Rev. Mat. Iberoamericana **26** (2010), no. 2, 551–589

A priori Hölder estimate, parabolic Harnack principle and heat kernel estimates for diffusions with jumps

Zhen-Qing Chen and Takashi Kumagai

Abstract

In this paper, we consider the following type of non-local (pseudo-differential) operators $\mathcal{L}$ on $\mathbb{R}^d$:

$$\mathcal{L}u(x) = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\Big(a_{ij}(x)\frac{\partial u(x)}{\partial x_j}\Big) + \lim_{\varepsilon\downarrow 0}\int_{\{y\in\mathbb{R}^d:\,|y-x|>\varepsilon\}} (u(y)-u(x))J(x,y)dy,$$

where $A(x)=(a_{ij}(x))_{1\le i,j\le d}$ is a measurable $d\times d$ matrix-valued function on $\mathbb{R}^d$ that is uniformly elliptic and bounded and J is a symmetric measurable non-trivial non-negative kernel on $\mathbb{R}^d\times\mathbb{R}^d$ satisfying certain conditions. Corresponding to $\mathcal{L}$ is a symmetric strong Markov process X on $\mathbb{R}^d$ that has both the diffusion component and pure jump component. We establish a priori Hölder estimate for bounded parabolic functions of $\mathcal{L}$ and parabolic Harnack principle for positive parabolic functions of $\mathcal{L}$. Moreover, two-sided sharp heat kernel estimates are derived for such operator $\mathcal{L}$ and jump-diffusion X. In particular, our results apply to the mixture of symmetric diffusion of uniformly elliptic divergence form operator and mixed stable-like processes on $\mathbb{R}^d$. To establish these results, we employ methods from both probability theory and analysis.

2000 Mathematics Subject Classification: Primary 60J35, 47G30, 60J45; Secondary 31C05, 31C25, 60J75.

Keywords: Symmetric Markov process, pseudo-differential operator, diffusion process, jump process, Lévy system, hitting probability, parabolic function, a priori Hölder estimate, parabolic Harnack inequality, transition density, heat kernel estimates.

1. Introduction

It is well-known that there is an intimate interplay between self-adjoint pseudo-differential operators on $\mathbb{R}^d$ and symmetric strong Markov processes on $\mathbb{R}^d$. For a large class of self-adjoint pseudo-differential operators $\mathcal{L}$ on $\mathbb{R}^d$ that enjoys maximum property, there is a jump-diffusion X on $\mathbb{R}^d$ associated with it so that $\mathcal{L}$ is the infinitesimal generator of X, and vice versa. The connection between $\mathcal{L}$ and X can also be seen as follows. The fundamental solution (also called heat kernel) for $\mathcal{L}$ is the transition density function of X. In this paper, we are interested in the a priori Hölder estimate for harmonic functions of such operator $\mathcal{L}$, parabolic Harnack principle and the sharp estimates on the heat kernel of $\mathcal{L}$.

Throughout this paper, $d \geq 1$ is an integer. Denote by m_d the d-dimensional Lebesgue measure in $\mathbb{R}^d$, and $C_c^1(\mathbb{R}^d)$ the space of C^1-functions on $\mathbb{R}^d$ with compact support. We consider the following type of non-local (pseudo-differential) operators $\mathcal{L}$ on $\mathbb{R}^d$:

$$(1.1) \qquad \mathcal{L}u(x) = \frac{1}{2}\sum_{i,j=1}^d \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u(x)}{\partial x_j}\right) + \lim_{\varepsilon \downarrow 0}\int_{\{y\in\mathbb{R}^d:\,|y-x|>\varepsilon\}} (u(y)-u(x))J(x,y)dy,$$

where $A(x) = (a_{ij}(x))_{1\leq i,j\leq d}$ is a measurable $d\times d$ matrix-valued function on $\mathbb{R}^d$ that is uniform elliptic and bounded in the sense that there exists a constant $c \geq 1$ such that

$$(1.2) \quad c^{-1}\sum_{i=1}^d \xi_i^2 \leq \sum_{i,j=1}^d a_{ij}(x)\xi_i\xi_j \leq c\sum_{i=1}^d \xi_i^2 \quad \text{for every } x, (\xi_1,\ldots,\xi_d)\in\mathbb{R}^d,$$

and J is a symmetric non-negative measurable kernel on $\mathbb{R}^d \times \mathbb{R}^d$ such that there are positive constants $\kappa_0 > 0$, and $\beta \in (0,2)$ so that

$$(1.3) \qquad J(x,y) \leq \kappa_0|x-y|^{-d-\beta} \qquad \text{for } |x-y| \leq \delta_0,$$

and that

$$(1.4) \qquad \sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d} (|x-y|^2\wedge 1)J(x,y)\,dy < \infty.$$

Clearly under condition (1.3), condition (1.4) is equivalent to

$$\sup_{x\in\mathbb{R}^d}\int_{\{y\in\mathbb{R}^d:|y-x|\geq 1\}} J(x,y)\,dy < \infty.$$

Associated with such a non-local operator $\mathcal{L}$ is an $\mathbb{R}^d$-valued symmetric strong Markov process X whose associated Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(\mathbb{R}^d; m_d)$ is given by

$$(1.5)\qquad \begin{cases} \mathcal{E}(u,v) = \dfrac{1}{2}\displaystyle\int_{\mathbb{R}^d} \nabla u(x)\cdot A(x)\nabla v(x)dx \\ \qquad\qquad\qquad + \displaystyle\int_{\mathbb{R}^d} (u(x)-u(y))(v(x)-v(y))J(x,y)dxdy, \\ \mathcal{F} = \overline{C_c^1(\mathbb{R}^d)}^{\mathcal{E}_1}, \end{cases}$$

where for $\alpha > 0$, $\mathcal{E}_\alpha(u,v) := \mathcal{E}(u,v) + \alpha \int_{\mathbb{R}^d} u(x)v(x)m_d(dx)$.

When the jumping kernel $J \equiv 0$ in (1.1) and (1.5), $\mathcal{L}$ is a uniform elliptic operator of divergence form and X is a symmetric diffusion on $\mathbb{R}^d$. It is well-known that X has a joint Hölder continuous transition density function $p(t,x,y)$, which enjoys the following celebrated Aronson's two-sided heat kernel estimate: there are constants $c_k > 0$, $k = 1, \dots, 4$, so that

$$c_1\, p^c(t, c_2|x-y|) \le p(t,x,y) \le c_3\, p^c(t, c_4|x-y|) \qquad \text{for } t > 0, x, y \in \mathbb{R}^d.$$

Here

$$(1.6)\qquad p^c(t,r) := t^{-d/2}\exp(-r^2/t).$$

It is also known that parabolic Harnack principle holds for such $\mathcal{L}$ and that every bounded parabolic function of $\mathcal{L}$ is locally Hölder continuous. See [15] for some history and a survey on this subject, where a mixture of analytic and probabilistic method is presented.

Let ϕ be a strictly increasing continuous function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\phi(0) = 0$, and $\phi(1) = 1$ such that there are constants $c \ge 1$, $0 < \beta_1 \le \beta_2 < 2$ such that

$$(1.7)\qquad c^{-1}\Big(\frac{R}{r}\Big)^{\beta_1} \le \frac{\phi(R)}{\phi(r)} \le c\,\Big(\frac{R}{r}\Big)^{\beta_2} \qquad \text{for every } 0 < r < R < \infty,$$

and

$$(1.8)\qquad \int_0^r \frac{s}{\phi(s)}ds \le c\,\frac{r^2}{\phi(r)} \qquad \text{for every } r > 0.$$

Observe that condition (1.7) implies that

$$c^{-1}r^{\beta_1} \le \phi(r) \le cr^{\beta_2} \qquad \text{for } r \ge 1$$

and

$$c^{-1}r^{\beta_2} \le \phi(r) \le cr^{\beta_1} \qquad \text{for } r \in (0,1].$$

In the sequel, if f and g are two functions defined on a set D, $f \asymp g$ means that there exists $C > 0$ such that $C^{-1}f(x) \le g(x) \le C\,f(x)$ for all $x \in D$.

When $A(x) \equiv 0$ in (1.5) and J is given by

$$J(x,y) \asymp \frac{1}{|x-y|^d\,\phi(|x-y|)}, \tag{1.9}$$

where ϕ satisfies the conditions (1.7)-(1.8), the corresponding process X is a mixed stable-like process on $\mathbb{R}^d$ studied in [9]. A typical example of J satisfying condition (1.9) is

$$J(x,y) = \int_{\alpha_1}^{\alpha_2} \frac{c(\alpha,x,y)}{|x-y|^{d+\alpha}}\,\nu(d\alpha),$$

where ν is a probability measure on $[\alpha_1,\alpha_2] \subset (0,2)$ and $c(\alpha,x,y)$ is a symmetric function in x and y is bounded between two positive constants that are independent of $\alpha \in [\alpha_1,\alpha_2]$. Under the above condition, a priori Hölder estimate and parabolic Harnack principle are established in [9] for parabolic functions of X. Moreover, it is proved in [9] that X has a jointly continuous transition density function $p(t,x,y)$ and that it has the following two-sided sharp estimates: there are positive constants $0 < c_1 < c_2$ so that

$$c_1 p^j(t,|x-y|) \le p(t,x,y) \le c_2 p^j(t,|x-y|) \qquad \text{for } t>0, x,y \in \mathbb{R}^d,$$

where

$$p^j(t,r) := \left(\phi^{-1}(t)^{-d} \wedge \frac{t}{r^d\phi(r)}\right) \tag{1.10}$$

with ϕ^{-1} being the inverse function of ϕ. Here and in the sequel, for two real numbers a and b, $a \wedge b := \min\{a,b\}$ and $a \vee b := \max\{a,b\}$. We point out that, in contrast to the diffusions (or differential operator) case, heat kernel estimates for pure jump processes (or non-local integro-differential operators) have been studied only quite recently. See the introduction part of [9] for a brief account of some history.

In this paper, we consider the case where both A and J are non-trivial in (1.1) and (1.5). Clearly the corresponding operators and jump diffusions take up an important place both in theory and in applications. However there are very limited work in literature for this mixture case on the topics of this paper, see [5], [6] and [14] though. One of the difficulties in obtaining fine properties for such an operator $\mathcal{L}$ and process X is that it exhibits different scales: the diffusion part has Brownian scaling $r \mapsto r^2$ while the pure jump part has a different type of scaling. Nevertheless, there is a folklore which says that with the presence of the diffusion part corresponding to $\frac{1}{2}\sum_{i,j=1}^d \frac{\partial}{\partial x_i}\big(a_{ij}(x)\frac{\partial}{\partial x_j}\big)$, better results can be expected under weaker assumptions on the jumping kernel J as the diffusion part helps to smooth things out. Our investigation confirms such an intuition. In fact we can establish a priori Hölder estimate and parabolic Harnack inequality under weaker conditions than (1.9). We now present the main results of this paper. Let $W^{1,2}(\mathbb{R}^d)$ denote the Sobolev space of order $(1,2)$ on $\mathbb{R}^d$; that is, $W^{1,2}(\mathbb{R}^d) := \{f \in L^2(\mathbb{R}^d; m_d) : \nabla f \in L^2(\mathbb{R}^d; m_d)\}$. It is not difficult to show the following.

Proposition 1.1 *Under the conditions* (1.2)-(1.4)*, the domain of the Dirichlet form of* (1.5) *is characterized by*

$$\mathcal{F} = W^{1,2}(\mathbb{R}^d) = \left\{ f \in L^2(\mathbb{R}^d; m_d) : \mathcal{E}(f,f) < \infty \right\}.$$

Let X be the symmetric Hunt process on $\mathbb{R}^d$ associated with the regular Dirichlet form $(\mathcal{E}, \mathcal{F})$. It will be shown in Theorem 2.2 below that X has infinite lifetime. Let $Z = \{Z_t := (V_0 - t, X_t), t \geq 0\}$ denote the space-time process of X. We say that a non-negative real valued Borel measurable function $h(t,x)$ on $[0,\infty)\times\mathbb{R}^d$ is *parabolic* (or *caloric*) on $D = (a,b)\times B(x_0,r)$ if there is a properly exceptional set $\mathcal{N} \subset \mathbb{R}^d$ such that for every relatively compact open subset D_1 of D,

$$h(t,x) = \mathbb{E}^{(t,x)}[h(Z_{\tau_{D_1}})]$$

for every $(t,x) \in D_1 \cap ([0,\infty)\times(\mathbb{R}^d \setminus \mathcal{N}))$, where $\tau_{D_1} = \inf\{s > 0 : Z_s \notin D_1\}$. We remark that in [8, 9] the space-time process is defined to be $(V_0 + t, X_t)$ but this is merely a notational difference. In this paper, we first show that any parabolic function of X is Hölder continuous. Recall that δ_0 is the positive constant in condition (1.3).

Theorem 1.2 *Assume that the Dirichlet form* $(\mathcal{E}, \mathcal{F})$ *given by* (1.5) *satisfies the conditions* (1.2)-(1.4) *and that for every* $0 < r < \delta_0$*,*

$$\inf_{\substack{x_0, y_0 \in \mathbb{R}^d \\ |x_0 - y_0| = r}} \; \inf_{x \in B(x_0, r/16)} \int_{B(y_0, r/16)} J(x,z)dz > 0. \tag{1.11}$$

Then for every $R_0 \in (0,1]$*, there are constants* $c = c(R_0) > 0$ *and* $\kappa > 0$ *such that for every* $0 < R \leq R_0$ *and every bounded parabolic function* h *in* $Q(0,x_0,2R) := (0,4R^2) \times B(x_0,2R)$*,*

$$|h(s,x) - h(t,y)| \leq c \, \|h\|_{\infty,R} \, R^{-\kappa} \left(|t-s|^{1/2} + |x-y| \right)^{\kappa} \tag{1.12}$$

holds for (s,x)*,* $(t,y) \in (R^2, 4R^2) \times B(x_0,R)$*, where*

$$\|h\|_{\infty,R} := \sup_{(t,y)\in[0,\,4R^2]\times\mathbb{R}^d\setminus\mathcal{N}} |h(t,y)|.$$

In particular, X has a jointly continuous transition density function $p(t,x,y)$ with respect to the Lebesgue measure.

Moreover, for every $t_0 \in (0,1)$ there are constants $c>0$ and $\kappa>0$ such that for any t, $s \in (t_0,\, 1]$ and $(x_i,y_i) \in \mathbb{R}^d \times \mathbb{R}^d$ with $i=1,2$,

$$\begin{aligned} &|p\,(s,x_1,y_1)-p(t,x_2,y_2)| \le \\ &\qquad\le c\, t_0^{-(d+\kappa)/2} \left(|t-s|^{1/2}+|x_1-x_2|+|y_1-y_2|\right)^{\kappa}. \end{aligned} \tag{1.13}$$

In addition to (1.2)-(1.4) and (1.11), if there is a constant $c>0$ such that

$$J(x,y) \le \frac{c}{r^d}\int_{B(x,r)} J(z,y)dz \qquad \text{whenever } r \le \tfrac{1}{2}|x-y| \wedge 1,\ x,y \in \mathbb{R}^d, \tag{1.14}$$

we show that the parabolic Harnack principle holds for non-negative parabolic functions of X. (Note that (1.14) was introduced in [3, 7] and it was denoted as $(\text{UJS})_{\le 1}$ there.)

Theorem 1.3 *Suppose that the Dirichlet form $(\mathcal{E},\mathcal{F})$ given by (1.5) satisfies the condition (1.2)-(1.4), (1.11) and (1.14). For every $\delta \in (0,1)$, there exist constants $c_1=c_1(\delta)$ and $c_2=c_2(\delta)>0$ such that for every $z \in \mathbb{R}^d$, $t_0 \ge 0$, $0<R\le c_1$ and every non-negative function u on $[0,\infty)\times\mathbb{R}^d$ that is parabolic on $(t_0, t_0+6\delta R^2)\times B(z,4R)$,*

$$\sup_{(t_1,y_1)\in Q_-} u(t_1,y_1) \le c_2 \inf_{(t_2,y_2)\in Q_+} u(t_2,y_2), \tag{1.15}$$

where

$$Q_- = (t_0+\delta R^2, t_0+2\delta R^2)\times B(x_0,R)$$

and

$$Q_+ = (t_0+3\delta R^2, t_0+4\delta R^2)\times B(x_0,R).$$

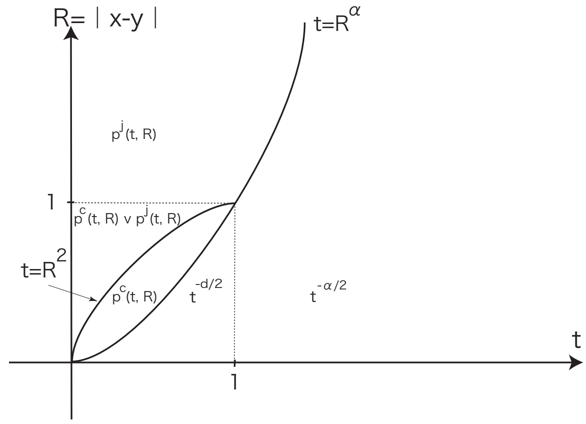
Note that elliptic versions of Theorem 1.2 and 1.3 are claimed in [12] under similar assumptions, however we have some difficulty to follow some of the arguments there. Clearly, our theorems imply the elliptic versions given in [12].

We next derive two-sided heat kernel estimate for X when $J(x,y)$ satisfies the condition (1.9). Clearly (1.3)-(1.4), (1.11) and (1.14) are satisfied when (1.9) holds. Recall that functions $p^c(t,x,y)$ and $p^j(t,x,y)$ are defined by (1.6) and (1.10), respectively.

Theorem 1.4 *Suppose that* (1.2) *holds and that the jumping kernel J of the Dirichlet form $(\mathcal{E}, \mathcal{F})$ given by* (1.5) *satisfies the condition* (1.9). *Denote by $p(t, x, y)$ the continuous transition density function of the symmetric Hunt process X associated with the regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ of* (1.5) *with the jumping kernel J given by* (1.9). *There are positive constants c_i, $i = 1, 2, 3, 4$ such that for every $t > 0$ and $x, y \in \mathbb{R}^d$,*

$$(1.16)\quad c_1 \left(t^{-d/2} \wedge \phi^{-1}(t)^{-d}\right) \wedge \left(p^c(t, c_2|x-y|) + p^j(t, |x-y|)\right) \le p(t, x, y) \\ \le c_3 \left(t^{-d/2} \wedge \phi^{-1}(t)^{-d}\right) \wedge \left(p^c(t, c_4|x-y|) + p^j(t, |x-y|)\right).$$

The following figure shows which term is the dominant term in each region when ϕ in (1.9) is given by $\phi(r) = r^\alpha$ with $0 < \alpha < 2$. It is worth mentioning that there is a short-time short-distance region in $t \le R^2 \le 1$ where the jump part is the dominant term.



When $A(x) \equiv I_{d\times d}$, the $d \times d$ identity matrix, and $J(x, y) = c|x - y|^{-d-\alpha}$ for some $\alpha \in (0, 2)$ in (1.5), that is, when X is the independent sum of a Brownian motion W on $\mathbb{R}^d$ and an isotropically symmetric α-stable process Y on $\mathbb{R}^d$, the transition density function $p(t, x, y)$ can be expressed as the convolution of the transition density functions of W and Y, whose two-sided estimates are known. In [14], heat kernel estimates for this Lévy process X are carried out by computing the convolution and the estimates are given in a form that depends on which region the point (t, x, y) falls into. Subsequently, the parabolic Harnack inequality (1.15) for such a Lévy process X is derived in [14] by using the two-sided Heat kernel estimate. Clearly such an approach is not applicable in our setting even when $\phi(r) = r^\alpha$, since in our case, the diffusion and jumping part of X are typically not independent. The two-sided estimate in this simple form of (1.16) is a new observation even in the independent sum of a Brownian motion and an isotropically symmetric α-stable process case considered in [14].

Our approach employs methods from both probability theory and analysis, but it is mainly probabilistic. It uses some ideas previously developed in [2, 4, 8, 9, 7]. To get a priori Hölder estimates for parabolic functions of X, we establish the following three key ingredients.

(i) Exit time upper bound estimate (Lemma 2.3):

$$\mathbb{E}_x[\tau_{B(x_0,r)}] \leq c_1 r^2 \qquad \text{for } x \in B(x_0, r),$$

where $\tau_{B(x_0,r)} := \inf\{t > 0 : X_t \notin B(x_0, r)\}$ is the first exit time from $B(x_0, r)$ by X.

(ii) Hitting probability estimate ((4.1) below):

$$\mathbb{P}_x \left(X_{\tau_{B(x,r)}} \notin B(x, s) \right) \leq \frac{c_2 r^2}{(s \wedge 1)^2} \qquad \text{for every } r \in (0, 1] \text{ and } s \geq 2r.$$

(iii) Hitting probability estimate for space-time process $Z_t = (V_0 - t, X_t)$ (Lemma 4.1): for every $x \in \mathbb{R}^d$, $r \in (0, 1]$ and any compact subset $A \subset Q(x, r) := (0, r^2) \times B(x, r)$,

$$\mathbb{P}^{(r^2,x)}(\sigma_A < \tau_r) \geq c_3 \frac{m_{d+1}(A)}{r^{d+2}},$$

where by slightly abusing the notation, $\sigma_A := \{t > 0 : Z_t \in A\}$ is the first hitting time of A, $\tau_r := \inf\{t > 0 : Z_t \notin Q(x, r)\}$ is the first exit time from $Q(x, r)$ by Z and m_{d+1} is the Lebesgue measure on $\mathbb{R}^{d+1}$.

Throughout this paper, we use the following notations. The probability law of the process X starting from x is denoted as $\mathbb{P}_x$ and the mathematical expectation under it is denoted as $\mathbb{E}_x$, while probability law of the space-time process $Z = (V, X)$ starting from (t, x), i.e. $(V_0, X_0) = (t, x)$, is denoted as $\mathbb{P}^{(t,x)}$ and the mathematical expectation under it is denoted as $\mathbb{E}^{(t,x)}$. To establish parabolic Harnack inequality, we need in addition the following.

(iv) Short time near-diagonal heat kernel estimate (Theorem 3.1): for every $t_0 > 0$, there is $c_4 = c_4(t_0) > 0$ such that for every $x_0 \in \mathbb{R}^d$ and $t \in (0, t_0]$,

$$p^{B(x_0,\sqrt{t})}(t, x, y) \geq c_4 t^{-d/2} \qquad \text{for } x, y \in B(x_0, \sqrt{t}/2).$$

Here $p^{B(x_0,\sqrt{t})}$ is the transition density function for the part process $X^{B(x_0,\sqrt{t})}$ of X killed upon leaving the ball $B(x_0, \sqrt{t})$.

(v) (Lemma 4.3): Let $R \le 1$ and $\delta < 1$. $Q_1 = [t_0 + 2\delta R^2/3,\, t_0 + 5\delta R^2] \times B(x_0, 3R/2)$, $Q_2 = [t_0 + \delta R^2/3,\, t_0 + 11\delta R^2/2] \times B(x_0, 2R)$ and define Q_- and Q_+ as in Theorem 1.3. Let $h : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}_+$ be bounded and supported in $[0,\infty) \times B(x_0, 3R)^c$. Then there exists $c_5 = c_5(\delta) > 0$ such that

$$\mathbb{E}^{(t_1,y_1)}[h(Z_{\tau_{Q_1}})] \le c_5 \mathbb{E}^{(t_2,y_2)}[h(Z_{\tau_{Q_2}})] \text{ for } (t_1,y_1) \in Q_- \text{ and } (t_2,y_2) \in Q_+.$$

The proof of (iv) uses ideas from [2], where a similar inequality is established for finite range pure jump process. However, some difficulties arise due to the presence of the diffusion part.

The upper bound heat kernel estimate in Theorem 1.4 is established by using method of scaling, by Meyer's construction of the process X based on finite range process $X^{(\lambda)}$, where the jumping kernel J is replaced by $J(x,y)\mathbb{1}_{\{|x-y|\le\lambda\}}$, and by Davies' method from [6] to derive an upper bound estimate for the transition density function of $X^{(\lambda)}$ through carefully chosen testing functions. Here we need to select the value of λ in a very careful way that depends on the values of t and $|x-y|$.

To get the lower bound heat kernel estimate in Theorem 1.4, we need a full scale parabolic Harnack principle that extends Theorem 1.3 to all $R > 0$ with the scale function $\widetilde{\phi}(R) := R^2 \wedge \phi(R)$ in place of $R \mapsto R^2$ there. To establish such a full scale parabolic Harnack principle, we show the following.

(iii') Strengthened version of (iii) (Lemma 6.5): for every $x \in \mathbb{R}^d$, $r > 0$ and any compact subset $A \subset Q(0,x,r) := [0, \gamma_0\widetilde{\phi}(r)] \times B(x,r)$,

$$\mathbb{P}^{(\gamma_0\widetilde{\phi}(r),x)}(\sigma_A < \tau_r) \ge c_3 \frac{m_{d+1}(A)}{r^d \widetilde{\phi}(r)}.$$

Here γ_0 denotes the constant $\gamma(1/2,1/2)$ in Proposition 6.2.

(vi) (Corollary 6.6): For every $\delta \in (0,\gamma_0]$, there is a constant $c_6 = c_6(\gamma)$ so that for every $0 < R \le 1$, $r \in (0, R/4]$ and $(t,x) \in Q(0,z,R/3)$ with $0 < t \le \gamma_0\widetilde{\phi}(R/3) - \delta\widetilde{\phi}(r)$,

$$\mathbb{P}^{(\gamma_0\widetilde{\phi}(R/3),z)}(\sigma_{U(t,x,r)} < \tau_{Q(0,z,R)}) \ge c_6 \frac{r^d \widetilde{\phi}(r)}{R^d \widetilde{\phi}(R)},$$

where $U(t,x,r) := \{t\} \times B(x,r)$.

With the full scale parabolic Harnack inequality, the lower bound heat kernel estimate can then be derived once the following estimate is obtained.

(*vii*) Tightness result (Proposition 6.3): there are constants $c_7 \geq 2$ and $c_8 > 0$ such that for every $t > 0$ and $x, y \in \mathbb{R}^d$ with $|x-y| \geq c_7\widetilde{\phi}(t)$,

$$\mathbb{P}_x\Big(X_t \in B(y, c_7\widetilde{\phi}^{-1}(t))\Big) \geq c_8 \frac{t(\widetilde{\phi}^{-1}(t))^d}{|x-y|^d\widetilde{\phi}(|x-y|)}.$$

Throughout the paper, we will define and use various Dirichlet forms, the corresponding processes and heat kernels. For the convenience of the reader, we list the notations here.

(Heat kernel)	(Process)	(Jump kernel)	(Dirichlet form)
$p(t,x,y)$	X	$J(x,y)$	$(\mathcal{E},\mathcal{F}) = (\mathcal{E}, W^{1,2}(\mathbb{R}^d))$
$p^B(t,x,y)$	X^B	$J(x,y)$	$(\mathcal{E},\mathcal{F}^B)$: X killed on exiting B
$p^{(\lambda)}(t,x,y)$	$X^{(\lambda)}$	$J(x,y)\mathbb{1}_{\{\lvert x-y\rvert\leq\lambda\}}$	$(\mathcal{E}^{(\lambda)}, W^{1,2}(\mathbb{R}^d))$
$p^{(\lambda;n)}(t,x,y)$	$X^{(\lambda;n)}$	$J(x,y)\mathbb{1}_{\{\lvert x-y\rvert\leq\lambda\}} \cdot\mathbb{1}_{B(n)\times B(n)}$	$(\mathcal{E}^{(\lambda;n)}, \mathcal{F}^{(\lambda;n)})$
$p_Y(t,x,y)$	Y	$\kappa(x,y)\lvert x-y\rvert^{-d-\beta}$	subordinated Dirichlet form $\dashrightarrow$ (A)
$q^\delta(t,x,y)$	Z^δ	$J_\delta(x,y)$ $\dashrightarrow$ (B)	$(\mathcal{E}^\delta, \mathcal{F}^\delta)$
$q^{\delta,B_r}(t,x,y)$	Z^{δ,B_r}	$J_\delta(x,y)$	$(\mathcal{E}^\delta, \mathcal{F}^{\delta,B_r})$: Z^δ killed on exiting B_r
$q_r^{\delta,B}(t,x,y)$	$r^{-1}Z^{\delta,B_r}_{r^2\cdot}$	$J_\delta^{\langle r\rangle}(x,y)$ $\dashrightarrow$ (C)	$(\mathcal{E}^{\langle r\rangle}, \mathcal{F}^{\langle r\rangle,B})$: $r^{-1}Z^\delta_{r^2\cdot}$ killed on exiting B
$p_r(t,x,y)$	$X^{\langle r\rangle}$	$J^{\langle r\rangle}(x,y)$ $\dashrightarrow$ (D)	$(\mathcal{E}^{\langle r\rangle}, \mathcal{F}^{\langle r\rangle}) = (\mathcal{E}^{\langle r\rangle}, W^{1,2}(\mathbb{R}^d))$
$p_r^{(\lambda)}(t,x,y)$	$X^{\langle r,\lambda\rangle}$	$J^{\langle r\rangle}(x,y)\mathbb{1}_{\{\lvert x-y\rvert\leq\lambda\}}$	$(\mathcal{E}^{\langle r,\lambda\rangle}, W^{1,2}(\mathbb{R}^d))$

where in the above,

(A) Y is the subordination of the symmetric diffusion for $\nabla(A\nabla)$, the local part of $\mathcal{E}$, by the subordinator $\eta = \{t + c_0\eta_t^{(1)}, t \geq 0\}$, where $\{\eta_t^{(1)}\}$ is a $(\beta/2)$-subordinator.

(B) $J_\delta(x,y) := J(x,y)\mathbb{1}_{\{|x-y|\geq\delta\}} + \kappa(x,y)|x-y|^{-d-\beta}\mathbb{1}_{\{|x-y|<\delta\}}$.

(C) $q_r^{\delta,B}(t,x,y) = q_r^B(t,x,y) := r^d q^{\delta,B_r}(r^2t, rx, ry)$,

$Z_t^{\langle r\rangle} := r^{-1}Z^\delta_{r^2t}$,

$J_\delta^{\langle r\rangle}(x,y) := r^{d+2}J_\delta(rx,ry)$ for $r \in (0,1]$.

(D) $p_r(t,x,y) := r^d p(\widetilde{\phi}(r)t, rx, ry)$,

$X_t^{\langle r\rangle} := r^{-1}X_{\widetilde{\phi}(r)t}$,

$J^{\langle r\rangle}(x,y) := \widetilde{\phi}(r)r^d J(rx,ry)$ for $r > 0$.

2. Heat kernel upper bound estimate and exit time estimate

Throughout this paper, we always assume the uniform elliptic condition (1.2) holds for the diffusion matrix A. Let $(\mathcal{E}, \mathcal{F})$ be the Dirichlet form in (1.5) with the jumping kernel J satisfying the conditions (1.3) and (1.4). We start this section by giving a

Proof of Proposition 1.1. For any $u \in C_0^1(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} \nabla u(x) \cdot A(x) \nabla u(x) dx + \|u\|_2^2 \asymp \int_{\mathbb{R}^d} |\nabla u(x)|^2 dx + \|u\|_2^2 =: \mathcal{C}_{1,c}(u,u),$$

and

$$\begin{aligned} (2.1) \qquad \int_{\mathbb{R}^d} &(u(x) - u(y))^2 J(x,y) dx dy \\ &\leq \int_{|x-y|\leq 1} (u(x) - u(y))^2 J(x,y) dx dy + c_1 \|u\|_2^2 \\ &\leq c_2 \Big(\int_{\mathbb{R}^d} \frac{(u(x) - u(y))^2}{|x-y|^{d+\beta}} dx dy + \|u\|_2^2 \Big) =: c_2 \mathcal{C}_{1,d}(u,u). \end{aligned}$$

Using Fourier transform, it is well-known that

$$\begin{aligned} (2.2) \qquad \mathcal{C}_{1,d}(u,u) &= c \int_{\mathbb{R}^d} (|\xi|^\beta + 1) |\widehat{u}(\xi)|^2 d\xi \\ &\leq 2c \int_{\mathbb{R}^d} (|\xi|^2 + 1) |\widehat{u}(\xi)|^2 d\xi = c_3 \mathcal{C}_{1,c}(u,u). \end{aligned}$$

Thus we have $\mathcal{E}(u,u) \asymp \mathcal{C}_{1,c}(u,u)$ for all $u \in C_0^1(\mathbb{R}^d)$. It follows then

$$\mathcal{F} = \overline{C_0^1(\mathbb{R}^d)}^{\mathcal{E}_1} = \overline{C_0^1(\mathbb{R}^d)}^{\mathcal{C}_{1,c}} = W^{1,2}(\mathbb{R}^d).$$

■

2.1. Preliminary heat kernel upper bound estimate

By the Nash's inequality

$$\begin{aligned} (2.3) \qquad \|f\|_2^{2+4/d} &\leq c_1 \int_{\mathbb{R}^d} |\nabla u(x)|^2 dx \cdot \|f\|_1^{4/d} \\ &\leq c_2 \mathcal{E}(f,f) \|f\|_1^{4/d} \quad \text{for } f \in W^{1,2}(\mathbb{R}^d), \end{aligned}$$

we have, by Theorem [6, Theorem 2.9] and [2, Theorem 3.1], that there is a properly $\mathcal{E}$-exceptional set $\mathcal{N} \subset \mathbb{R}^d$ of X and a positive symmetric kernel

$p(t,x,y)$ defined on $[0,\infty)\times(\mathbb{R}^d\backslash\mathcal{N})\times(\mathbb{R}^d\backslash\mathcal{N})$ such that for every $x\in\mathbb{R}^d\backslash\mathcal{N}$ and $t>0$,

$$\mathbb{E}_x\left[f(X_t)\right]=\int_{\mathbb{R}^d}p(t,x,y)f(y)m_d(dy),$$
$$p(t+s,x,y)=\int_{\mathbb{R}^d}p(t,x,z)p(s,z,y)\quad\text{for every } t,s>0 \text{ and } x,y\in\mathbb{R}^d\setminus\mathcal{N},$$

and

$$p(t,x,y)\le ct^{-d/2}\qquad\text{for } t>0 \text{ and every } x,y\in\mathbb{R}^d\setminus\mathcal{N}. \tag{2.4}$$

Moreover, there is an $\mathcal{E}$-nest $\{F_k,k\ge 1\}$ of compact subsets of $\mathbb{R}^d$ so that $\mathcal{N}=\mathbb{R}^d\setminus\cup_{k=1}^{\infty}F_k$ and that for every $t>0$ and $y\in\mathbb{R}^d\setminus\mathcal{N}$, $x\mapsto p(t,x,y)$ is continuous on each F_k. Later, as a consequence of the Hölder continuity result for parabolic functions, $p(t,x,y)$ in fact has a continuous version so the exceptional set $\mathcal{N}$ can be taken to be an empty set.

Now, for $\lambda\in\mathbb{Q}_+$, where $\mathbb{Q}_+$ is the set of positive rational numbers, let $(\mathcal{E}^{(\lambda)},W^{1,2}(\mathbb{R}^d))$ be the Dirichlet form defined by (1.5) but with the jumping kernel $J(x,y)\mathbb{1}_{\{|x-y|\le\lambda\}}$ in place of $J(x,y)$. Let $X^{(\lambda)}$ be the symmetric strong Markov process associated with $(\mathcal{E}^{(\lambda)},W^{1,2}(\mathbb{R}^d))$, and let $p^{(\lambda)}(t,x,y)$ be its transition density function.

Proposition 2.1 *Let*

$$\delta(\lambda):=\sup_{\xi\in\mathbb{R}^d}\int_{\{\eta\in\mathbb{R}^d:\,|\eta-\xi|\le\lambda\}}|\xi-\eta|^2J(\eta,\xi)d\eta.$$

Then, there exist $c_1,c_2>0$ (independent of $\lambda\in\mathbb{Q}_+$) such that for any $s>0$, the following holds for all $t>0$ and q.e. x,y,

$$p^{(\lambda)}(t,x,y)\le c_1t^{-d/2}\exp\left(-s|x-y|+c_2s^2\left(1+e^{2\lambda s}\delta(\lambda)\right)t\right). \tag{2.5}$$

Proof. First, note that by condition (1.3), we have

$$\lim_{\lambda\to 0}\delta(\lambda)=0. \tag{2.6}$$

We use Davies' method to derive the desired heat kernel upper bound. From Nash's inequality (2.3), by the same reasoning as that for X at the beginning of this section, the symmetric process $X^{(\lambda)}$ has a quasi-continuous transition density function $p^{(\lambda)}(t,x,y)$ defined on $[0,\infty)\times(\mathbb{R}^d\setminus\mathcal{N}_\lambda)\times(\mathbb{R}^d\setminus\mathcal{N}_\lambda)$ such that

$$p^{(\lambda)}(t,x,y)\le c_1\,t^{-d/2}\qquad\text{for every } t>0 \text{ and } x,y\in\mathbb{R}^d\setminus\mathcal{N}_\lambda. \tag{2.7}$$

Note that the above constant $c_1>0$ is independent of $\lambda>0$. By (2.2), we have $\mathcal{E}_1^{(\lambda)}(u,u)\asymp\mathcal{C}_{1,c}(u,u)\asymp\mathcal{E}_1(u,u)$, so a set is $\mathcal{E}_1^{(\lambda)}$-exceptional if and only if it is $\mathcal{E}_1$-exceptional.

Thus, letting $\mathcal{N} = \cup_{\lambda\in\mathbb{Q}_+}\mathcal{N}_\lambda$, $\mathcal{N}$ is a $\mathcal{E}_1$-exceptional set. (2.7) together with [6, Theorem 3.25] and [2, Theorem 3.2] implies that there exist constants $C > 0$ and $c > 0$, such that

$$p^{(\lambda)}(t,x,y) \le c_1\, t^{-d/2}\ \exp\left(-|\psi(y)-\psi(x)| + C\,\Lambda_\lambda(\psi)^2\, t\right) \tag{2.8}$$

for all $t > 0$, $x, y \in \mathbb{R}^d \setminus \mathcal{N}$, and for any function ψ having $\Lambda_\lambda(\psi) < \infty$. Here

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

where for $\xi \in \mathbb{R}^d$,
(2.9)

$$\Gamma_\lambda[v](\xi) := \sum_{i,j=1}^{d} a_{ij}(\xi)\frac{\partial v}{\partial x_i}(\xi)\frac{\partial v}{\partial x_j}(\xi) + \int_{\{\eta\in\mathbb{R}^d:\,|\eta-\xi|\le\lambda\}}(v(\eta)-v(\xi))^2 J(\eta,\xi)d\eta,$$

For $s > 0$, take

$$\psi(\xi) := s\ (|\xi - x| \wedge |x-y|) \qquad \text{for } \xi \in \mathbb{R}^d.$$

Note that $|\psi(\eta)-\psi(\xi)| \le s\,|\eta-\xi|$ for all $\xi,\eta \in \mathbb{R}^d$. So for $\xi \in \mathbb{R}^d$,

$$\begin{aligned} e^{-2\psi(\xi)}\Gamma_\lambda[e^{\psi}](\xi) &\le c_2|\nabla\psi(\xi)|^2 + \int_{|\eta-\xi|\le\lambda}(1-e^{\psi(\eta)-\psi(\xi)})^2 J(\eta,\xi)d\eta \\ &\le c_2 s^2 + \int_{|\eta-\xi|\le\lambda}(\psi(\eta)-\psi(\xi))^2\, e^{2|\psi(\eta)-\psi(\xi)|} J(\eta,\xi)d\eta \\ &\le c_2 s^2 + s^2\, e^{2\lambda s}\int_{|\eta-\xi|\le\lambda}|\eta-\xi|^2 J(\eta,\xi)d\eta \\ &\le c_2 s^2\left(1 + e^{2\lambda s}\delta(\lambda)\right). \end{aligned}$$

Here $c_2 > 0$ is independent of $\lambda \in \mathbb{Q}_+$. The same estimate holds for $e^{2\psi(\xi)}\Gamma_\lambda[e^{-\psi}](\xi)$. So we have the desired estimate. ■

2.2. Conservativeness

Theorem 2.2 *The process X is conservative; that is, X has infinite lifetime.*

Proof. Recall the process $X^{(\lambda)}$ defined in the previous subsection. X can be obtained from $X^{(\lambda)}$ through Meyer's construction by adding all the jumps whose size is larger than λ (see Remarks 3.4-3.5 of [2] and Lemma 3.1 of [4]). Note that by (1.3) and (1.4), there is a constant $b_0 > 0$ such that

$$\sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbb{1}_{\{|x-y|>\lambda\}}J(x,y)dy \le b_0\lambda^{-\beta} \qquad \text{for every } \lambda \in (0,1]. \tag{2.10}$$

Thus, it suffices to show that $X^{(\lambda)}$ is conservative. To show this, we look at reflected jump-diffusions with jumping kernel $J(x,y)\mathbb{1}_{\{|x-y|\le\lambda\}}$ in big balls, as in [9, Theorem 4.7]. In the following, we fix $\lambda\in\mathbb{Q}_+$. Let $x_0\in\mathbb{R}^d$, $r_n\ge 100\lambda$. Define $B(n)=\overline{B(x_0,r_n)}$ and

$$\mathcal{E}^{(\lambda;n)}(f,f)=\int_{B(n)}\nabla f(x)\cdot A(x)\nabla f(x)dx + \int_{B(n)}\int_{B(n)}(f(x)-f(y))^2 J(x,y)\mathbb{1}_{\{|x-y|\le\lambda\}}dxdy,$$

$$\mathcal{F}^{(\lambda;n)}=\overline{\{f\in C^1(\overline{B(n)}):\ \mathcal{E}^{(\lambda;n)}(f,f)<\infty\}}^{\mathcal{E}_1^{(\lambda;n)}},$$

where $\mathcal{E}_1^{(\lambda;n)}(u,u):=\mathcal{E}^{(\lambda;n)}(u,u)+\int_{B(n)}u(x)^2dx$. Clearly $(\mathcal{E}^{(\lambda;n)},\mathcal{F}^{(\lambda;n)})$ is a regular symmetric Dirichlet form on $L^2(\overline{B(n)};dx)$. Let $X^{(\lambda;n)}$ be the Hunt process on $\overline{B(n)}$ associated with $(\mathcal{E}^{(\lambda;n)},\mathcal{F}^{(\lambda;n)})$. Since a constant function $1\in\mathcal{F}^{(\lambda;n)}$ with $\mathcal{E}^{(\lambda;n)}(1,1)=0$, $X^{(\lambda;n)}$ is recurrent and so $X^{(\lambda;n)}$ is conservative. Let $p^{(\lambda;n)}(t,x,y)$ be the transition density function of $X^{(\lambda;n)}$. Then, similarly to the proof of Proposition 2.1, we see that $p^{(\lambda;n)}(t,x,y)$ exists for all $t>0$, $x,y\in B(n)\setminus\mathcal{N}_n$, where $\mathcal{N}_n$ is a properly exceptional set for $X^{(\lambda;n)}$, and moreover it enjoys the estimate (2.5) with constants independent of n. Using (2.5) with $s=1$, for $x\in B(n)\setminus\mathcal{N}_n$, $t\in[1,2]$ and $R\le r_n$, we have

$$\begin{aligned}\mathbb{P}_x\left(|X_s^{(\lambda;n)}-x|\ge R\right)&=\int_{B(n)\setminus B(x,R)}p^{(\lambda;n)}(t,x,y)dy\\&\le c_1\int_{B(n)\setminus B(x,R)}e^{-|x-y|}dy\le c_2e^{-R},\end{aligned}$$

where c_1,c_2 may depend on λ, but they are independent of n and R. Given this estimate, the rest is the same as that of [9, Theorem 4.7]. We will sketch the argument. Note that for $x\in B_{r_n-\lambda}\setminus\mathcal{N}_n$, $X^{(\lambda;n)}$ has the same distribution as that of $X^{(\lambda)}$ before $X^{(\lambda;n)}$ leaves the ball $B_{r_n-\lambda}$. Thus, estimating as in [9, (4.23)], we have for a.e. $x\in B_{r_0}$,

$$\begin{aligned}\mathbb{P}_x\left(\zeta>1\text{ and }\sup_{s\le1}|X_s^{(\lambda)}-x|\le R\right)&\ge\mathbb{P}_x\left(\sup_{s\le1}|X_s^{(\lambda;n)}-x|\le R\right)\\&\ge 1-2c_2e^{-R/2}\qquad\text{for every }R>0,\end{aligned}$$

where ζ is the lifetime of $X^{(\lambda)}$. Passing $R\to\infty$, we have for a.e. $x\in B_{r_0}$,

$$\mathbb{P}_x(X_1^{(\lambda)}\in\mathbb{R}^d)=1. \tag{2.11}$$

Taking $r_0\uparrow\infty$, (2.11) holds for a.e. $x\in\mathbb{R}^d$; by the Markov property, $\mathbb{P}_x(X_t^{(\lambda)}\in\mathbb{R}^d)=1$ for every rational $t>0$. Since for each rational $t>0$, $P_t^{(r)}1$ is finely continuous and $P_t^{(r)}1=1$ a.e. on $\mathbb{R}^d$, we must have $P_t^{(r)}1=1$ q.e. on $\mathbb{R}^d$, so that $\mathbb{P}_x(\zeta=\infty)=1$ for q.e. $x\in\mathbb{R}^d$. ∎

2.3. Exit time estimate

For $A \subset \mathbb{R}^d$, denote by

$$\tau_A := \inf\{t > 0 : X_t \notin A\}$$

the first exit time from A by X.

Lemma 2.3 *For every $x_0 \in \mathbb{R}^d$ and $r > 0$, $\mathbb{E}_x\left[\tau_{B(x_0,r)}\right] \le c_1 r^2$ for every $x \in B(x_0, r) \setminus \mathcal{N}$.*

Proof. The proof for this is nowadays standard, see for example [10]. For reader's convenience, we spell out the details here. Let $c > 0$ be the constant in (2.4). Take $c_2 > 0$ be large enough so that

$$c\, m_d(B(0,1))\, c_2^{-d/2} \le \tfrac{1}{2}.$$

Then for every $r > 0$, $x_0 \in \mathbb{R}^d$ and $x \in B(x_0, r) \setminus \mathcal{N}$, with $t := c_2 r^2$ we have by (2.4),

$$\mathbb{P}_x(X_t \in B(x_0, r)) = \int_{B(x_0,r)} p(t, x, z)dz \le c\, t^{-d/2} m_d(B(x_0, r)) \le \tfrac{1}{2}.$$

Since X is conservative, this implies that for every $x \in B(x_0, r) \setminus \mathcal{N}$,

$$\mathbb{P}_x(\tau_{B(x_0,r)} \le t) \ge \mathbb{P}_x(X_t \notin B(x_0, r)) \ge 1/2.$$

In other words, we have $\mathbb{P}_x(\tau_{B(x_0,r)} > t) \le \frac{1}{2}$. By the Markov property of X, for integer $k \ge 1$,

$$\mathbb{P}_x\big(\tau_{B(x_0,r)} > (k+1)t\big) \le \mathbb{E}_x\big[\mathbb{P}_{X_{kt}}(\tau_{B(x_0,r)} > t); \tau_{B(x_0,r)} > mt\big] \le \tfrac{1}{2}\mathbb{P}_x\big(\tau_{B(x_0,r)}\mathbb{1} > kt\big).$$

Using mathematical induction, we can conclude that for every $k \ge 1$,

$$\mathbb{P}_x(\tau_{B(x_0,r)} > kt) \le 2^{-k},$$

which yields the desired estimate

$$\mathbb{E}_x\left[\tau_{B(x_0,r)}\right] \le \sum_{k=0}^{\infty} t\mathbb{P}_x(\tau_{B(x_0,r)} > kt) \le c_1 r^2.$$

■

Lemma 2.4 *There is are constants $a_0, r_0 \in (0, 1)$ so that for every $x \in \mathbb{R}^d \setminus \mathcal{N}$,*

$$\mathbb{P}_x\left(\sup_{s \le a_0 r^2} |X_s - X_0| \le r\right) \ge 1/4 \qquad \textit{for every } r \in (0, r_0].$$

Consequently, there exists a constant $a_1 > 0$ so that for every $x \in \mathbb{R}^d \setminus \mathcal{N}$,

$$\mathbb{E}_x\left[\tau_{B(x,r)}\right] \ge a_1 r^2 \qquad \textit{for every } r \in (0, r_0].$$

Proof. By Lemma 3.6 of [2] and (2.10), we have for $0 < r \le 1$,

$$\mathbb{P}_x\Big(\sup_{s\le a_0r^2}|X_s - X_0| \le r\Big) \ge e^{-(b_0r^{-\beta})(a_0r^2)}\,\mathbb{P}_x\Big(\sup_{s\le a_0r^2}|X^{(r)}_s - X^{(r)}_0| \le r\Big)$$
$$\ge e^{-a_0b_0}\,\mathbb{P}_x\Big(\sup_{s\le a_0r^2}|X^{(r)}_s - X^{(r)}_0| \le r\Big).$$

So it suffices to show that there is a positive constant $a_0 \in (0,1)$ small so that

$$a_0b_0 < b_0a_0^{\beta/2} < \log(8/7) \tag{2.12}$$

and that

$$\mathbb{P}_x\Big(\sup_{s\le a_0r^2}|X^{(r)}_s - X^{(r)}_0| \le r\Big) \ge 1/2 \quad \text{for every } r \in (0, r_0]\cap\mathbb{Q} \text{ and } x \in \mathbb{R}^d\setminus\mathcal{N}.$$

Taking $s = 1/\sqrt{t}$ in (2.5), we have

$$p^{(r)}(t,x,y) \le c_0t^{-d/2}\exp\left(-\frac{|x-y|}{\sqrt{t}} + c_2\left(1 + e^{2r/\sqrt{t}}\delta(r)\right)\right). \tag{2.13}$$

Using polar coordinate,

$$\int_{\{|x-y|\ge r/2\}} c_0t^{-d/2}e^{2c_2}\exp\left(-\frac{|x-y|}{\sqrt{t}}\right)dy = \omega_dc_0e^{2c_1}\int_{\frac{r}{2\sqrt{t}}}^{\infty} e^{-v}dv, \tag{2.14}$$

where ω_d is a positive constant that depends only on dimension d. Let $a_0 > 0$ be small enough so that

$$\omega_dc_0e^{2c_2}\int_{1/(2\sqrt{a_0})}^{\infty} e^{-v}dv < 1/8.$$

Due to (2.6), there exists $r_0 \in (0,1)$ so that

$$e^{2/\sqrt{a_0}}\delta(r) \le 1 \qquad \text{for every } r \in (0, r_0].$$

This together with (2.13) and (2.14) implies that for every $r \in (0, r_0]\cap\mathbb{Q}$ and $x \in \mathbb{R}^d$,

$$\mathbb{P}_x\left(|X^{(r)}_{a_0r^2} - X^{(r)}_0| \ge r/2\right) = \int_{\{|y-x|\ge r/2\}} p^{(r)}(a_0r^2, x, y)dy \le 1/8.$$

Moreover, by [2, Lemma 3.6], we have for every $s \le a_0r^2$ with $r \in (0, r_0]\cap\mathbb{Q}$,

$$\mathbb{P}_x\left(|X^{(r)}_s - x| < r/2\right) \ge \mathbb{P}_x\left(|X^{(r)}_s - x| < \sqrt{s/a_0}/2\right)$$
$$\ge e^{-s\,J_{s,r}}\,\mathbb{P}_x\left(\left|X^{(\sqrt{s/a_0})}_s - x\right| < \sqrt{s/a_0}/2\right) \ge \frac{7}{8}e^{-s\,J_{s,r}},$$

where

$$J_{s,r} = \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_{\{\sqrt{s/a_0}<|x-y|\le r\}} J(x,y)dy.$$

By (2.10) and (2.12),

$$sJ_{s,r} \le b_0 a_0^{\beta/2} s^{(2-\beta)/2} \le b_0 a_0^{\beta/2} < \log(8/7)$$

and so

$$\inf_{x\in\mathbb{R}^d\setminus\mathcal{N}} \mathbb{P}_x\left(|X_s^{(r)} - x| < r/2\right) \ge (7/8)^2 > 3/4.$$

In other words, we have

$$\sup_{x\in\mathbb{R}^d\setminus\mathcal{N}} \mathbb{P}_x\left(|X_s^{(r)} - x| \ge r/2\right) < 1/4 \qquad \text{for every } s \le a_0 r^2.$$

Now, since $X^{(r)}$ is conservative, by Lemma 3.8 of [2],

$$\sup_{x\in\mathbb{R}^d\setminus\mathcal{N}} \mathbb{P}_x\left(\sup_{s\le a_0 r^2} |X_s^{(r)} - X_0^{(r)}| \ge r\right) < 1/2,$$

for every $r \in (0, r_0] \cap \mathbb{Q}$. This proves the lemma. ■

3. Short time near-diagonal heat kernel lower bound estimate

Let X be the strong Markov process associated with the Dirichlet form $(\mathcal{E}, \mathcal{F})$ of (1.5) with the jumping kernel satisfying the condition (1.3)-(1.4) and (1.11). Recall that $p(t,x,y)$ is the transition density function for X. For a ball $B \subset \mathbb{R}^d$, denote by $p^B(t,x,y)$ the transition density function of the subprocess X^B of X killed upon exiting B. In this section we will establish the following.

Theorem 3.1 *For each $t_0 > 0$, there exists $c = c(t_0) > 0$ such that for every $x_0 \in \mathbb{R}^d$ and $t \le t_0$,*

$$p^{B(x_0,\sqrt{t})}(t,x,y) \ge c\, t^{-d/2} \qquad \text{for q.e. } x,y \in B(x_0, \sqrt{t}/2)$$

and

$$p(t,x,y) \ge c\, t^{-d/2} \qquad \text{for q.e. } x,y \text{ with } |x-y|^2 \le t.$$

This result will be used in later sections with $t_0 = 1$. For its proof, we adopt an approach from [2] that deals with finite range pure jump processes. But there are some new technical difficulties to overcome in our setting.

Fix $x_0 \in \mathbb{R}^d$ and let $a_1 = 12/(2-\beta)$. (In fact, the following argument works for any fixed a_1 bigger than $4 \vee (6/(2-\beta))$.) For $r > 0$, define

$$\Psi_r(x) = c((1 - r^{-1}|x - x_0|)_+)^{a_1},$$

where $c > 0$ is the normalizing constant such that $\int_{\mathbb{R}^d} \Psi_r(x)dx = 1$. Then the following weighted Poincaré inequality holds. (See, for example, [13, Theorem 5.3.4] for the proof.)

Proposition 3.2 *There is a positive constant $c_1 = c_1(d)$ independent of r, such that*

$$\int_{B(x_0,r)} (u(x) - u_{\Psi_r})^2 \Psi_r(x)dx \le c_1 r^2 \int_{B(x_0,r)} |\nabla u(x)|^2 \Psi_r(x)dx \quad \text{for } u \in C_b^\infty(\mathbb{R}^d).$$

Here $u_{\Psi_r} := \int_{B(x_0,r)} u(x)\Psi_r(x)dx$.

Let W be the symmetric diffusion that corresponds to the divergence form operator $\nabla(A\nabla)$, the local part of $\mathcal{E}$. Let $\eta^{(1)} = \{\eta_t^{(1)}, t \ge 0\}$ be an $(\beta/2)$-subordinator and define $\eta_t = t + c_0 \eta_t^{(1)}$, where $c_0 > 0$ is a large constant to be chosen at the end of this paragraph. Define Y to be the subordination of W by the subordinator $\eta = \{\eta_t; t \ge 0\}$. Note that Y is a symmetric strong Markov process, whose continuous part has the same law as W, and its jumping part comes from the subordination of W by $c_0\eta^{(1)}$. By the uniform ellipticity (1.2) of the diffusion matrix $A(x)$, the heat kernel of W enjoys Aronson-type two-sided Gaussian estimate. It follows that (see [16]) the jump kernel of Y is of the form $\kappa(x,y)/|x-y|^{d+\beta}$, where $\kappa(x,y)$ is a symmetric measurable function that is bounded between two positive constants. By taking $c_0 > 0$ sufficiently large, we can and do assume that

$$J(x,y) \le \frac{\kappa(x,y)}{|x-y|^{d+\beta}} \qquad \text{for all } |x-y| \le 1.$$

For $\delta \in (0,1)$, set

$$J_\delta(x,y) = \begin{cases} J(x,y) & \text{for } |x-y| \ge \delta; \\ \kappa(x,y)|y-x|^{-d-\beta} & \text{for } |x-y| < \delta, \end{cases} \tag{3.1}$$

and define $(\mathcal{E}^\delta, \mathcal{F}^\delta)$ with J_δ in place of J in the definition of $(\mathcal{E}, \mathcal{F})$.

For $\delta \in (0,1)$, let Z^δ be the symmetric Markov process associated with $(\mathcal{E}^\delta, \mathcal{F}^\delta)$. Note that the jumping kernel for Z^δ differs from that of Y by a bounded and integrable kernel. So Z^δ can be constructed from Y through Meyer's construction (see Remarks 3.4 and 3.5 of [2] and Lemma 3.1 of [4]). Consequently, the process Z^δ can be modified to start from every point in $\mathbb{R}^d$ and Z^δ is conservative. Moreover by a similar proof to that in [2], we can show that Z^δ has a quasi-continuous transition density function $q^\delta(t,x,y)$ defined on $[0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d$, with respect to the Lebesgue measure on $\mathbb{R}^d$. Since Y is a subordination of W, we can readily get a two-sided kernel estimate on $p_Y(t,x,y)$ of Y from that of W. In fact, since the heat kernel of W is comparable to that of Brownian motion, $p_Y(t,x,y)$ is comparable to that of the independent sum of Brownian motion and a rotationally symmetric β-stable process. So by [14],

$$\begin{aligned} &c_1 \left(t^{-d/2} \wedge t^{-d/\beta}\right) \left(t^{-d/2} e^{-c_2|x-y|^2/t} + t^{-d/\beta} \left(1 \wedge \frac{t}{|x-y|^{d+\beta}}\right)\right) \le p_Y(t,x,y) \\ &\quad \le c_3 \left(t^{-d/2} \wedge t^{-d/\beta}\right) \left(t^{-d/2} e^{-4_2|x-y|^2/t} + t^{-d/\beta} \left(1 \wedge \frac{t}{|x-y|^{d+\beta}}\right)\right) \end{aligned} \tag{3.2}$$

for all $t > 0$ and $x, y \in \mathbb{R}^d$. Consequently, parabolic Harnack principle holds for Y (see [14, Theorem 4.5]). On the other hand, as a consequence of Meyer's construction (see the proof of Proposition 2.1 of [7]) and (3.2), there are constant $t_0, r \in (0,1)$ and $c > 1$, which depend on δ, so that

$$c^{-1} p_Y(t,x,y) \le q^\delta(t,x,y) \le c\, p_Y(t,x,y) \text{ for } t \in (0,t_0] \text{ and } |x-y| \le r_0. \tag{3.3}$$

From (3.3), we can easily show that parabolic Harnack principle holds at small-size scale for Z^δ and that its parabolic functions are jointly continuous (see [7, Remark 4.3 (*ii*)]). In particular, $q^\delta(t,x,y)$ is jointly continuous on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d$.

For $r \in (0,1]$, let $B_r = B(0,r)$ and let $(\mathcal{E}^\delta, \mathcal{F}^{\delta,B_r})$ be the Dirichlet form corresponding to the process Z^δ killed on leaving the ball B_r. Let $q^{\delta,B_r}(t,x,y)$ be its heat kernel with respect to the Lebesgue measure in B_r. We first prove the following, which corresponds to Lemmas 4.5, 4.6 and 4.7 in [2]. The latter can be traced back to Fabes and Stroock's simplified version [11] of Nash's lower bound approach to the heat kernel estimates for symmetric diffusions. Due to the non-local nature of the operator $\mathcal{L}$ of (1.1) in this paper, certain regularity issues need to be addressed before the aforementioned method can be employed.

Proposition 3.3 *(i) For each $t > 0$ and $y_0 \in B_r$, we have*

$$q^{\delta,B_r}(t,\cdot,y_0), \quad \frac{\Psi_r(\cdot)}{q^{\delta,B_r}(t,\cdot,y_0)} \in \mathcal{F}^{\delta,B_r}.$$

(ii) Fix $y_0 \in B$ and let

$$\overline{G}(t) = \int_{B_r} \Psi_r(x) \log q^{\delta,B_r}(t,x,y_0)\, dx.$$

Then for every $t > 0$,

$$\overline{G}'(t) = -\mathcal{E}\left(q^{\delta,B_r}(t,\cdot,y_0), \frac{\Psi_r(\cdot)}{q^{\delta,B_r}(t,\cdot,y_0)} \right).$$

The following lemma plays a key role in our proof of above proposition.

Lemma 3.4 *Assume $0 < \delta < 1/16$. Let $0 < t_1 < t_2 < \infty$ and $r \in (16\delta, 1]$. There is a constant $c_1 = c_1(\delta, r, t_0, t_1) > 0$ such that*

$$q^{\delta,B_r}(t,x,y) \geq c_1(r-|x|)^2(r-|y|)^2 \qquad \text{for every } t \in [t_1,t_2] \text{ and } x,y \in B_r.$$

Proof. Due to the Chapman-Kolmogorov equation, without loss of generality, we can and do assume that

$$t_1 < 3a_0 \min\{\delta_0 r, r_0\}^2/16,$$

where $\delta_0 \in (0,1)$ is the constant in (1.3) and (1.11). and a_0 and r_0 are the constant in Lemma 2.4.

First, since as mentioned above Z^δ enjoys parabolic Harnack principle at the small-size scale, we have by the same proof as that for Lemma 4.2 of [2] that for every $\gamma \in (0,1)$, there is a constant $c_\gamma > 0$ so that

$$q^{\delta,B_r}(t,x,y) \geq c_\gamma \qquad \text{for } t \in [t_1/12,\, t_2] \text{ and } x,y \in B(0,\gamma r). \tag{3.4}$$

So it suffices to prove the lemma for $x, y \in B_r$ with

$$\max\{r-|x|, r-|y|\} < r_1 := \min\{r_0,\, \delta_0 r/8,\, t_1/(4a_0)\}.$$

Let $y \in B_r$ with $\delta(y) := r - |y| < r_1$. Take $y_0 \in B(0,(1-3\delta_0/4)r)$ with $|y-y_0| = \delta_0 r$. Define $T := \inf\{t > 0 :\ |Z_t^\delta - Z_{t-}^\delta| \geq \delta_0 r\}$ and set $s_0 = t_1/3$.

By the strong Markov property of Z^δ,

(3.5)

$$\begin{aligned}
&\mathbb{P}_y\left(Z^\delta_{s_0}\in B(0,(1-\delta_0/2)r) \text{ and } \tau_{B_r}>s_0)\right)\\
&\geq \mathbb{P}_y\Bigg(T\leq a_0\delta(y)^2/4,\ Z^\delta_T\in B(y_0,\ \delta_0 r/16),\ \sup_{s<T}|Z^\delta_s-y|\leq \delta(y)/2\\
&\qquad \text{and } \sup_{s\in[T,s_0+T]}|Z^\delta_s-Z^\delta_T|\leq \delta_0 r/4\Bigg)\\
&\geq \mathbb{P}_y\left(T\leq a_0\delta(y)^2/4,\ Z^\delta_T\in B(y_0,\ \delta_0 r/16) \text{ and } \sup_{s<T}|Z^\delta_s-y|\leq \delta(y)/2\right)\\
&\qquad \cdot \inf_{y\in\mathbb{R}^d\setminus\mathcal{N}}\mathbb{P}_x\left(\sup_{s\in[0,s_0]}|Z^\delta_s-x|\leq\delta_0 r/4\right).
\end{aligned}$$

Note that by conditions (1.3)-(1.4) and (1.11),

$$\kappa_1:=\sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbb{1}_{\{|x-z|>\delta_0 r\}}J_\delta(x,z)dz<\infty \quad \text{and}$$

$$\kappa_2:=\inf_{y\in B_r}\ \inf_{x\in B(y,\,\delta_0 r/16)\setminus\mathcal{N}}\int_{B(y_0,\,\delta_0 r/16)}J_\delta(x,z)dz>0.$$

As T is the first time the process Z^δ makes a jump of size no less than $\delta_0 r$, T is stochastically dominated from above by the exponential random variable with parameter κ_1 and at time T, process Z^δ jumps to position z according to the probability kernel

$$\frac{J_\delta(Z^\delta_{T-},dz)}{\int_{\{w:\,|w-Z^\delta_{T-}|\geq\delta_0 r)\}}J_\delta(Z^\delta_{T-},dw)}.$$

Thus we have

$$\begin{aligned}
&\mathbb{P}_y\left(T\leq a_0\delta(y)^2/4 \text{ and } Z^\delta_T\in B(y_0,\ \delta_0 r/16)\Big|\sup_{s<T\wedge(a_0\delta(y)^2/4)}|Z^\delta_s-y|\leq\delta(y)/2\right)\\
&\qquad\geq\left(1-e^{-\kappa_1 a_0\delta(y)^2/4}\right)(\kappa_2/\kappa_1)\geq c\,a_0\,\delta(y)^2. \qquad (3.6)
\end{aligned}$$

By Meyer's construction [2, Lemma 3.6] and Lemma 2.4,

$$\begin{aligned}
&\mathbb{P}_y\left(\sup_{s<T\wedge(a_0\delta(y)^2/4)}|Z^\delta_s-y|\leq\delta(y)/2\right)\\
&\qquad\geq\mathbb{P}_y\left(\sup_{s\leq a_0\delta(y)^2/4}|Z^\delta_s-y|\leq\delta(y)/2 \text{ and } T\geq a_0\delta(y)^2/4\right)\\
&\qquad\geq e^{-\kappa\cdot a_0\delta(y)^2/4}\,\mathbb{P}_y\left(\sup_{s\leq a_0\delta(y)^2/4}|Z^\delta_s-y|\leq\delta(y)/2\right)\geq 1/(4e^\kappa).
\end{aligned}$$

This together with (3.6) yields that

$$\mathbb{P}_y\left(T \le a_0\delta(y)^2/4, Z^\delta_T \in B(y_0,\delta_0 r/16) \text{ and } \sup_{s<T}|Z^\delta_{s_0}-y| \le \delta(y)/2\right)$$
$$\ge c\,\delta(y)^2. \tag{3.7}$$

Since $s_0 = t_1/3 < a_0(\delta_0 r)^2/16$, we have from Lemma 2.4 that

$$\inf_{x\in\mathbb{R}^d\setminus\mathcal{N}} \mathbb{P}_x\left(\sup_{s\le s_0}|Z^\delta_s - Z^\delta_0| \le \delta_0 r/4\right) \ge 1/4.$$

Therefore we have by (3.5) and (3.7) that

$$\mathbb{P}_y\left(Z^\delta_{s_0} \in B(0,(1-\delta_0/2)r) \text{ and } \tau_{B_r} > s_0\right) \ge c(r-|y|)^2.$$

Now for $t\in[t_1/2,t_2]$, $y\in B_r$ and $z\in B(0,(1-\delta_0/2)r)$, by (3.4)

$$\begin{aligned} q^{\delta,B_r}(t,y,z) &\ge \int_{B(0,(1-\delta_0/2)r)} q^{\delta,B_r}(s_0,y,w)q^{\delta,B_r}(t-s_0,w,z)dw \\ &\ge c\int_{B(0,(1-\delta_0/2)r)} q^{\delta,B_r}(s_0,y,w)dw \\ &= c\,\mathbb{P}_y\left(Z^\delta_{s_0}\in B(0,(1-\delta_0/2)r) \text{ and } \tau_{B_r}>s_0\right) \\ &\ge c(r-|y|)^2. \end{aligned}$$

This together with the Chapman-Kolmogorov's equation

$$q^{\delta,B_r}(t,x,y) \ge \int_{B(0,(1-\delta_0/2)r)} q^{\delta,B_r}(t/2,x,z)q^{\delta,B_r}(t/2,z,y)dz$$

proves the lemma. ■

Proof of Proposition 3.3. (i) First, similarly to the proof of [2, Lemma 4.1], we have

$$q^{\delta,B_r}(t,x,y) \le c_1 t^{-d/2} \quad \text{and} \quad \left|\frac{\partial q^{\delta,B_r}(t,x,y)}{\partial t}\right| \le c_1 t^{-1-d/2} \tag{3.8}$$

for every $x,y\in B_r$ and $t>0$. Using this, $q^{\delta,B_r}(t,\cdot,y_0)\in\mathcal{F}^{\delta,B_r}$ can be proved in the same way as the proof of [2, Lemma 4.5]. Next, by Lemma 3.4 and by the choice of a_1, for every $y_0\in B_r$, $\varepsilon\in(0,1)$ and $\gamma\in\left(\frac{2-\beta}{6},1\right]$, there is a constant $C=C(y_0,\beta,\delta,\varepsilon)>0$ such that

$$\Psi_r(x)^\gamma/q^{\delta,B_r}(t,x,y_0) \le C, \qquad \text{for every } t\in(\varepsilon,\varepsilon^{-1}] \text{ and } x\in B_r. \tag{3.9}$$

Using this, $\Psi_r(\cdot)^{1/2}/q^{\delta,B_r}(t,\cdot,y_0)$ is bounded on B_r. By extending the function

$$x \mapsto \frac{\Psi_r(x)}{q^{\delta,B_r}(t,x,y_0)}$$

to be zero on B_r^c, we see that it vanishes continuously on B_r^c. Similar to the proof of Proposition 1.1,

$$\mathcal{F}^{\delta,B_r} = \left\{f \in L^2(\mathbb{R}^d; m_d) : f|_{B_r^c} \equiv 0 \text{ and } \mathcal{E}^\delta(f,f) < \infty\right\}.$$

So, in order to prove $h_t(\cdot) := \Psi_r(\cdot)/q^{\delta,B_r}(t,\cdot,y_0) \in \mathcal{F}^{\delta,B_r}$, it is enough to prove $\mathcal{E}^\delta(h_t,h_t) < \infty$. Let $u_t(\cdot) = q^{\delta,B_r}(t,\cdot,y_0)$. In order to show

$$\int_{B_r} \nabla h_t(x) A(x) \nabla h_t(x) dx < \infty,$$

it is enough to prove

$$\int_{B_r} |u_t(x)\nabla\Psi_r(x) - \Psi_r(x)\nabla u_t(x)|^2/u_t(x)^4 dx < \infty,$$

since $a(\cdot)$ is uniform elliptic. Computing this,

$$\begin{aligned}
&\int_{B_r} \frac{|u_t(x)\nabla\Psi_r(x) - \Psi_r(x)\nabla u_t(x)|^2}{u_t(x)^4} dx \\
&\qquad \le 2\left(\int_{B_r} \frac{|\nabla\Psi_r(x)|^2}{u_t(x)^2} dx + \int_{B_r} \frac{|\Psi_r(x)\nabla u_t(x)|^2}{u_t(x)^4} dx\right) \\
&\qquad \le 2\left(c_1 c_2^2\, m_d(B_r) + c_2^4 \int_{B_r} |\nabla u_t(x)|^2 dx\right) < \infty,
\end{aligned}$$

where $|\nabla\Psi_r|^2/\Psi_r \le c_1$ and $\Psi_r^{1/2}/u_t \le c_2$ (due to (3.9)) are used in the second inequality. The proof of

$$\int_{B_r}\int_{B_r} (u_t(x) - u_t(y))^2 J_\delta(x,y) dx dy + 2\int_{B_r} u_t(x)^2 \left(\int_{B_r^c} J_\delta(x,y) dy\right) dx < \infty$$

can be done similarly to that of [2, Lemma 4.6] (with a suitable change due to the shape of J_δ, for example $\gamma = (2-\beta)/3$ in the proof). We thus obtain $\mathcal{E}^\delta(h_t,h_t) < \infty$.

(*ii*) Given (*i*), (3.8) and (3.9), this can be proved in the same way as the proof of [2, Lemma 4.7]. ■

The idea of the proof of the following theorem is motivated by that of Theorem 3.4 in [7] and Proposition 4.9 in [2]. However, due to the existence of the divergence form part, various non-trivial changes are required.

Theorem 3.5 *For each $t_0 > 0$, there exists $c = c(t_0) > 0$, independent of $\delta \in (0,1)$ such that for every $x_0 \in \mathbb{R}^d$, $t \le t_0$,*

$$q^{\delta, B(x_0, t^{1/2})}(t,x,y) \ge c\, t^{-d/2} \qquad \text{for q.e. } x, y \in B(x_0, \sqrt{t}/2) \tag{3.10}$$

and

$$q^{\delta}(t,x,y) \ge c\, t^{-d/2} \qquad \text{for q.e. } x, y \text{ with } |x-y|^2 \le t. \tag{3.11}$$

Proof. Fix $\delta \in (0,1)$ and, for simplicity, in this proof we sometimes drop the superscript "δ" from Z^δ and $q^\delta(t,x,y)$. Also, for notational convenience, let $x_0 = 0$. For ball $B_r := B(0,r) \subset \mathbb{R}^d$, let $q^{B_r}(t,x,y)$ denote the transition density function of the subprocess Z^{B_r} of Z killed on leaving the ball B_r.

Define $B := B(0,1)$ and for $r \le 1$, let $(\mathcal{E}^{\langle r\rangle}, \mathcal{F}^{\langle r\rangle, B})$ be the Dirichlet form corresponding to $\{r^{-1} Z^{\delta, B_r}_{r^2 t}, t \ge 0\}$, which is the subprocess of $\{Z^{\langle r\rangle}_t := r^{-1} Z^{\delta}_{r^2 t}, , t \ge 0\}$ killed on leaving the unit ball B. Define

$$q^B_r(t,x,y) = q^{\delta,B}_r(t,x,y) := r^d q^{B_r}(r^2 t, rx, ry). \tag{3.12}$$

It is easy to see $q^B_r(t,x,y)$ is the transition density function for process $r^{-1} Z^{\delta, B_r}_{r^2 t}$.

Set $\Psi(x) = c((1-|x|)_+)^{a_1}$, where $c > 0$ is the normalizing constant. Let $x_0 \in B(0,1)$, $r \le 1$, and define

$$\begin{aligned}
u(t,x) &:= q^B_r(t,x,x_0), \\
v(t,x) &:= q^B_r(t,x,x_0)/\Psi(x)^{1/2}, \\
H(t) &:= \int_B \Psi(y) \log u(t,y) dy, \\
G(t) &:= \int_B \Psi(y) \log v(t,y) dy = \int_B \Psi(y) \log u(t,y) dy \\
&\qquad\qquad\qquad - \frac{1}{2} \int_B \Psi(x) \log \Psi(x) dx = H(t) + c_1.
\end{aligned}$$

By Proposition 3.3 and the scaling, we have

$$G'(t) = -\mathcal{E}^{\langle r\rangle}\Big(u(t,\cdot), \frac{\Psi}{u(t,\cdot)}\Big) =: -(J_1 + J_2), \tag{3.13}$$

where J_1 is the diffusion part and J_2 is the jump part of the Dirichlet form.

We first estimate the jump part. Write $J^{\langle r\rangle}_\delta(x,y) := r^{d+2} J_\delta(rx, ry)$. By the same argument as in the proof of Proposition 4.9 of [2] (up to the formula four lines after (4.15) there), we have

$$
\begin{aligned}
J_2 &= \mathcal{E}^{\langle r\rangle,j}\Big(u(t,\cdot),\frac{\Psi}{u(t,\cdot)}\Big)\\
&\le \int_B\int_B \{(\Psi(x)^{1/2}-\Psi(y)^{1/2})^2-(\Psi(x)\wedge\Psi(y))(\log\frac{v(t,y)}{v(t,x)})^2\}J_\delta^{\langle r\rangle}(x,y)dxdy\\
&\quad+\int_B\Psi(x)\Big(2\int_{B^c}J_\delta^{\langle r\rangle}(x,y)dy\Big)dx\\
&\le \int_B\int_B(\Psi(x)^{1/2}-\Psi(y)^{1/2})^2J_\delta^{\langle r\rangle}(x,y)dxdy+\int_B\Psi(x)\Big(2\int_{B^c}J_\delta^{\langle r\rangle}(x,y)dy\Big)dx\\
&= \mathcal{E}^{\langle r\rangle,j}(\Psi^{1/2},\Psi^{1/2})\le c_2r^{2-\beta}\mathcal{E}(\Psi^{1/2},\Psi^{1/2})\le c_2\mathcal{E}(\Psi^{1/2},\Psi^{1/2})<\infty,
\end{aligned}
$$

where the last inequality is due to the shape of J and the Lipschitz continuity of Ψ (note that $c_2\mathcal{E}(\Psi,\Psi)$ is independent of r).

We next estimate the diffusion part.

$$
\begin{aligned}
J_1 &= \mathcal{E}^{\langle r\rangle,c}\Big(u(t,\cdot),\frac{\Psi}{u(t,\cdot)}\Big)\le\int_B\nabla u(t,x)a(rx)\nabla\Big(\frac{\Psi(x)}{u(t,x)}\Big)dx\\
&= \int_B\nabla\log u(t,x)a(rx)\nabla\Psi(x)dx\\
&\quad-\int_B\nabla\log u(t,x)a(rx)\nabla\log u(t,x)\Psi(x)dx. \qquad (3.14)
\end{aligned}
$$

Note that

$$
\begin{aligned}
0&\le\int_B((\nabla\log u)\sqrt{\Psi}-\frac{\nabla\Psi}{\sqrt{\Psi}})a^{\langle r\rangle}\cdot((\nabla\log u)\sqrt{\Psi}-\frac{\nabla\Psi}{\sqrt{\Psi}})dx\\
&=\int_B\nabla\Psi a^{\langle r\rangle}\cdot\nabla\Psi\Psi^{-1}dx+\int_B(\nabla\log u)a^{\langle r\rangle}\cdot(\nabla\log u)\Psi dx\\
&\qquad-2\int_B(\nabla\log u)a^{\langle r\rangle}\cdot\nabla\Psi dx,
\end{aligned}
$$

where $a^{\langle r\rangle}(\cdot)=a(r\cdot)$. Using this and (1.2) in (3.14), we obtain

$$
\begin{aligned}
J_1&\le c_3\int_B\frac{|\nabla\Psi(x)|^2}{\Psi(x)}dx-c_4\int_B|\nabla\log u(t,x)|^2\Psi(x)dx\\
&=c_5-c_4\int_B|\nabla\log u(t,x)|^2\Psi(x)dx,
\end{aligned}
$$

where the last equality is due to the fact $|\nabla\Psi(x)|^2/\Psi(x)\le c_{5.5}$ for $x\in B$, which is because $a_1\ge 2$ in the definition of Ψ. Thus, using Proposition 3.2,

$$
J_1\le c_6-c_7\int_B(\log u(t,x)-H(t))^2\Psi(x)dx.
$$

Combining these, we obtain from (3.13),

$$G'(t) = H'(t) \geq -c_8 + c_7 \int_B (\log u(t,y) - H(t))^2 \Psi(y)\, dy. \tag{3.15}$$

Given this inequality, (2.4) and Lemma 2.4, the rest of the proof is the same as that of [2, Proposition 4.9] (cf. also [7, Theorem 3.4]). ■

Proof of Theorem 3.1. For any ball $B \subset \mathbb{R}^d$, let $(\mathcal{E}^{\delta,B}, \mathcal{F}^{\delta,B})$ denote the Dirichlet form of the subprocess $Z^{\delta,B}$ of Z^δ killed upon leaving the ball B. Similarly to the proof of [2, Theorem 1.5 and Theorem 2.6], we can show that $(\mathcal{E}^\delta, \mathcal{F}^\delta)$ and $(\mathcal{E}^{\delta,B}, \mathcal{F}^{\delta,B})$ converge as $\delta \to 0$ to $(\mathcal{E}, \mathcal{F})$ and $(\mathcal{E}^B, \mathcal{F}^B)$, respectively in the sense of Mosco, where B is a ball in $\mathbb{R}^d$. Therefore the semigroup of Z^δ and $Z^{\delta,B}$ converge in L^2 to that of X and X^B, respectively. Theorem 3.1 follows from Theorem 3.5 by a similar argument as that for [2, Theorem 1.3]. ■

4. Hölder continuity and Parabolic Harnack inequality

4.1. Hölder continuity

In this subsection, the Dirichlet form $(\mathcal{E}, \mathcal{F})$ is given by (1.5) with the jumping kernel satisfying the conditions (1.3)-(1.4), and X is its associated strong Markov process in $\mathbb{R}^d$.

For $r \in (0, 1]$, define

$$Q(x,r) := (0, r^2] \times B(x,r).$$

For each $A \subset [0,\infty) \times \mathbb{R}^d$, denote $\sigma_A := \inf\{t > 0 : Z_t \in A\}$ and $A_s := \{y \in \mathbb{R}^d : (s,y) \in A\}$.

Lemma 4.1 *There exists $C_2 > 0$ such that for all $x \in \mathbb{R}^d$, $0 < r \leq 1$ and any compact subset $A \subset Q(0,x,r)$,*

$$\mathbb{P}^{(r^2,x)}(\sigma_A < \tau_r) \geq C_2 \frac{m_{d+1}(A)}{r^{d+2}},$$

where $\tau_r = \tau_{Q(x,r)}$ and m_{d+1} is the Lebesgue measure on $\mathbb{R}^{d+1}$.

Proof. For $0 < r \leq 1$,

$$\begin{aligned} r^2\, \mathbb{P}^{(r^2,x)}(\sigma_A < \tau_r) &\geq \int_0^{r^2} \mathbb{P}^{(r^2,x)}\left((r^2 - s, X_s^{B(x,r)}) \in A\right) ds \\ &= \int_0^{r^2} \int_{A_{r^2-s}} p^{B(x,r)}(s,x,y)\,dyds \geq \int_0^{r^2} \int_{A_{r^2-s}} \frac{c}{r^d}\,dyds = c\,\frac{m_{d+1}(A)}{r^d}, \end{aligned}$$

where Theorem 3.1 is used in the last inequality. ■

We can now establish the Hölder continuity for parabolic functions of X. First, recall the following well-known formula (see, for example [9, Appendix A]).

Lemma 4.2 (Lévy system formula) *Let f be a non-negative measurable function on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d$ that vanishes along the diagonal. Then for every $t \geq 0$, $x \in \mathbb{R}^d \setminus \mathcal{N}$ and stopping time T (with respect to the filtration of X),*

$$\mathbb{E}_x \left[\sum_{s \leq T} f(s, X_{s-}, X_s) \right] = \mathbb{E}_x \left[\int_0^T \left(\int_{\mathbb{R}^d} f(s, X_s, y) J(X_s, y) dy \right) ds \right].$$

Proof of Theorem 1.2. For $x \in \mathbb{R}^d \setminus \mathcal{N}$ and $r < 1$, apply Lemma 4.2 to $f(s, y, z) = \mathbb{1}_{B(x,r)}(y) \mathbb{1}_{B(x,2r)}(z)$ and $T = \tau_{B(x,r)}$. Then it follows from (1.4) and Lemma 2.3, for every $s \geq 2r$,

$$\begin{aligned} \text{(4.1)} \quad \mathbb{P}_x \left(X_{\tau_{B(x,r)}} \notin B(x, s) \right) &= \mathbb{E}_x \left[\int_0^{\tau_{B(x,r)}} \left(\int_{\mathbb{R}^d \setminus B(x,s)} J(X_t, y) dy \right) dt \right] \\ &\leq 4(s \wedge 1)^{-2} \, \mathbb{E}_x \left[\int_0^{\tau_{B(x,r)}} \left(\int_{\mathbb{R}^d} (|X_t - y|^2 \wedge 1) J(X_t, y) dy \right) dt \right] \\ &\leq c \, (s \wedge 1)^{-2} \, \mathbb{E}_x \left[\tau_{B(x,r)} \right] \\ &\leq c \, r^2 / (s \wedge 1)^2. \end{aligned}$$

Using this and Lemma 4.1, the rest of the proof is the same as that for the proof of Theorem 4.14 in [8] except that the estimate for

$$\text{(4.2)} \quad \sum_{i=1}^{\infty} \mathbb{E}_{z_1} \left[q(Z_{\tau_{k+1}}) - q(z_2); \, \sigma_A > \tau_{k+1} \text{ and } Z_{\tau_{k+1}} \in Q_{k-i} \setminus Q_{k+1-i} \right]$$

at the bottom of page 57 of [8] should be bound as follows. Take $\rho < \eta$, then

$$\begin{aligned} \text{(4.2)} \quad &\leq \sum_{i=1}^{k} (b_{k-i} - a_{k-i}) \mathbb{P}_{z_1}(X_{\tau_{k+1}} \notin Q_{k+1-i}) + \|h\|_{\infty,R} \mathbb{P}_{z_1}(X_{\tau_{k+1}} \notin Q_0) \\ &\leq \sum_{i=1}^{k} c \, \eta^k (\rho^2/\eta)^i + c \, \|h\|_{\infty,R} \, \rho^{k+1} \\ &\leq c \, \eta^{k-1} \rho^2 + c \, \rho^{k+1} \\ &\leq c \eta^{k+1}. \end{aligned}$$

■

4.2. Parabolic Harnack inequality

In this subsection, the Dirichlet form $(\mathcal{E}, \mathcal{F})$ is given by (1.5) with the jumping kernel satisfying the conditions (1.3)-(1.4) and (1.14), and X is its associated strong Markov process in $\mathbb{R}^d$.

Recall that $Z_s := (V_s, X_s)$ is the space-time process of X, where $V_s = V_0 - s$. The following lemma corresponds to [7, Lemma 4.2]. Noting that the continuous component of the process does not play any role since the function h is supported in $[0,\infty) \times B(x_0, 3R)^c$, the proof is almost the same as that of [7, Lemma 4.2]. We point out that condition (1.14) is used in a crucial way in the proof of this lemma.

Lemma 4.3 *Let $R \le 1$ and $\delta < 1$. $Q_1 = [t_0 + 2\delta R^2/3,\, t_0 + 5\delta R^2] \times B(x_0, 3R/2)$, $Q_2 = [t_0 + \delta R^2/3,\, t_0 + 11\delta R^2/2] \times B(x_0, 2R)$ and define Q_- and Q_+ as in Theorem* 1.3. *Let $h : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}_+$ be bounded and supported in $[0,\infty) \times B(x_0, 3R)^c$. Then there exists $C_1 = C_1(\delta) > 0$ such that the following holds:*

$$\mathbb{E}^{(t_1,y_1)}[h(Z_{\tau_{Q_1}})] \le C_1 \mathbb{E}^{(t_2,y_2)}[h(Z_{\tau_{Q_2}})] \quad \text{for } (t_1, y_1) \in Q_- \text{ and } (t_2, y_2) \in Q_+.$$

Proof of Theorem 1.3. With the above lemma, Lemma 4.1 and the heat kernel estimates in the previous sections, the proof is almost the same as that of the proof of [7, Theorem 4.1] for $R \le 1$. ■

5. Heat kernel upper bound estimate

For the remaining two sections, we assume that the jumping kernel J for the Dirichlet form $(\mathcal{E}, \mathcal{F})$ of (1.5) satisfies condition (1.9). For simplicity, define

$$\widetilde{\phi}(r) := r^2 \wedge \phi(r).$$

Note that $r \to \widetilde{\phi}(r)$ is a strictly increasing function on $[0,\infty)$ so it has an inverse function $\widetilde{\phi}^{-1}(r)$. Clearly,

$$\widetilde{\phi}^{-1}(r) = r^{1/2} \vee \phi^{-1}(r),$$

where ϕ^{-1} is the inverse function of ϕ. Note that

$$\widetilde{\phi}^{-1}(t)^{-d} = t^{-d/2} \wedge \phi^{-1}(t)^{-d}.$$

Theorem 5.1 *There are positive constants c_1 and c_2 such that for every $x, y \in \mathbb{R}^d$ and $t > 0$, we have*

$$p(t,x,y) \le c_1\, \widetilde{\phi}^{-1}(t)^{-d} \wedge \left(p^c(t, c_2|x-y|) + p^j(t, |x-y|)\right). \tag{5.1}$$

Before proving this theorem, we make some preparations. For $r > 0$, let $(\mathcal{E}^{\langle r\rangle}, \mathcal{F}^{\langle r\rangle})$ be the Dirichlet form corresponding to $\{X_t^{\langle r\rangle} := r^{-1}X_{\widetilde{\phi}(r)t}, t \geq 0\}$. By simple computations, we see that $\mathcal{F}^{\langle r\rangle} = W^{1,2}(\mathbb{R}^d)$ and for $u, v \in \mathcal{F}^{\langle r\rangle}$,

$$\mathcal{E}^{\langle r\rangle}(u,v) = \frac{\widetilde{\phi}(r)}{r^2}\int_{\mathbb{R}^d} \nabla u(x) \cdot a(rx)\nabla v(x)dx \\ + \int_{\mathbb{R}^d} (u(x)-u(y))(v(x)-v(y))J^{\langle r\rangle}(x,y)dxdy,$$

where $J^{\langle r\rangle}(x,y) = \widetilde{\phi}(r)r^d J(rx, ry)$. Note that

$$J^{\langle r\rangle}(x,y) \asymp \frac{\widetilde{\phi}(r)}{|x-y|^d\phi(r|x-y|)} = \frac{1}{|x-y|^d\phi_r(|x-y|)},$$

where $\phi_r(s) := \phi(rs)/\widetilde{\phi}(r)$ (note that ϕ_r enjoys the properties (1.7) and (1.8) with the constant $c > 0$ independent of r). Clearly the transition density function $p_r(t,x,y)$ of $X^{\langle r\rangle}$ with respect to m_d is given by

$$p_r(t,x,y) := r^d p(\widetilde{\phi}(r)t, rx, ry). \tag{5.2}$$

The following on-diagonal estimate holds for $p(t,x,y)$:

$$p(t,x,y) \leq c\left(t^{-d/2} \wedge \phi^{-1}(t)^{-d}\right), \qquad \forall t > 0. \tag{5.3}$$

If follows from the Nash inequality for the stable-type Dirichlet form obtained in [9, Theorem 3.1], we have $p(t,x,y) \leq c\phi^{-1}(t)^{-d}$, so that (5.3) holds. Thus, using (5.2), we have

$$p_r(t,x,y) \leq r^d\left(\widetilde{\phi}^{-1}(\widetilde{\phi}(r)t)\right)^{-d} =: g(r,t). \tag{5.4}$$

Clearly $g(r,1) = 1$ and

$$\begin{aligned} g(r,t) &\leq c\left(r^d(\widetilde{\phi}(r)t)^{-d/2}\mathbb{1}_{\{\widetilde{\phi}(r)t\leq 1\}} + r^d(\phi^{-1}(\widetilde{\phi}(r)t))^{-d}\mathbb{1}_{\{\widetilde{\phi}(r)t>1\}}\right) \\ &\leq c\left(r^d\,\widetilde{\phi}(r)^{-d/2}t^{-d/2}\mathbb{1}_{\{\widetilde{\phi}(r)t\leq 1\}} + r^d\,\widetilde{\phi}(r)^{-d/\beta_2}t^{-d/\beta_2}\mathbb{1}_{\{\widetilde{\phi}(r)t>1\}}\right). \end{aligned}$$

For $\lambda > 0$, define

$$J^{\langle r,\lambda\rangle}(x,y) := J^{\langle r\rangle}(x,y)\mathbb{1}_{\{|x-y|\leq\lambda\}}$$

and let $(\mathcal{E}^{\langle r,\lambda\rangle}, W^{1,2}(\mathbb{R}^d))$ be defined as $(\mathcal{E}^{\langle r\rangle}, \mathcal{F}^{\langle r\rangle})$ but with jumping kernel $J^{\langle r,\lambda\rangle}$ in place of $J^{\langle r\rangle}$. Let $X^{\langle r,\lambda\rangle}$ be the symmetric strong Markov process associated with $(\mathcal{E}^{\langle r,\lambda\rangle}, W^{1,2}(\mathbb{R}^d))$. The process $X^{\langle r,\lambda\rangle}$ can be obtained from $X^{\langle r\rangle}$ by removing all the jumps whose size is larger than λ. We will apply Davies' method to derive heat kernel estimate for process $X^{\langle r,\lambda\rangle}$. On-diagonal estimate (5.4) together with Theorem 3.25 of [6] implies that there exist constants $C>0$ and $c>0$, independent of $\lambda>0$ and $\delta>0$ such that

$$p_r^{(\lambda)}(t,x,y) \le g(r,t)\ \exp\big(-|\psi(y)-\psi(x)| + C\ \Lambda_{r,\lambda}(\psi)^2\ t\big) \tag{5.5}$$

for all $t>0$, $x,y\in\mathbb{R}^d\setminus\mathcal{N}$ and every $\lambda>0$, and for some ψ satisfying $\Lambda_{r,\lambda}(\psi)<\infty$, where

$$\Lambda_{r,\lambda}(\psi)^2 = \|e^{-2\psi}\Gamma_{r,\lambda}[e^{\psi}]\|_\infty \vee \|e^{2\psi}\Gamma_{r,\lambda}[e^{-\psi}]\|_\infty.$$

Here

$$\begin{aligned}\Gamma_{r,\lambda}[v](\xi) = {} & \frac{\widetilde{\phi}(r)}{r^2}\sum_{i,j=1}^d a_{ij}(r\xi)\frac{\partial v}{\partial x_i}(\xi)\frac{\partial v}{\partial x_j}(\xi) \\ & + \int_{|\eta-\xi|\le\lambda}(v(\eta)-v(\xi))^2 J^{\langle r\rangle}(\eta,\xi)d\eta,\quad \xi\in\mathbb{R}^d.\end{aligned} \tag{5.6}$$

Define

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

A key observation is that $\mathcal{H}(\Gamma_{r,\lambda})$ contains the *cut-off distance function* ψ given by

$$\psi(\xi) := \frac{s}{3}\left(|\xi-x|\wedge|x-y|\right) \qquad \text{for } \xi\in\mathbb{R}^d, \tag{5.7}$$

where $s>0$ is a parameter to be chosen later. Note that $|\psi(\eta)-\psi(\xi)|\le (s/3)|\eta-\xi|$ for all $\xi,\eta\in\mathbb{R}^d$. So

$$\begin{aligned} e^{-2\psi(\xi)}\Gamma_{r,\lambda}[e^{\psi}](\xi) &\le c_1|\nabla\psi(\xi)|^2 + \int_{|\eta-\xi|\le\lambda}(1-e^{\psi(\eta)-\psi(\xi)})^2 J^{\langle r\rangle}(\eta,\xi)d\eta \\ &\le c_1\frac{s^2}{9} + \int_{|\eta-\xi|\le\lambda}(\psi(\eta)-\psi(\xi))^2\ e^{2|\psi(\eta)-\psi(\xi)|}J^{\langle r\rangle}(\eta,\xi)d\eta \\ &\le c_1\frac{s^2}{9} + (\frac{s}{3})^2\ e^{2s\lambda/3}\int_{|\eta-\xi|\le\lambda}|\eta-\xi|^2 J^{\langle r\rangle}(\eta,\xi)d\eta \\ &\le c_1\frac{s^2}{9} + cs^2e^{2s\lambda/3}\int_0^\lambda \frac{t}{\phi_r(t)}\,dt \\ &\le c_1\frac{s^2}{9} + cs^2e^{2s\lambda/3}\frac{\lambda^2}{\phi_r(\lambda)} \\ &\le c_2(s^2+\frac{e^{s\lambda}}{\phi_r(\lambda)}), \end{aligned}$$

for every $\xi\in\mathbb{R}^d$. Here we used Lemma 2.1 (ii) of [9] for the fourth inequality and the fifth inequality is by (1.8).

The same estimate holds for $e^{2\psi(\xi)}\Gamma_{r,\lambda}[e^{-\psi}](\xi)$. Denote the constant $c_2 > 0$ by C_* and define

$$F(r,\lambda,s,t,R) := \exp\left(-\frac{sR}{3} + C_*\left(s^2 + \frac{e^{s\lambda}}{\phi_r(\lambda)}\right)t\right). \tag{5.8}$$

Then, by (5.5), with $R = |x-y|$, we have

$$p_r^{(\lambda)}(t,x,y) \le g(r,t)F(r,\lambda,s,t,R). \tag{5.9}$$

Note that there is a freedom to choose $s > 0$ properly. We are now ready to prove Theorem 5.1.

Proof of Theorem 5.1. By (5.3), it suffices to show that

$$p(t,x,y) \le c_1\left(p^c(t,c_2|x-y|) + p^j(t,|x-y|)\right). \tag{5.10}$$

Our proof consists of considering 5 cases. Recall that $R := |x-y|$.

Case 1: $R^2 < t < \phi(R) \le 1$.

Take $r = 1$, $\lambda = R$ and $s = \frac{1}{\sqrt{t}}$ in (5.9). Note that in this case, $g(1,t) = ct^{-d/2}$ and

$$\frac{e^{sR}}{\phi(R)} = \frac{e^{R/\sqrt{t}}}{\phi(R)} < \frac{e}{t} = es^2.$$

So

$$p_1^{(\lambda)}(t,x,y) \le c_1\, t^{-d/2} e^{-\frac{sR}{3} + C_*(1+e)s^2 t} = c_2\, t^{-d/2}\, e^{-\frac{R}{3\sqrt{t}}}.$$

(In fact, $p_1^{(\lambda)}(t,x,y) \le c_1 t^{-d/2}$ in this case.) It follows by Meyer's construction that

$$\begin{aligned}
p(t,x,y) &\le p_1^{(\lambda)}(t,x,y) + t \sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d} J(x,y)\mathbb{1}_{\{|x-y|>\lambda\}}dy \\
&\le c_1\, t^{-d/2}\, e^{-R/(3\sqrt{t})} + c_1\frac{t}{R^d\phi(R)} \\
&\le c_1\, t^{-d/2}\, e^{-R^2/(3t)} + c_1\frac{t}{R^d\phi(R)}.
\end{aligned}$$

The last inequality is due to the assumption that $R^2 < t$. So (5.10) holds in this case.

Case 2: $\phi(R) \le t$.

This is a free lunch as $p^j(t,x,y) \approx c\phi^{-1}(t)^{-d}$ in this case and (5.10) follows.

Let $K = \beta_1/(72C_*(d+\beta_1))$ and let $a = eK/c$, where C_* and c are the positive constants in (5.8) and (1.7), respectively. Before we consider the remaining three cases, let us first do estimate on $F := F(r, \lambda, s, t, R)$ under two situations:

$$(i)\ \ e^{KR^2/t} \geq \frac{a\phi_r(R)}{t} \text{ with } R^2 \geq t, \qquad \text{and} \qquad (ii)\ \ e^{KR^2/t} < \frac{a\phi_r(R)}{t}.$$

Since $\min_{x>0} e^x/x = e$, we have

$$\frac{1}{K} \cdot \frac{t}{\phi_r(R)} e^{KR^2/t} = \frac{R^2}{\phi_r(R)} \cdot \frac{t}{KR^2} e^{KR^2/t} \geq \frac{\widetilde{\phi}(r)R^2}{\phi(rR)} \cdot e,$$

which, by (1.7) is no less than $1/c$ if $\min\{r, R\} \geq 1$ or if $r \leq 1$ but $rR \geq 1$. So Situation (ii) may happen only when $r < 1 \leq R$ and $rR < 1$.

Situation (i)**:** $e^{KR^2/t} \geq \frac{a\phi_r(R)}{t}$ and $R^2 \geq t$.

Let $H = \beta_1/(12(d+\beta_1))$. We take $\lambda = HR$ and $s = (HR)^{-1}\log(e\phi_r(R)/t) > 0$ in (5.8). By (1.7), there is a constant $c_1 > 0$ such that

$$\frac{e^{s\lambda}}{\phi_r(\lambda)} t \leq c_1 \frac{e^{s\lambda}}{\phi_r(R)} t = c_1 e.$$

Moreover, using the assumption,

$$\begin{aligned} C_* s^2 t &= C_* \frac{st}{HR} \log \frac{e\phi_r(R)}{t} = C_* \frac{st}{HR} \log \frac{e}{a} + C_* \frac{st}{HR} \log \frac{a\phi_r(R)}{t} \\ &\leq c_2 \frac{st}{R} + C_* \frac{st}{HR} \frac{KR^2}{t} = s\left(c_2 \frac{t}{R} + \frac{R}{6}\right) \leq \frac{sR}{4} + c_3, \end{aligned}$$

since $K = \beta_1/(72C_*(d+\beta_1)) = H/(6C_*)$. The last inequality is due to that fact that when $R^2/t \geq 12c_2$,

$$s\left(c_2 \frac{t}{R} + \frac{R}{6}\right) \leq s\left(\frac{R}{12} + \frac{R}{6}\right) = \frac{sR}{4},$$

while for $1 \leq R^2/t < 12c_2$,

$$c_2 \frac{st}{R} = c_2 \frac{t}{HR^2} \log\left(e\frac{\phi_r(R)}{t}\right) \leq \frac{c_2}{H} \log\left(\frac{e}{a} e^{12c_2K}\right) =: c_3.$$

So, by (5.8), we have

$$\begin{aligned} F &\leq \exp\left(-\frac{sR}{12} + c_3 + C_* c_1 e\right) \\ &= c_4 \left(\frac{t}{\phi_r(R)e}\right)^{1/(12H)} = c_5 \left(\frac{t}{\phi_r(R)}\right)^{d/\beta_1 + 1}. \end{aligned} \tag{5.11}$$

Situation (ii)**:** $e^{KR^2/t} < \frac{a\phi_r(R)}{t}$.

We take $\lambda = KR/(6C_*)$, $s = R/(6C_* t)$ in (5.8). By (1.7), there is a constant $c > 0$ such that

$$\frac{e^{s\lambda}}{\phi_r(\lambda)} t \le c\,\frac{e^{s\lambda}}{\phi_r(R)}\,t = c\,\frac{e^{KR^2/t}}{\phi_r(R)}\,t \le ca.$$

So

$$\begin{aligned}(5.12)\quad F &\le \exp\left(-\frac{sR}{3} + C_* s^2 t + C_* ca\right)\\ &= c_6 \exp\left(-\frac{sR}{3} + C_* \frac{sR}{6C_*}\right) = c_6 \exp\left(-\frac{sR}{6}\right) = c_6 \exp\left(-\frac{R^2}{6C_* t}\right).\end{aligned}$$

Case 3: $t \le 1 \le R$.

We will take $r = 1$ in this case so by (5.4),

$$g(t,1) = ct^{-d/2} \le ct^{-d/\beta_1}.$$

This case falls into Situation (i) and so we have from (5.9) and (5.11)

$$p_1^{(\lambda)}(t,x,y) \le ct^{-d/\beta_1}\left(\frac{t}{\phi(R)}\right)^{d/\beta_1+1} = c\frac{t}{\phi(R)^{d/\beta_1+1}} \le \frac{c_7 t}{R^d \phi(R)},$$

where we used (1.7) in the last inequality. By Meyer's construction, we conclude

$$(5.13)\qquad p(t,x,y) \le c_8\left(\frac{t}{R^d\,\phi_r(R)} + \frac{t}{R^d \phi(R)}\right) \le \frac{c_8 t}{R^d\,\phi(R)}.$$

This establishes (5.10) in this case.

Case 4: $\phi(R) \ge t \ge 1$.

Let $r = \phi^{-1}(t) \ge 1$, $x' = x/r$ and $y' = y/r$. Since $R \ge r$, $|x'-y'| \ge 1$ so the estimate for $p_r(1,x',y')$ falls into Situation (i). As $g(r,1) = 1$, we have from (5.9), (5.11) and Meyer's construction

$$\begin{aligned} r^d p(\phi(r),x,y) &= p_r(1,x',y')\\ &\le p_r^{(\lambda)}(1,x',y') + \sup_{x\in\mathbb{R}^d}\int_{\mathbb{R}^d} J^{\langle r\rangle}(x,y)\mathbb{1}_{\{|x-y|>\lambda\}}\\ &\le c\left(\frac{1}{\phi_r(|x'-y'|)}\right)^{d/\beta_1+1} + \frac{c}{|x'-y'|^d\phi_r(|x'-y'|)}\\ &\le c_9\,\frac{1}{\phi_r(1)^{d/\beta_1}|x'-y'|^d\,\phi_r(|x'-y'|)} + \frac{c_9}{|x'-y'|^d\phi_r(|x'-y'|)}\\ &\le \frac{c_{10}\phi(r)}{|x'-y'|^d\phi_r(|x-y|)}.\end{aligned}$$

Here we used (1.7) in the second to the last inequality and the fact that $\phi_r(1) \geq 1$ in the last inequality. Since $t = \phi(r)$, we conclude that

$$p(t, x, y) \leq \frac{c_{10} t}{|x-y|^d \, \phi(|x-y|)}.$$

This proves (5.10) in this case.

Case 5: $t < R^2 (\leq \phi(R)) \leq 1$.

Let $r = R = |x-y|$, $x' = x/r$, $y' = y/r$. Note that $\widetilde{\phi}(r) = r^2$ as $r \leq 1$ and $|x' - y'| = 1$. Let $t' = t/r^2 \leq 1$. Note that

$$g(r, t') \leq c(t')^{-d/2} \leq c(t')^{-d/\beta_1}.$$

If $e^{K/t'} \geq a\phi_r(1)/t'$, then we are in Situation (i) for $p_r(t', x', y')$. By the same calculation as that for Case 3, we have

$$r^d p(r^2 t', x, y) = p_r(t', x', y') \leq \frac{c_{11} t'}{|x'-y'|^d \, \phi_r(|x'-y'|)} = \frac{c_{11} t' r^2}{|x'-y'|^d \, \phi(|x-y|)}.$$

Noting $t = t' r^2$, we obtain

$$p(t, x, y) \leq \frac{c_{11} t}{|x-y|^d \, \phi(|x-y|)}.$$

If $e^{K/t'} < a\phi_r(1)/t'$, then we are in Situation (ii) for $p_r(t', x', y')$. So by (5.9), (5.12) and Meyer's construction

$$\begin{aligned} r^d p(r^2 t', x, y) &= p_r(t', x', y') \\ &\leq p_r(t', x', y') + t' \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} J^{\langle r \rangle}(x, y) \mathbb{1}_{\{|x-y| > \lambda\}} dy \\ &\leq c_{12} t'^{-d/2} \exp\left(-\frac{c_{13}|x'-y'|^2}{t'} \right) + \frac{c_{14} t'}{|x'-y'|^d \phi_r(|x'-y'|)}. \end{aligned}$$

Noting $t = t' r^2$, we obtain

$$p(t, x, y) \leq c_{15} t^{-d/2} \exp(-\frac{c_{16}|x-y|^2}{t}) + \frac{c_{17} t}{|x-y|^d \, \phi(|x-y|)}.$$

This proves the claim (5.10).

The upper bound estimate in (5.1) is now established for every $t > 0$ and $x, y \in \mathbb{R}^d$. ∎

6. Heat kernel lower bound estimate

Recall that $\widetilde{\phi}(t) := t^2 \wedge \phi(t)$ and so $\widetilde{\phi}^{-1}(t)^{-d} = t^{-d/2} \wedge \phi^{-1}(t)^{-d}$. In this section, we will establish the following.

Theorem 6.1 *There exist positive constants c_1 and c_2 such that*

$$p(t,x,y) \geq c_1 \, \widetilde{\phi}^{-1}(t)^{-d} \wedge \left(p^c(t, c_2|x-y|) + p^j(t, |x-y|)\right) \tag{6.1}$$

for each $x, y \in \mathbb{R}^d$ and $t > 0$.

To prove it, we need first establish some tightness results and extend Lemma 4.1 to all $r > 0$ and Theorem 1.3 to all $R > 0$.

6.1. Tightness and some lower bound estimate

Using the heat kernel upper bound, we can prove the following estimate of the exit time from a ball.

Proposition 6.2 *For each $A > 0$ and $0 < B < 1$, there exists $\gamma = \gamma(A, B) \in (0, 1/2)$ such that for every $r > 0$ and $x \in \mathbb{R}^d \setminus \mathcal{N}$,*

$$\mathbb{P}_x \left(\tau_{B(x,\, Ar)} < \gamma \, \widetilde{\phi}(r) \right) \leq B.$$

Proof. Let $x \in \mathbb{R}^d \setminus \mathcal{N}$. By the upper bound estimate in (5.1), for every $s > 0$ and $t > 0$,

$$\begin{aligned}
\mathbb{P}_x \left(|X_t - x| \geq s \right) &= \int_{B(x,s)^c} p(t,x,y)dy \\
&\leq \int_{B(x,s)^c} \frac{c_1 \, t dy}{|x-y|^d \phi(c_1|x-y|)} + c_2 t^{-d/2} \int_{B(x,s)^c} \exp(-\frac{c_3|x-y|^2}{t}) dy \\
&\leq \frac{c_4 t}{\phi(s)} + c_5 \exp(-\frac{c_6 s^2}{t}) \leq \frac{c_4 t}{\phi(s)} + \frac{c_7 t}{s^2} \leq \frac{c_8 t}{\widetilde{\phi}(s)}.
\end{aligned}$$

The above computation is standard; see Lemma 2.1 (i) in [9] for the estimate of the stable part in the second inequality, and [1] Lemma 3.9 (a) for the estimate of the Gaussian part in the second inequality. Given this inequality, the rest of the proof is the same as that of Proposition 4.9 in [9] with $\widetilde{\phi}$ in place of ϕ for the case of $\gamma_1 = \gamma_2 = 0$ there. ■

Using Proposition 6.2, one can prove the following proposition in the same way as the proof of Proposition 4.11 in [9] but with $\widetilde{\phi}$ in place of ϕ for the case of $\gamma_1 = \gamma_2 = 0$ there.

Proposition 6.3 *There exist constants* $c_1 \geq 2$ *and* $c_2 > 0$ *such that for every* $t > 0$ *and every* $x, y \in \mathbb{R}^d \setminus \mathcal{N}$ *with*

$$\mathbb{P}_x\left(X_t \in B(y, c_1\widetilde{\phi}^{-1}(t))\right) \geq c_2 \frac{t(\widetilde{\phi}^{-1}(t))^d}{|x-y|^d \widetilde{\phi}(|x-y|)}. \tag{6.2}$$

6.2. Parabolic Harnack Inequality

Denote $\gamma(1/2, 1/2)$ in Proposition 6.2 by γ_0. For each $r, t > 0$, we define

$$Q(t, x, r) := [t, t + \gamma_0\widetilde{\phi}(r)] \times B(x, r).$$

The following is an extension of Lemma 4.3 to all $r > 0$.

Lemma 6.4 *There exists* $C_1 > 0$ *such that for every* $x \in \mathbb{R}^d$*,* $r > 0$*,* $y \in B(x, r/3)$ *and a bounded nonnegative function* h *on* $[0, \infty) \times \mathbb{R}^d$ *that is supported in* $[0, \infty) \times B(x, 2r)^c$*,*

$$\mathbb{E}^{(\gamma_0\widetilde{\phi}(r), x)}\left[h(\tau_r, X_{\tau_r})\right] \leq C_1 \mathbb{E}^{(\gamma_0\widetilde{\phi}(r), y)}\left[h(\tau_r, X_{\tau_r})\right], \tag{6.3}$$

where $\tau_r = \tau_{Q(0,x,r)}$.

Proof. The proof is the same as Lemma 6.1 in [9]. Note that the continuous component of the process does not play any role since the function h is supported in $[0, \infty) \times B(x, 2r)^c$. (Note that in [9] the space-time process is running forward in the sense that $V_t = V_0 + t$ there while in this paper $V_t = V_0 - t$ is defined to run backward. Clearly there is one-to-one correspondence between these two situations. Thus the estimate in Lemma 6.1 in [9] is under probability law $\mathbb{P}^{(0,x)}$ while here it is under $\mathbb{P}^{(\gamma_0\widetilde{\phi}(r), x)}$. The same remark applies in the following when [9] is cited, for example, in the proof of the next three results.) ■

For each $A \subset [0, \infty) \times \mathbb{R}^d$, denote $\sigma_A := \inf\{t > 0 : Z_t \in A\}$.

Lemma 6.5 *There exists* $C_2 > 0$ *such that for all* $x \in \mathbb{R}^d$*,* $r > 0$ *and any compact subset* $A \subset Q(0, x, r)$*,*

$$\mathbb{P}^{(\gamma_0\widetilde{\phi}(r), x)}(\sigma_A < \tau_r) \geq C_2 \frac{m_{d+1}(A)}{r^d \widetilde{\phi}(r)},$$

where $\tau_r = \tau_{Q(0,x,r)}$.

Proof. When $r \leq 1$, this is proved in Lemma 4.1. When $r \geq 1$, we have $\widetilde{\phi}(r) = \phi(r)$ so the desired inequality can be proved similarly to Lemma 6.2 in [9]. ■

Define $U(t,x,r) := \{t\} \times B(x,r)$.

Corollary 6.6 *For every* $0 < \delta \le \gamma_0$*, there exists* $C_3 > 0$ *such that for every* $R \in (0,1]$*,* $r \in (0, R/4]$ *and* $(t,x) \in Q(0,z,R/3)$ *with* $0 < t \le \gamma_0 \widetilde{\phi}(R/3) - \delta\widetilde{\phi}(r)$*,*

$$\mathbb{P}^{(\gamma_0\widetilde{\phi}(R/3),z)}\left(\sigma_{U(t,x,r)} < \tau_{Q(0,z,R)}\right) \ge C_3 \frac{r^d \widetilde{\phi}(r)}{R^d \widetilde{\phi}(R)}.$$

Proof. Given Lemma 6.5 and Proposition 6.2, the proof is the same as Corollary 6.3 in [9] but with $\widetilde{\phi}$ in place of ϕ there. ■

The following extends the parabolic Harnack principle in Theorem 1.3 to all $R > 0$.

Theorem 6.7 *For every* $0 < \delta \le \gamma_0$*, there exists* $c_1 > 0$ *such that for every* $z \in \mathbb{R}^d$*,* $R > 0$ *and every non-negative function* h *on* $[0,\infty) \times \mathbb{R}^d$ *that is parabolic on* $[0, \gamma\widetilde{\phi}(2R)] \times B(z, 2R)$*,*

$$\sup_{(t,y)\in Q(\delta\widetilde{\phi}(R),z,R)} h(t,y) \le c_1 \inf_{y\in B(z,R)} h(0,y).$$

In particular, the following holds for $t > 0$*.*

$$\sup_{(s,y)\in Q((1-\gamma)t,z,\widetilde{\phi}^{-1}(t))} p(s,x,y) \le c \inf_{y\in B(z,\widetilde{\phi}^{-1}(t))} p((1+\gamma)t,x,y). \tag{6.4}$$

Proof. Given Lemma 6.4, Lemma 6.5 and Corollary 6.6, the proof of this PHI is the same as that of Theorem 4.12 in [9] plus the last paragraph of that for Theorem 4.1 in [7]. (See also the proof of Theorem 4.5 in [14].) ■

6.3. Lower bound

Lemma 6.8 *There exist* c_1*,* $c_2 > 0$ *such that*

$$p(t,x,y) \ge c_1 \left(\widetilde{\phi}^{-1}(t)\right)^{-d}$$

for all $t > 0$ *and* $x, y \in \mathbb{R}^d \setminus \mathcal{N}$ *with* $|x-y| \le c_2\, \widetilde{\phi}^{-1}(t)$*.*

Proof. This is already proved in Theorem 3.1 for $t \le 1$. Given (5.1), Proposition 6.2, and Theorem 6.7, the proof is the same as that of Lemma 4.13 in [9] but with $\widetilde{\phi}$ in place of ϕ there. ■

Proof of Theorem 6.1. Let $t > 0$. Due to Lemma 6.8, it is enough to prove the theorem for $|x-y| \geq c_2\,\widetilde{\phi}^{-1}(t)$. Applying Proposition 6.3 with $t_* = (1-\gamma)t$ in place of t, we have

$$\mathbb{P}_x(X_{t_*} \in B(y,\, c_1\widetilde{\phi}^{-1}(t_*))) \geq c_2 \frac{t_*(\widetilde{\phi}^{-1}(t_*))^d}{|x-y|^d \phi(c_3|x-y|)}.$$

As $m_d(B(y,\, c_1\phi^{-1}(t_*))) \leq c_4\,(\phi^{-1}(t_*))^d$, the above implies $p\,(t_*, x, z) \geq c_5 t/(|x-y|^d \phi(c_3|x-y|))$ for some $z \in B(y,\, c_1\phi^{-1}(t_*))$. By applying (6.4) as before, we have

$$p(t,x,y) \geq c\frac{t}{|x-y|^d \phi(|x-y|)}.$$

For (6.1), the exponential decay appears on its right hand side only when $t < r^2 (\leq \phi(r)) \leq 1$ (Case 4 in the upper bound), where $r = |x-y|$. So, the only case left is this case. In this case, choose $N \in \mathbb{N}$ so that $s := t/N \asymp (r/N)^2$ (so $N \asymp r^2/t$). Then, $p(s,x,y) \geq cs^{-d/2}$, by Lemma 6.8. Thus the usual chain argument gives $p(t,x,y) \geq ct^{-d/2}\exp(-c'r^2/t)$. ■

References

[1] Barlow, M. T.: Diffusions on fractals. In *Lectures on probability theory and statistics* (*Saint-Flour, 1995*), 1–21. Lectures Notes in Math **1690**. Springer, Berlin, 1998.

[2] Barlow, M. T., Bass, R. F., Chen, Z. Q. and Kassmann, M.: Nonlocal Dirichlet forms and symmetric jump processes. *Trans. Amer. Math. Soc.* **361** (2009), no. 4, 1963–1999.

[3] Barlow, M. T., Bass, R. F. and Kumagai, T.: Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps. *Math. Z.* **261** (2009), no. 2, 297–320.

[4] Barlow, M. T., Grigor'yan, A. and Kumagai, T.: Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.* **626** (2009), 135–157.

[5] Bass, R. F., Kumagai, T. and Uemura, T.: Convergence of symmetric Markov chains on $\mathbb{Z}^d$. To appear in *Probab. Theory Relat. Fields.*

[6] Carlen, E. A., Kusuoka, S. and Stroock, D. W.: Upper bounds for symmetric Markov transition functions. *Ann. Inst. H. Poincaré Probab. Statist.* **23** (1987), no. 2, suppl., 245–287.

[7] Chen, Z. Q., Kim, P. and Kumagai, T.: Weighted Poincaré inequality and heat kernel estimates for finite range jump processes. *Math. Ann.* **342** (2008), no. 4, 833–883.

[8] Chen, Z. Q. and Kumagai, T.: Heat kernel estimates for stable-like processes on d-sets. *Stochastic Process. Appl.* **108** (2003), no. 1, 27–62.

[9] Chen, Z. Q. and Kumagai, T.: Heat kernel estimates for jump processes of mixed types on metric measure spaces. *Probab. Theory Related Fields* **140** (2008), no. 1-2, 277–317.

[10] Chung, K. L.: Greenian bounds for Markov processes. *Potential Anal.* **1** (1992), no. 1, 83–92.

[11] Fabes, E. B. and Stroock, D. W.: A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Rational Mech. Anal.* **96** (1986), no. 4, 327–338.

[12] Foondun, M.: Heat kernel estimates and Harnack inequalities for some Dirichlet forms with non-local part. *Electron. J. Probab.* **14** (2009), no. 11, 314–340.

[13] Saloff-Coste, L.: *Aspects of Sobolev-type inequalities.* London Mathematical Society Lecture Note Series, **289**. Cambridge University Press, Cambridge, 2002.

[14] Song, R. and Vondraček, Z.: Parabolic Harnack inequality for the mixture of Brownian motion and stable process. *Tohoku Math. J.* (*2*) **59** (2007), no. 1, 1–19.

[15] Stroock, D. W.: Diffusion semigroup corresponding to uniformly elliptic divergence form operator. In *Séminaire de Probabilités, XXII*, 316–347. Lecture Notes in Math. **1321**. Springer, Berlin, 1988.

[16] Stós, A.: Symmetric α-stable processes on d-sets. *Bull. Polish Acad. Sci. Math.* **48** (2000), no. 3, 237–245.

Recibido: 18 de agosto de 2008
Revisado: 26 de enero de 2009

Zhen-Qing Chen
Department of Mathematics
University of Washington
Seattle, WA 98195, USA
zchen@math.washington.edu

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

The first author is partially supported by NSF Grant DMS-0600206. The second author is partially supported by the Grant-in-Aid for Scientific Research (B) 18340027.