Heat kernel upper bounds for jump processes and the first exit time

Martin T. Barlow [*] Alexander Grigor'yan [†] Takashi Kumagai [‡]

September 2007

Contents

0 Introduction

Let $\{P_t\}_{t\geq 0}$ be a Markovian semigroup acting in $L^2(M,\mu)$ where (M,d,μ) is a metric measure space, and assume that P_t has a continuous integral kernel $p_t(x,y)$ so that

$$P_t f(x) = \int_M p_t(x,y) f(y) \mu(dy),$$

for all $t > 0$, $x \in M$, and $f \in L^2(M,\mu)$. The function $p_t(x,y)$ can be considered as the transition density of the associated Markov process $X = \{X_t\}_{t\geq 0}$, and the question of estimating of $p_t(x,y)$, which is the main subject of this paper, is closely related to the properties of X.

The function $p_t(x,y)$ is also referred to as a heat kernel, and this terminology goes back to the classical Gauss-Weierstrass heat kernel associated with the heat semigroup $\{e^{t\Delta}\}_{t\geq 0}$ in $\mathbb{R}^n$, whose Markov process is Brownian motion. A somewhat more general but still well treated case is when (M,d,μ) is a Riemannian metric measure space, that is, when M is a Riemannian

[*] Research partially supported by NSERC (Canada) and EPSRC (UK)

[†] Supported by SFB 701 of the German Research Council (DFG).

[‡] Research partially supported by the Grant-in-Aid for Scientific Research (B) 18340027 (Japan).

manifold, d is the geodesic distance, and μ is the Riemannian measure. The Laplace-Beltrami operator Δ on M generates the heat semigroup $\{e^{t\Delta}\}_{t\geq 0}$ possessing a smooth heat kernel $p_t(x,y)$, which is associated with the Brownian motion on M. One of the most interesting questions is to determine whether the heat kernel on a given manifold satisfies the following *Gaussian estimate*:

$$p_t(x,y) \leq Ct^{-\gamma} \exp\left(-\frac{d^2(x,y)}{Ct}\right), \tag{0.1}$$

for all $t > 0$ and $x, y \in M$, where γ and C are positive constants. Obviously, if (0.1) holds then it implies the *on-diagonal* estimate

$$p_t(x,x) \leq Ct^{-\gamma}, \tag{0.2}$$

for all $t > 0$ and $x \in M$. Surprisingly enough, the converse is true as well.

Theorem 0.1 ([8], [10], [17]) *On any Riemannian manifold, the on-diagonal estimate* (0.2) *implies the Gaussian estimate* (0.1).

The proof of this theorem is based on the property of the geodesic distance that $|\nabla d| \leq 1$, which is true on any Riemannian manifold. On a general metric measure space, the analogue of this property would typically fail.

In the general setting, obtaining proper off-diagonal estimates from the on-diagonal one requires some additional conditions providing a link between the distance function and the process. For a large variety of self-similar fractal sets, the heat kernel of the corresponding self-similar diffusion process admits the upper bound

$$p_t(x,y) \leq \frac{C}{t^{\alpha/\beta}} \exp\left(-\left(\frac{d^\beta(x,y)}{Ct}\right)^{\frac{1}{\beta-1}}\right), \tag{0.3}$$

where $\alpha > 0$ and $\beta > 1$ are parameters related to the geometry of the underlying space (see [1]). Typically, a matching lower bound (with a different value of C) holds as well, but in this paper we are concerned only with upper bounds.

Let $B(x,r)$ denote a metric ball of radius r centered at $x \in M$, and assume that, for some $\alpha > 0$ and $c > 0$,

$$c^{-1}r^\alpha \leq \mu(B(x,r)) \leq cr^\alpha, \tag{0.4}$$

for all $x \in M$ and $r > 0$. For any open set $U \subset M$, let τ_U be the first exit time of the process X from U. The following result is known.

Theorem 0.2 *Let* (M, d, μ) *satisfy* (0.4) *and let* X *be a stochastically complete diffusion on* M *such that the heat kernel of* X *is continuous and satisfies the on-diagonal estimate*

$$p_t(x,x) \leq Ct^{-\alpha/\beta} \text{ for all } x \in M, t > 0, \tag{0.5}$$

where $\beta > 1$. *Then the following conditions are equivalent:*

(1) *The off-diagonal estimate* (0.3).

(2) *The estimate of the mean exit time:*

$$\mathbb{E}^x \tau_{B(x,r)} \simeq r^\beta \text{ for all } x \in M,\ r > 0. \tag{0.6}$$

(3) *The tail estimate of the first exit time:*

$$\mathbb{P}^x(\tau_{B(x,r)} \le t) \le C \exp\left((- \left(\frac{r^\beta}{Ct}\right)^{\frac{1}{\beta-1}} \right), \quad \textit{for all } x \in M,\ t,\ r > 0. \tag{0.7}$$

The sign $\simeq$ in (0.6) means that the ration of the both sides is bounded from above and below by positive constants.

The implication (1) $\Rightarrow$ (2) was proved in [15], while (2) $\Rightarrow$ (3) $\Rightarrow$ (1) is contained in [1]; see also [15] and [11] for more general results of this kind. Let us give the proof of (3) $\Rightarrow$ (1), which is the easiest part of Theorem 0.2. By the semigroup property,

$$p_t(x, z) \le \sqrt{p_t(x, x)p_t(z, z)} \le Ct^{-\alpha/\beta}. \tag{0.8}$$

Now (0.7) implies that

$$\int_{M\setminus B(x,r)} p_t(x, z)\mu(dz) = \mathbb{P}^x(X_t \notin B(x, r)) \le C \exp\left(- \left(\frac{r^\beta}{Ct}\right)^{\frac{1}{\beta-1}} \right). \tag{0.9}$$

Setting $r = \frac{1}{2}d(x, y)$, using an elementary estimate,

$$\begin{aligned} p_{2t}(x, y) &= \int_M p_t(x, z)p_t(z, y)\mu(dz) \\ &\le \int_{M\setminus B(x,r)} p_t(x, z)p_t(z, y)\mu(dz) + \int_{M\setminus B(y,r)} p_t(x, z)p_t(z, y)\mu(dz) \\ &\le \sup_{z\in M} p_t(z, y) \int_{M\setminus B(x,r)} p_t(x, z)\mu(dz) + \sup_{z\in M} p_t(x, z) \int_{M\setminus B(y,r)} p_t(z, y)\mu(dz), \end{aligned} \tag{0.10}$$

we obtain (0.3) by substituting (0.9) and (0.8) into (0.10).

Note that the crucial estimate (0.7) is very much related to the fact that X is a diffusion. It is natural to ask if there is an analogue of Theorem 0.2 when X is a Markov process with jumps. Certainly, Theorem 0.2 can fail for jump processes. For example, if X is the symmetric stable process in $\mathbb{R}^n$ of index $\beta < 2$ then the heat kernel of this process satisfies the estimate

$$p_t(x, y) \le C \min\left(t^{-\alpha/\beta}, \frac{t}{d(x, y)^{\alpha+\beta}}\right) \simeq Ct^{-\alpha/\beta}\left(1 + \frac{d(x, y)}{t^{1/\beta}}\right)^{-(\alpha+\beta)}, \tag{0.11}$$

where $\alpha = n$ and $d(x, y) = |x - y|$, and the matching lower bound is true as well. Obviously, if the heat kernel satisfies (0.11) and the matching lower bound, then the conditions (0.5) and (0.6) of Theorem 0.2 are satisfied while (0.7) and (0.3) fail.

The first purpose of this paper is to provide some conditions in terms of the first exit time, which are equivalent to the heat kernel bound of the form (0.11). As far as we know this is the first result of this type. Here is a simplified version of our main result, Theorem 1.2.

Theorem 0.3 *Let X be a stochastically complete Hunt process on a metric measure space (M, d, μ) with a continuous heat kernel $p_t(x, y)$. Assuming that* (0.4) *and* (0.5) *are satisfied for some $\alpha, \beta > 0$, the following are equivalent:*

(*a*) *The off-diagonal estimate* (0.11).

(*b*) *For all* $x_0 \in M$, $r > 0$, $t > 0$, *writing* $\tau = \tau_{B(x_0,r)}$,

$$\mathbb{P}^{x_0}(\tau \le t) \le C\frac{t}{r^\beta}, \tag{0.12}$$

Furthermore, for all $x \in B(x_0, r/2)$, $y \in M \setminus B(x_0, 2r)$, *and* $0 < R \le r$,

$$\mathbb{P}^x(\tau \le t,\ X_\tau \in B(y,R)) \le C\frac{tR^\alpha}{r^{\alpha+\beta}}. \tag{0.13}$$

The condition (0.12) can be regarded as an analogue of (0.7). The condition (0.13) is specific for jump processes and estimates the probability that at the moment of exit the process jumps from the ball $B(x_0, r)$ to some other ball $B(y, R)$.

Note that if one repeats the argument (0.10), but using (0.12) instead of (0.7) then one obtains a weaker estimate than (0.11). We use a more complicated bootstrapping argument enabling self-improvement of the heat kernel estimate. Namely, we prove by induction in q the estimate

$$p_t(x,y) \le \frac{C}{t^{\alpha/\beta}}\left(\frac{t}{d(x,y)^\beta}\right)^q, \tag{0.14}$$

which bridges (0.5) and (0.11): for $q = 0$ (0.14) is equivalent to (0.5), while for $q = \alpha/\beta + 1$ it is equivalent to (0.11).

Under some additional assumptions on the space M and the process X we obtain the upper bound (0.11) under certain hypotheses in terms of the jumping density of the process – see Theorem 1.4. A number of previous papers have obtained heat kernel upper bounds for jump processes under similar conditions – see in particular [3, 5, 13]. The main contribution here is to introduce a new decomposition of the heat kernel (see Lemma 3.1), which simplifies the argument.

1 Framework and Main Theorem

Let (M_0, d) be a locally compact separable metric space, μ be a Radon measure on M_0 with full support, and $(\mathcal{E}, \mathcal{F})$ be a regular Dirichlet form on $L^2(M_0, \mu)$. We denote the associated Hunt process as $X = (X_t, t \ge 0, \mathbb{P}^x, x \in M_0)$ and its transition probability as $P_t(x, dy)$. It is well known (see Chapter 7 in [9]) that there is a properly exceptional[1] set $\mathcal{N}_0 \subset M_0$ of X such that the associated Hunt process is uniquely determined up to the ambiguity of starting points from $\mathcal{N}_0$. We write Δ for the cemetery state, and ζ for the lifetime of the process X, and as usual take $X_t = \Delta$ for $t \ge \zeta$.

The transition probability P_t can be regarded as an operator on non-negative Borel functions on $M_0 \setminus \mathcal{N}_0$ by means of the identity

$$P_t f(x) = \int_{M_0\setminus\mathcal{N}_0} f(y) P_t(x, \mu(dy)) = \mathbb{E}^x(f(X_t)),\ \ x \in M_0 \setminus \mathcal{N}_0.$$

The family of operators $\{P_t\}_{t\ge0}$ is called the *heat semigroup* of X.

[1]A set $N \subset M$ is called properly exceptional if N is Borel, $\mu(N) = 0$, and

$$\mathbb{P}_x(X_t \in N \text{ or } X_{t-} \in N \text{ for some } t \ge 0) = 0$$

for all $x \in M \setminus N$.

Definition 1.1 *A* heat kernel *(called also a* transition density*) of X is a non-negative measurable function $p_t(x,y)$ defined on $\mathbb{R}_+ \times M \times M$ where $M \subset M_0$, with the following properties:*

1. *The set $M_0 \setminus M$ is a properly exceptional subset of M_0 containing $\mathcal{N}_0$.*

2. *For any non-negative Borel function f on M and for all $t > 0$, $x \in M$,*

$$P_t f(x) = \int_M p_t(x,y) f(y) \mu(dy).$$

3. *For all $t > 0$ and $x, y \in M$,*

$$p_t(x,y) = p_t(y,x).$$

4. *For all $t, s > 0$ and $x, y \in M$,*

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

The set M is called the domain of the heat kernel. If in addition set M can be represented in the form

$$M = \bigcup_{n=1}^{\infty} F_n, \tag{1.1}$$

*where $\{F_n\}_{n=1}^{\infty}$ is an $\mathcal{E}$-*regular nest[2] *and the function $p_t(x,\cdot)$ is continuous on each F_n for every $t > 0$ and $x \in M$, then the heat kernel $p_t(x,y)$ is called* regular.

Set

$$B(x,r) := \{y \in M_0 : d(x,y) < r\}$$

and consider the following hypotheses:

$(H1)$ There exist $\alpha > 0$ and $C > 0$ and such that

$$\mu(B(x,r)) \le C r^{\alpha}, \quad \text{for all } x \in M_0, r > 0. \tag{1.2}$$

$(H2)$ (*The Nash inequality*) There exist $\beta > 0$ and $C > 0$ such that

$$||f||_2^{2+(2\beta/\alpha)} \le C\mathcal{E}(f,f)||f||_1^{2\beta/\alpha}, \quad \text{for all } f \in \mathcal{F}, \tag{N}$$

where $\|\cdot\|_p$ stands for the norm in $L^p(M_0,\mu)$.

It is well known (cf. [4], [7]) that (N) is equivalent to the $L^1 \to L^\infty$ *ultracontractivity* of the heat semigroup:

$$\|P_t f\|_\infty \le C t^{-\alpha/\beta} \|f\|_1, \tag{1.3}$$

for all $f \in L^1(M_0,\mu)$ and $t > 0$. Further, (1.3) is equivalent to the fact that a heat kernel $p_t(x,y)$ of X exists and its domain M can be chosen so that

$$p_t(x,y) \le C t^{-\alpha/\beta} \text{ for all } t > 0 \text{ and } x, y \in M, \tag{1.4}$$

[2]This means that $\{F_n\}$ is an increasing sequence of closed sets such that $\mathrm{Cap}(M \setminus F_n) \to 0$ as $n \to \infty$ and $\mu(F_n \cap U) > 0$ for any open set U such that $F_n \cap U$ is non-empty (cf. [9, p.67]).

(cf. [2, Theorem 2.1] and [11, Lemma 2.2 and Corollary 8.4]). Moreover, by the proof of [2, Theorem 2.1], the heat kernel can be made regular. Combining the results cited above, we conclude that under assumption $(H2)$, a regular heat kernel exists. The regularity of $p_t(x,y)$ implies that function $p_t(x,\cdot)$ is quasi-continuous[3] for all $t>0$ and $x\in M$.

In what follows, let us fix a regular heat kernel $p_t(x,y)$ with the domain M. We may and will consider the Dirichlet form $(\mathcal{E},\mathcal{F})$ and the process X on M rather than on M_0. Our purpose is to establish equivalent conditions for upper bounds for the heat kernel that are typical for certain jump processes.

It is known (see [9, Theorem 3.2.1]) that any regular Dirichlet form admits a unique representation in the following form:

$$\mathcal{E}(u,v)=\mathcal{E}^{(c)}(u,v)+\int_{M\times M\setminus diag}(u(x)-u(y))(v(x)-v(y))n(dx,dy)+\int_M u(x)v(x)k(dx), \quad (1.5)$$

for all $u,v\in\mathcal{F}\cap C_0(M)$. Here $\mathcal{E}^{(c)}$ is a symmetric form that satisfies the strong local property, n is a symmetric positive Radon measure on $M\times M$ off the diagonal $diag$, and k is a positive Radon measure on M. The measure n is called the *jumping measure* and k is called the *killing measure*.

For any set $U\subset M$, let τ_U be the *first exit time* from U, that is,

$$\tau_U=\inf\{t>0: X_t\notin U\}. \quad (1.6)$$

Note that since $\Delta\notin U$, we have $\tau_U\le\zeta$. If U is open then, by the right continuity of the process, we have $X_{\tau_U}\notin U$.

We will discuss the equivalence of the following three properties.

(a) X is stochastically complete, and for all $x,y\in M$ and $t>0$,

$$p_t(x,y)\le C\min\left(t^{-\alpha/\beta},\frac{t}{d(x,y)^{\alpha+\beta}}\right). \quad (UHKP)$$

(b) For all $x_0\in M$, $r>0$, $t>0$,

$$\mathbb{P}^{x_0}(\tau\le t)\le C\frac{t}{r^\beta}, \quad (1.7)$$

where $\tau=\tau_{B(x_0,r)}$. Furthermore, for all $x\in B(x_0,r/2)$, $y\in M\setminus B(x_0,2r)$, and $0<R\le r$,

$$\mathbb{P}^x(\tau\le t,\ X_\tau\in B(y,R))\le C\frac{tR^\alpha}{r^{\alpha+\beta}}, \quad (1.8)$$

(see Fig. 1).

(c) There exists a jumping density $n(x,y)$ w.r.t. μ, i.e.

$$n(dx,dy)=n(x,y)\mu(dx)\mu(dy),$$

such that, for μ-a.e. $x,y\in M$,

$$n(x,y)\le\frac{C}{d(x,y)^{\alpha+\beta}}. \quad (UJ)$$

[3] A function u is called quasi-continuous if, for any $\varepsilon>0$, there exists an open set G in the domain of u such that $\mathrm{Cap}(G)<\varepsilon$ and $u|_{G^c}$ is continuous.

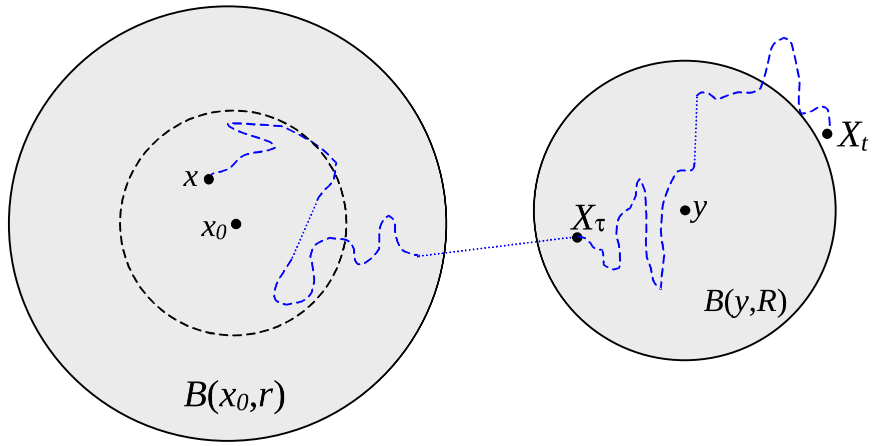

Figure 1: Illustration to (1.8)

Our first main results is:

Theorem 1.2 *If hypotheses* $(H1)$ *and* $(H2)$ *hold then*

$$(a) \Leftrightarrow (b) \Rightarrow (c).$$

By a result of [12], (a) implies the lower bound for the volume of balls: there exists $c > 0$ such that

$$cr^\alpha \le \mu(B(x,r)), \quad \text{for all } x \in M,\ r > 0. \tag{1.9}$$

Combining with Theorem 1.2, we obtain

Corollary 1.3 *Assume that* $(H1)$ *and* $(H2)$ *hold. Then* (b) *implies* (1.9).

An alternative proof of this statement will be given in Section 2.4.

Remark. Under the assumptions of Theorem 1.2, the implication $(c) \Rightarrow (a)$ does not hold in general. Indeed, let $M_0 = \mathbb{R}$, $\beta = 1$ and consider the Dirichlet form:

$$\mathcal{E}(u,v) = \int_{\mathbb{R}} (\nabla u(x), \nabla v(x))dx + \int\int_{\mathbb{R}\times\mathbb{R}\setminus diag} \frac{(u(x)-u(y))(v(x)-v(y))}{|x-y|^2} dxdy.$$

This is the sum of the Dirichlet forms for the Brownian motion and the Cauchy process (i.e. a stable process of index 1). The associated process X can be written as $X = B + Z$, where B is a standard Brownian motion of $\mathbb{R}$, Z is a Cauchy process, and B and Z are independent. Let $t \in (0,1)$, and take $Z_0 = B_0 = X_0 = 0$. Since the transition density of Z does satisfy $(UHKP)$ with $\beta = 1$, we have

$$\mathbb{P}(|Z_t| > t^{1/2}) \le c \int_{t^{1/2}}^{\infty} tr^{-2} dr \le c' t^{1/2}. \tag{1.10}$$

On the other hand

$$\mathbb{P}(|X_t| \ge t^{1/2}) \ge \mathbb{P}(|B_t| \ge 2t^{1/2}) - \mathbb{P}(|Z_t| > t^{1/2}) \ge c_1 - ct^{1/2}.$$

So there exists $c_2 > 0$ such that, for all sufficiently small t, we have

$$\mathbb{P}(|X_t| \ge t^{1/2}) \ge c_2.$$

Thus by (1.10) the density of X cannot satisfy (*UHKP*) with $\beta = 1$.

We now turn to the question of obtaining heat kernel upper bounds on $p_t(x,y)$ given an upper bound on the jump density $n(x,y)$. We restrict ourself to the case when the Dirichlet form is given by

$$\mathcal{E}(u,v) = \int_{M\times M\setminus diag} (u(x)-u(y))(v(x)-v(y))n(x,y)\mu(dx)\mu(dy), \tag{1.11}$$

and the jumping density $n(x,y)$ satisfies (UJ); in particular, the condition (c) is satisfied. Let Lip_0 be the set of compactly supported Lipschitz functions on M. It is easy to check that if $0 < \beta < 2$ then $\mathcal{E}(f,f) < \infty$ for any $f \in \mathrm{Lip}_0$. Hence, it is natural to assume that $\mathrm{Lip}_0 \subset \mathcal{F}$.

Theorem 1.4 *Suppose that* $(H1)$ *and* $(H2)$ *hold. Assume in addition that* $0 < \beta < 2$*,* $\mathcal{E}$ *is given by* (1.11)*,* $\mathrm{Lip}_0 \subset \mathcal{F}$*, and* (c) *is satisfied. Then* $(UHKP)$ *holds, that is,*

$$p_t(x,y) \le C \min\left(t^{-\alpha/\beta}, \frac{t}{d(x,y)^{\alpha+\beta}}\right) \qquad \text{for all } x,y \in M,\ t>0.$$

In order to obtain the implication $(c) \Rightarrow (a)$ we still need to ensure that X is stochastically complete, which can be proved under additional assumptions as in the next statement.

Theorem 1.5 *Suppose, in addition to the hypotheses of Theorem* 1.4 *that d is a geodesic metric, and that* μ *satisfies* (1.9)*. Then* X *is stochastically complete, and so* (a) *and* (b) *hold.*

In both Theorems 1.4, 1.5, we assume the Nash inequality (N) and the upper bound (UJ) for the jump density. It was shown in [13] that the Nash inequality (N) follows from the two sided estimate of $n(x,y)$:

$$\frac{C_2}{d(x,y)^{\alpha+\beta}} \le n(x,y) \le \frac{C_1}{d(x,y)^{\alpha+\beta}}. \tag{1.12}$$

Furthermore, it was proved in [5] and [13] that if measure μ satisfies (1.2) and (1.9) and $n(x,y)$ satisfies (1.12) with $0<\beta<2$ then the heat kernel admits the upper bound $(UHKP)$ as well as a matching lower bound.

2 Proofs

2.1 Some tools for heat kernel estimates

For any two non-negative μ-measurable functions f,g on M, set

$$(f,g) = \int_M fg d\mu.$$

In the next statement, we assume only the conditions from the first paragraph of Section 1 and set $M = M_0 \setminus \mathcal{N}_0$.

Lemma 2.1 *Let* U *and* V *be two disjoint non-empty open subsets of* M *and* f,g *be non-negative Borel functions on* M*. Let* $\tau = \tau_U$ *and* $\tau' = \tau_V$ *be the first exit times from* U *and* V*, respectively. Then, for all* $a,b,t>0$ *such that* $a+b=t$*, we have*

$$(P_t f, g) \le (\mathbb{E}^{\cdot}(\mathbf{1}_{\{\tau \le a\}} P_{t-\tau} f(X_\tau)), g) + (\mathbb{E}^{\cdot}(\mathbf{1}_{\{\tau' \le b\}} P_{t-\tau'} g(X_{\tau'})), f). \tag{2.1}$$

Proof. We have the obvious inequality

$$P_t f = \mathbb{E}^{\cdot} f(X_t) \leq \mathbb{E}^{\cdot}(1_{(X_a \notin U)} f(X_t)) + \mathbb{E}^{\cdot}(1_{(X_a \notin V)} f(X_t)). \tag{2.2}$$

By definition, $X_a \notin U$ implies $\tau_U \leq a$. Hence, using the strong Markov property, we can estimate the first term in (2.2) as follows:

$$\mathbb{E}^{\cdot}(1_{(X_a \notin U)} f(X_t)) \leq \mathbb{E}^{\cdot}(1_{(\tau_U \leq a)} f(X_t)) = \mathbb{E}^{\cdot}(1_{(\tau_U \leq a)} P_{t-\tau_U} f(X_{\tau_U})), \tag{2.3}$$

which matches the first term in the right hand side of (2.1).

Notice that the second term in (2.2) can be written in the form $\mathbb{E}^{\cdot}(h(X_a) f(X_t))$, where $h = 1_{\{x \notin V\}}$. Using the identity

$$(\mathbb{E}^{\cdot}(h(X_a) f(X_t)), g) = (\mathbb{E}^{\cdot}(h(X_{t-a}) g(X_t)), f)$$

(see [9, Lemma 4.1.2]) and $a + b = t$, we obtain

$$(\mathbb{E}^{\cdot}(1_{(X_a \notin V)} f(X_t)), g) = (\mathbb{E}^{\cdot}(1_{(X_b \notin V)} g(X_t)), f).$$

Estimating the right hand side here similarly to (2.3) and combining all the lines above, we obtain (2.1). □

Remark. The most interesting case of (2.1), which occurs in applications, is when f is supported in V and g is supported in U. An intuitive explanation of (2.1) is given by noting that $(P_t f, g) = (\mathbb{E}^{\cdot} f(X_t), g)$ is symmetric in f, g and can be represented as a integral in the space of paths between two points $x \in U$ and $y \in V$. Let τ be the first exit time from U starting at $x \in U$ and τ' be the first exit time from V starting at $y \in V$, we have on the same path that $\tau + \tau' \leq t$ (see Fig. 2), which implies that either $\tau \leq a$ or $\tau' \leq b$.

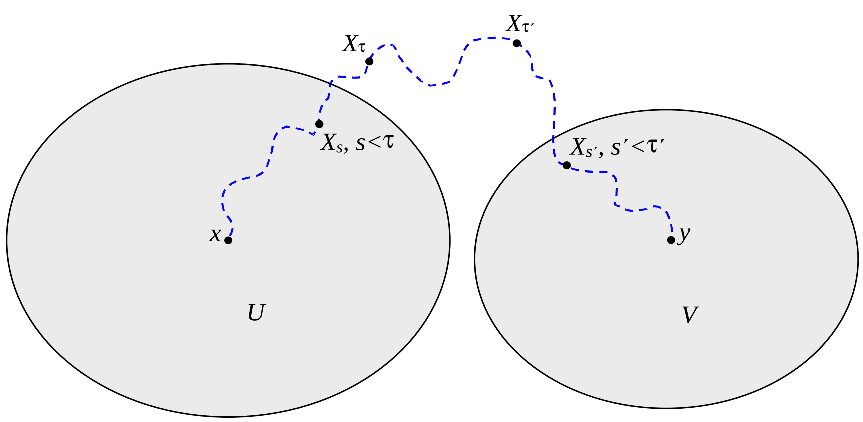

Figure 2: Illustration to $\tau + \tau' \leq t$.

Lemma 2.2 *Let $p_t(x, y)$ be a regular heat kernel of X with the domain M and $\varphi(x, y)$ be a continuous function on $M \times M$. Suppose that, for some fixed $t > 0$, the following inequality holds for $\mu \times \mu$-almost all $(x, y) \in M \times M$:*

$$p_t(x, y) \leq \varphi(x, y). \tag{2.4}$$

Then (2.4) holds for all $(x, y) \in M \times M$.

Proof. Fix some $x \in M$ and first show that if (2.4) holds for all $y \in S$ where $S \subset M$ is a set of full measure then (2.4) extends to all $y \in M$. Indeed, let $p_t(x, y) > \varphi(x, y)$ for some $y \in M$. Choose an index n such that $y \in F_n$, where $\{F_n\}_{n=1}^{\infty}$ is a regular nest from the definition of a regular heat kernel (see Section 1). Since the function $p_t(x, \cdot) - \varphi(x, \cdot)$ is continuous on F_n, there is an open set $U \ni y$ in M such that $p_t(x, \cdot) > \varphi(x, \cdot)$ on $F_n \cap U$. Since $F_n \cap U$ is non-empty, we obtain by the definition of a regular nest that $\mu(F_n \cap U) > 0$. Hence, $p_t(x, \cdot) > \varphi(x, \cdot)$ on a set of positive measure, which contradicts the hypothesis that $\mu(M \setminus S) = 0$ (cf. the proof of Theorem 2.1.2 (ii) in [9]).

Let E be a set of full measure in $M \times M$ such that (2.4) holds for all $(x, y) \in E$. Define the sets

$$\begin{aligned} E_x &= \{y \in M : (x, y) \in E\} \\ M' &= \{x \in M : \mu(M \setminus E_x) = 0\}. \end{aligned}$$

By Fubini's theorem $\mu(M \setminus M') = 0$. Let us show that (2.4) holds for all $x \in M'$ and $y \in M$. Indeed, by the definition of E_x, (2.4) holds for all $y \in E_x$. Since $x \in M'$, the set E_x has full measure, which implies by the above argument that (2.4) extends to all $y \in M$.

Using the symmetry of the heat kernel, we can switch the arguments x and y and continue as follows. Since for any $y \in M$ the inequality (2.4) holds for all $x \in M'$ and M' has full measure, (2.4) extends by the above argument to all $x \in M$, which finishes the proof. □

2.2 Proof of Theorem 1.2: $(b) \Rightarrow (a)$

We begin by proving that X is stochastically complete. Let ζ denote the lifetime of X. Then for any x, r, the definition of exit times gives that $\tau_{B(x,r)} \leq \zeta$. By (1.7),

$$\mathbb{P}^x(\zeta \leq t) \leq \mathbb{P}^x(\tau_{B(x,r)} \leq t) \leq c\frac{t}{r^\beta},$$

and so, letting $r \to \infty$, we have $\mathbb{P}^x(\zeta \leq t) = 0$ for all t.

We now turn to the proof of $(UHKP)$. Since (N), and so (1.4) holds, it is sufficient to prove that, for all distinct $x, y \in M$ and $t > 0$,

$$p_t(x, y) \leq \frac{Ct}{d(x, y)^{\alpha+\beta}}. \tag{2.5}$$

For a parameter $q \geq 0$, consider the following condition, which will be called (H_q): there exists C_q such that, for all $x, y \in M$ and $t > 0$,

$$p_t(x, y) \leq \frac{C_q}{t^{\alpha/\beta}} \left(\frac{t}{d(x, y)^\beta}\right)^q. \tag{2.6}$$

Observe that (1.4) is equivalent to (H_0), whereas (2.5) is equivalent to $(H_{1+\alpha/\beta})$. Note also that the condition (H_q) gets stronger when q increases. Indeed, if (H_q) holds and $q' < q$ then $(H_{q'})$ holds for the following reason: if $t \geq d(x, y)^\beta$ then (2.6) trivially follows from (1.4), whereas if $t < d(x, y)^\beta$ then the exponent q in (2.6) can be replaced by a smaller value without violating the inequality.

We will prove the following implications under the hypotheses of (b):

(i) If (H_q) holds with $q < \alpha/\beta$ then (H_{q+1}) holds.

(ii) If (H_q) holds with $q > \alpha/\beta$ then (2.5) holds.

These two claims will finish the proof. Indeed, set

$$q_0 = \sup\{q : (H_q) \text{ holds}\}.$$

Then (H_q) holds for $q \in [0, q_0)$ and fails for $q \in (q_0, \infty)$. By (i) and the fact that (H_0) holds we have that $q_0 \geq \alpha/\beta + 1$. Hence (H_q) holds with $q = \alpha/\beta + \frac{1}{2}$, and so by (ii) (2.5) holds.

Proof of (i)**.** Assume that (2.6) holds for some $q < \alpha/\beta$ and prove that, for all distinct $x, y \in M$ and $t > 0$,

$$p_t(x,y) \leq \frac{C}{t^{\alpha/\beta}} \left(\frac{t}{d(x,y)^\beta} \right)^{q+1}. \tag{2.7}$$

In what follows, fix $t > 0$ and set $\rho = t^{1/\beta}$. Observe that if $d(x,y) \leq 4\rho$ (or, more generally, $d(x,y) \leq \text{const}\, \rho$) then (2.7) trivially follows from (1.4).

Fix some distinct points $x_0, y_0 \in M$, such that $d(x_0, y_0) > 4\rho$ and set $r = \frac{1}{2} d(x_0, y_0)$ so that $r > 2\rho$. Applying Lemma (2.1) with $U = B(x_0, r)$ and $V = B(y_0, r)$, we obtain, for all non-negative Borel functions f and g on M,

$$(P_t f, g) \leq (\mathbb{E}^{\cdot}(\mathbf{1}_{\{\tau \leq t/2\}} P_{t-\tau} f(X_\tau)), g) + (\mathbb{E}^{\cdot}(\mathbf{1}_{\{\tau' \leq t/2\}} P_{t-\tau'} g(X_{\tau'})), f), \tag{2.8}$$

where $\tau = \tau_{B(x_0,r)}$ and $\tau' = \tau_{B(y_0,r)}$.

Let f be supported in $B(y_0, \rho)$ and g be supported in $B(x_0, \rho)$. In particular, we have

$$(\mathbb{E}^{\cdot}(\mathbf{1}_{\{\tau \leq t/2\}} P_{t-\tau} f(X_\tau)), g) = \int_{B(x_0,\rho)} \mathbb{E}^x(\mathbf{1}_{\{\tau \leq t/2\}} P_{t-\tau} f(X_\tau)) g(x) \mu(dx), \tag{2.9}$$

and a similar identity holds for the second term in (2.8). In order to estimate the integral in (2.9), set $\rho_k = 2^k \rho$ where $k = 1, 2, ...,$ and consider the annuli

$$\begin{aligned} A_1 &: = B(y_0, \rho_1) \\ A_k &: = B(y_0, \rho_k) \setminus B(y_0, \rho_{k-1}),\ k > 1 \end{aligned}$$

(see Fig. 3).

Since the annuli $\{A_k\}_{k=1}^\infty$ form a partition of M, we have

$$\mathbb{E}^x(\mathbf{1}_{\{\tau \leq t/2\}} P_{t-\tau} f(X_\tau)) = \sum_{k=1}^\infty \mathbb{E}^x(\mathbf{1}_{\{\tau \leq t/2\}} \mathbf{1}_{\{X_\tau \in A_k\}} P_{t-\tau} f(X_\tau)). \tag{2.10}$$

To estimate the first term in the sum (2.10), with $k = 1$, observe that

$$t/2 \leq t - \tau \leq t$$

whence by (1.4)

$$P_{t-\tau} f(X_\tau) \leq C t^{-\alpha/\beta} \|f\|_1.$$

Applying (1.8) with $R = \rho_1 = 2t^{1/\beta} < r$ and using $t < r^\beta$ and $q < \alpha/\beta$, we obtain

$$\begin{aligned} \mathbb{E}^x(\mathbf{1}_{\{\tau \leq t/2\}} \mathbf{1}_{\{X_\tau \in A_1\}} P_{t-\tau} f(X_\tau)) &\leq \mathbb{P}^x(\tau \leq t/2,\ X_\tau \in B(y_0, \rho_1)) C t^{-\alpha/\beta} \|f\|_1 \\ &\leq \frac{C t \rho_1^\alpha}{r^{\alpha+\beta} t^{\alpha/\beta}} \|f\|_1 \\ &= \frac{Ct}{r^{\alpha+\beta}} \|f\|_1 \\ &\leq \frac{C}{t^{\alpha/\beta}} (\frac{t}{r^\beta})^{q+1} \|f\|_1. \end{aligned} \tag{2.11}$$

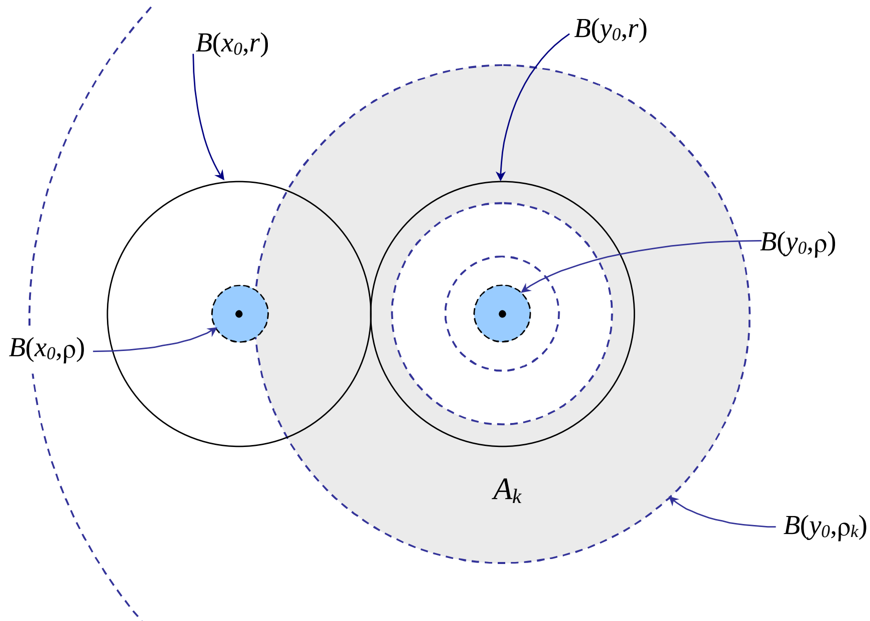

Figure 3: Annuli A_k

Consider now the terms in (2.10) with $k > 1$. If $X_\tau \in A_k$ then $d(X_\tau, y_0) \geq \rho_{k-1}$ and, hence, for all $y \in B(y_0, \rho)$,

$$d(X_\tau, y) \geq \rho_{k-1} - \rho \geq \frac{1}{2}\rho_{k-1} = \frac{1}{4}\rho_k.$$

Then (2.6) yields

$$P_{t-\tau} f(X_\tau) = \int_{B(y_0,\rho)} p_{t-\tau}(X_\tau, y) f(y)\mu(dy) \leq \frac{C}{t^{\alpha/\beta}} (\frac{t}{\rho_k^\beta})^q \|f\|_1. \tag{2.12}$$

Next, consider separately the terms with $\rho_k > r$ and with $\rho_k \leq r$. Using $\rho < r/2$, we obtain from (1.7), for any $x \in B(x_0, \rho)$,

$$\mathbb{P}^x(\tau \leq t/2) \leq \mathbb{P}^x(\tau_{B(x,r/2)} \leq t/2) \leq C\frac{t}{r^\beta}.$$

Using this estimate and (2.12), we obtain

$$\begin{aligned}
& \sum_{\{k:\ \rho_k > r\}} \mathbb{E}^x(\mathbf{1}_{\{\tau \leq t/2\}} \mathbf{1}_{\{X_\tau \in A_k\}} P_{t-\tau} f(X_\tau)) \\
\leq & \sum_{\{k:\rho_k > r\}} \mathbb{P}^x(\tau \leq t/2) \frac{C}{t^{\alpha/\beta}} (\frac{t}{\rho_k^\beta})^q \|f\|_1 \\
\leq & \ C \sum_{\{k:\rho_k > r\}} \frac{t}{r^\beta} \frac{1}{t^{\alpha/\beta}} (\frac{t}{\rho_k^\beta})^q \|f\|_1 \\
\leq & \ C \frac{t}{r^\beta} \frac{1}{t^{\alpha/\beta}} (\frac{t}{r^\beta})^q \|f\|_1 \\
= & \ \frac{C}{t^{\alpha/\beta}} (\frac{t}{r^\beta})^{q+1} \|f\|_1.
\end{aligned} \tag{2.13}$$

Similarly, using (2.12) and (1.8) with $R = \rho_k$, we obtain

$$\begin{aligned} &\sum_{\{k>1:\ \rho_k \le r\}} \mathbb{E}^x(\mathbf{1}_{\{\tau\le t/2\}}\mathbf{1}_{\{X_\tau\in A_k\}}P_{t-\tau}f(X_\tau)) \\ \le &\sum_{\{k>1:\rho_k\le r\}} \mathbb{P}^x(\tau \le t/2,\ X_\tau \in B(y_0,\rho_k))\frac{C}{t^{\alpha/\beta}}\left(\frac{t}{\rho_k^\beta}\right)^q \|f\|_1 \\ \le &\sum_{\{k:\ \rho_k\le r\}} \frac{Ct\rho_k^\alpha}{r^{\alpha+\beta}}\frac{C}{t^{\alpha/\beta}}\left(\frac{t}{\rho_k^\beta}\right)^q \|f\|_1 \\ \le &\ \frac{C}{t^{\alpha/\beta}}\frac{t^{q+1}}{r^{\alpha+\beta}}\|f\|_1 \sum_{\{k:\rho_k\le r\}} \rho_k^{\alpha-\beta q}. \end{aligned} \tag{2.14}$$

Since $\alpha - \beta q > 0$, the sum in (2.14) is comparable to the largest term, that is, to $r^{\alpha-\beta q}$, whence it follows that

$$\sum_{\{k>1:\ \rho_k\le r\}} \mathbb{E}^x(\mathbf{1}_{\{\tau\le t/2\}}\mathbf{1}_{\{X_\tau\in A_k\}}P_{t-\tau}f(X_\tau)) \le \frac{C}{t^{\alpha/\beta}}\frac{t^{q+1}}{r^{\alpha+\beta}}r^{\alpha-\beta q}\|f\|_1 = \frac{C}{t^{\alpha/\beta}}\left(\frac{t}{r^\beta}\right)^{q+1}\|f\|_1.$$

Thus, we have shown that, for any $x \in B(x_0,\rho)$,

$$\mathbb{E}^x(\mathbf{1}_{\{\tau\le t/2\}}P_{t-\tau}f(X_\tau)) \le \frac{C}{t^{\alpha/\beta}}\left(\frac{t}{r^\beta}\right)^{q+1}\|f\|_1,$$

whence by (2.9)

$$(\mathbb{E}^{\cdot}(\mathbf{1}_{\{\tau\le a\}}P_{t-\tau}f(X_\tau)), g) \le \frac{C}{t^{\alpha/\beta}}\left(\frac{t}{r^\beta}\right)^{q+1}\|f\|_1\|g\|_1.$$

Estimating similarly the second term in (2.8), we obtain

$$(P_t f, g) \le \frac{C}{t^{\alpha/\beta}}\left(\frac{t}{r^\beta}\right)^{q+1}||f||_1||g||_1.$$

It follows that, for $\mu\times\mu$-almost all $(x,y)\in B(x_0,\rho)\times B(y_0,\rho)$,

$$p_t(x,y) \le \frac{C}{t^{\alpha/\beta}}\left(\frac{t}{d(x,y)^\beta}\right)^{q+1}. \tag{2.15}$$

Consider the set

$$\mathcal{M}_\rho = \{(x,y)\in M\times M : d(x,y) > 4\rho\}.$$

Since $\mathcal{M}_\rho$ is a separable metric space, it can be covered by a countable family of subsets $\{B(x_k,\rho)\times B(y_k,\rho)\}_{k=1}^\infty$ where $(x_k,y_k)\in\mathcal{M}_\rho$. By the above argument, (2.15) holds for $\mu\times\mu$-almost all $(x,y)\in B(x_k,\rho)\times B(y_k,\rho)$ for any k, whence it follows that (2.15) holds for $\mu\times\mu$-almost all $(x,y)\in\mathcal{M}_\rho$. As it was already mentioned at the beginning of the proof, (2.15) trivially holds if $d(x,y)\le 4\rho$, that is, if $(x,y)\notin\mathcal{M}_\rho$. Combining (2.15) with (1.4), we obtain that, for $\mu\times\mu$-almost all $(x,y)\in M\times M$,

$$p_t(x,y) \le \frac{C}{t^{\alpha/\beta}}\min\left(1, \left(\frac{t}{d(x,y)^\beta}\right)^{q+1}\right).$$

Since the right hand side is a continuous function of $(x, y) \in M \times M$ for any fixed $t > 0$, we conclude by Lemma 2.2 that the same inequality holds for all $x, y \in M$, which proves (2.7). □

Proof of (ii)**.** Assuming that (2.6) holds for some $q > \alpha/\beta$, we need to prove that, for all distinct $x, y \in M$ and $t > 0$,

$$p_t(x,y) \le \frac{Ct}{d(x,y)^{\alpha+\beta}}. \tag{2.16}$$

Using the same setting and notation as in the part (i), let us estimate again the sum (2.10) as follows. For the term with $k = 1$ use the upper bound (2.11), which is already in the required form. For the terms with $\rho_k > r$, using (2.13) and $q > \alpha/\beta$ and $t < r^\beta$, we obtain

$$\sum_{\{k:\ \rho_k > r\}} \mathbb{E}^x(\mathbf{1}_{\{\tau \le t/2\}}\mathbf{1}_{\{X_\tau \in A_k\}} P_{t-\tau} f(X_\tau)) \le \frac{C}{t^{\alpha/\beta}} \left(\frac{t}{r^\beta}\right)^{q+1} \|f\|_1 \le \frac{Ct}{r^{\alpha+\beta}} \|f\|_1.$$

For the terms with $\rho_k \le r$, use the estimate (2.14) but then argue as follows. Since $\alpha - \beta q < 0$, the largest term in the sum in (2.14) is of the order $\rho^{\alpha-\beta q} = t^{\alpha/\beta - q}$, whence

$$\sum_{\{k>1:\ \rho_k \le r\}} \mathbb{E}^x(\mathbf{1}_{\{\tau \le t/2\}}\mathbf{1}_{\{X_\tau \in A_k\}} P_{t-\tau} f(X_\tau)) \le \frac{C}{t^{\alpha/\beta}} \frac{t^{q+1}}{r^{\alpha+\beta}} t^{\alpha/\beta - q} \|f\|_1 = \frac{Ct}{r^{\alpha+\beta}} \|f\|_1.$$

Combining the above estimates, we finish the proof of (ii) in the same way as in part (i). □

2.3 Proof of Theorem 1.2: $(a) \Rightarrow (c)$ and $((a) + (c)) \Rightarrow (b)$

We use the following Lévy system formula (see, for example, [5, Lemma 4.7]).

Lemma 2.3 *Assume that the jumping measure has a density $n(x, y)$ for μ-a.e. $x, y \in M$. Let f be a non-negative measurable function on $\mathbb{R}_+ \times M \times M$, vanishing on the diagonal. Then for every $t \ge 0$, $x \in M$ and every stopping time T (with respect to the filtration of $\{X_t\}$),*

$$\mathbb{E}^x[\sum_{s \le T} f(s, X_{s-}, X_s)] = \mathbb{E}^x[\int_0^T \int_M f(s, X_s, y) n(X_s, y) \mu(dy) ds].$$

Proof of $(a) \Rightarrow (c)$**.** Consider the form $\mathcal{E}_t(f, g) := (f - P_t f, g)/t$. Since X is stochastically complete, we can write

$$\mathcal{E}_t(f,g) = \frac{1}{2t} \int_M \int_M (f(x) - f(y))(g(x) - g(y)) p_t(x,y) \mu(dx) \mu(dy).$$

It is well known (see [9]) that $\lim_{t\to 0} \mathcal{E}_t(f, g) = \mathcal{E}(f, g)$ for all $f, g \in \mathcal{F}$. Let A, B be disjoint compact sets and take $f, g \in \mathcal{F}$ such that Supp$f \subset A$ and Supp$g \subset B$. Then

$$\mathcal{E}_t(f,g) = -\frac{1}{t} \int_A \int_B f(x) g(y) p_t(x,y) \mu(dy) \mu(dx) \overset{t\to 0}{\to} -\int_A \int_B f(x) g(y) n(dx, dy).$$

Using $(UHKP)$, we obtain

$$\int_A \int_B f(x) g(y) n(dx, dy) \le C \int_A \int_B \frac{f(x) g(y)}{d(x,y)^{\alpha+\beta}} \mu(dy) \mu(dx),$$

for all $f, g \in \mathcal{F}$ such that Supp$f \subset A$ and Supp$g \subset B$. Since A, B are arbitrary disjoint compact sets, we see that $n(dx, dy)$ is absolutely continuous w.r.t. $\mu(dx)\mu(dy)$ and (UJ) holds μ-a.e. for $x, y \in M$. We thus obtain (c). □

Proof of $(a)+(c) \Rightarrow (b)$**.** We first prove (1.7). By taking $C \geq 1$, it is enough to prove it for $t < r^\beta$. Using (*UHKP*), (1.2), and the stochastic completeness of X, we have

$$\mathbb{P}^x(d(X_t,x) \geq r) = \int_{B(x,r)^c} p_t(x,z)\mu(dz) \leq ct \int_{B(x,r)^c} \frac{\mu(dy)}{d(x,y)^{\alpha+\beta}} \leq c\frac{t}{r^\beta}. \tag{2.17}$$

By (2.17) and the strong Markov property of $\{X_t\}$ at time τ,

$$\begin{aligned} \mathbb{P}^x(\tau \leq t) &\leq \mathbb{P}^x(\tau \leq t \text{ and } d(X_{2t},x) \leq r/2) + \mathbb{P}^x(d(X_{2t},x) > r/2) \\ &\leq \mathbb{P}^x(\tau \leq t \text{ and } d(X_{2t},X_\tau) \geq r/2) + c_1 t/r^\beta \\ &= \mathbb{P}^x(1_{\tau \leq t}\mathbb{P}^{X_\tau}(d(X_{2t-\tau},X_0) \geq r/2)) + c_1 t/r^\beta \\ &\leq \sup_{y \in B(x,r)^c} \sup_{s \leq t} \mathbb{P}^y(d(X_{2t-s},y) \geq r/2) + c_1 t/r^\beta \\ &\leq c_2 t/r^\beta. \end{aligned} \tag{2.18}$$

Here in the second and the last lines, we used (2.17). The stochastic completeness is used in the first line of the calculation; without it we would have to add a third term $\mathbb{P}^x(\zeta \leq t)$ to (2.18).

Next we prove (1.8). If $r/2 \leq R \leq r$ then using (a) we obtain, for all $x' \in B(x,r/2)$ and $y \notin B(x,2r)$,

$$\mathbb{P}^{x'}(\tau_{B(x,r)} \leq t,\ X_\tau \in B(y,R)) \leq \mathbb{P}^{x'}(\tau_{B(x',r/2)} \leq t) \leq C\frac{t}{(r/2)^\beta} \leq C\frac{tR^\alpha}{r^{\alpha+\beta}}.$$

Assume now $R < r/2$ so that the distance between the balls $B(x,r)$ and $B(y,R)$ is at least $r/2$. Applying Lemma 2.3 with the function

$$f(s,\xi,\eta) = 1_{(0,t]}(s)1_{\overline{B(x,r)}}(\xi)1_{B(y,R)}(\eta)$$

and noticing that $f(s,X_{s-},X_s)$ can be equal to 1 for $s \leq \tau$ only when $s=\tau$, we obtain

$$\mathbb{P}^{x'}(\tau \leq t,\ X_\tau \in B(y,R)) = \mathbb{E}^{x'}\big[\sum_{s \leq \tau} f(s,X_{s-},X_s)\big] = \mathbb{E}^{x'}\left[\int_0^{\tau\wedge t}\int_{B(y,R)} n(X_s,z)\mu(dz)ds\right].$$

Noticing that $X_s \in B(x,r)$, $z \in B(y,R)$ and using (UJ) and (1.2), we obtain

$$\mathbb{P}^{x'}(\tau \leq t,\ X_\tau \in B(y,R)) \leq \mathbb{E}^{x'}\left[\int_0^{\tau\wedge t}\int_{B(y,R)} \frac{C}{d(X_s,z)^{\alpha+\beta}}\mu(dz)ds\right] \leq C\frac{tR^\alpha}{r^{\alpha+\beta}},$$

which finishes the proof. □

2.4 Proof of Corollary 1.3

Since (a) and (b) are equivalent by Theorem 1.2, we can assume that both (a) and (b) are satisfied. By (1.7), we have

$$1 - C\frac{t}{r^\beta} \leq \mathbb{P}^x(\tau_{B(x,r)} > t) \leq \int_{B(x,r)} p_t(x,y)\mu(dy), \tag{2.19}$$

for all $t > 0, r \geq 0$, and $x \in M$. Taking $t = \varepsilon r^\beta$ in (2.19) where $\varepsilon > 0$ is so small that $1 - \varepsilon C > \frac{1}{2}$, and using (0.5), we obtain

$$\frac{1}{2} < \int_{B(x,r)} p_{\varepsilon r^\beta}(x,y)\mu(dy) \leq C\frac{\mu(B(x,r))}{r^\alpha},$$

whence (1.9) follows.

3 Obtaining upper bounds from the jump kernel

3.1 Splitting the jump kernel

We use the following construction of Meyer [16] for jump processes. Let $n(x,y) = n'(x,y) + n''(x,y)$, and suppose there exists C_1 such that

$$N(x) = \int n''(x,y)\mu(dy) \le C_1 \text{ for all } x.$$

Let $(Y_t, t \ge 0)$ be a process corresponding to the jump kernel n'. Then we can construct a process X corresponding to the jump kernel n by the following procedure. Let ξ_i, $i \ge 1$, be i.i.d. exponential random variables of parameter 1 independent of Y. Set

$$H_t = \int_0^t N(Y_s)ds, \quad T_1 = \inf\{t \ge 0 : H_t \ge \xi_1\},$$

and

$$q(x,y) = \frac{n''(x,y)}{N(x)}. \tag{3.1}$$

We remark that Y is a.s. continuous at T_1. We let $X_t = Y_t$ for $0 \le t < T_1$, and then define X_{T_1} with law $q(X_{T_1-},\cdot) = q(Y_{T_1},\cdot)$. (More formally, X_{T_1} should be defined as a function of X_{T_1-} and a random variable η_1 which is independent of ξ_i and Y). The construction now proceeds in the same way from the new space-time starting point (T_1, X_{T_1}). Since N is bounded, there can be (a.s.) only finitely many extra jumps added in any bounded time interval. In [16] (see also [14]) it is proved that the resulting process corresponds to the jump kernel n.

Now let

$$r_t(x,y) = \int q(x,z)p_t(z,y)\mu(dz). \tag{3.2}$$

The density $r_s(x,y)$ corresponds to first jumping according the law $q(x,\cdot)$ and then running the process X for time t.

Let $\mathcal{F}_t^Y = \sigma(Y_s, s \in [0,t])$, and write $p_t^Y(x,y)$ for the transition density of Y.

Lemma 3.1 *Let $n = n' + n''$, X and Y be as above.*
(a)

$$\mathbb{E}^x(f(T_1)|\mathcal{F}_\infty^Y) = \int_0^\infty f(t)e^{-H_t}N(Y_t)dt.$$

(b) For any Borel set B

$$\mathbb{P}^x(X_t \in B) = \mathbb{P}^x(Y_t \in B, T_1 > t) + \mathbb{E}^x \int_0^t \int_B r_{t-s}(Y_s,z)N(Y_s)\mu(dz)ds. \tag{3.3}$$

(c) If $\|n''\|_\infty < \infty$ then

$$p_t(x,y) \le p_t^Y(x,y) + t\|n''\|_\infty \qquad \text{for } \mu\text{-a.a. } y \in M. \tag{3.4}$$

Proof. Write $T = T_1$.
(a) Since ξ_1 is independent of $\mathcal{F}_\infty^Y$ we have

$$\mathbb{P}^x(T > t|\mathcal{F}_\infty^Y) = e^{-H_t}.$$

So the density of T conditional on $\mathcal{F}_\infty^Y$ is $e^{-H_t}N(Y_t)$ and the first assertion is clear.

(b) Since $X = Y$ on $[0, T)$ we have

$$\mathbb{P}^x(X_t \in B) = \mathbb{P}^x(Y_t \in B, T > t) + \mathbb{P}^x(X_t \in B, T \leq t).$$

Let $r_t(x, B) = \int_B r_t(x, y)\mu(dy)$. Then by the construction of X

$$\mathbb{P}^x(X_{T+t} \in B | X_{T-}) = r_t(Y_T, B).$$

So

$$\mathbb{P}^x(X_t \in B | \mathcal{F}^Y_\infty) = 1_{\{X_t \in B\}} e^{-H_t} + \int_0^t r_{t-s}(Y_s, B) N(Y_s) ds.$$

Taking expectations now gives (3.3).
(c) Since (3.3) holds for any Borel set B, and every $x \in M$, we obtain

$$p_t(x, y) \leq p^Y_t(x, y) + \mathbb{E}^x \int_0^t r_{t-s}(Y_s, y) N(Y_s) ds \quad \text{for } \mu\text{-a.a.} y. \tag{3.5}$$

Now as $N(x)q(x, y) = n''(x, y)$,

$$\begin{aligned} N(x) r_s(x, z) &= \int n''(x, y) p_s(y, z) \mu(dy) \\ &\leq \|n''\|_\infty \int p_s(y, z) \mu(dy) = \|n''\|_\infty. \end{aligned}$$

This bounds the second term in (3.5) by $t\|n''\|_\infty$, proving (c). □

3.2 Proof of Theorem 1.4

First note that (UJ) (with $\beta < 2$) and (1.2) gives:

$$\begin{aligned} \int_{B(x,r)^c} n(x, y)\mu(dy) &\leq \sum_{n=1}^\infty \int_{B(x,2^n r) - B(x, 2^{n-1} r)} n(x, y) \mu(dy) \\ &\leq \sum_{n=1}^\infty c(2^n r)^\alpha (2^{(n-1)r})^{-\alpha-\beta} \leq c r^{-\beta}. \end{aligned}$$

Similarly we have

$$\int_{B(x,r)} d(x, y)^2 n(x, y) \mu(dy) \quad \leq \quad c r^{2-\beta}.$$

Let $K > 0$ and let

$$n_K(x, y) = n(x, y) 1_{\{d(x,y) \leq K\}}.$$

Let $n'' = n - n_K$, and let $q(x, y)$, $r_t(x, y)$ be given by (3.1), (3.2). We write $\mathcal{E}_K$, $p^{(K)}$ etc. for quantities associated with n_K. We have

$$\begin{aligned} \mathcal{E}(f, f) - \mathcal{E}_K(f, f) &= \int\int 1_{(d(x,y) > K)} (f(x) - f(y))^2 n(x, y) \mu(dx) \mu(dy) \\ &\leq \int\int 1_{(d(x,y) > K)} 4 f(x)^2 n(x, y) \mu(dx) \mu(dy) \\ &\leq \int f(x)^2 \mu(dx) \sup_x \int_{B(x,K)^c} n(x, y) \mu(dy) \\ &\leq c||f||_2^2 K^{-\beta}. \end{aligned}$$

Hence from (N) one gets

$$||f||_2^{2+(2\beta/\alpha)} \le C(\mathcal{E}_K(f,f) + cK^{-\beta}||f||_2^2)||f||_1^{2\beta/\alpha}. \tag{3.6}$$

So by Theorem 3.25 in [4] and by the assumption $\mathrm{Lip}_0 \subset \mathcal{F}$,

$$p_t^{(K)}(x,y) \le ct^{-\alpha/\beta} e^{c_1 tK^{-\beta} - E_K(2t,x,y)} \qquad \text{for } \mu\text{-a.a. } x,y \in M. \tag{3.7}$$

Here $E_K(2t,x,y)$ is given by the following:

$$\begin{aligned}
\Gamma_K(\psi)(x) &= \int (e^{\psi(x)-\psi(y)} - 1)^2 n_K(x,y)dy, \\
\Lambda(\psi)^2 &= \|\Gamma_K(\psi)\|_\infty \vee \|\Gamma_K(-\psi)\|_\infty, \\
E_K(t,x,y) &= \sup\{|\psi(x) - \psi(y)| - t\Lambda(\psi)^2 : \ \psi \in \mathrm{Lip}_0 \text{ with } \Lambda(\psi) < \infty\}.
\end{aligned}$$

Let $H_t \subset M \times M$ be a set such that $(\mu \times \mu)((M \times M) \setminus H_t) = 0$ and (UJ), (3.7) hold for $(x,y) \in H_t$. Fix $x_0, y_0 \in H_t$ with $d(x_0,y_0) = R$ and let $t > 0$. Let $K = R/\theta$, where $\theta = 3(\beta+\alpha)/\beta$. If $t \ge K^\beta$ then (*UHKP*) is immediate, so we will assume that $t < K^\beta$. Let

$$\psi(x) = \lambda(R - d(x_0,x))_+.$$

So $|\psi(x) - \psi(y)| \le \lambda d(x,y)$. Note that $|e^t - 1|^2 \le t^2 e^{2|t|}$. Hence

$$\begin{aligned}
\Gamma_K(e^\psi)(x) &= \int (e^{\psi(x)-\psi(y)} - 1)^2 n_K(x,y)dy \\
&\le e^{2\lambda K}\lambda^2 \int d(x,y)^2 n_K(x,y)dy \\
&\le c(\lambda K)^2 e^{2\lambda K} K^{-\beta} \le ce^{3\lambda K} K^{-\beta}.
\end{aligned}$$

So we have

$$-E_K(2t,x_0,y_0) \le -\lambda R + c_1 te^{3\lambda K} K^{-\beta}. \tag{3.8}$$

Set

$$\lambda = \frac{1}{3K}\log(\frac{K^\beta}{t})$$

Then

$$\begin{aligned}
-E_K(2t,x_0,y_0) &\le -\frac{R}{3K}\log(\frac{K^\beta}{t}) + c_1 tK^{-\beta}(\frac{K^\beta}{t}) \\
&= c_1 - (\frac{\alpha+\beta}{\beta})\log(\frac{K^\beta}{t}).
\end{aligned}$$

So,

$$\begin{aligned}
p_t^{(K)}(x_0,y_0) &\le ct^{-\alpha/\beta} e^{c_1 tK^{-\beta} - E_K(2t,x_0,y_0)} \\
&\le c' t^{-\alpha/\beta}(\frac{t}{K^\beta})^{(\beta+\alpha)/\beta} = c'\frac{t}{K^{\beta+\alpha}} = c''\frac{t}{R^{\beta+\alpha}}.
\end{aligned} \tag{3.9}$$

Since by (UJ) $n''(x,y) \le cK^{-\beta-\alpha}$, by (3.4) we obtain

$$p_t(x_0,y_0) \le ctR^{-\beta-\alpha} + c'tK^{-\beta-\alpha} \le ctR^{-\beta-\alpha},$$

which gives the proof of (*UHKP*) for $(x_0,y_0) \in H_t$. Since the right hand side of (*UHKP*) is a continuous function on $M \times M$, by Lemma 2.2, we obtain (*UHKP*) for all $x_0, y_0 \in M$.

3.3 Stochastic completeness

In this subsection, we note that under a stronger assumption on the space (M_0, d, μ), we can prove the stochastic completeness from $(H2)$ and (UJ).

Proof of Theorem 1.5. For a symmetric measurable function $J(\cdot,\cdot)$, let

$$\begin{aligned}\mathcal{E}^J(f,f) &= \int_M\int_M (f(x)-f(y))^2 J(x,y)\mu(dx)\mu(dy),\\ \mathcal{F}^J &= \overline{\{f\in C(M):\mathcal{E}^J(f,f)<\infty\}}^{\mathcal{E}_1^J},\end{aligned}$$

where $\mathcal{E}_1^J(u,u) := \mathcal{E}^J(u,u)+\int_M u(x)^2\mu(dx)$. Define $J_*(x,y) = d(x,y)^{-\alpha-\beta}$. Then, under the above assumption for (M_0,d,μ), the results in [6] imply that $(\mathcal{E}^{J_*},\mathcal{F}^{J_*})$ is a regular Dirichlet form and it is stochastically complete. Denote the corresponding process as Y. For each $\delta>0$, let $J_\delta^1(x,y) = J_*(x,y)1_{\{d(x,y)<\delta\}}$, and define $J_\delta^2(x,y) = J_\delta^1(x,y)+n(x,y)1_{\{d(x,y)\geq\delta\}}$. Then, for $i=1,2$, $(\mathcal{E}^{J_\delta^i},\mathcal{F}^{J_\delta^i})$ is a regular Dirichlet form; denote the corresponding process as $Y^{i,\delta}$. Using (UJ), we have for every $x\in M$

$$\begin{aligned}\int_M (J_\delta^2(x,y)-J_\delta^1(x,y))\mu(dy) &\leq \int_M (J_*(x,y)-J_\delta^1(x,y))\mu(dy)\\ &= \int_{\{y\in M:\, d(x,y)\geq\delta\}} d(x,y)^{-\alpha-\beta}\mu(dy)\leq c_\delta<\infty.\end{aligned}$$

Thus we see that $Y^{1,\delta}, Y^{2,\delta}$ are stochastically complete. This is because the process Y and $Y^{2,\delta}$ can be obtained from $Y^{1,\delta}$ through Meyer's construction as discussed in §3.1, and therefore the stochastic completeness of Y implies that of $Y^{1,\delta}$, and then that of $Y^{2,\delta}$. Moreover, since $\mathcal{E}^{J_\delta^2}$ is larger than or equal to $\mathcal{E}$ (due to (UJ)), by $(H2)$ we have

$$p_t^{Y^{2,\delta}}(x,y)\leq c_1 t^{-\alpha/\beta}\qquad \text{for all } x,y\in M, t>0,$$

where $c_1>0$ is independent of δ. Here $p_t^{Y^{2,\delta}}(x,y)$ is the heat kernel of $Y^{2,\delta}$. Then, by Theorem 3.25 in [4] (as in §3.2 up to (3.8)),

$$p_t^{Y^{2,\delta}}(x,y)\leq c_2 t^{-\alpha/\beta}\exp(-c_3 d(x,y))\qquad \text{for all } x,y\in M, t\in(0,1], \tag{3.10}$$

where $c_2,c_3>0$ are independent of δ. Let $\{P_t^{Y^{2,\delta}}\}_t$ be the transition semigroup of $Y^{2,\delta}$ and let $x_0\in M$ be fixed. By (3.10), for each $\varepsilon>0$, there exists R_ε such that

$$P_1^{Y^{2,\delta}}1_{B(x_0,r)^c}(x) = \int_{B(x_0,r)^c} p_1^{Y^{2,\delta}}(x,y)\mu(dy)<\varepsilon\qquad \text{for all } x\in B(x_0,1), r\geq R_\varepsilon, \delta>0.$$

By Theorem 4.3 in [2], $Y^{2,\delta}$ converges to X in the Mosco sense as $\delta\to 0$. This implies (see [2, Proposition 4.2]) that for each $r>0$, $P_1^{Y^{2,\delta}}1_{B(x_0,r)}$ converges in L^2 to $P_1 1_{B(x_0,r)}$, which implies

$$P_1 1_{B(x_0,r)}(x)\geq 1-\varepsilon\qquad \text{for all } r\geq R_\varepsilon, x\in B(x_0,1).$$

Since $\varepsilon>0$ and $x_0\in M$ are arbitrary, we have $P_1 1 = 1$, which proves the stochastic completeness of X. □

Remark. Instead of assuming d to be a geodesic metric, a weaker assumption [6, (1.1)] suffices. See §4.6 in [6].

Acknowledgments. The authors are grateful to Zhenqing Chen for valuable comments concerning the quasi everywhere existence of the transition density and stochastic completeness. The first and second authors gratefully acknowledge the financial support of RIMS Kyoto, where this work was initiated.

References

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

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

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

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

[6] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for jump processes of mixed types on metric measure spaces. Probab. Theory Relat. Fields, to appear.

[7] T. Coulhon. Ultracontractivity and Nash type inequalities. J. Funct. Anal. **141** (1996) 510-539.

[8] E.B. Davies. Explicit constants for Gaussian upper bounds on heat kernels. Amer. J. Math. **109** (1987) 319-334.

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

[10] A. Grigor'yan. Gaussian upper bounds for the heat kernel on arbitrary manifolds. J. Diff. Geom. **45** (1997) 33-52.

[11] A. Grigor'yan. Heat kernel upper bounds in fractal spaces, preprint, 2004.

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

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

[14] N. Ikeda, N. Nagasawa, and S. Watanabe. A construction of Markov processes by piecing out. Proc. Japan Acad. **42** (1966), 370–375.

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

[16] P.-A. Meyer. Renaissance, recollements, mélanges, ralentissement de processus de Markov. Ann. Inst. Fourier **25** (3–4) (1975), 464–497.

[17] V.I. Ushakov. Stabilization of solutions of the third mixed problem for a second order parabolic equation in a non-cylindric domain. Math. USSR Sb. **39** (1981) 87-105.

Martin T. Barlow
Department of Mathematics
University of British Columbia
Vancouver, V6T 1Z2, Canada
E-mail: barlow@math.ubc.ca

Alexander Grigor'yan
Fakultät für Mathematik
Universität Bielefeld
D-33501 Bielefeld, Germany
E-mail: grigor@math.uni-bielefeld.de

Takashi Kumagai
Department of Mathematics
Kyoto University
Kyoto 606-8502, Japan
E-mail: kumagai@math.kyoto-u.ac.jp