *functional dimension* of (F,μ) as

$$d_w := 2 \sup\left\{\alpha : \mathcal{D}(\mathcal{E}_2^{(\alpha)}) \text{ is dense in } L^2(\mu)\right\}. \tag{2.4}$$

That is, the $d_w/2$ is the supreme of all numbers α such that the space $\mathcal{D}(\mathcal{E}_2^{(\alpha)})$ is *dense* in $L^2(\mu)$ (Note that $\mathcal{D}(\mathcal{E}_2^{(\alpha)})$ may only consist of constant functions, if α is large enough). For a regular Euclidean domain F equipped with the Lebesgue measure, it is well-known that $d_w = 2$. However, for a non-regular domain F, we generally have that $d_w \ge 2$; for instance, if F is the Sierpinski gasket in $\mathbb{R}^n\,(n\ge 2)$ equipped with the $\log(n+1)/\log 2-$dimensional Hausdorff measure, then $d_w = \log(n+3)/\log 2$. The functional dimension coincides with the walk dimension of (F,μ), see [15, 18, 21].

3. Nash-type inequalities

In this section, we will show that the space $\mathcal{D}(\mathcal{E}_p)$ defined as above satisfies the Nash-type inequality for any $1 < p < \infty$.

Theorem 3.1. *Let μ be a d-measure satisfying (2.1), and let $\mathcal{E}_p$ be defined as above for $1 < p < \infty$. Then the Nash-type inequality holds:*

$$\|u\|_p^{p\left(1+\frac{p}{p-1}\cdot\frac{\alpha}{2d}\right)} \leq C\,(r_0^{-p\alpha/2}\|u\|_p^p + \mathcal{E}_p[u])\,\|u\|_1^{\frac{p}{p-1}\cdot\frac{p\alpha}{2d}}, \tag{3.1}$$

for $u \in L^1(\mu)\cap\mathcal{D}(\mathcal{E}_p)$, where C is independent of r_0 and u.

Proof. The proof given here is motivated by [1, 20]. Let $u \in L^1(\mu)\cap\mathcal{D}(\mathcal{E}_p)$. Define the average function u_r of u by

$$u_r(x) := \frac{1}{\mu(B(x,r))}\int_{B(x,r)} u(z)\,d\mu(z),\ x \in F.$$

By (2.1), we have that

$$\|u_r\|_\infty \leq c_0\, r^{-d}\|u\|_1,\ 0 < r < r_0.$$

Using (2.1) again, we see that

$$\|u_r\|_1 \leq C\,\|u\|_1.$$

Thus

$$\|u_r\|_p^p = \int_F |u_r(x)|^p\,d\mu(x) \leq \|u_r\|_\infty^{p-1}\|u_r\|_1 \leq C\,r^{-d(p-1)}\|u\|_1^p,\ 0 < r < r_0, \tag{3.2}$$

for some C independent of r and u. On the other hand, we use Hölder's inequality, and (2.1), (2.2) to have that

$$\begin{aligned}
\|u-u_r\|_p^p &= \int_F \left|\frac{1}{\mu(B(x,r))}\int_{B(x,r)}(u(x)-u(y))\,d\mu(y)\right|^p d\mu(x)\\
&\leq c_0\, r^{-d}\int_F\int_{B(x,r)} |u(x)-u(y)|^p\,d\mu(y)d\mu(x)\\
&= c_0\, r^{-d}\int_F\int_{B(x,r)} \frac{|u(x)-u(y)|^p}{\rho(x,y)^{d+p\alpha/2}}\cdot\rho(x,y)^{d+p\alpha/2}\,d\mu(y)d\mu(x)\\
&\leq c_0\, r^{p\alpha/2}\mathcal{E}_p[u],\ 0 < r < r_0. \qquad (3.3)
\end{aligned}$$

Therefore, using the elementary inequality $(a+b)^p \leq 2^{p-1}(a^p+b^p)$ for any $a, b \geq 0$ and $1 \leq p < \infty$, it follows from (3.2) and (3.3) that

$$\begin{aligned}
\|u\|_p^p &\leq 2^{p-1}(\|u_r\|_p^p + \|u-u_r\|_p^p)\\
&\leq C\left(r^{-d(p-1)}\|u\|_1^p + r^{p\alpha/2}\mathcal{E}_p[u]\right),\ \ 0 < r < r_0. \qquad (3.4)
\end{aligned}$$

Noting that

$$\|u\|_p^p \leq \left(\frac{r}{r_0}\right)^{p\alpha/2}\|u\|_p^p$$

for $r \geq r_0$, we see from (3.4) that

$$\|u\|_p^p \leq C\left(r^{-d(p-1)}\|u\|_1^p + r^{p\alpha/2}\left(r_0^{-p\alpha/2}\|u\|_p^p + \mathcal{E}_p[u]\right)\right) \tag{3.5}$$

for all $r > 0$, where C is independent of u, r and r_0. Minimizing the right-hand side of (3.5), we arrive at (3.1). ∎

Remark 3.1.

a) *If F is unbounded and (2.1) holds for all $r > 0$ (that is $r_0 = \infty$), then by Theorem 3.1,*

$$\|u\|_p^{p\left(1+\frac{p}{p-1}\cdot\frac{\alpha}{2d}\right)} \le C\,\mathcal{E}_p[u]\cdot\|u\|_1^{\frac{p}{p-1}\cdot\frac{p\alpha}{2d}},\ 1<p<\infty. \tag{3.6}$$

b) *Fix $\lambda > 0$, and let*

$$E_{p,\lambda}[u] := \int_F\int_{\rho(y,x)\le\lambda}\frac{|u(x)-u(y)|^p}{\rho(x,y)^{d+p\alpha/2}}d\mu(y)d\mu(x). \tag{3.7}$$

A similar argument leads to (3.1) with r_0 replaced by $\min\{\lambda, \mathrm{diam}(F)\}$, and $\mathcal{E}_p$ by $E_{p,\lambda}$.

c) *The case $p = 2$ and $\alpha < 2$ in the Euclidean setting was addressed in [8] by using Sobolev-type embedding theorems and the interpolation theorem. (Under additional assumption $\alpha < d \wedge 2$, it was also addressed in [11].)*

Next we shall restrict ourselves to the interesting case $p = 2$. We fix $0 < \alpha < d_w$ and write $\mathcal{E}[u] := \mathcal{E}_2^{(\alpha)}[u]$, $\mathcal{F} := \mathcal{D}(\mathcal{E}_2^{(\alpha)})$. It is easy to see that $\mathcal{E}$ is a symmetric bilinear form on $\mathcal{F}$, and satisfies the *Markov property*

$$\mathcal{E}[(0\vee u)\wedge 1] \le \mathcal{E}[u]$$

for $u \in \mathcal{F}$. Furthermore, since $0 < \alpha < d_w$, $\mathcal{F}$ is dense in $L^2(\mu)$. Thus $(\mathcal{E},\mathcal{F})$ is a *Dirichlet form* (cf [10]). However $(\mathcal{E},\mathcal{F})$ is *non-local*, that is $\mathcal{E}(u,v) \ne 0$ for some $u, v \in \mathcal{F}$ with disjoint support. (Here $\mathcal{E}(u,v)$ is defined in an obvious way.) Observe that if $d < \alpha$, then the space $\mathcal{F}$ is embedded into the Hölder space with exponent $(\alpha - d)/2$, see for example [12, 13, 16]. Thus the Dirichlet form $(\mathcal{E},\mathcal{F})$ is *regular* if $d < \alpha$. Hence, we have

Proposition 3.2. *For any $0 < \alpha < d_w$, the quadratic form $(\mathcal{E},\mathcal{F})$ defined as above is a non-local Dirichlet form. Moreover $(\mathcal{E},\mathcal{F})$ is regular if $d < \alpha$.*

Since $(\mathcal{E},\mathcal{F})$ is a Dirichlet form, there is a strongly continuous, μ-symmetric semigroup $\{T_t\}_{t\ge0} := \{T_t^{(\alpha)}\}_{t\ge0}$ on $L^2(\mu)$, see for example [10, p. 18]. By Theorem 3.1 ($p = 2$), we have that

$$\|u\|_2^{2\left(1+\frac{\alpha}{d}\right)} \le C\,(r_0^{-\alpha}\|u\|_2^2 + \mathcal{E}[u])\ \|u\|_1^{\frac{2\alpha}{d}} \tag{3.8}$$

for all $u \in L^1(\mu)\cap\mathcal{F}$, where C is independent of r_0. Hence, by Theorem (2.1) in [7, Sect. 2], the semigroup $\{T_t\}_{t\ge0}$ has a kernel $p(t,x,y) := p^{(\alpha)}(t,x,y)$:

$$T_tu(x) = \int_F p(t,x,y)u(y)d\mu(y) \quad (x\in F,\ t>0), \tag{3.9}$$

which is quasi continuous and satisfies the upper-bound estimate

$$p(t,x,y) \le C\ t^{-d/\alpha}\exp\left(r_0^{-\alpha}t\right) \tag{3.10}$$

for all $t > 0$ and $x, y \in F\setminus N$, where N is a properly exceptional set with respect to $(\mathcal{E},\mathcal{F})$ and C is independent of r_0.

4. Heat kernels estimates for $0 < \alpha < 2$

In this section we shall obtain two-sided heat kernel estimates for $0 < \alpha < 2$. The result was already obtained by Chen and Kumagai [8] when the d-set is embedded in a Euclidean space. We will follow the patterns of [4, 8], and partly simplify the proof.

Let (F,ρ,μ) be a locally compact separable metric measure space where the Borel measure μ satisfies (2.1). We assume that $B(x,r)$ is relatively compact for all $x \in F$ and

$r > 0$. Assume further that there is a separable metric space $X \supset F$, and $\rho(\cdot,\cdot)$ can be extended to be a metric on X with dilation for F, i.e. for each $x, y \in F$ and each $\delta < 1$, there exist $\delta^{-1}x, \delta^{-1}y \in X$ such that

$$\rho(\delta^{-1}x, \delta^{-1}y) = \delta^{-1}\rho(x,y). \tag{4.1}$$

If $F \subset \mathbb{R}^n$, these conditions always hold by taking $X = \mathbb{R}^n$. On the other hand, there are examples where F is not embedded in $\mathbb{R}^n$, but the conditions hold. One of such examples is the following. Consider $\mathbb{R}^{\mathbb{N}} := \{(x_1, x_2, \cdots) : x_i \in \mathbb{R}, i \geq 1\}$. Define $I_n \subset \mathbb{R}^{\mathbb{N}}$ by

$$I_n := \left\{(0, \cdots, x_n, 0, \cdots) : x_n \in [0, 2^{-n}]\right\}, \quad n \geq 1.$$

Let $F = \cup_{n \geq 1} I_n$, and $\rho(x,y) = \sqrt{\sum_{n=1}^{\infty}(x_n - y_n)^2}$ for $x = (x_1, x_2, \cdots), y = (y_1, y_2, \cdots)$ and let μ be the usual 1-dimensional Lebesgue measure. Then (F, ρ, μ) satisfies all the assumptions above with $d = 1$, but F is not embedded in any Euclidean space since F is infinite dimensional. One may also construct similar examples using fractal sets.

Let $0 < \alpha < 2$. Define $(\mathcal{E}, \mathcal{F})$ as before in (2.2), (2.3) with $\mathcal{F} = \mathcal{D}(\mathcal{E}_2)$. For each fixed $x_0 \in F$ and for each $r > 0$, denote $B_r := B(x_0, r)$. Note that for $0 < r < r'$,

$$\int_{\overline{B_r}} \int_{\overline{B_r}} \frac{d\mu(x)d\mu(y)}{\rho(x,y)^{d+\alpha-2}} < \infty \text{ and } \int_{\overline{B_r}} \int_{F \setminus B_{r'}} \frac{d\mu(x)d\mu(y)}{\rho(x,y)^{d+\alpha}} < \infty,$$

which can be proved by simple calculations using (2.1) and $(0, 2r] = \cup_{k=0}^{\infty}(r/2^k, r/2^{k-1}]$. Thus,

$$\varphi(x) := 1 - \frac{\rho(x, \overline{B_r}) \wedge \rho(\overline{B_r}, B_{r'}^c)}{\rho(\overline{B_r}, B_{r'}^c)}, \ x \in F$$

is a Lipschitz continuous function and $\varphi \in \mathcal{F}$. Using this fact and Stone-Weierstrass' theorem, it is easy to see that $d_w \geq 2$ and $\mathcal{F} \cap C_0(F)$ is dense in $C_0(F)$, where $C_0(F)$ is a space of continuous compact supported functions on F (see detailed proof in [21]; the proof of this part works even if F is not in a Euclidean space). In general, we do not know whether or not $\mathcal{F} \cap C_0(F)$ is dense in $\mathcal{F}$. (If F is in a Euclidean space and $0 < \alpha < 2$, then this can be proved using the embedding theorem and the properties of Besov spaces in the Euclidean space.) Alternatively, define

$$\mathcal{F}_* := \overline{\mathcal{F} \cap C_0(F)},$$

i.e. $(\mathcal{E} + \|\cdot\|_2^2)^{1/2}$-closure of $\mathcal{F} \cap C_0(F)$. Then $(\mathcal{E}, \mathcal{F}_*)$ is a regular Dirichlet form on $L^2(F, \mu)$. We denote by Y the corresponding Hunt process.

The main result of this section reads as follows.

Theorem 4.1. *For $0 < \alpha < 2$, let $p(t,x,y)$ be the heat kernel corresponding to $(\mathcal{E}, \mathcal{F}_*)$. Then there exist $C > 0$ and $t_0 > 0$ such that*

$$\begin{aligned} C^{-1}\, t^{-d/\alpha} \left(1 + t^{-1/\alpha}\rho(y,x)\right)^{-(d+\alpha)} &\leq p(t,x,y) \\ &\leq C\, t^{-d/\alpha} \left(1 + t^{-1/\alpha}\rho(y,x)\right)^{-(d+\alpha)} \end{aligned} \tag{4.2}$$

for all $0 < t < t_0$ and all $x, y \in F$.

Remark that when $0 < d < \alpha < 2$ or when F is in a Euclidean space and $0 < \alpha < 2$, then this result holds for $(\mathcal{E}, \mathcal{F})$ without any change of the proof below.

In the first two subsections, we will prepare several estimates needed. We then prove the upper and lower bounds in subsection 4.3 for $x, y \in F \setminus N$ for some properly exceptional set. In subsection 4.4 we use the Hölder continuity and obtain the heat kernel estimates for all $x, y \in F$.

4.1. **Tightness and hitting probability estimates.** For any $\delta \in (0,1]$, let $G^{(\delta)} := \delta^{-1}F = \{\delta^{-1}x : x \in F\}$, and $\nu^{(\delta)}$ be the measure determined by $\nu^{(\delta)}(J) := \delta^{-d}\mu(\delta J)$ for any measurable set $J \subset G^{(\delta)}$. It is trivial to see that $\nu^{(\delta)}$ is a d-measure for any $\delta \in (0,1]$

$$c_0^{-1}\, r^d \le \nu^{(\delta)}(B(\xi, r)) \le c_0\, r^d \tag{4.3}$$

for all $\xi \in G^{(\delta)}$ and all $0 < r \le \delta^{-1}r_0$, where $B(\xi, r)$ is a ball in $G^{(\delta)}$ with radius r and center ξ (keeping in mind that $r_0 = \mathrm{diam}(F)$). Since our arguments below will not involve δ so much, we will denote by $G := G^{(\delta)}$ and $\nu := \nu^{(\delta)}$, etc.

Instead of estimating $p(t, x, y)$ directly on $(F, \rho(\cdot,\cdot), \mu)$, we shall give an upper-bound of $q(t, \xi, \eta) := q^{(\delta)}(t, \xi, \eta)$ at time $t = 1/2$ on $(G, \rho(\cdot,\cdot), \nu)$, where

$$q(t, \xi, \eta) := \delta^d\, p(\delta^\alpha t, \delta\,\xi, \delta\,\eta), \quad t > 0 \text{ and } \xi, \eta \in G. \tag{4.4}$$

Note that $q(t, \xi, \eta)$ is a *heat kernel* (or *transition density*) on (G, ν); this can be easily verified by using the fact that

$$\int_J h(\xi) d\nu(\xi) = \delta^{-d} \int_{\delta\, J} h(\delta^{-1}x) d\mu(x), \quad J \subset G \tag{4.5}$$

for any integrable function h on J.

The following is easy by scaling argument (cf. [8]).

Lemma 4.1. *Let $(W, \mathcal{D}_*(W))$ be the Dirichlet form corresponding to the heat kernel $q(t, \xi, \eta)$. Then, it is regular and*

$$W[v] = \int_G \int_G \frac{|v(\xi) - v(\eta)|^2}{\rho(\xi, \eta)^{d+\alpha}} d\nu(\eta) d\nu(\xi), \tag{4.6}$$

that is W has the same pattern as $\mathcal{E}$, but on (G, ν).

By this lemma, we see that the Nash-type inequality (3.8) holds for W, with r_0 replaced by $\delta^{-1}r_0$ and $\mathcal{E}$ by W. Thus

$$q(t, \xi, \eta) \le C\, t^{-\frac{d}{\alpha}} \exp\left(r_0^{-\alpha}\delta^\alpha t\right) \tag{4.7}$$

for all $t > 0$ and ν-almost all $\xi, \eta \in G$.

In the following, we denote by $Y^{(\delta)}$ the Hunt process which corresponds to $(W, \mathcal{D}_*(W))$. As in Theorem 3.2 in [8], there is a properly exceptional set N with respect to $(\mathcal{E}, \mathcal{F})$ such that $q(t, \xi, \eta)$ is a kernel that satisfies Chapman-Kolmogorov equation and (4.7) for every $t > 0$, every $\xi, \eta \in G \setminus \delta^{-1}N$ and $\delta \in (0,1] \cap \mathbb{Q}$. In the following of this paper, we will slightly abuse notation and write F, G for $F \setminus N, G \setminus \delta^{-1}N$ respectively.

The next formula of Lévy system for Y is given in Lemma 4.7 in [8]. We will use it implicitly and explicitly throughout the paper.

Lemma 4.2. *Let f be a non-negative measurable function on $\mathbb{R}_+ \times F \times F$, vanishing on the diagonal. Then for every $t \ge 0$, $x \in F$ and every stopping time T (with respect to the filtration of Y),*

$$\mathbb{E}^x\left[\sum_{s \le T} f(s, Y_{s-}, Y_s)\right] = \mathbb{E}^x\left[\int_0^T \int_F \frac{f(s, Y_s, y)}{\rho(Y_s, y)^{d+\alpha}} d\mu(y) ds\right].$$

Remark 4.3. *In* [8] *Lemma 4.7, this lemma is stated for predictable stopping time T, but in fact it holds for any stopping time T. This fact is known for experts; see for example* [9] *for the proof.*

We now decompose the Dirichlet form into two forms, a form with small jumps and with large jumps, and use the perturbation theory of a semigroup by a bounded linear operator on $L^2(\nu)$ as in [4, 8]. Let $\lambda \geq 1$ and consider a quadratic form W_λ on (G, ν)

$$W_\lambda[v] := \int_G \int_{\rho(\eta,\xi) \leq \lambda^{\frac{1}{2}}} \frac{(v(\xi) - v(\eta))^2}{\rho(\xi,\eta)^{d+\alpha}} d\nu(\eta) d\nu(\xi). \tag{4.8}$$

For $v \in L^2(G, \nu)$,

$$\begin{aligned} 0 \ &\leq \ W[v] - W_\lambda[v] = \int_G \int_{\rho(\eta,\xi) > \lambda^{\frac{1}{2}}} \frac{(v(\eta) - v(\xi))^2}{\rho(\eta,\xi)^{d+\alpha}} d\nu(\eta) d\nu(\xi) \\ &\leq \ c_1 \int_G v(\xi)^2 \Big(\int_{\rho(\eta,\xi) > \lambda^{\frac{1}{2}}} \frac{d\nu(\eta)}{\rho(\eta,\xi)^{d+\alpha}} \Big) d\nu(\xi) \leq c_2 \|v\|_{L^2}^2. \end{aligned}$$

So, the $\big(W + \|\cdot\|_2^2\big)^{1/2}$-norm is comparable to the $\big(W_\lambda + \|\cdot\|_2^2\big)^{1/2}$-norm. In particular, $(W_\lambda, \mathcal{D}_*(W))$ is the regular Dirichlet form. We denote by $X = X^{(\lambda)}$ the corresponding process. Note that $W_\lambda[v] = \delta^{\alpha-d} E_{2,\delta^2\lambda}[v_\delta]$ where $E_{p,\lambda}$ is defined in Remark 3.1 b) and $v_\delta(\cdot) = v(\delta^{-1}\cdot)$. Thus, the heat kernel of W_λ, which we will denote by $q^{(\lambda)}$, can be written using that of $E_{2,\delta^2\lambda}$ in the same way as in (4.4).

We summarize the notation here.

p	heat kernel for Y,	corresponding Dirichlet form $(\mathcal{E}, \mathcal{F}_*)$
q	heat kernel for $Y^{(\delta)}$,	corresponding Dirichlet form $(W, \mathcal{D}_*(W))$
$q^{(\lambda)}$	heat kernel for $X = X^{(\lambda)}$,	corresponding Dirichlet form $(W_\lambda, \mathcal{D}_*(W))$

Let B_λ be a bounded linear operator on $L^2(\nu)$

$$B_\lambda v(\xi) := 2 \int_{\rho(\eta,\xi) > \lambda^{\frac{1}{2}}} \frac{(v(\eta) - v(\xi))}{\rho(\eta,\xi)^{d+\alpha}} d\nu(\eta). \tag{4.9}$$

Let $\mathcal{L}$ and $\mathcal{L}^{(\lambda)}$ be the *generators* corresponding to W and W_λ, respectively. Since

$$W(u, v) = W_\lambda(u, v) - \int_F B_\lambda u(\xi) v(\xi) \, d\nu(\xi)$$

for $u, v \in \mathcal{D}_*(W)$, we see that

$$\mathcal{L} = \mathcal{L}^{(\lambda)} + B_\lambda.$$

Hence, if $\{T(t)\}_{t \geq 0}$ and $\{T_\lambda(t)\}_{t \geq 0}$ are the semigroups associated with W and W_λ respectively, we have that

$$T(t)v = T_\lambda(t)v + \sum_{k=1}^{\infty} S_k^{(\lambda)}(t)v, \qquad v \in L^\infty(\nu), \tag{4.10}$$

where

$$S_k^{(\lambda)}(t)v = \int_0^t T_\lambda(t-s) B_\lambda S_{k-1}^{(\lambda)}(s) v ds, \qquad k \geq 1 \tag{4.11}$$

with $S_0^{(\lambda)}(t) := T_\lambda(t)$, see [19, Theorem 3.5, p. 205]. We note that the series in (4.10) defines a bounded linear operator on $L^\infty(\nu)$ for each $t > 0$; this can be seen as follows.

First, from (4.3) and $(\lambda^{1/2},\infty)=\cup_{k=0}^{\infty}(2^k\lambda^{1/2},2^{k+1}\lambda^{1/2}]$, we have

$$\int_{\rho(\eta,\xi)>\lambda^{\frac{1}{2}}}\frac{d\nu(\eta)}{\rho(\eta,\xi)^{d+\alpha}}\le c\lambda^{-\frac{\alpha}{2}}. \tag{4.12}$$

Using this, we see that there exists $c_1>0$ independent of λ such that

$$\|B_\lambda v\|_\infty\le c_1\,\lambda^{-\frac{\alpha}{2}}\|v\|_\infty.$$

Noting that $\|T_\lambda(t)v\|_\infty\le\|v\|_\infty$, by induction we have from (4.11) that

$$S_k^{(\lambda)}(t)v(\xi)\le\frac{(c_1\lambda^{-\frac{\alpha}{2}}\,t)^k}{k!}\|v\|_\infty,\quad t>0,\ k\ge1,\ \nu-\text{a.e. }\xi \tag{4.13}$$

and so the series above is bounded from $L^\infty(\nu)$ to $L^\infty(\nu)$ for each $t>0$. Further, we can take a quasi continuous modification so that (4.11), (4.13) hold for $G\setminus N^{(\delta,\lambda)}$ where $N^{(\delta,\lambda)}$ is a properly exceptional set with respect to $(W_\lambda,\mathcal{D}_*(W))$. Define $N^{(\delta)}:=\cup_{\lambda\in\mathbb{Q},\lambda\ge1}N^{(\delta,\lambda)}$.

We now estimate the heat kernel $q^{(\lambda)}(t,\xi,\eta)$ that corresponds to the Dirichlet form W_λ, by using the result in [7]. Recall that we write G for $G\setminus\delta^{-1}N$.

Lemma 4.4. *There exist $c_1,c_2>0$ such that*

$$q^{(\lambda)}(t,x,y)\le c_1\,t^{-\frac{d}{\alpha}}\,\exp\left(c_2\lambda^{-\frac{\alpha}{2}}\,t\right)\,\exp\left(-\lambda^{-\frac{1}{2}}\rho(x,y)\right) \tag{4.14}$$

for all $t\in(0,1]$, $x,y\in G$ and $\lambda\ge1$.

Proof. By Remark 3.1 b), we have that the Nash-type inequality (3.8) holds, with r_0 replaced by $\lambda^{\frac{1}{2}}\wedge(\delta^{-1}r_0)$ and $\mathcal{E}$ by W_λ. Thus, by [7, Theorem (3.25), p. 272], we obtain that there exists $C>0$ independent of λ (and of $\delta\in(0,1]$) such that

$$q^{(\lambda)}(t,x,y)\le C\,t^{-\frac{d}{\alpha}}\,\exp\left(\lambda^{-\frac{\alpha}{2}}\,t\right)\,\exp\left(-|\psi(y)-\psi(x)|+C\,\Gamma(\psi)^2\,t\right) \tag{4.15}$$

for all $t\in(0,1]$ and all $x,y\in G$, and for any ψ satisfying $\Gamma(\psi)<\infty$, where

$$\Gamma(\psi)^2:=\|e^{-2\psi}\Gamma_\lambda[e^\psi]\|_\infty\vee\|e^{2\psi}\Gamma_\lambda[e^{-\psi}]\|_\infty$$

and Γ_λ is the *density* of W_λ determined by

$$\Gamma_\lambda[v](\xi):=\int_{\rho(\eta,\xi)\le\lambda^{\frac{1}{2}}}\frac{(v(\eta)-v(\xi))^2}{\rho(\eta,\xi)^{d+\alpha}}d\nu(\eta),\ \xi\in G. \tag{4.16}$$

Define

$$\mathcal{H}(\Gamma_\lambda):=\{v:G\to\mathbb{R}\mid\sup_{\xi\in G}\Gamma_\lambda[v](\xi)<\infty\Big\}.$$

Fix $\xi_0\in G$. The key is that $\mathcal{H}(\Gamma_\lambda)$ contains the *cut-off distance function* ψ given by

$$\psi(\xi):=\begin{cases}\lambda^{-\frac{1}{2}}\,\rho(\xi,x), & \text{if }\rho(\xi,x)\le\rho(x,y),\\ \lambda^{-\frac{1}{2}}\,\rho(x,y), & \text{if }\rho(\xi,x)\ge\rho(x,y),\end{cases} \tag{4.17}$$

for $\xi\in G$. (This is the reason why we restrict ourselves to the case that the walk dimension $\alpha<2$; if $\alpha\ge2$, then ψ given as above is not in $\mathcal{H}(\Gamma_\lambda)$, and in fact, $\mathcal{H}(\Gamma_\lambda)$ may contain only constant functions. If so, the Davies approach can not apply to yield the upper off-diagonal estimates, as indicated in [8]).

Noting that $|\psi(\eta)-\psi(\xi)|\le\lambda^{-\frac12}\,\rho(\eta,\xi)$ for all $\xi,\eta\in G$, we see that

$$\begin{aligned}
e^{-2\psi(\xi)}\Gamma_\lambda[e^{\psi}](\xi) &= \int_{\rho(\eta,\xi)\le\lambda^{\frac12}}\frac{(1-e^{\psi(\eta)-\psi(\xi)})^2}{\rho(\eta,\xi)^{d+\alpha}}d\nu(\eta)\\
&\le \int_{\rho(\eta,\xi)\le\lambda^{\frac12}}\frac{(\psi(\eta)-\psi(\xi))^2\,e^{2|\psi(\eta)-\psi(\xi)|}}{\rho(\eta,\xi)^{d+\alpha}}d\nu(\eta)\\
&\le c\lambda^{-1}\int_{\rho(\eta,\xi)\le\lambda^{\frac12}}\rho(\eta,\xi)^{-d+(2-\alpha)}d\nu(\eta)\le C\,\lambda^{-\frac{\alpha}{2}}
\end{aligned}$$

for all $\xi\in G$. The same estimate holds for $e^{2\psi(\xi)}\Gamma_\lambda[e^{-\psi}](\xi)$. Thus $\Gamma(\psi)^2\le C\,\lambda^{-\alpha/2}$. Noting that $|\psi(y)-\psi(x)|=\lambda^{-\frac12}\,\rho(x,y)$, we see that (4.14) follows from (4.15). □

Note that a similar conclusion to Lemma 4.4 was obtained in [8, (4.13)] and in [4, Proposition 2.4] when $F\subset\mathbb{R}^n$.

It follows from Lemma 4.4 that for $t\in[1/4,\,1]$ and $r>0$,

$$\mathbb{P}^x\left(\rho(X_t^{(\lambda)},x)\ge r\right)=\int_{\{y\in G:\,\rho(y,x)\ge r\}}q^{(\lambda)}(t,x,y)\nu(dy)\le c\,I_{r,\lambda}, \tag{4.18}$$

where $I_{r,\lambda}:=e^{-\frac{r}{2}\lambda^{-\frac12}+c_2\lambda^{-\frac{\alpha}{2}}}$. Define $\sigma_r=\sigma_r(X^{(\lambda)}):=\inf\{t\ge0:\,\rho(X_t^{(\lambda)},X_0^{(\lambda)})\ge r\}$. Then by (4.18) and the strong Markov property of $X^{(\lambda)}$,

$$\begin{aligned}
\mathbb{P}^x(\sigma_r\le1/2) &\le \mathbb{P}^x\left(\sigma_r\le1/2\text{ and }\rho(X_1^{(\lambda)},x)\le r/2\right)+\mathbb{P}^x\left(\rho(X_1^{(\lambda)},x)>r/2\right)\\
&\le \mathbb{P}^x\left(\sigma_r\le1/2\text{ and }\rho(X_1^{(\lambda)},X_{\sigma_r}^{(\lambda)})\ge r/2\right)+c\,I_{r/2,\lambda}\\
&= \mathbb{P}^x\left(\mathbf{1}_{\{T:=\sigma_r(X^{(\lambda)})\le1/2\}}\mathbb{P}^{X_T^{(\lambda)}}\left(\rho(X_{1-T}^{(\lambda)},X_0^{(\lambda)})\ge r/2\right)\right)+c\,I_{r/2,\lambda}\\
&\le \sup_{y\in B(x,r)^c}\sup_{s\le1/2}\mathbb{P}^y\left(\rho(X_{1-s}^{(\lambda)},y)\ge r/2\right)+c\,I_{r/2,\lambda}\\
&\le c'\,I_{r/2,\lambda}.
\end{aligned}$$

Here in the second and the last lines, we used (4.18). In the third line, we set $T:=\sigma_r(X^{(\lambda)})$ and used the strong Markov property for the random variable T. By the strong Markov property of $X^{(\lambda)}$, for every $r>0$,

$$\begin{aligned}
\mathbb{P}^x\left(\sup_{s\le1}\rho(X_s^{(\lambda)},X_0^{(\lambda)})>r\right) &\le \mathbb{P}^x(\sigma_r\le1/2)+\mathbb{P}^x(1/2<\sigma_r\le1)\\
&\le c'\,I_{r/2,\lambda}+\mathbb{P}^x(\sigma_{r/2}\le1/2)+\mathbb{P}^x(\sigma_{r/2}>1/2,\sigma_r\le1)\\
&\le c'\,I_{r/2,\lambda}+\mathbb{P}^x(\sigma_{r/2}\le1/2)+\mathbb{E}^x\left[\mathbb{P}^{X_{1/2}^{(\lambda)}}(\sigma_{r/2}\le1/2)\right]\\
&\le c''\,I_{r/4,\lambda}.
\end{aligned}\tag{4.19}$$

The constants $c,c',c''>0$ above are independent of $\delta>0$, $x\in G$ and $\lambda\ge1$.

Fix $\xi_0,\eta_0\in G\setminus N^{(\delta)}$ such that $\rho(\xi_0,\eta_0)\ge6$ and let $\lambda\in[\rho(\xi_0,\eta_0),\frac43\rho(\xi_0,\eta_0)]\cap\mathbb{Q}$. Note that ξ_0,η_0 depend on $\delta>0$. Define

$$\varphi(\xi):=\mathbf{1}_{B(\eta_0,1)}(\xi),\quad \xi\in G. \tag{4.20}$$

We claim that

$$|T_\lambda(t)\varphi(\xi)|\le c\,\left(\lambda^{-(d+\alpha)}+H_0(\xi)\right) \tag{4.21}$$

for all $t \in (0,1]$ and all $\xi \in G \setminus N^{(\delta)}$, where $H_0(\xi) := H_0(\xi, t)$ is a non-negative function on G for each t with support in $B(\xi_0, 2^{-1}\lambda)^c$ and satisfying $\|H_0\|_1 \vee \|H_0\|_\infty \le 1$, and c is independent of λ and ξ_0. In fact,

$$\begin{aligned}
|T_\lambda(t)\varphi(\xi)| &= \int_{B(\eta_0,1)} q^{(\lambda)}(t,\xi,\eta)d\nu(\eta) \\
&= \left(\mathbf{1}_{B(\xi_0,2^{-1}\lambda)^c}(\xi) + \mathbf{1}_{B(\xi_0,2^{-1}\lambda)}(\xi)\right) \int_{B(\eta_0,1)} q^{(\lambda)}(t,\xi,\eta)d\nu(\eta) \\
&=: H_0(\xi)(\nu(B(\eta_0,1)) \vee 1) + \mathbf{1}_{B(\xi_0,2^{-1}\lambda)}(\xi) \int_{B(\eta_0,1)} q^{(\lambda)}(t,\xi,\eta)d\nu(\eta).
\end{aligned}$$

It is easy to see from (4.14) that H_0 has the desired property. Note that, since $\lambda \ge 6$, the right hand side of (4.19) is smaller than $c\lambda^{-(d+\alpha)}$. Thus, it follows from (4.19) that

$$\mathbf{1}_{B(\xi_0,2^{-1}\lambda)}(\xi) \int_{B(\eta_0,1)} q^{(\lambda)}(t,\xi,\eta)d\nu(\eta) \le \mathbb{P}^\xi \left(\sup_{s\le 1} \rho(X_s^{(\lambda)}, X_0^{(\lambda)}) > \lambda/12\right) \le C\, \lambda^{-(d+\alpha)}$$

for all $t \in (0,1]$ and all $\xi \in G \setminus N^{(\delta,\lambda)}$, where C is independent of $\lambda \ge 6$ and ξ_0. Thus (4.21) follows.

Let $\lambda \ge 6$. We say that a function f belongs to $\Lambda(k,c)$ ($k \ge 0$) for some $c > 0$, if

$$|f(\xi)| \le c\left(\lambda^{-(d+\alpha)} + H_k(\xi)\right) \tag{4.22}$$

for all $\xi \in G \setminus N^{(\delta,\lambda)}$, where H_k is a function on G with support in $B(\xi_0, 2^{-(k+1)}\lambda)^c$ and satisfying $\|H_k\|_1 \vee \|H_k\|_\infty \le 1$.

Let M be the least integer greater than or equal to $2(d+\alpha)/\alpha$. The following result simplifies the corresponding results in [8, Lemma 4.8] and [4, Proposition 4.4, Lemma 4.5].

Proposition 4.5. *There exists $C_1 > 0$ such that for each $k \le M$ and $\lambda \ge 6$, if f belongs to $\Lambda(k,c)$, then the function $B_\lambda f$ belongs to $\Lambda(k+1, C_1 c)$. The same holds for the function $T_\lambda(t) f$ for each $t \in (0,1]$.*

Proof. Let $f \in \Lambda(k,c)$. By (4.9), (4.12) and (4.22), we have that

$$\begin{aligned}
|B_\lambda f(\xi)| &\le 2\int_{\rho(\eta,\xi)>\lambda^{\frac12}} \frac{|f(\eta)|+|f(\xi)|}{\rho(\eta,\xi)^{d+\alpha}} d\nu(\eta) \\
&\le C\left(\lambda^{-(d+\alpha)} + H_k(\xi) + \int_{\rho(\eta,\xi)>\lambda^{\frac12}} \frac{H_k(\eta)}{\rho(\eta,\xi)^{d+\alpha}} d\nu(\eta)\right) \\
&= C\left(\lambda^{-(d+\alpha)} + H_k(\xi) + \mathbf{1}_{B(\xi_0,2^{-(k+2)}\lambda)^c}(\xi) \int_{\rho(\eta,\xi)>\lambda^{\frac12}} \frac{H_k(\eta)}{\rho(\eta,\xi)^{d+\alpha}} d\nu(\eta)\right. \\
&\qquad \left. + \mathbf{1}_{B(\xi_0,2^{-(k+2)}\lambda)}(\xi) \int_{\rho(\eta,\xi)>\lambda^{\frac12}} \frac{H_k(\eta)}{\rho(\eta,\xi)^{d+\alpha}} d\nu(\eta)\right) \\
&=: C\left(\lambda^{-(d+\alpha)} + H^*_{k+1}(\xi)\right. \\
&\qquad \left. + \mathbf{1}_{B(\xi_0,2^{-(k+2)}\lambda)}(\xi) \int_{\substack{\rho(\eta,\xi) > \lambda^{\frac12} \\ \rho(\eta,\xi_0) \ge 2^{-(k+1)}\lambda}} \frac{H_k(\eta)}{\rho(\eta,\xi)^{d+\alpha}} d\nu(\eta)\right).
\end{aligned} \tag{4.23}$$

Clearly, the function H^*_{k+1} is supported in $B(\xi_0, 2^{-(k+2)}\lambda)^c$ and satisfies that $\|H^*_{k+1}\|_\infty \le C'$, $\|H^*_{k+1}\|_1 \le C'$ for some C'. So, letting $H_{k+1} := \frac{1}{C'} H^*_{k+1}$, we have $\|H_{k+1}\|_1 \vee \|H_{k+1}\|_\infty \le$

1. On the other hand, the last term in (4.23) is bounded by $C\lambda^{-(d+\alpha)}$, by noting that $\|H_k\|_1 \le 1$ and

$$\rho(\eta,\xi) \ge \rho(\eta,\xi_0) - \rho(\xi,\xi_0) \ge 2^{-(k+2)}\lambda \ge 2^{-(M+2)}\lambda$$

since $\rho(\eta,\xi_0) \ge 2^{-(k+1)}\lambda$ and $\rho(\xi,\xi_0) \le 2^{-(k+2)}\lambda$. Therefore (4.23) gives that $B_\lambda f \in \Lambda(k+1, C_1 c)$ for some C_1 independent of λ.

Next, by (4.19) and the fact $\|H_k\|_\infty \le C$,

$$\begin{aligned}
\int_{\rho(\eta,\xi_0)\ge 2^{-(k+1)}\lambda} q^{(\lambda)}(t,\xi,\eta)|H_k(\eta)|d\nu(\eta) &\le C\mathbb{P}^\xi\left(\rho(X_t^{(\lambda)},\xi_0) \ge 2^{-(k+1)}\lambda\right)\\
&\le C\mathbb{P}^\xi\left(\sup_{s\le 1}\rho(X_s^{(\lambda)},\xi) \ge 2^{-(k+2)}\lambda\right)\\
&\le CI_{\lambda/2^{k+4},\lambda} = Ce^{-c_1\lambda^{1/2}+c_2\lambda^{-\alpha/2}} \le c_3\lambda^{-(d+\alpha)},
\end{aligned}$$

for $\lambda \ge 6$. Thus, we similarly have that $T_\lambda(t)f \in \Lambda(k+1, C_1 c)$ for each $t \in (0,1]$ if $f \in \Lambda(k,c)$. $\square$

The following corresponds to (4.31) in [8], but we have a simpler proof.

Proposition 4.6. *There exist $c_1, c_2 > 0$ and $t_0 > 0$ such that the following holds.*

$$\mathbb{P}^\xi(Y_t \in B(\eta, c_1 t^{1/\alpha})) \le c_2\frac{t^{1+d/\alpha}}{\rho(\eta,\xi)^{d+\alpha}} \qquad \text{for all } \xi,\eta \in F,\ t \le t_0. \tag{4.24}$$

Proof. Let $Y^{(\delta)}$ be the process corresponding to $(W, \mathcal{D}_*(W))$ on (G,ν). Then, by scaling it is enough to prove

$$\mathbb{P}^{\xi_0}(Y_{c_3}^{(\delta)} \in B^{(\delta)}(\eta_0, 1)) \le c_4\, \rho(\eta_0,\xi_0)^{-(d+\alpha)} \tag{4.25}$$

for all $\xi_0, \eta_0 \in G \setminus N^{(\delta)}$, where c_4 is independent of ξ_0, η_0 and $\delta \in \mathbb{Q}\cap(0,1]$. Indeed, as $Y_t^{(\delta)} = \delta^{-1} Y_{\delta^\alpha t}$, $\delta N^{(\delta)}$ is an exceptional set of Y and so by enlarging the properly exceptional set N if necessary, we may assume $N \supset \cup_{\delta\in\mathbb{Q}\cap(0,1]}\delta N^{(\delta)}$. It is enough to prove this when $\rho(\eta_0,\xi_0) > c_4^{1/(d+\alpha)}$, since otherwise it is a trivial inequality.

Fix $\xi_0, \eta_0 \in G$, and let $\lambda \in \left[\rho(\eta_0,\xi_0), \frac{4}{3}\rho(\eta_0,\xi_0)\right] \cap \mathbb{Q}$. Let φ be defined as in (4.20). It follows from (4.21) that $T_\lambda(t)\varphi \in \Lambda(0,c)$ for each $t \in (0,1]$. By Proposition 4.5, we see that $T_\lambda(t-s)B_\lambda T_\lambda(s)\varphi \in \Lambda(2, C_1^2 c)$ for each $t \in (0,1]$ and $0 < s < t$. Thus, we have from (4.11) that $S_1^{(\lambda)}(t)\varphi \in \Lambda(2, C_1^2 c)$ for each $t \in (0,1]$. Inductively, we see that $S_k^{(\lambda)}(t)\varphi \in \Lambda(2k, C_1^{2k} c)$ for each $t \in (0,1]$ for $1 \le k \le M$ where M is (as in Proposition 4.5) the least integer such that $\frac{\alpha}{2}M \ge d+\alpha$. Noting that $H_k(\xi_0) = 0$ for each $0 \le k \le M$ since the support of H_k is contained in $B(\xi_0, 2^{-(k+1)}\lambda)^c$, it follows that

$$\begin{aligned}
\sum_{k=0}^{M-1} |S_k^{(\lambda)}(c_3)\varphi(\xi_0)| &\le \sum_{k=0}^{M-1} C_1^{2k} c\left(\lambda^{-(d+\alpha)} + H_k(\xi_0)\right)\\
&= \sum_{k=0}^{M-1} C_1^{2k} c\, \lambda^{-(d+\alpha)} \le c_6\, \lambda^{-(d+\alpha)}
\end{aligned}$$

for sufficiently small $c_3 > 0$ and for some $c_6 > 0$ independent of ξ_0, η_0. On the other hand, we see from (4.13) that

$$\sum_{k=M}^{\infty} |S_k^{(\lambda)}(c_3)\varphi(\xi_0)| \le \sum_{k=M}^{\infty} \frac{(c_7\lambda^{-\frac{\alpha}{2}})^k}{k!} \le c_8\ \lambda^{-\frac{\alpha}{2}M} \le c_8\ \lambda^{-(d+\alpha)}$$

for some $c_8 > 0$ independent of ξ_0, η_0. Therefore,

$$\begin{aligned} T(c_3)\varphi(\xi_0) &= \int_{B(\eta_0,1)} q\left(c_3, \xi_0, \eta\right)\ d\nu(\eta) \\ &= \sum_{k=0}^{\infty} S_k^{(\lambda)}(c_3)\varphi(\xi_0) \\ &\le \sum_{k=0}^{M-1} |S_k^{(\lambda)}(c_3)\varphi(\xi_0)| + \sum_{k=M}^{\infty} |S_k^{(\lambda)}(c_3)\varphi(\xi_0)| \\ &\le c_9\ \lambda^{-(d+\alpha)} \le c_{10}\ \rho(\eta_0, \xi_0)^{-(d+\alpha)} \end{aligned}$$

for some $c_{10} > 0$ independent of ξ_0 and η_0, proving (4.25) for $\xi_0, \eta_0 \in G\setminus N^{(\delta)}$ with $\rho(\xi_0, \eta_0)$ large. $\square$

Using the above proof, we can prove the following tightness of the process (Proposition 4.1 in [8]) rather easily. For $A \subset F$, let $\tau_A := \inf\{t \ge 0 : Y_t \notin A\}$.

Proposition 4.7. *For each $r_0 > 0$, $A > 0$ and $0 < B < 1$, there exists some $\gamma := \gamma(r_0, A, B) \in (0,1)$ such that for every $0 < r \le r_0$ and $x \in F$,*

$$\mathbb{P}^x\left(\tau_{B(x,\,Ar)} < \gamma\, r^{\alpha}\right) \le B.$$

Proof. We will apply the above argument with $\lambda = 1$. Denote by $\{Q_t\}_{t\ge 0}$ the transition semigroup of $X^{(1)}$. First, by (4.13), for any bounded function f on G, we have

$$\|P_t^{(\delta)} f - Q_t f\|_\infty \le \sum_{k=1}^{\infty} \frac{(c_1\, t)^k}{k!}\, \|f\|_\infty \le c_1\, t\, e^{c_1 t}\, \|f\|_\infty.$$

Combining this and (4.18), it follows that there is a constant $c_2 > 0$ independent of $\delta > 0$ such that for ν-a.e. $\xi \in G$, every $t \le 1$ and $\delta > 0$,

$$\mathbb{P}^\xi\left(\rho(Y_t^{(\delta)}, \xi) > l\right) \le \mathbb{P}^\xi\left(\rho(X_t^{(1)}, \xi) > l\right) + c_2\, t \le c\, I_{l,1} + c_2\, t. \tag{4.26}$$

Thus for each $\eta \in G$ with $\rho(\xi, \eta) < l$,

$$\mathbb{P}^\xi\left(\rho(Y_t^{(\delta)}, \eta) > 2l\right) \le \mathbb{P}^\xi\left(\rho(Y_t^{(\delta)}, \xi) > l\right) \le c\, I_{l,1} + c_2\, t, \tag{4.27}$$

for ν-a.e. $\xi \in G$, and every $t \le 1$ and $\delta > 0$. As $\xi \mapsto \mathbb{P}^\xi\left(\rho(Y_t^{(\delta)}, \eta) > l\right)$ for each fixed $\eta \in G$ is a quasi-continuous function with respect to the process $Y^{(\delta)}$, by looking at rational $l > 0$ and rational $t > 0$ first, we conclude that there is a properly exceptional set $N_\eta^{(\delta)}$ of $Y^{(\delta)}$ such that (4.27) holds for all $l > 0$ and for $\xi \in G \setminus N_\eta^{(\delta)}$. Since G is separable, it can be covered with countably many balls $B(\eta_k, l)$. The above exceptional set can be choosen uniform for all $k \ge 1$, which we denote by $N_*^{(\delta)}$. Now for $\xi \in G \setminus N_*^{(\delta)}$,

$$\mathbb{P}^\xi\left(\rho(Y_t^{(\delta)}, \xi) > 3l\right) \le \mathbb{P}^\xi\left(\rho(Y_t^{(\delta)}, \eta_k) > 2l\right) \le c\, I_{l,1} + c_2\, t$$

for all $l > 0$ where η_k is chosen to satisfy $\rho(\xi, \eta_k) < l$. Applying the same argument to derive (4.19) from (4.18), we conclude there are positive constants c_3, c_4 such that for $\xi \in G \setminus N_*^{(\delta)}$,

$$\mathbb{P}^{\xi}\left(\sup_{s\le t} \rho(Y_s^{(\delta)}, \xi) > l\right) \le c_3 e^{-c_4 l} + c_3\, t \qquad \text{for every } l \ge 1 \text{ and } t \le 1. \tag{4.28}$$

As before, we may assume $N \supset \cup_{\delta\in\mathbb{Q}\cap(0,1]} \delta N_*^{(\delta)}$. Recall that F here is taken to be $F \setminus N$, so (4.28) implies that for every $x \in F$, $\delta \in \mathbb{Q} \cap (0, 1]$ and $l \ge 1$,

$$\mathbb{P}^{x}\left(\sup_{s\le \delta^\alpha\, t} \rho(Y_s, x) > \delta\, l\right) \le c_3 e^{-c_4 l} + c_3\, t \qquad \text{for every } l \ge 1 \text{ and } t \le 1. \tag{4.29}$$

By the continuity in δ, the above holds for every $\delta \in (0, 1]$ as well.

For $r_0 > 0$, $A > 0$ and $B \in (0, 1)$, we choose $l > A\, r_0$ and $s_0 < (r_0^\alpha/2) \wedge r_0$ so that $c_3 e^{-c_4 l} + c_3\, s_0 < B$ and therefore by (4.29),

$$\mathbb{P}^{x}\left(\sup_{s\le \delta^\alpha\, s_0} \rho(Y_s, x) \ge \delta\, l\right) \le B \qquad \text{for every } 0 < \delta \le 1.$$

Now for $r \in (0, r_0]$, letting $\delta = A\, r/l$, we have

$$\mathbb{P}^{x}\left(\sup_{s\le \gamma\, r^\alpha} \rho(Y_s, x) > A\, r\right) \le B,$$

where $\gamma = s_0\, A^\alpha/l^\alpha$. Thus the result holds. □

In the following, we will denote $\gamma_{r_0} = \gamma(r_0, 1/2, 1/2)$.

When $\rho(\xi, \eta) \ge 2t^{1/\alpha}$, the estimate in Proposition 4.6 is optimal and the following holds. (The corresponding estimate is given in the proof of Theorem 1.1 in [8].)

Proposition 4.8. *There exist $c_1 \ge 2, c_2 > 0$ such that the following holds for all $\xi, \eta \in F$ with $\rho(\xi, \eta) \ge c_1 t^{1/\alpha}$ and $t \le t_0$,*

$$\mathbb{P}^{\xi}(Y_t \in B(\eta, c_1 t^{1/\alpha})) \ge c_2 \frac{t^{1+d/\alpha}}{\rho(\eta, \xi)^{d+\alpha}}. \tag{4.30}$$

Proof. By Proposition 4.7, starting at $z \in B(y, t^{1/\alpha})$, there is a positive probability (independent of z and t) such that the process Y does not move more than $c_1 t^{1/\alpha}$ by time t. Thus, it is sufficient to show

$$\mathbb{P}^{x}\left(Y_s \text{ hits ball } B(y, t^{1/\alpha}) \text{ before } t\right) \ge c_2 \left(\frac{t^{1/\alpha}}{\rho(x, y)}\right)^{d+\alpha}, \tag{4.31}$$

for all $t \in (0, t_0]$ and $\rho(x, y) \ge 3t^{1/\alpha}$. (This corresponds to Proposition 4.2 (ii) in [8].)

Now with $B_x := B(x,\, t^{1/\alpha})$, $B_y := B(y,\, t^{1/\alpha})$ and $\tau_x := \tau_{B_x}$, it follows from Proposition 4.7 (with $r = t^{1/\alpha}, r_0 = 2^{1/\alpha}$ and $A = B = 1/2$),

$$\mathbb{E}^{x}\left[t \wedge \tau_x\right] \ge \gamma_{2^{1/\alpha}}\, t\, \mathbb{P}^{x}\left(\tau_x \ge \gamma_{2^{1/\alpha}}\, t\right) \ge \gamma_{2^{1/\alpha}}\, t/2, \tag{4.32}$$

for each $t \le 2$. Thus, from Lemma 4.2 with $T = t \wedge \tau_x$ and $f(s,u,v) = 1_{B_x}(u)1_{B_y}(v)$,

$$
\begin{aligned}
& \mathbb{P}^x \Big(Y_s \text{ hits ball } B(y,\, t^{1/\alpha}) \text{ before } t \Big) \\
\ge\ & \mathbb{P}^x (Y_{t\wedge\tau_x} \in B(y, t^{1/\alpha})) = \mathbb{E}^x \left[\int_0^{t\wedge\tau_x} \int_{B_y} \frac{\mu(du)ds}{\rho(Y_s,u)^{d+\alpha}} \right] \\
\ge\ & c_3 \mathbb{E}^x \left[t \wedge \tau_x \right] \int_{B_y} \frac{\mu(du)}{\rho(x,y)^{d+\alpha}} \ge c_4\, t \frac{\mu(B(y,t^{1/\alpha}))}{\rho(x,y)^{d+\alpha}} \ge c_5 \frac{t^{1+d/\alpha}}{\rho(x,y)^{d+\alpha}}.
\end{aligned}
$$

Here in the second to the third inequality, (4.32) is used. This establishes (4.31). □

4.2. **Harnack inequality.** In [8], the parabolic Harnack inequality (Proposition 4.3 in [8]) is proved in order to obtain heat kernel estimates. Here we will follow the strategy.

We first introduce a space-time process $Z_s := (V_s, Y_s)$, where $V_s := V_0 + s$. The filtration generated by Z satisfying the usual condition will be denoted as $\{\widetilde{\mathcal{F}}_s;\ s \ge 0\}$. The law of the space-time process $s \mapsto Z_s$ starting from (t,x) will be denoted as $\mathbb{P}^{(t,x)}$.

We say that a non-negative Borel measurable function $h(t,x)$ on $[0,\infty) \times G$ is *parabolic* in a relatively open subset D of $[0,\infty)\times G$ if for every relatively compact open subset D_1 of D, $h(t,x) = \mathbb{E}^{(t,x)}[h(Z_{\tau_{D_1}})]$ for every $(t,x) \in D_1$, where $Z_s = (V_s, Y_s)$ is the space-time process with $V_s = V_0 + s$ and $\tau_{D_1} = \inf\{s > 0 : Z_s \notin D_1\}$.

For each $R_0 > 0$, $t \le 1$ and $r \le R_0$, we define

$$
Q_r := Q_{R_0}(t,x,r) = [t, t + \gamma_{R_0} r^\alpha] \times B(x,r).
$$

Theorem 4.2. *For every $R_0 > 0$, $0 < \delta \le \gamma_{R_0}$, there exists $c_1 > 0$ such that for every $z \in F$, $0 < R \le R_0$ and every non-negative function h on $[0,\infty)\times G$ that is parabolic and bounded on $[0, \gamma_{R_0}(2R)^\alpha] \times B(z, 2R)$,*

$$
\sup_{(t,y)\in Q_{R_0}(\delta R^\alpha, z, R)} h(t,y) \le c_1 \inf_{y \in B(z,R)} h(0,y).
$$

In particular, the following holds.

$$
\sup_{(s,y)\in Q_{t_0^{1/\alpha}}((1-\gamma_{t_0^{1/\alpha}})t, z, t^{1/\alpha})} p(s,x,y) \le c \inf_{y\in B(z,t^{1/\alpha})} p((1+\gamma_{t_0^{1/\alpha}})t, x, y). \tag{4.33}
$$

The proof is the same as that of Proposition 4.3 in [8]. Since the proof is relatively long, we omit it here.

4.3. **Heat kernel bounds.** In this subsection, we prove Theorem 4.1 for $x, y \in F \setminus N$. Though the proof is the same as the corresponding proof in [8], we will rewrite it since it is short. As before, we abuse notation and write F for $F \setminus N$.

Proof of Theorem 4.1 (Upper bound). Let $r = \rho(x,y)$. By (3.10), we only need to consider the case that $r > t^{1/\alpha}$. Let $t_* = (1 + \gamma_{t_0^{1/\alpha}})t$. By Proposition 4.6,

$$
\int_{B(y,ct^{1/\alpha})} p(t_*,\, x,\, z)\, \mu(dz) \le c_1 \frac{t^{1+d/\alpha}}{r^{d+\alpha}}. \tag{4.34}
$$

We can take $c \le 1$ in (4.34). It follows that, using the fact that $\mu(B(y, ct^{1/\alpha})) \ge c_2\, t^{d/\alpha}$,

$$
\inf_{z\in B(y,t^{1/\alpha})} p(t_*, x, z) \le \inf_{z \in B(y, ct^{1/\alpha})} p(t_*, x, z) \le \frac{c_3\, t}{r^{d+\alpha}}.
$$

Now by (4.33),

$$p(t,x,y) \leq \sup_{(s,z)\in Q_{t_0^{1/\alpha}}((1-\gamma_{t_0^{1/\alpha}})t,y,t^{1/\alpha})} p(s,x,z) \leq c \inf_{z\in B(y,t^{1/\alpha})} p(t_*,x,z) \leq \frac{Ct}{r^{d+\alpha}},$$

and the proof is completed. □

Concerning the lower bound, we first prove it for the "near diagonal" case.

Lemma 4.9. *There exist c_1, $c_2 > 0$ such that*

$$p(t,x,y) \geq c_1\, t^{-d/\alpha}$$

for all $0 < t \leq t_0$ and $x, y \in F$ with $\rho(x,y) \leq c_2\, t^{1/\alpha}$.

Proof. By Proposition 4.7, there exists $c_3 > 1$ such that

$$\mathbb{P}^x(\sup_{s\leq t} \rho(Y_s,x) > c_3 t^{1/\alpha}) \leq 1/2 \qquad \text{for all } 0 < t \leq t_0.$$

On the other hand, using the upper bound in Theorem 4.1, there exists $0 < c_4 < c_3/2$ such that

$$\mathbb{P}^x(Y_t \in B(x, c_4 t^{1/\alpha})) \leq 1/4 \qquad \text{for all } 0 < t \leq t_0.$$

Thus, if we set $E(t) = B(x, c_3 t^{1/\alpha}) \setminus B(x, c_4 t^{1/\alpha})$, then

$$\mathbb{P}^x(Y_t \in E(t)) \geq 1/4 \qquad \text{for all } 0 < t \leq t_0.$$

Now for an arbitrary but fixed $t \in (0, t_0]$, define $t_* := (1-\gamma_{t_0^{1/\alpha}})t$. Applying to above with t_* in place of t yields $\sup_{z\in E(t_*)} p(t_*,x,z) \geq c_5 {t_*}^{-d/\alpha}$, since $\mu(E(t_*)) \leq c_6 {t_*}^{d/\alpha}$. Now by Theorem 4.2, we have that for all y with $\rho(y,x) < c_2 t^{1/\alpha}$, where $c_2 := (1-\gamma_{t_0^{1/\alpha}})^{1/\alpha}/2$,

$$p(t,x,y) \geq \inf_{w\in B(x,\, c_2 t^{1/\alpha})} p(t,x,w) \geq c \sup_{(s,w)\in Q_{t_0^{1/\alpha}}(\gamma_{t_0^{1/\alpha}} t,\, x,\, c_2 t^{1/\alpha})} p(t-s,x,w) \geq c_1 t^{-d/\alpha}.$$

where $c_1 := c_5\, c\,(1-\gamma_{t_0^{1/\alpha}})^{-d/\alpha}$. □

Proof of Theorem 4.1 (Lower bound). Due to Lemma 4.9, it is enough to prove the theorem for $\rho(x,y) \geq c_2\, t^{1/\alpha}$. Applying Proposition 4.8 with $t_* := (1-\gamma_{t_0^{1/\alpha}})t$ in place of t, we have

$$\mathbb{P}^x(Y_{t_*} \in B(y,\, C_1 t_*^{1/\alpha})) \geq c_6 \frac{t^{1+d/\alpha}}{\rho(x,y)^{d+\alpha}}.$$

As $\mu(B(y,\, C_1 t_*^{1/\alpha})) \leq c_7 t^{d/\alpha}$, the above implies $\sup_{z\in B(y,\, C_1 t_*^{1/\alpha})} p(t_*,x,z) \geq c_8\, t/\rho(x,y)^{d+\alpha}$. By applying (4.33) as in the proof of the upper bound in this subsection, we obtain the desired result. □

4.4. **Hölder continuity.** So far, we have proved Theorem 4.1 for $x, y \in F \setminus N$. Now we can prove the following Hölder continuity of the heat kernel by applying the estimates obtained above. As a result, we obtain Theorem 4.1 for $x, y \in F$.

Proposition 4.10. *For every $R_0 > 0$, there are constants $c = c(R_0) > 0$ and $\beta > 0$ such that for every $0 < R \leq R_0$ and every bounded parabolic function h in $Q_{R_0}(0, x_0, 2R)$,*

$$|h(s,x) - h(t,y)| \leq c\, \|h\|_{\infty,F}\, R^{-\beta} \left(|t-s|^{1/\alpha} + \rho(x,y)\right)^{\beta} \tag{4.35}$$

holds for (s,x), $(t,y) \in Q_{R_0}(0, x_0, R)$, *where* $\|h\|_{\infty,F} := \sup_{(t,y)\in[0,\,\gamma_{R_0}(2R)^\alpha]\times F} |h(t,y)|$. *In particular, for the transition density function* $p(t,x,y)$ *of* Y, *there are constants* $c > 0$ *and* $\beta > 0$ *such that for any* $0 < t_0 < 1$, t, $s \in [t_0, 2]$ *and* $(x_i, y_i) \in F \times F$ *with* $i = 1, 2$,

$$|p(s,x_1,y_1) - p(t,x_2,y_2)| \le c\, t_0^{-(d+\beta)/\alpha} \left(|t-s|^{1/\alpha} + \rho(x_1,x_2) + \rho(y_1,y_2) \right)^{\beta}. \tag{4.36}$$

The proof is the same as that of Theorem 4.14 in [8]. As before, we omit it here since it relatively long.

Acknowledgement. The authors are grateful to Zhen-Qing Chen, Alexander Grigor'yan and Martina Zähle for valuable comments and suggestions. Remark 4.3 was pointed out by Zhen-Qing Chen. The authors also thank an anonymous referee for detailed and helpful comments.

References

[1] D. Bakry and T. Coulhon and M. Ledoux and L. Saloff-Coste, Sobolev inequalities in disguise, Indiana Univ. Math. J. **44** (1995), 1033-1074.

[2] M.T. Barlow, *Diffusions on fractals*, Lectures on Probability Theory and Statistics, Lect. Notes Math., Springer, **1690** (1998), 1-121.

[3] M.T. Barlow and R.F. Bass, Brownian motion and harmonic analysis on Sierpínski carpets, Canad. J. Math. **51**, no. 4 (1999), 673-744.

[4] R.F. Bass and D.A. Levin, Transition probabilities for symmetric jump processes, Trans. Amer. Math. Soc. **354**, no. 7 (2002), 2933–2953.

[5] R.F. Bass and D.A. Levin, Harnack inequalities for jump processes, Potential Anal. **17** (2002), 375–388.

[6] K. Bogdan, A. Stós and P. Sztonyk, Harnack inequality for stable processes on d-sets, Studia Math. **158** (2003), 163–198.

[7] E.A. Carlen and S. Kusuoka and D.W. Stroock, Upper bounds for symmetric Markov transition functions, Ann. Inst. Heri. Poincaré-Probab. Statist. **23** (1987), 245-287.

[8] Z.Q. Chen and T. Kumagai, Heat kernel estimates for stable-like processes on d-sets, Stochastic Process Appl. **108** (2003), 27-62.

[9] Z.Q. Chen and T. Kumagai, Heat kernel estimates for jump processes of mixed types on metric measure spaces, in preparation.

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

[11] M. Fukushima and T. Uemura, On Sobolev and capacitary inequalities for contractive Besov spaces over d-sets, Potential Anal. **18** (2003), 59–77.

[12] A. Grigor'yan, Heat kernels and function theory on metric measure spaces, In: Heat kernels and analysis on manifolds, graphs, and metric spaces (P. Auscher et al. (eds.)), Contemp. Math. **338**, pp. 143–172, Amer. Math. Soc., 2003.

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

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

[15] J. Hu, A note on Hajłasz-Sobolev spaces on fractals, J. Math. Anal. Appl. **280** (2003), 91–101.

[16] J. Hu and M. Zähle, Jump processes and nonlinear fractional heat equations on fractals, Math. Nachr. (to appear).

[17] T. Kumagai, Some remarks on stable-like jump processes on fractals, In: Trends in Math., Fractals in Graz 2001 (P. Grabner and W. Woess (eds.)), pp. 185–196, Birkhäuser, 2002.

[18] T. Kumagai, Function spaces and stochastic processes on fractals, In: Fractal geometry and stochastics III (C. Bandt et al. (eds.)), Progr. Probab. **57**, pp. 221–234, Birkhäuser, 2004.

[19] R.S. Phillips, Perturbation theory for semi-groups of linear operators, Trans. Amer. Math. Soc. **74** (1953), 199–221.

[20] L. Saloff-Coste, A note on Poincaré, Sobolev, and Harnack inequalities, Internat. Math. Res. Notices 1992, no. 2, 27–38.
[21] A. Stós, Symmetric α-stable processes on d-sets, Bull. Pol. Acad. Sci. Math. **48** (2000), 237-245.

Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China
E-mail address: `hujiaxin@mail.tsinghua.edu.cn`

Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan
E-mail address: `kumagai@kurims.kyoto-u.ac.jp`