On the dichotomy in the heat kernel two sided estimates

Alexander Grigor'yan and Takashi Kumagai

Abstract. We study the off-diagonal estimates for transition densities of diffusions and jump processes in a setting when they depend essentially only on the time and distance. We state and prove the dichotomy for the tail of the transition density.

1. Preliminaries

Let (M, d) be a locally compact separable metric space and μ be a Radon measure on M with full support.

Definition 1.1. A family $\{p_t\}_{t>0}$ of measurable functions $p_t(x, y)$ on $M \times M$ is called a *heat kernel* if the following conditions are satisfied, for μ-almost all $x, y \in M$ and all $s, t > 0$:

(*i*) Positivity: $p_t(x, y) \geq 0$.

(*ii*) The total mass inequality

$$\int_M p_t(x, y) d\mu(y) \leq 1. \tag{1.1}$$

(*iii*) Symmetry: $p_t(x, y) = p_t(y, x)$.

(*iv*) Semigroup property:

$$p_{s+t}(x, y) = \int_M p_s(x, z) p_t(z, y) d\mu(z). \tag{1.2}$$

(*v*) Approximation of identity: for any $f \in L^2 := L^2(M, \mu)$,

$$\int_M p_t(x, y) f(y) d\mu(y) \xrightarrow{L^2} f(x) \quad \text{as } t \to 0+. \tag{1.3}$$

Any heat kernel gives rise to the *heat semigroup* $\{P_t\}_{t>0}$ where P_t is the operator on functions defined by

$$P_t u(x) = \int_M p_t(x, y) u(y) d\mu(y). \tag{1.4}$$

2000 *Mathematics Subject Classification.* 47D07, 60G52, 60J35, 60J60.

Supported by SFB 701 of the German Research Council (DFG).
Supported by the Grant-in-Aid for Scientific Research (B) 18340027 (Japan).

The conditions $(i)-(iii)$ of Definition 1.1 imply that P_t is a bounded self-adjoint operator in L^2 and, moreover, is a contraction (see, for example, [**10**, Page 28]). The semigroup identity (1.2) implies that $P_tP_s = P_{t+s}$, that is, the family $\{P_t\}_{t>0}$ is a *semigroup.* It follows from (1.3) that

$$s\text{-}\lim_{t\to 0} P_t = \mathrm{I},$$

where I is the identity operator in L^2 and s-lim stands for *strong* limit. Hence, $\{P_t\}_{t>0}$ is a strongly continuous, self-adjoint, contraction semigroup in L^2.

Given the semigroup $\{P_t\}_{t>0}$, define the *infinitesimal generator* $\mathcal{L}$ of the semigroup by

$$\mathcal{L}f := \lim_{t\to 0} \frac{f - P_t f}{t}, \tag{1.5}$$

where the limit is understood in the L^2-norm. The *domain* $\mathrm{dom}(\mathcal{L})$ of the generator $\mathcal{L}$ is the space of functions $f \in L^2$ for which the limit in (1.5) exists. By the Hille–Yosida theorem, $\mathrm{dom}(\mathcal{L})$ is dense in L^2. Furthermore, $\mathcal{L}$ is a self-adjoint, positive definite operator, which immediately follows from the fact that the semigroup $\{P_t\}$ is self-adjoint and contractive. Moreover, we have

$$P_t = \exp(-t\mathcal{L}), \tag{1.6}$$

where the right hand side is understood in the sense of spectral theory.

The notion of the heat kernel is closely linked to Markov processes. Let $\left(\{X_t\}_{t\geq 0}, \{\mathbb{P}_x\}_{x\in M}\right)$ be a reversible Hunt process on M, and assume that it has the transition density $p_t(x,y)$, that is, a function such that, for all $x \in M$, $t > 0$, and all Borel sets $A \subset M$,

$$\mathbb{P}_x(X_t \in A) = \int_M p_t(x,y)\, d\mu(y).$$

Then $p_t(x,y)$ is a heat kernel in the sense of Definition 1.1.

Given a heat kernel $p_t(x,y)$ on (M,d,μ), let us ask whether it satisfies a two-sided estimate of the following type:

$$\frac{c_1}{t^{\alpha/\beta}}\Phi\left(C_1\frac{d(x,y)}{t^{1/\beta}}\right) \leq p_t(x,y) \leq \frac{C_2}{t^{\alpha/\beta}}\Phi\left(c_2\frac{d(x,y)}{t^{1/\beta}}\right), \tag{1.7}$$

where $\alpha, \beta, c_1, c_2, C_1, C_2$ are positive constants, Φ is a non-negative monotone decreasing function on $[0,\infty)$, and (1.7) is supposed to hold for all $t > 0$ and μ-almost all $x, y \in M$. There are two important classes of heat kernels that actually satisfy (1.7). The first class contains the heat kernels of diffusions on various fractals, where the function Φ is of the form

$$\Phi(s) = \exp(-s^{\frac{\beta}{\beta-1}}), \tag{1.8}$$

(see [**1**] and the references therein). The classical Gauss-Weierstrass heat kernel on $\mathbb{R}^n$, given by

$$p_t(x,y) = \frac{1}{(4\pi t)^{n/2}}\exp\left(-\frac{|x-y|^2}{4t}\right),$$

is included in this class with $\alpha = n$ and $\beta = 2$. For this heat kernel, the generator $\mathcal{L}$ coincides with the Friedrichs extension in $L^2(\mathbb{R}^n)$ of $-\Delta$, where Δ is the classical Laplace operator in $\mathbb{R}^n$.

The second class contains the heat kernels of stable-like processes, where the function Φ is of the form

$$\Phi(s) = (1+s)^{-(\alpha+\beta)}, \tag{1.9}$$

(see [**9**] and the references therein). For example, for any $\beta \in (0,2)$, the heat kernel of the β-stable processes in $\mathbb{R}^n$ is included in this class with $\alpha = n$ (recall that the generator of the β-stable process is given by $\mathcal{L} = (-\Delta)^{\beta/2}$). In particular, the heat kernel of the 1-stable process coincides with the Poisson kernel:

$$p_t(x,y) = \frac{c_n}{t^n}\left(1 + \frac{|x-y|^2}{t^2}\right)^{-\frac{n+1}{2}}, \tag{1.10}$$

where $c_n = \Gamma\left(\frac{n+1}{2}\right)/\pi^{(n+1)/2}$, which clearly satisfies (1.7) with function (1.9).

If a heat kernel is stochastically complete and satisfies (1.7) with a function $\Phi(s)$ that decays fast enough as $s \to \infty$ then the parameters α and β in (1.7) are the invariants of the metric measure space (M,d,μ) and do not depend on a particular choice of the heat kernel (cf. [**14**, Corollary 4.7]). The nature of the parameters α and β is of great interest. The parameter α turns out to be the Hausdorff dimension of M. The parameter β is called the *walk dimension* of the heat kernel p_t. This terminology comes from the following observation: if the heat kernel p_t is the transition density of a Markov process X_t on M, then, under mild assumptions about Φ, (1.7) implies that the average time t needed for the process X_t to move away to a distance r from the origin is of the order r^β (see [**1**, Lemma 3.9]).

In this note we are concerned with the following question:

What functions Φ can actually occur in the estimate (1.7)?

Somewhat unexpectedly, it turns out that the functions (1.8) and (1.9) essentially exhaust all possibilities. Indeed, our main result, Theorem 4.1, says that, under mild additional assumptions, if a heat kernel p_t satisfies (1.7) then Φ is equivalent to one of the functions (1.8) or (1.9).

Our approach to the heat kernel estimates is purely analytic and does use the associated Hunt process. For probabilistic approaches to the heat kernel estimates for diffusions and jump processes see [**1**], [**3**], [**4**], [**5**], [**6**], [**7**], [**8**], [**9**], [**12**], [**15**], [**16**], [**17**] and references therein.

2. The Dirichlet form associated with a heat kernel

Let (M,d,μ) be a metric measure space with a heat kernel $\{p_t\}_{t>0}$, and let $\{P_t\}_{t>0}$ be the heat semigroup defined by (1.4). For any $t>0$, we define a quadratic form $\mathcal{E}_t$ on L^2 by

$$\mathcal{E}_t[u] := \left(\frac{u - P_t u}{t}, u\right), \tag{2.1}$$

where $(\cdot,\cdot)$ is the inner product in L^2. An easy computation shows that $\mathcal{E}_t$ can be equivalently defined by

(2.2)
$$\mathcal{E}_t[u] = \frac{1}{2t}\int_M\int_M |u(x)-u(y)|^2\, p_t(x,y)d\mu(y)d\mu(x) + \frac{1}{t}\int_M (1 - P_t 1(x))\, u^2(x)d\mu(x).$$

In terms of the spectral resolution $\{E_\lambda\}$ of the generator $\mathcal{L}$, the form $\mathcal{E}_t$ can be expressed as follows

$$\mathcal{E}_t [u] = \int_0^\infty \frac{1 - e^{-t\lambda}}{t} d\|E_\lambda u\|_2^2,$$

which implies that $\mathcal{E}_t [u]$ is decreasing in t (indeed, this is an elementary exercise to show that the function $t \mapsto \frac{1-e^{-t\lambda}}{t}$ is decreasing).

Let us define a quadratic form $\mathcal{E}$ by

$$\mathcal{E}[u] := \lim_{t \to 0+} \mathcal{E}_t [u] = \int_0^\infty \lambda d\|E_\lambda u\|_2^2 \tag{2.3}$$

(where the limit may be $+\infty$ since $\mathcal{E}[u] \geq \mathcal{E}_t [u]$) and its *domain* $\mathcal{D}(\mathcal{E})$ by

$$\mathcal{D}(\mathcal{E}) := \{u \in L^2 : \mathcal{E}[u] < \infty\}.$$

It is clear from (2.2) and (2.3) that $\mathcal{E}_t$ and $\mathcal{E}$ are non-negative definite. It is easy to see from (2.3) that $\mathcal{D}(\mathcal{E}) = \mathrm{dom}(\mathcal{L}^{1/2})$. In particular, the domain $\mathcal{D}(\mathcal{E})$ is dense in L^2. Note that $\mathcal{D}(\mathcal{E})$ contains $\mathrm{dom}(\mathcal{L})$. Indeed, if $u \in \mathrm{dom}(\mathcal{L})$ then using (1.5) and (2.1), we obtain

$$\mathcal{E}[u] = \lim_{t \to 0} \mathcal{E}_t [u] = (\mathcal{L}u, u) < \infty. \tag{2.4}$$

The quadratic form $\mathcal{E}[u]$ extends to a bilinear form $\mathcal{E}(u, v)$ by the polarization identity

$$\mathcal{E}(u, v) = \frac{1}{4} \left(\mathcal{E}[u + v] - \mathcal{E}[u - v]\right).$$

It follows from (2.4) that $\mathcal{E}(u, v) = (\mathcal{L}u, v)$ for all $u, v \in \mathrm{dom}(\mathcal{L})$.

The space $\mathcal{D}(\mathcal{E})$ is naturally endowed with the inner product

$$\mathcal{E}_1 (u, v) := (u, v) + \mathcal{E}(u, v). \tag{2.5}$$

It is possible to show that the form $\mathcal{E}$ is *closed*, that is, the space $\mathcal{D}(\mathcal{E})$ is *Hilbert* with the $\mathcal{E}_1$ inner product.

It is easy to see from (1.4) and the definition of a heat kernel that the semigroup $\{P_t\}$ is *Markovian*, that is $0 \leq u \leq 1$ implies $0 \leq P_t u \leq 1$. This implies that the form $\mathcal{E}$ satisfies the Markov property, that is $u \in \mathcal{D}(\mathcal{E})$ implies $v := \min(u_+, 1) \in \mathcal{D}(\mathcal{E})$ and $\mathcal{E}[v] \leq \mathcal{E}[u]$. Hence, $\mathcal{E}$ is a *Dirichlet form* (cf. [**10**]).

We say that $\mathcal{E}$ is *local* if $\mathcal{E}(u, v) = 0$ whenever u, v are functions from $\mathcal{D}(\mathcal{E})$ such that the supports $\mathrm{supp} u$ and $\mathrm{supp} v$ are disjoint compact sets. The form $\mathcal{E}$ (or the heat kernel p_t) is called *stochastically complete* if $P_t 1 = 1$ for all $t > 0$, that is, the equality holds in (1.1).

The Dirichlet form $\mathcal{E}$ is said to be *regular* if there exists a subspace $\mathcal{C} \subset \mathcal{D}(\mathcal{E}) \cap C_0(M)$ such that $\mathcal{C}$ is dense in $\mathcal{D}(\mathcal{E})$ with $\mathcal{E}_1$-norm and dense in $C_0(M)$ with uniform norm. (Here $C_0(M)$ is the space of continuous compactly supported functions on M.) When $\mathcal{E}$ is regular, there is a corresponding Markov process X_t which is, furthermore, a Hunt process.

3. Two lemmas

Fix two positive parameters α and β and a monotone decreasing function $\Phi : [0, +\infty) \to [0, +\infty)$ such that $\Phi(s) > 0$ for some $s > 0$.

Lemma 3.1. *Assume that* $\{p_t\}$ *is a heat kernel on* (M, d, μ) *such that, for all* $t > 0$ *and almost all* $x, y \in M$,

$$p_t(x, y) \leq \frac{C}{t^{\alpha/\beta}} \Phi\left(\frac{d(x, y)}{t^{1/\beta}}\right), \tag{3.1}$$

for some $C > 0$. *Then either the associated Dirichlet form* $\mathcal{E}$ *is local or*

$$\Phi(s) \geq \frac{c}{(1+s)^{\alpha+\beta}} \tag{3.2}$$

for all $s > 0$ *and some* $c > 0$.

Proof. Let $u, v \in L^2(M, \mu)$ be two non-negative functions with compact disjoint supports $A = \text{supp} u$ and $B = \text{supp} v$, and set

$$r = d(A, B) > 0 \tag{3.3}$$

(see Fig. 1).

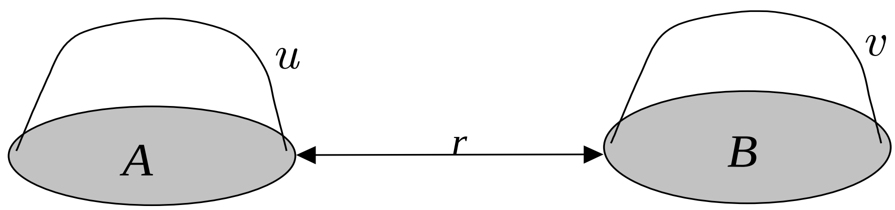

Figure 1. Functions u and v

Consider the bilinear form $\mathcal{E}_t$ on $L^2(M, \mu)$, which is given by (2.1). Since $(u, v) = 0$, we obtain

$$\mathcal{E}_t(u, v) = \left(\frac{u - P_t u}{t}, v\right) = -\frac{1}{t}(P_t u, v),$$

that is,

$$\mathcal{E}_t(u, v) = -\frac{1}{t} \int_A \int_B p_t(x, y) u(x) v(y) \, d\mu(y) d\mu(x). \tag{3.4}$$

If $x \in A$ and $y \in B$ then $d(x, y) \geq r$. Therefore, for almost all $x \in A$ and $y \in B$,

$$p_t(x, y) \leq \frac{C}{t^{\alpha/\beta}} \Phi\left(\frac{r}{t^{1/\beta}}\right),$$

which together with (3.4) implies

$$|\mathcal{E}_t(u, v)| \leq \frac{C}{t^{1+\alpha/\beta}} \Phi\left(\frac{r}{t^{1/\beta}}\right) \|u\|_{L^1} \|v\|_{L^1} \tag{3.5}$$

(note that $\|u\|_{L^1} \leq \mu(A)^{1/2} \|u\|_{L^2} < \infty$ and the same holds for v). If (3.2) fails then there exists a sequence $\{s_k\} \to \infty$ such that

$$s_k^{\alpha+\beta} \Phi(s_k) \to 0 \quad \text{as } k \to \infty.$$

Define a sequence $\{t_k\}$ from the condition

$$s_k = \frac{r}{t_k^{1/\beta}}.$$

Then

$$s_k^{\alpha+\beta}\Phi\left(s_k\right)=\frac{r^{\alpha+\beta}}{t_k^{1+\alpha/\beta}}\Phi\left(\frac{r}{t_k^{1/\beta}}\right)\to 0\quad\text{as } k\to\infty,$$

and (3.5) implies that

$$\mathcal{E}_{t_k}\left(u,v\right)\to 0\quad\text{as } k\to\infty. \tag{3.6}$$

Therefore, if in addition $u,v\in\mathcal{D}(\mathcal{E})$ then, by (2.3) and (3.6),

$$\mathcal{E}\left(u,v\right)=\lim_{k\to\infty}\mathcal{E}_{t_k}\left(u,v\right)=0,$$

whence the locality of $\mathcal{E}$ follows. ■

Lemma 3.2. *Assume that $\{p_t\}$ is a heat kernel on (M,d,μ) such that, for all $t>0$ and almost all $x,y\in M$,*

$$p_t\left(x,y\right)\geq\frac{c}{t^{\alpha/\beta}}\Phi\left(\frac{d\left(x,y\right)}{t^{1/\beta}}\right), \tag{3.7}$$

for some $c>0$. Then

$$\Phi\left(s\right)\leq\frac{C}{\left(1+s\right)^{\alpha+\beta}} \tag{3.8}$$

for all $s>0$ and some $C>0$.

Proof. Let u be a non-constant function from $L^2\left(M,\mu\right)$. Choose a ball $Q\subset M$ where u is non-constant and let $a>b$ be two real values such that the sets

$$A=\{x\in Q:u\left(x\right)\geq a\}\ \text{ and }\ B=\{x\in Q:u\left(x\right)\leq b\}$$

have positive measures (see Fig. 2).

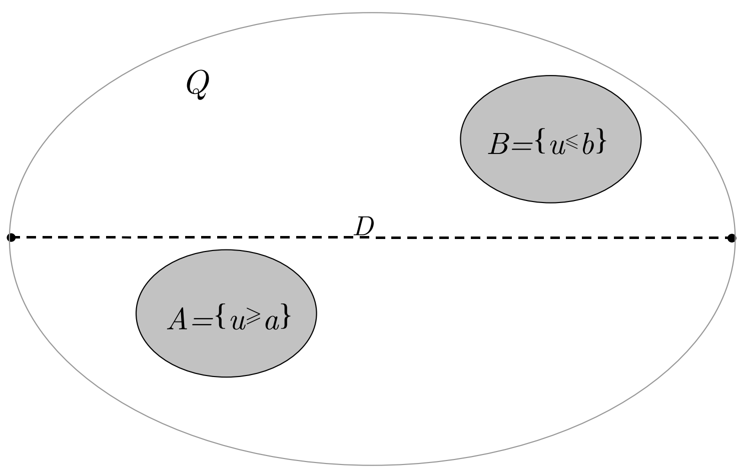

Figure 2. Sets A and B

If the diameter of Q is D then, by (3.7), we have, for almost all $x,y\in Q$,

$$p_t\left(x,y\right)\geq\frac{c}{t^{\alpha/\beta}}\Phi\left(\frac{D}{t^{1/\beta}}\right),$$

whence by (2.2)

$$\begin{aligned}\mathcal{E}[u] &\geq \mathcal{E}_t[u] \geq \frac{1}{2t}\int_A\int_B (u(x)-u(y))^2 p_t(x,y)d\mu(y)d\mu(x)\\ &\geq (a-b)^2\mu(A)\mu(B)\frac{c}{2t^{1+\alpha/\beta}}\Phi\left(\frac{D}{t^{1/\beta}}\right)\\ &= \frac{c'}{t^{1+\alpha/\beta}}\Phi\left(\frac{D}{t^{1/\beta}}\right), \end{aligned} \tag{3.9}$$

where $c'>0$. If (3.8) fails then there exists a sequence $\{s_k\}\to\infty$ such that

$$s_k^{\alpha+\beta}\Phi(s_k)\to\infty \quad \text{as } k\to\infty.$$

Define a sequence $\{t_k\}$ from the condition

$$s_k=\frac{D}{t_k^{1/\beta}}.$$

Then

$$\frac{1}{t_k^{1+\alpha/\beta}}\Phi\left(\frac{D}{t_k^{1/\beta}}\right)=D^{-(\alpha+\beta)}s_k^{\alpha+\beta}\Phi(s_k)\to\infty \quad \text{as } k\to\infty,$$

and (3.9) yields $\mathcal{E}(u,u)=\infty$.

Hence, we have arrived at the conclusion that the domain of the form $\mathcal{E}$ contains only constants. Since $\mathcal{D}(\mathcal{E})$ is dense in $L^2(M,\mu)$, it follows that $L^2(M,\mu)$ consists only of constants. Since μ is a Radon measure on M with full support, it follows that M consists of a single point, say, $M=\{x\}$. Then (1.1) implies that, for all $t>0$,

$$p_t(x,x)\leq\frac{1}{\mu(\{x\})},$$

while by (3.7) $p_t(x,x)\to\infty$ as $t\to 0$. This contradiction finishes the proof. ∎

Definition 3.3. We say that the metric space (M,d) satisfies the *chain condition* if there exists a (large) constant C such that, for any two points $x,y\in M$ and for any positive integer n, there exists a sequence $\{x_i\}_{i=0}^n$ of points in M such that $x_0=x$, $x_n=y$, and

$$d(x_i,x_{i+1})\leq C\frac{d(x,y)}{n}, \qquad \text{for all } i=0,1,\cdots,n-1.$$

In what follows, we write

$$h(s)\simeq f(s)$$

if there exist constants $c_1,c_2>0$ such that

$$c_1f(s)\leq h(s)\leq c_2f(s),$$

for the specified range of the argument s. Similarly, we write

$$h(s)\asymp f(Cs)g(cs)$$

if there exist constants $C_1,c_1,C_2,c_2>0$ such that

$$f(C_1s)g(c_1s)\leq h(s)\leq f(C_2s)g(c_2s),$$

for the specified range of s.

Corollary 3.4. *Assume that the following estimate holds for all $t > 0$ and almost all $x, y \in M$:*

$$p_t(x,y) \asymp \frac{C}{t^{\alpha/\beta}} \Phi\left(c\frac{d(x,y)}{t^{1/\beta}}\right). \tag{3.10}$$

Then either the Dirichlet form $\mathcal{E}$ is local or

$$\Phi(s) \simeq \frac{1}{(1+s)^{\alpha+\beta}}. \tag{3.11}$$

Proof. Indeed, if $\mathcal{E}$ is non-local then, by Lemmas 3.1 and 3.2, the function Φ must satisfy (3.2) and (3.8), whence (3.11) follows. ■

4. Main result

Now we can state and prove our main result.

Theorem 4.1. *Assume that the metric space (M,d) satisfies the chain condition and all metric balls are relatively compact*[1]. *Let $p_t(x,y)$ be a heat kernel in a metric measure space (M,d,μ). Assume that the heat kernel is stochastically complete, the associated Dirichlet form $\mathcal{E}$ is regular, and* (3.10) *holds with some $\alpha, \beta > 0$ and Φ. Then $\beta \le \alpha + 1$,*

$$\mu(B(x,r)) \simeq r^{\alpha}, \tag{4.1}$$

and the following dichotomy holds:

- *either the Dirichlet form $\mathcal{E}$ is local, $\beta \ge 2$, and*
$$\Phi(s) \asymp C \exp\left(-cs^{\frac{\beta}{\beta-1}}\right).$$
- *or the Dirichlet form $\mathcal{E}$ is non-local and*
$$\Phi(s) \simeq (1+s)^{-(\alpha+\beta)}.$$

The main new point of this theorem is the dichotomy of the function Φ. The other claims such as the estimate (4.1) and the inequalities for α and β follow from the results of [**14**] are included here for the sake of completeness of the statement.

Proof. By Lemma 3.2, we have the upper bound

$$\Phi(s) \le \frac{C}{(1+s)^{\alpha+\beta}}, \tag{4.2}$$

which, in particular, implies

$$\int_0^\infty s^{\alpha-1}\Phi(s)\,ds < \infty. \tag{4.3}$$

By [**14**, Theorem 3.2] (see also [**11**]), the estimate (3.10) with a function Φ satisfying (4.3) and the stochastic completeness of the heat kernel imply (4.1). By [**14**, Theorem 4.8(ii)], (see also [**11**]), (3.10) with (4.3) and the chain condition imply that $\beta \le \alpha+1$. If the form $\mathcal{E}$ is non-local, then, by Corollary 3.4, Φ satisfies (3.11), which finishes the proof in this case.

[1] Note that the relative compactness of balls does not follows from the fact that (M,d) is locally compact; for example, take $M = (0,1)$.

Assume now that the form $\mathcal{E}$ is local. As it follows from the proof of [**14**, Theorem 3.2], (3.10) and (4.3) imply that, for any $\varepsilon > 0$ there is $\delta > 0$ such that, for all $r > 0$ and almost all $x \in M$,

$$\int_{B(x,r)^c} p_t(x,y)\,d\mu(y) \le \varepsilon \qquad \text{for all } t \le \delta r^\beta. \tag{4.4}$$

Alternatively, one can easily obtain (4.4) directly from the upper bounds in (3.10) and (4.1). Now we use a result of [**13**, Theorem 4.3], which on top of the locality of the Dirichlet form and (4.4), requires the regularity of the Dirichlet form, the stochastic completeness, and the relative compactness of the metric balls – all that we have by hypotheses. By [**13**, Theorem 4.3], we obtain that, for all positive r, t, λ and almost all $x \in M$,

$$\int_{B(x,r)^c} p_t(x,y)\,d\mu(y) \le C \exp\left(\lambda t - c\lambda^{1/\beta} r\right), \tag{4.5}$$

with some positive constants C, c.

If $\beta < 1$ then letting in (4.5) $\lambda \to \infty$, we obtain that the right hand side in (4.5) goes to 0. It follows that, for almost all $x \in M$,

$$\int_{M\setminus\{x\}} p_t(x,y)\,d\mu(y) = 0.$$

The stochastic completeness implies that, for some $x \in M$, $\mu(\{x\}) > 0$ and

$$p_t(x,x)\,\mu(\{x\}) = 1,$$

which however contradicts (3.10) (cf. the proof of Lemma 3.2). Hence, we conclude that $\beta \ge 1$.

Setting in (4.5)

$$\lambda = \begin{cases} \left(\dfrac{cr}{2t}\right)^{\frac{\beta}{\beta-1}}, & \text{if } \beta > 1, \\ t^{-1}, & \text{if } \beta = 1 \end{cases}$$

we obtain that, for all positive r, t and almost all $x \in M$,

$$\int_{B(x,r)^c} p_t(x,y)\,d\mu(y) \le \begin{cases} C \exp\left(-c\left(\frac{r^\beta}{t}\right)^{\frac{1}{\beta-1}}\right), & \text{if } \beta > 1 \\ C\exp\left(-c\frac{r}{t}\right), & \text{if } \beta = 1 \end{cases} \tag{4.6}$$

(where the constants constants c, C may be different from those of (4.5)).

By (1.2), we have, for all $t > 0$, almost all $x, y \in M$, and $r := \frac{1}{2} d(x,y)$,

$$\begin{aligned} p_t(x,y) &= \int_M p_{\frac{t}{2}}(x,z)\,p_{\frac{t}{2}}(z,y)\,d\mu(z) \\ &\le \left(\int_{B(x,r)^c} + \int_{B(y,r)^c}\right) p_{\frac{t}{2}}(x,z)\,p_{\frac{t}{2}}(z,y)\,d\mu(z) \\ &\le \operatorname{essup}_{z\in M} p_{\frac{t}{2}}(z,y) \int_{B(x,r)^c} p_{\frac{t}{2}}(x,z)\,d\mu(z) \\ &\quad + \operatorname{essup}_{z\in M} p_{\frac{t}{2}}(x,z) \int_{B(y,r)^c} p_{\frac{t}{2}}(y,z)\,d\mu(z). \end{aligned}$$

Since by (3.10) $\operatorname{essup} p_t \le C t^{-\alpha/\beta}$, combining this with (4.6) we obtain, for almost all $x, y \in M$,

$$p_t(x,y) \le \begin{cases} \dfrac{C}{t^{\alpha/\beta}} \exp\left(-c\left(\dfrac{d^\beta(x,y)}{t}\right)^{\frac{1}{\beta-1}}\right), & \text{if } \beta > 1 \\ \dfrac{C}{t^{\alpha}} \exp\left(-c\dfrac{d(x,y)}{t}\right), & \text{if } \beta = 1 \end{cases} \tag{4.7}$$

Now we use [**14**, Theorem 4.8(i)] that says the following: if the heat kernel satisfies for all $t > 0$ and almost all $x, y \in M$ the estimates

$$\frac{1}{t^{\alpha/\beta}} \Phi_1\left(\frac{d(x,y)}{t^{1/\beta}}\right) \le p_t(x,y) \le \frac{1}{t^{\alpha/\beta}} \Phi_2\left(\frac{d(x,y)}{t^{1/\beta}}\right), \tag{4.8}$$

where α, β are positive constants, and Φ_1 and Φ_2 are non-negative monotone decreasing functions on $[0, +\infty)$, such that $\Phi_1(1) > 0$ and $\Phi_2(s) = o\left(s^{-N}\right)$ as $s \to \infty$ for any $N > 0$, then $\beta \ge 2$. Clearly, in our setting the lower bound in (4.8) follows from (3.10) with $\Phi_1(s) = C\Phi(cs)$ and for the upper bound we use the function

$$\Phi_2(s) = \begin{cases} C \exp\left(-cs^{\frac{\beta}{\beta-1}}\right), & \text{if } \beta > 1, \\ C \exp(-cs), & \text{if } \beta = 1, \end{cases}$$

that comes from (4.7). Since Φ_1 and Φ_2 satisfy the cited above conditions, we conclude that $\beta \ge 2$.

On the other hand, when $\beta \ge 2$ (in fact, $\beta > 1$ is enough), the standard chain argument (see [**14**, Corollary 3.5]) shows that the lower bound in (3.10) implies the lower bounds

$$p_t(x,y) \ge \frac{C}{t^{\alpha/\beta}} \exp\left(-c\left(\frac{d^\beta(x,y)}{t}\right)^{\frac{1}{\beta-1}}\right). \tag{4.9}$$

Combining (4.7) and (4.9) with (3.10), we obtain

$$\Phi(s) \asymp C \exp\left(-cs^{\frac{\beta}{\beta-1}}\right),$$

which finishes the proof. ∎

If the heat kernel bounds (3.10) holds only for $t \in (0, T)$ for some $T > 0$ then the statement of Theorem 4.1 remains true with obvious modification of the conclusions. On the contrary, we do not know if the localization is possible for large t, that is, when (3.10) holds for all $t > T$. For example, this case occurs for a continuous time simple random walk on $\mathbb{Z}^d$, which satisfies (3.10) with $\alpha = d$, $\beta = 2$, and

$$\Phi(s) \asymp C \exp(-cs^2) \quad \text{for} \quad t \ge \max(d(x,y), 1),$$

but (3.10) does not hold for $t \ll 1$.

In the setting of Theorem 4.1, the assumption of the locality of the form $\mathcal{E}$ leads to the following relations between α and β:

$$2 \le \beta \le \alpha + 1. \tag{4.10}$$

These relations were proved in [**14**] assuming in addition to (3.10) that function $\Phi(s)$ tends sufficiently fast to 0 as $s \to \infty$. Here we do not need the latter hypothesis at expense of using the locality of $\mathcal{E}$.

By [**2**], any couple of α, β satisfying (4.10), can be realized for the heat kernel estimates (3.10) with a local form. In the case of a non-local form, we have instead the range

$$0 < \beta \leq \alpha + 1.$$

Any couple in the range $0 < \beta < \alpha + 1$ can be realized for the estimate (3.10). Indeed, if $\mathcal{L}$ is the generator of a diffusion with parameters α and β from the range (4.10) then $\mathcal{L}^{\delta}$, $\delta \in (0, 1)$, generated a jump process with the walk dimension $\beta' = \delta\beta$ and the same α, so that β' can take any value from $(0, \alpha + 1)$. We do not know whether $\beta = \alpha + 1$ can actually occur for non-local processes.

Acknowledgements. The authors thank K. Kuwae for valuable comments to the first draft of this paper. They are grateful to the referee for careful reading of the manuscript. The first author gratefully acknowledges the financial support and the hospitality of RIMS Kyoto, where this work was initiated.

References

[1] **Barlow M.T.**, Diffusions on fractals, *in:* "Lectures on Probability Theory and Statistics, Ecole d'été de Probabilités de Saint-Flour XXV - 1995", Lecture Notes Math. 1690, Springer, 1998. 1-121.

[2] **Barlow M.T.**, Which values of the volume growth and escape time exponent are possible for graphs?, *Revista Matemática Iberoamericana*, **40** (2004) 1-31.

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

[4] **Barlow M.T., Bass R.F., Chen Z-Q., Kassmann M.**, Non-local Dirichlet forms and symmetric jump processes, to appear in *Trans. Amer. Math. Soc.*

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

[6] **Barlow M.T., Bass R.F., Kumagai T.**, Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps, to appear in *Math. Z.*

[7] **Barlow M.T., Coulhon T., Kumagai T.**, Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs, *Comm. Pure Appl. Math.*, **58** (2005) no.12, 1642–1677.

[8] **Barlow M.T., Grigor'yan A., Kumagai T.**, Heat kernel upper bounds for jump processes and the first exit time, to appear in *J. Reine Angew. Math.*

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

[10] **Fukushima M., Oshima Y., Takeda M.**, "Dirichlet forms and symmetric Markov processes", Studies in Mathematics 19, De Gruyter, 1994.

[11] **Grigor'yan A.**, Heat kernels and function theory on metric measure spaces, *in:* "Heat kernels and analysis on manifolds, graphs, and metric spaces", *Contemporary Mathematics*, **338** (2003) 143-172.

[12] **Grigor'yan A.**, Heat kernel upper bounds on fractal spaces, preprint

[13] **Grigor'yan A., Hu J.**, Off-diagonal upper estimates for the heat kernel of the Dirichlet forms on metric spaces, preprint

[14] **Grigor'yan A., Hu J., Lau K.S.**, Heat kernels on metric-measure spaces and an application to semi-linear elliptic equations, *Trans. Amer. Math. Soc.*, **355** (2003) no.5, 2065-2095.

[15] **Hambly B.M., Kumagai T.**, Transition density estimates for diffusion processes on post critically finite self-similar fractals, *Proc. London Math. Soc. (3)*, **78** (1999) 431–458.

[16] **Kigami J.**, Local Nash inequality and inhomogeneity of heat kernels, *Proc. London Math. Soc. (3)*, **89** (2004) 525-544.

[17] **Telcs A.**, Random walks on graphs with volume and time doubling, *Revista Matemática Iberoamericana*, **22** (2006) 17-54.

Department of Mathematics, University of Bielefeld, Postfach 100131, 33501 Bielefeld, Germany

E-mail address: grigor@math.uni-bielefeld.de

Department of Mathematics, Faculty of Science, Kyoto University, Kyoto 606-8502, Japan

E-mail address: kumagai@math.kyoto-u.ac.jp