# Weighted Poincaré Inequality and Heat Kernel Estimates for Finite Range Jump Processes 

Zhen-Qing Chen, ${ }^{*}$ Panki Kim ${ }^{\dagger}$ and Takashi Kumagai ${ }^{\ddagger}$


#### Abstract

It is well-known that there is a deep interplay between analysis and probability theory. For example, for a Markovian infinitesimal generator $\mathcal{L}$, the transition density function $p(t, x, y)$ of the Markov process associated with $\mathcal{L}$ (if it exists) is the fundamental solution (or heat kernel) of $\mathcal{L}$. A fundamental problem in analysis and in probability theory is to obtain sharp estimates of $p(t, x, y)$. In this paper, we consider a class of non-local (integro-differential) operators $\mathcal{L}$ on $\mathbb{R}^{d}$ of the form $$
\mathcal{L} u(x)=\lim _{\varepsilon \not 00} \int_{\left\{y \in \mathbb{R}^{d}:|y-x|>\varepsilon\right\}}(u(y)-u(x)) J(x, y) d y,
$$ where $J(x, y)=\frac{c(x, y)}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y| \leq \kappa\}}$ for some constant $\kappa>0$ and a measurable symmetric function $c(x, y)$ that is bounded between two positive constants. Associated with such a nonlocal operator $\mathcal{L}$ is an $\mathbb{R}^{d}$-valued symmetric jump process of finite range with jumping kernel $J(x, y)$. We establish sharp two-sided heat kernel estimate and derive parabolic Harnack principle for them. Along the way, some new heat kernel estimates are obtained for more general finite range jump processes that were studied in [BBCK]. One of our key tools is a new form of weighted Poincaré inequality of fractional order, which corresponds to the one established by Jerison [J] for differential operators. Using Meyer's construction of adding new jumps, we also obtain various a priori estimates such as Hölder continuity estimates for parabolic functions of jump processes (not necessarily of finite range) where only a very mild integrability condition is assumed for large jumps. To establish these results, we employ methods from both probability theory and analysis extensively.


AMS 2000 Mathematics Subject Classification: Primary 60J75, 60J35, Secondary 31C25, 31C05.

## Running title: Heat Kernel Estimate for Finite Range Jump Process.

[^0]
## 1 Introduction and Main Results

The second order elliptic differential operators and diffusion processes take up, respectively, an central place in the theory of partial differential equations (PDE) and in probability theory, see [GT] and $[\mathrm{IW}]$ for example. There are close relationships between these two subjects. For a large class of second order elliptic differential operators $\mathcal{L}$ on $\mathbb{R}^{d}$, there is a diffusion process $X$ 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 $X$.

Recently there are intense interests in studying discontinuous Markov processes, due to their importance both in theory and in application. See, for example, [B, JW, KSZ] and the references therein. The infinitesimal generator of a discontinuous Markov process in $\mathbb{R}^{d}$ is no longer a differential operator but rather a non-local (or, integro-differential) operator. For example, the infinitesimal generator of a isotropically symmetric $\alpha$-stable process in $\mathbb{R}^{d}$ with $\alpha \in(0,2)$ is a fractional Laplacian operator $c(-\Delta)^{\alpha / 2}$. Recently there are also many interests from the theory of PDE (such as singular obstacle problems) to study non-local operators; see, for example, [CSS, S] and the references therein.

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

$$
\mathcal{L} u(x)=\lim _{\varepsilon \downarrow 0} \int_{\left\{y \in \mathbb{R}^{d}:|y-x|>\varepsilon\right\}}(u(y)-u(x)) J(x, y) d y,
$$

where

$$
\begin{equation*}
J(x, y)=\frac{c(x, y)}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y| \leq \kappa\}} \tag{1.1}
\end{equation*}
$$

for some constant $\kappa>0$ and a measurable symmetric function $c(x, y)$ that is bounded between two positive constants. Associated with such a non-local operator $\mathcal{L}$ is an $\mathbb{R}^{d}$-valued finite range symmetric jump process $X$ with jumping kernel $J(x, y)$. We will be concerned with obtaining sharp two-sided heat kernel estimates for $\mathcal{L}$ (or, equivalently, for $X$ ), as well as establishing parabolic Harnack inequality and a priori joint Hölder continuity estimate for parabolic functions of $\mathcal{L}$. Our approach employs a combination of probabilistic and analytic techniques.

Two-sided heat kernel estimates for diffusions (or second order elliptic differential operators) have a long history and many beautiful results have been established. But two-sided heat kernel estimates for jump processes in $\mathbb{R}^{d}$ have only been studied recently. In $[\mathrm{K}]$, Kolokoltsov obtained two-sided heat kernel estimates for certain stable-like processes in $\mathbb{R}^{d}$, whose infinitesimal generators are a class of pseudo-differential operators having smooth symbols. Bass and Levin [BL] used a completely different approach to obtain similar estimates for discrete time Markov chain on $\mathbb{Z}^{d}$ where the conductance between $x$ and $y$ is comparable to $|x-y|^{-d-\alpha}$ for $\alpha \in(0,2)$. In [CK1], two-sided heat kernel estimates and a scale-invariant parabolic Harnack inequality for symmetric $\alpha$-stable-like processes on $d$-sets are obtained. (See [HK] for some extensions.) Very recently in [CK2], parabolic Harnack inequality and two-sided heat kernel estimates are even established for non-local operators of variable order. But so far the two-sided heat kernel estimates for non-local
operators have been established only for the case that the jumping kernel has full support on the state space. See [BBCK] for some result on parabolic Harnack inequality and heat kernel estimate for non-local operators of variable order on $\mathbb{R}^{d}$, whose jumping kernel is supported on jump size less than or equal to 1 .

Throughout this paper, $d \geq 1$ and $\alpha \in(0,2)$. Let the jump kernel $J$ be defined by (1.1) and let

$$
\begin{align*}
\mathcal{Q}(u, v) & :=\frac{1}{2} \int_{\mathbb{R}^{d}} \int_{\mathbb{R}^{d}}(u(x)-u(y))(v(x)-v(y)) J(x, y) d x d y  \tag{1.2}\\
\mathcal{D} & :=\left\{f \in L^{2}\left(\mathbb{R}^{d}, d x\right): \mathcal{Q}(f, f)<\infty\right\} \tag{1.3}
\end{align*}
$$

It is easy to check that $(\mathcal{Q}, \mathcal{D})$ is a regular Dirichlet form on $\mathbb{R}^{d}$ and so there is a Hunt process $X$ associated with it. When the jumping kernel $J(x, y)$ is the unrestricted $\frac{c(x, y)}{|x-y|^{d+\alpha}}$, the associated process is the symmetric stable-like process $Y$ on $\mathbb{R}^{d}$ studied in [CK1]. Among other things, it is shown in [CK1] that $Y$ has Hölder continuous transition density function and so $Y$ can be modified to start from every $x \in \mathbb{R}^{d}$. Since $X$ can be constructed from $Y$ by removing jumps of size larger than $\kappa$ via Meyer's construction (see [BBCK, BGK]), $X$ is conservative and can be modified to start from every point in $\mathbb{R}^{d}$. For this reason, in the sequel, we will call such $X$ a finite range (or truncated) stable-like process. It is proved in [BBCK, Theorem 3.1] that there is a properly exceptional set $\mathcal{N} \subset \mathbb{R}^{d}$ and a positive symmetric kernel $p(t, x, y)$ defined on $(0, \infty) \times\left(\mathbb{R}^{d} \backslash \mathcal{N}\right) \times\left(\mathbb{R}^{d} \backslash \mathcal{N}\right)$ such that $p(t, x, y)$ is the transition density function of $X$ (starting from $x \in \mathbb{R}^{d} \backslash \mathcal{N}$ ) with respect to the Lebesgue measure on $\mathbb{R}^{d}$, and for each $y \in \mathbb{R}^{d} \backslash \mathcal{N}$ and $t>0, x \mapsto p(t, x, y)$ is quasi-continuous. It is this version of the transition density function of $X$ we will take throughout this paper. Here a set $\mathcal{N} \subset \mathbb{R}^{d}$ is called properly exceptional with respect to the process $X$ if it has zero Lebesgue measure and

$$
\mathbb{P}^{x}\left(\left\{X_{t}, X_{t-}\right\} \subset \mathbb{R}^{d} \backslash \mathcal{N} \text { for every } t>0\right)=1 \quad \text { for } x \in \mathbb{R}^{d} \backslash \mathcal{N}
$$

It is well-known (see [FOT]) that every exceptional set is $\mathcal{Q}$-polar and every $\mathcal{Q}$-polar set is contained in a properly exceptional set. Later we will show in Theorem 4.3, $p(t, x, y)$ in fact has a Hölder continuous version and so we can take $\mathcal{N}=\emptyset$. The purpose of this paper is to obtain sharp upper and lower estimates on $p(t, x, y)$. The jump size cutoff constant $\kappa$ in (1.1) plays no special role, so for convenience we will simply take $\kappa=1$ for the rest of this paper.

When $c(x, y)$ is a constant, $X$ is a finite range (also called truncated) isotropically symmetric $\alpha$-stable process in $\mathbb{R}^{d}$ with jumps of size larger than 1 removed. The potential theory of this Lévy process is studied in [KS1, KS2]. One interesting fact is that, even though scale-invariant elliptic Harnack principle is true for such a process, the boundary Harnack principle is only valid for the positive harmonic functions of this process in bounded convex domains (see the last section of [KS1] for a counterexample). Since the parabolic Harnack principle implies elliptic Harnack principle, our Theorem 4.1 extends the result on Harnack principle in [KS1] to the case that $c(x, y)$ is not necessarily constant.

Finite range stable processes, more generally finite range jump processes, are very important both in theory and in application. Finite range jump processes are very natural in applications
where jumps only up to a certain size are allowed. Moreover, in some aspects, finite range jump processes have nicer behaviors and are more preferable than unrestricted jump processes. For instance, in [CR], to show certain property of Schramm-Loewner evolution driven by symmetric stable processes, finite range (or truncated) stable process has been used as a tool. However, as we shall see below, in some other respects, finite range jump processes are much more delicate to study than unrestricted jump processes.

In the sequel, for two non-negative functions $f$ and $g$, the notation $f \asymp g$ means that there are positive constants $c_{1}, c_{2}, c_{3}$ and $c_{4}$ so that $c_{1} g\left(c_{2} x\right) \leq f(x) \leq c_{3} g\left(c_{4} x\right)$ in the common domain of definition for $f$ and $g$. The Euclidean distance between $x$ and $y$ will be denoted as $|x-y|$. For $a, b \in \mathbb{R}, a \wedge b:=\min \{a, b\}$ and $a \vee b:=\max \{a, b\}$. We will use $\mu_{d}$ or $d x$ to denote the Lebesgue measure in $\mathbb{R}^{d}$. A statement that is said to be hold quasi-everywhere (q.e. in abbreviation) on a set $A \subset \mathbb{R}^{d}$ if there is an $\mathcal{Q}$-polar set $\mathcal{N}$ such that the statement holds for every point in $A \backslash \mathcal{N}$.

Our theorems on the heat kernel estimate on $p(t, x, y)$ can be stated as follows (in the figure, $R_{*}$ is a constant in $\left.(0,1)\right)$ :
(i) [Proposition 2.1 and Theorem 3.6] In the regions $D_{1}$ and $D_{2}$, we have

$$
p(t, x, y) \asymp\left(t^{-d / \alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\right) .
$$

(More precisely, $p(t, x, y) \asymp t^{-d / \alpha}$ in $D_{1}$ and $p(t, x, y) \asymp \frac{t}{|x-y|^{d+\alpha}}$ in $D_{2}$.)
(ii) [Theorem 2.3 and Theorem 3.6] In the region $D_{3}$, we have

$$
p(t, x, y) \asymp\left(\frac{t}{|x-y|}\right)^{c|x-y|}=\exp \left(-c|x-y| \log \frac{|x-y|}{t}\right) .
$$

(iii) [Theorem 2.3 and Theorem 3.6] In the regions $D_{4}$ and $D_{5}$, we have

$$
p(t, x, y) \asymp t^{-d / 2} \exp \left(-\frac{c|x-y|^{2}}{t}\right) .
$$

(More precisely, $p(t, x, y) \asymp t^{-d / 2}$ in $D_{4}$ and $p(t, x, y) \asymp t^{-d / 2} \exp \left(-\frac{c|x-y|^{2}}{t}\right)$ in $D_{5}$.)


As we see, the heat kernel estimate is of $\alpha$-stable type in (i), of Poisson type in (ii) and of Gaussian type in (iii). Such behavior of the heat kernel, in particular (i) and (iii), may be useful in applications. For example, in mathematical finance, it has been observed that even though discontinuous stable processes provide better representations of financial data than Gaussian processes (cf. [HPR]), financial data tend to become more Gaussian over a longer time-scale (see [M] and the references therein). Our heat kernel estimates show that finite range stable-like processes have this type of property: they behave like discontinuous stable processes in small scale and behave like Brownian motion in large scale. Furthermore, they avoid large sizes of jumps which can be considered as impossibly huge changes of financial data in short time.

In fact, some of our heat kernel estimates for $t \geq 1$ will be stated and proved for a more general class of finite range jump processes that is studied in [BBCK] (see (2.16), Theorems 2.4 and 3.5 below). These heat kernel estimates improve the estimates given in [BBCK, Theorems 1.2 and 1.3] significantly. They are also used in Section 4 to show the two-sided estimates for Green functions of these processes for $|x-y| \geq 1$.

To get the near diagonal lower bound of the heat kernel $p(t, x, y)$, we introduce and prove a general scaling version of weighted Poincaré inequality of fractional order (see Theorem 5.1 below). This inequality may be of independent interest. (For the details on (weighted) Poincaré inequality and lower bound estimate of heat kernels for diffusions, we refer our readers to [FS, J, SC, SS] and the references therein.) The proof of our weighted Poincaré inequality is quite long and involved. To keep the flow of the main ideas of our proof for the heat kernel estimates, we put the proof of the weighted Poincaré inequality in the last section. We hope that the establishing of such a scaling version of weighted Poincaré inequality and its usage in getting the heat kernel lower bound estimate will shed new light on our understanding of the heat kernel behavior of more general Markov processes.

Using the heat kernel estimates, we derive the parabolic Harnack inequality for the finite range jump processes. Our proof uses a combination of the techniques developed in [CK1, CK2] and in $[\mathrm{BBK}, \mathrm{BBCK}]$. As a direct consequence of the heat kernel estimates, we derive a two-sided sharp estimate for Green functions in $\mathbb{R}^{d}$ for $d \geq 3$. From the heat kernel estimates and the parabolic Harnack inequality, we also obtain the Hölder continuity of the parabolic functions of finite range stable-like processes. In particular, we note that the Hölder continuity for bounded parabolic functions is a consequence of the local heat kernel estimate, while the parabolic Harnack inequality at small size scale can be obtained from the local heat kernel estimate and some mild condition on the jumping kernel for large jumps. This allows us to establish the parabolic Harnack inequality and the joint Hölder continuity for parabolic functions for a larger class of symmetric processes that can be obtained from finite range stable-like process by adding larger jumps with uniformly bounded (total) jumping intensity for those jumps of size larger than 1 through Meyer's construction. See Theorem 4.5 for details.

The remainder of this paper is organized as follows. In Section 2, we prove the upper bound estimates of the heat kernel. Section 3 contains the results on the lower bound estimates of the heat kernel. In Section 4, we establish parabolic Harnack principle and the two-sided estimates for Green functions of the finite range jump processes as well as Hölder continuity of heat kernels. In
the last section, we give the proof of weighted Poincaré inequality of fractional order.

## 2 Heat Kernel Upper Bound Estimate

In this section, we will state the results on the upper bound estimates of the heat kernel for the finite range symmetric $\alpha$-stable-like process $X$ more precisely and present proofs. Most of the heat kernel estimates in this section and next one are established for quasi-everywhere (q.e.) point in $\mathbb{R}^{d}$. However in Theorem 4.3 of Section 4, we will show that the heat kernels of finite range stable processes are Hölder continuous and therefore these estimates hold for every point in $\mathbb{R}^{d}$.

Proposition 2.1 (i) For each $T^{*}>0$, there exists $c_{1}=c_{1}\left(T^{*}\right)>0$ such that

$$
p(t, x, y) \leq c_{1}\left(t^{-d / \alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\right)
$$

for all $t \in\left(0, T^{*}\right]$ and q.e. $x, y \in \mathbb{R}^{d}$.
(ii) There exist $0<R_{*}<1$ and $c_{2}>0$ such that

$$
c_{2}\left(t^{-d / \alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\right) \leq p(t, x, y)
$$

for all $t \in\left(0, T_{*}\right]$ and q.e. $x, y \in \mathbb{R}^{d}$ with $|x-y| \in\left(0, R_{*}\right]$ where $T_{*}:=R_{*}^{\alpha}$.

Proof. The estimates on these regions can be deduced from the existing results. Let $p_{0}(t, x, y)$ be the transition density function of stable-like process $Y$ on $\mathbb{R}^{d}$ whose jumping kernel is $\frac{c(x, y)}{|x-y|^{d+\alpha}}$. Since $X$ can be constructed from $Y$ by removing jumps of size larger than 1 via Meyer's construction, by [BBCK, Lemma 3.6] and [BGK, Lemma 3.1(c)] we have

$$
p(t, x, y) \leq e^{t\|\mathcal{J}\|_{\infty}} p_{0}(t, x, y) \quad \text { and } \quad p_{0}(t, x, y) \leq p(t, x, y)+t\left\|J_{1}\right\|_{\infty}
$$

where

$$
J_{1}(x, y):=\frac{c(x, y)}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y|>1\}} \quad \text { and } \quad \mathcal{J}(x):=\int_{\mathbb{R}^{d}} J_{1}(x, y) d y
$$

Applying the estimates on $p_{0}(t, x, y)$ in [CK1] to the above two inequalities, we have

$$
\begin{equation*}
p(t, x, y) \leq c_{1} e^{t\|\mathcal{J}\|_{\infty}}\left(t^{-d / \alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\right) \tag{2.1}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{1}{c_{1}}\left(t^{-d / \alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\right)-t\left\|J_{1}\right\|_{\infty} \leq p(t, x, y) \tag{2.2}
\end{equation*}
$$

Now (i) follows immediately from (2.1). Since

$$
\frac{1}{2 c_{1}} t^{-d / \alpha} \leq \frac{1}{c_{1}} t^{-d / \alpha}-t\left\|J_{1}\right\|_{\infty} \quad \text { if } t \leq\left(2 c_{1}\left\|J_{1}\right\|_{\infty}\right)^{-\frac{\alpha}{d+\alpha}}
$$

and

$$
\frac{1}{2 c_{1}} \frac{t}{|x-y|^{d+\alpha}} \leq \frac{1}{c_{1}} \frac{t}{|x-y|^{d+\alpha}}-t\left\|J_{1}\right\|_{\infty} \quad \text { if }|x-y| \leq\left(2 c_{1}\left\|J_{1}\right\|_{\infty}\right)^{-\frac{1}{d+\alpha}}
$$

we get (ii) from (2.2).
The discrete Markov chain analogue of the following result is established in [BL, Proposition 2.1]. See also [CKS, Section 2] where the following result is discussed when $c(x, y)$ is a constant and $\alpha=1$.

Proposition 2.2 There exist $c_{1}, c_{2}>0$ such that

$$
p(t, x, y) \leq \begin{cases}c_{1} t^{-d / \alpha} & \text { for } t \in(0,1]  \tag{2.3}\\ c_{2} t^{-d / 2} & \text { for } t \in[1, \infty)\end{cases}
$$

Proof. By Proposition 2.1(i), we only need to show (2.3) for $t \in[1, \infty)$.
Let $\left(\mathcal{E}^{0}, \mathcal{F}^{0}\right)$ be the Dirichlet form for the finite range isotropically symmetric $\alpha$-stable process with jumps of size larger than 1 removed. That is,

$$
\begin{aligned}
\mathcal{E}^{0}(u, u) & =\int_{\mathbb{R}^{d} \times \mathbb{R}^{d}}(u(x)-u(y))^{2} \frac{c_{0}(d, \alpha)}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y| \leq 1\}} d x d y \\
\mathcal{F}^{0} & =\left\{u \in L^{2}\left(\mathbb{R}^{d}, d x\right): \mathcal{E}^{0}(u, u)<\infty\right\}
\end{aligned}
$$

where $c_{0}(d, \alpha)>0$ is a constant. Note that $\mathcal{D} \subset \mathcal{F}^{0}$ and there is a constant $\kappa:=\kappa(d, \alpha)>0$ such that

$$
\begin{equation*}
\mathcal{E}^{0}(u, u) \leq \kappa \mathcal{Q}(u, u) \quad \text { for } u \in \mathcal{F} \tag{2.4}
\end{equation*}
$$

By the Fourier transform, we have

$$
\mathcal{E}^{0}(f, g)=c_{0} \int_{\mathbb{R}^{d}} \hat{g}(\xi) \overline{\hat{f}}(\xi) \phi(\xi) d \xi
$$

where $\hat{f}(\xi):=(2 \pi)^{-d / 2} \int_{\mathbb{R}^{d}} e^{i \xi \cdot y} f(y) d y$ is the Fourier transform of $u$ and

$$
\begin{equation*}
\phi(\xi):=\int_{\{|y|<1\}} \frac{1-\cos (\xi \cdot y)}{|y|^{d+\alpha}} d y \tag{2.5}
\end{equation*}
$$

By the change of variable $y=x /|\xi|$, we have from (2.5)

$$
\begin{equation*}
\phi(\xi)=|\xi|^{\alpha} \int_{\{|x|<|\xi|\}} \frac{1-\cos \left(\frac{\xi}{|\xi|} \cdot x\right)}{|x|^{d+\alpha}} d x . \tag{2.6}
\end{equation*}
$$

Note that $1-\cos \left(\frac{\xi}{|\xi|} \cdot x\right)$ behaves like $|x|^{2}$ for small $|x|$. Moreover, as $|\xi|$ goes to infinity, the integral in the above equation goes to a positive constant. Thus it is easy to see that there exist $M>1$ and $c_{1}>0$ such that

$$
\phi(\xi) \geq \begin{cases}c_{1}|\xi|^{\alpha}, & \text { for all }|\xi|>M \\ c_{1}|\xi|^{2}, & \text { for all }|\xi| \leq M\end{cases}
$$

Thus for every $r \leq 1$, we have

$$
\begin{aligned}
\int_{\{|\xi|>r\}}|\hat{f}(\xi)|^{2} d \xi & \leq \int_{\{M>|\xi|>r\}}\left(\frac{|\xi|}{r}\right)^{2}|\hat{f}(\xi)|^{2} d \xi+\int_{\{|\xi| \geq M\}}\left(\frac{|\xi|}{r}\right)^{\alpha}|\hat{f}(\xi)|^{2} d \xi \\
& \leq c_{2}\left(r^{-2} \int_{\{M>|\xi|>r\}} \phi(\xi)|\hat{f}(\xi)|^{2} d \xi+r^{-\alpha} \int_{\{|\xi| \geq M\}} \phi(\xi)|\hat{f}(\xi)|^{2} d \xi\right) \\
& \leq c_{2} r^{-2} \int_{\mathbb{R}^{d}} \phi(\xi)|\hat{f}(\xi)|^{2} d \xi=c_{3} r^{-2} \mathcal{E}^{0}(f, f) \leq c_{3} \kappa r^{-2} \mathcal{Q}(f, f),
\end{aligned}
$$

where the last inequality is due to (2.4). Using the above inequality, we get

$$
\begin{equation*}
\|f\|_{2}^{2}=\int_{\{|\xi|>r\}}|\hat{f}(\xi)|^{2} d \xi+\int_{\{|\xi| \leq r\}}|\hat{f}(\xi)|^{2} d \xi \leq c_{4}(\kappa)\left(r^{-2} \mathcal{Q}(f, f)+2 r^{d}\|f\|_{1}^{2}\right), \quad r \leq 1 . \tag{2.7}
\end{equation*}
$$

Note that, if $a \leq b$, the function $r \rightarrow h(r):=a r^{-2}+2 b r^{d}$ has a local minimum at

$$
r=\left(\frac{a}{d b}\right)^{\frac{1}{d+2}} \leq 1 .
$$

Thus by minimizing the right-hand side of $(2.7)$ for $\mathcal{Q}(f, f) \leq\|f\|_{1}^{2}$, we get

$$
\|f\|_{2}^{2} \leq c_{5} \mathcal{E}^{0}(f, f)^{\frac{d}{d+2}}\|f\|_{1}^{\frac{4}{d+2}} \leq c_{6} \mathcal{Q}(f, f)^{\frac{d}{d+2}}\|f\|_{1}^{\frac{4}{d+2}} .
$$

Therefore by Theorem 2.9 in [CKS], we conclude that

$$
p(t, x, y) \leq c_{7} t^{-d / 2} \quad \text { for all } t \in[1, \infty) .
$$

Theorem 2.3 There exist $C_{*}<1$ and $c_{1}, c_{2}, c_{3}, c_{4}>0$ such that

$$
\begin{equation*}
p(t, x, y) \leq c_{1}\left(\frac{t}{|x-y|}\right)^{c_{2}|x-y|}=c_{1} \exp \left(-c_{2}|x-y| \log \frac{|x-y|}{t}\right) \tag{2.8}
\end{equation*}
$$

for q.e. $x, y \in \mathbb{R}^{d}$ with $(t,|x-y|) \in\left\{(t, R): R \geq \max \left\{t / C_{*}, R_{*}\right\}\right\}$ and

$$
\begin{equation*}
p(t, x, y) \leq c_{3} t^{-d / 2} \exp \left(-\frac{c_{4}|x-y|^{2}}{t}\right) \tag{2.9}
\end{equation*}
$$

for q.e. $x, y \in \mathbb{R}^{d}$ with $(t,|x-y|) \in\left\{(t, R): R_{*} \leq R \leq t / C_{*}\right\}$, where $R_{*}$ is given in Proposition 2.1.

Proof. Using Proposition 2.3 above, [CKS, Corollary 3.28] and [BBCK, Theorem 3.1], we have

$$
\begin{equation*}
p(t, x, y) \leq c\left(t^{-d / \alpha} \vee t^{-d / 2}\right) \exp (-E(2 t, x, y)) \quad \text { for q.e. } x, y \in \mathbb{R}^{d} \text {. } \tag{2.10}
\end{equation*}
$$

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

$$
\begin{aligned}
\Gamma(\psi)(x) & =\int\left(e^{\psi(x)-\psi(y)}-1\right)^{2} J(x, y) d y \\
\Lambda(\psi)^{2} & =\|\Gamma(\psi)\|_{\infty} \vee\|\Gamma(-\psi)\|_{\infty} \\
E(t, x, y) & =\sup \left\{|\psi(x)-\psi(y)|-t \Lambda(\psi)^{2}: \psi \in \operatorname{Lip}_{0} \text { with } \Lambda(\psi)<\infty\right\},
\end{aligned}
$$

where $\operatorname{Lip}_{0}$ is a space of compactly supported Lipschitz continuous functions on $\mathbb{R}^{d}$.
Fix $x_{0}, y_{0} \in \mathbb{R}^{d}$ and let $R=\left|x_{0}-y_{0}\right| \geq R_{*}$. Define

$$
\psi(x)=\lambda\left(R-\left|x_{0}-x\right|\right)^{+} .
$$

So $|\psi(x)-\psi(y)| \leq \lambda|x-y|$. Note that $\left|e^{t}-1\right|^{2} \leq t^{2} e^{2|t|}$. Hence

$$
\Gamma(\psi)(x)=\int\left(e^{\psi(x)-\psi(y)}-1\right)^{2} J(x, y) d y \leq e^{2 \lambda} \lambda^{2} \int|x-y|^{2} J(x, y) d y \leq c_{1} \lambda^{2} e^{2 \lambda}
$$

So we have

$$
\begin{equation*}
-E\left(2 t, x_{0}, y_{0}\right) \leq-\lambda R+c_{1} t \lambda^{2} e^{2 \lambda} \tag{2.11}
\end{equation*}
$$

For each $t$ and $R$, take $\lambda_{0}>0$ such that

$$
\begin{equation*}
\lambda_{0} e^{2 \lambda_{0}}=\frac{R}{2 c_{1} t} . \tag{2.12}
\end{equation*}
$$

Since $x e^{2 x}$ is strictly increasing, it is easy to check that such $\lambda_{0}$ exists uniquely. Then the right hand side of (2.11) is equal to $-\lambda_{0} R / 2$. Let $C_{*}=\left(2 c_{1} e\right)^{-1}$ which is less than 1 by taking $c_{1}$ large. When $R /\left(2 c_{1} t\right) \geq e$ (i.e. $\left.t \leq C_{*} R\right),(2.12)$ holds with $\lambda_{0} \asymp \log (R / t)$, and when $R /\left(2 c_{1} t\right)<e$ (i.e. $t \geq C_{*} R$ ), (2.12) holds with $\lambda_{0} \asymp R / t$. Putting these into (2.10), we obtain the following; In the region $\left\{(t, R): t \leq C_{*} R, R \geq R_{*}\right\}$,

$$
\begin{equation*}
p(t, x, y) \leq c^{\prime}\left(t^{-d / \alpha} \vee t^{-d / 2}\right) \exp \left(-c_{2} R \log \frac{R}{t}\right)=c^{\prime}\left(t^{-d / \alpha} \vee t^{-d / 2}\right)\left(\frac{t}{R}\right)^{c_{2} R} \tag{2.13}
\end{equation*}
$$

and in the region $\left\{(t, R): t \geq C_{*} R, R \geq R_{*}\right\}$,

$$
p(t, x, y) \leq c^{\prime} t^{-d / 2} \exp \left(-c^{\prime \prime} R^{2} / t\right)
$$

which gives (2.9).
To complete the proof, we need to discuss the former case more. When $t \geq 1$, the right hand side of (2.13) is bounded from above by $c^{\prime}(t / R)^{c_{2} R}$, and when $t \leq 1$ and $R \geq R^{*}$ for some large $R^{*}>1$, it is bounded from above by $c^{\prime}(t / R)^{c_{2} R-d / \alpha} \leq c^{\prime}(t / R)^{c_{3} R}$, both of which give (2.8). So all we need is to consider the case $t \leq 1$ and $R_{*} \leq R \leq R^{*}$. But in this case, the desired estimate is already established in Proposition 2.1(i).

Now let's consider a more general non-local Dirichlet form $(\mathcal{E}, \mathcal{F})$. Set

$$
\begin{align*}
\mathcal{E}(f, f) & =\int_{\mathbb{R}^{d}} \int_{\mathbb{R}^{d}}(f(y)-f(x))^{2} J(x, y) d x d y  \tag{2.14}\\
\mathcal{F} & =\overline{C_{c}^{1}\left(\mathbb{R}^{d}\right)^{2}} \tag{2.15}
\end{align*}
$$

where the jump kernel $J(x, y)$ is a symmetric non-negative function of $x$ and $y$ such that $J(x, y)=0$ for $|x-y| \geq 1$ and there exist $\alpha, \beta \in(0,2), \beta>\alpha$ and positive $\kappa_{1}, \kappa_{2}$ such that

$$
\begin{equation*}
\kappa_{1}|y-x|^{-d-\alpha} \leq J(x, y) \leq \kappa_{2}|y-x|^{-d-\beta} \quad \text { for }|y-x|<1 . \tag{2.16}
\end{equation*}
$$

Here $\mathcal{E}_{1}(f, f):=\mathcal{E}(f, f)+\|f\|_{2}^{2}, C_{c}^{1}\left(\mathbb{R}^{d}\right)$ denotes the space of $C^{1}$ functions on $\mathbb{R}^{d}$ with compact support, and $\mathcal{F}$ is the closure of $C_{c}^{1}\left(\mathbb{R}^{d}\right)$ with respect to the metric $\mathcal{E}_{1}(f, f)^{1 / 2}$. The Dirichlet form $(\mathcal{E}, \mathcal{F})$ is regular on $\mathbb{R}^{d}$ and so it associates a Hunt process $Z$, starting from quasi-everywhere in $\mathbb{R}^{d}$. It is proved in $[\mathrm{BBCK}]$ that $Z$ is conservative and has quasi-continuous transition density function $q(t, x, y)$ with respect to the Lebesgue measure on $\mathbb{R}^{d}$.

When $t \in[1, \infty)$, only the upper bound of the jumping kernel played a role in the proofs of Proposition 2.2 and Theorem 2.3. Thus, combining with Theorem 1.2 in [BBCK], the following is true for $Z$.

Theorem 2.4 There is a constant $c>0$ such that

$$
q(t, x, y) \leq c\left(t^{-d / \alpha} \vee t^{-d / 2}\right) \quad \text { for q.e. } x, y \in \mathbb{R}^{d} .
$$

Moreover, there exist $C_{1}<1, R_{1} \leq \frac{1}{4}$ and $c_{1}, c_{2}, c_{3}, c_{4}>0$ such that

$$
\begin{equation*}
q(t, x, y) \leq c_{1}\left(\frac{t}{|x-y|}\right)^{c_{2}|x-y|}=c_{1} \exp \left(-c_{2}|x-y| \log \frac{|x-y|}{t}\right) \quad \text { for q.e. } x, y \in \mathbb{R}^{d} \tag{2.17}
\end{equation*}
$$

with $(t,|x-y|) \in\left\{(t, R): t \geq 1, R \geq \max \left\{t / C_{1}, R_{1}\right\}\right\}$ and

$$
\begin{equation*}
q(t, x, y) \leq c_{3} t^{-d / 2} \exp \left(-\frac{c_{4}|x-y|^{2}}{t}\right) \quad \text { for q.e. } x, y \in \mathbb{R}^{d} \tag{2.18}
\end{equation*}
$$

with $(t,|x-y|) \in\left\{(t, R): t \geq 1, R_{1} \leq R \leq t / C_{1}\right\}$.
The above theorem will be used in the next section to prove the near-diagonal lower bound for $q(t, x, y)$.

## 3 Heat Kernel Lower Bound Estimate

In this section, we give the proof of the lower bound estimate of the heat kernel. We first record a simple observation, which sheds lights on the different heat kernel behaviors at small (stable) and large (Gaussian) scale. Recall that a finite range isotropically symmetric $\alpha$-stable process in $\mathbb{R}^{d}$ with jumps of size larger than 1 removed is the Lévy process with Lévy measure $c_{0}(d, \alpha)|h|^{-d-\alpha} \mathbf{1}_{\{|h| \leq 1\}} d h$.

Lemma 3.1 Let $X$ be finite range isotropically symmetric $\alpha$-stable process in $\mathbb{R}^{d}$ with jumps of size larger than 1 removed. For $\lambda>0$, define

$$
Y_{t}^{(\lambda)}:=Y_{0}^{(\lambda)}+\lambda^{-1 / 2}\left(X_{\lambda t}-X_{0}\right) \quad \text { and } \quad Z_{t}^{(\lambda)}:=Z_{0}^{(\lambda)}+\lambda^{-1 / \alpha}\left(X_{\lambda t}-X_{0}\right) .
$$

Then the process $Y^{(\lambda)}$ converges in finite-dimensional distributions to a Brownian motion on $\mathbb{R}^{d}$ as $\lambda \rightarrow \infty$ and $Z^{(\lambda)}$ converges in finite-dimensional distributions to the isotropically symmetric $\alpha$-stable process as $\lambda \rightarrow 0$.

Proof. Recall that the Lévy exponent $\phi$ of $X$ is given by (2.5). Clearly $Y^{(\lambda)}$ and $Z^{(\lambda)}$ are Lévy processes as well, with

$$
\mathbb{E}\left[e^{i \xi \cdot\left(Y_{t}^{(\lambda)}-Y_{0}^{(\lambda)}\right)}\right]=\mathbb{E}\left[e^{i \xi \cdot \lambda^{-1 / 2}\left(X_{\lambda t}-X_{0}\right)}\right]=e^{\lambda t \phi\left(\lambda^{-1 / 2} \xi\right)}, \quad \xi \in \mathbb{R}^{d}
$$

and

$$
\mathbb{E}\left[e^{i \xi \cdot\left(Z_{t}^{(\lambda)}-Z_{0}^{(\lambda)}\right)}\right]=\mathbb{E}\left[e^{i \xi \cdot \lambda^{-1 / \alpha}\left(X_{\lambda t}-X_{0}\right)}\right]=e^{\lambda t \phi\left(\lambda^{-1 / \alpha} \xi\right)}, \quad \xi \in \mathbb{R}^{d}
$$

Let $\phi_{\lambda}(\xi)$ and $\psi_{\lambda}(\xi)$ denote the Lévy exponents of $Y^{(\lambda)}$ and $Z^{(\lambda)}$, respectively. Then we have by above and (2.6) that

$$
\begin{equation*}
\phi_{\lambda}(\xi)=\lambda \phi\left(\lambda^{-1 / 2} \xi\right)=\lambda^{1-\alpha / 2}|\xi|^{\alpha} \int_{\left\{x \in \mathbb{R}^{d}:|x| \leq \lambda^{-1 / 2}|\xi|\right\}} \frac{1-\cos x_{1}}{|x|^{d+\alpha}} d x \tag{3.1}
\end{equation*}
$$

which converges to $c|\xi|^{2}$ as $\lambda \rightarrow \infty$. Moreover, there is $c_{1}>0$ so that

$$
\begin{equation*}
\left|\phi_{\lambda}(\xi)\right| \leq c_{1}|\xi|^{2} \quad \text { for every } \xi \in \mathbb{R}^{d} \text { and } \lambda>0 \tag{3.2}
\end{equation*}
$$

Similarly,

$$
\begin{equation*}
\psi_{\lambda}(\xi)=\lambda \phi\left(\lambda^{-1 / \alpha} \xi\right)=|\xi|^{\alpha} \int_{\left\{x \in \mathbb{R}^{d}:|x| \leq \lambda^{-1 / \alpha}|\xi|\right\}} \frac{1-\cos x_{1}}{|x|^{d+\alpha}} d x \tag{3.3}
\end{equation*}
$$

which increases to $c_{2}|\xi|^{\alpha}$ as $\lambda \downarrow 0$, where $c_{2}=\int_{\mathbb{R}^{d}} \frac{1-\cos \left(x_{1}\right)}{|x|^{d+\alpha}} d x$. This proves the lemma.
Inequality (3.2) will be used later in the proof of Theorem 3.4.
Now let's consider the more general non-local Dirichlet form $(\mathcal{E}, \mathcal{F})$ in (2.14)-(2.15). Recall that the jump kernel $J(x, y)$ for $(\mathcal{E}, \mathcal{F})$ is zero for $|x-y| \geq 1$ and satisfies the condition (2.16), and $q(t, x, y)$ is the transition density function for the associated Hunt process $Z$ with respect to the Lebesgue measure on $\mathbb{R}^{d}$.

Define

$$
\phi(x)=c\left(1-|x|^{2}\right)^{12 /(2-\beta)} \mathbf{1}_{B(0,1)}(x),
$$

where $c>0$ is the normalizing constant so that $\int_{\mathbb{R}^{d}} \phi(x) d x=1$.
The following proposition is an immediate consequence of the assumption (2.16) and Theorem 5.1 in Section 5 below. As mentioned earlier, to keep the flow of our proof for heat kernel estimates, we will postpone its proof to Section 5 .

Proposition 3.2 There is a positive constant $c_{1}=c_{1}(d, \alpha, \beta)$ independent of $r>1$, such that for every $u \in L^{1}(B(0,1), \phi d x)$,

$$
\begin{aligned}
& \int_{B(0,1)}\left(u(x)-u_{\phi}\right)^{2} \phi(x) d x \\
\leq & c_{1} \int_{B(0,1) \times B(0,1)}(u(x)-u(y))^{2} r^{d+2} J(r x, r y) \sqrt{\phi(x) \phi(y)} d x d y
\end{aligned}
$$

Here $u_{\phi}:=\int_{B(0,1)} u(x) \phi(x) d x$.

Remark 3.3 The above weighted Poincaré inequality in fact holds for more general weight function $\phi$. See Section 5 for the details.

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

$$
J_{\delta}(x, y)= \begin{cases}J(x, y) & \text { for }|x-y| \geq \delta  \tag{3.4}\\ \kappa_{2}|y-x|^{-d-\beta} & \text { for }|x-y|<\delta\end{cases}
$$

and define $\left(\mathcal{E}^{\delta}, \mathcal{F}^{\delta}\right)$ in the same way as we $\operatorname{defined}(\mathcal{E}, \mathcal{F})$ in (2.14)-(2.15).
For $\delta \in(0,1)$, let $Z^{\delta}$ be the symmetric Markov process associated with $\left(\mathcal{E}^{\delta}, \mathcal{F}^{\delta}\right)$. By $[\mathrm{BBCK}]$, the process $Z^{\delta}$ can be modified to start from every point in $\mathbb{R}^{d}$ and is conservative; moreover $Z^{\delta}$ 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 $\mathbb{R}^{d}$.

The idea of the proof of the following theorem is motivated by that of Proposition 4.9 in [BBCK]. For ball $B\left(x_{0}, r\right) \subset \mathbb{R}^{d}$, let $q^{\delta, B\left(x_{0}, r\right)}(t, x, y)$ denote the transition density function of the subprocess $Z^{B\left(x_{0}, r\right)}$ of $Z$ killed upon leaving the ball $B\left(x_{0}, r\right)$.

Theorem 3.4 Suppose the Dirichlet form $(\mathcal{E}, \mathcal{F})$ is given by (2.14)-(2.15) with the jumping kernel $J$ satisfying the condition (2.16) and $J_{\delta}$ is given by (3.4). For each $\delta_{0}>0$, there exists $c=c\left(\delta_{0}\right)>0$, independent of $\delta \in(0,1)$ such that for every $x_{0} \in \mathbb{R}^{d}, t \geq \delta_{0}$,

$$
\begin{equation*}
q^{\delta, B\left(x_{0}, t^{1 / 2}\right)}(t, x, y) \geq c t^{-d / 2} \quad \text { for every } t \geq \delta_{0} \text { and q.e. } x, y \in B\left(x_{0}, \sqrt{t} / 2\right) \tag{3.5}
\end{equation*}
$$

and

$$
\begin{equation*}
q^{\delta}(t, x, y) \geq c t^{-d / 2} \quad \text { for every } t \geq \delta_{0} \text { and q.e. } x, y \text { with }|x-y|^{2} \leq t . \tag{3.6}
\end{equation*}
$$

Proof. In view of [BBCK, Theorem 4.10], it suffices to prove that there are $t_{0}<1 / 2$ and $c>0$, independent of $\delta \in(0,1)$, such that (3.5)-(3.6) hold for $t \geq t_{0}^{-1}$. In fact, if $\delta_{0}<t_{0}^{-1}$ and $\delta_{0} \leq t \leq t_{0}^{-1}$, we let $n_{0}=1+\left[2 / \sqrt{t_{0} \delta_{0}}\right]$, where $[a]$ is the largest integer which is no larger than $a$. By [BBCK, Theorem 4.10], we have

$$
\begin{equation*}
q^{\delta, B\left(x_{0}, \delta_{0}^{1 / 2}\right)}(s, x, y) \geq c_{0}, \quad \text { for every } \frac{\delta_{0}}{n_{0}} \leq s \leq t_{0}^{-1} \text { and } x, y \in B\left(x_{0}, 3 \delta_{0}^{1 / 2} / 4\right) \tag{3.7}
\end{equation*}
$$

where the constant $c_{0}$ is independent of $\delta$ and $x_{0} \in \mathbb{R}^{d}$. Given $x, y \in B\left(x_{0}, \sqrt{t} / 2\right)$, let $z_{1} \cdots z_{n_{0}-1}$ be equally spaced points on the line segment joining $x$ and $y$ such that $x \in B\left(z_{1}, 3 \delta_{0}^{1 / 2} / 4\right) \subset$ $B\left(z_{1}, \delta_{0}^{1 / 2}\right) \subset B\left(x_{0}, t^{1 / 2}\right)$ and $y \in B\left(z_{n_{0}-1}, 3 \delta_{0}^{1 / 2} / 4\right) \subset B\left(z_{n_{0}-1}, \delta_{0}^{1 / 2}\right) \subset B\left(x_{0}, t^{1 / 2}\right)$. Using (3.7) and the semigroup property, we have

$$
\begin{aligned}
& q^{\delta, B\left(x_{0}, t^{1 / 2}\right)}(t, x, y) \\
= & \int_{B\left(x_{0}, t^{1 / 2}\right)} \ldots \int_{B\left(x_{0}, t^{1 / 2}\right)} q^{\delta, B\left(x_{0}, t^{1 / 2}\right)}\left(t / n_{0}, x, w_{1}\right) \ldots q^{\delta, B\left(x_{0}, t^{1 / 2}\right)}\left(t / n_{0}, w_{n_{0}-1}, y\right) d w_{1} \ldots d w_{n_{0}-1} \\
\geq & \int_{B\left(z_{1}, 3 \delta_{0}^{1 / 2} / 4\right)} \ldots \int_{B\left(z_{n_{0}-1}, 3 \delta_{0}^{1 / 2} / 4\right)} q^{\delta, B\left(z_{1}, \delta_{0}^{1 / 2}\right)}\left(t / n_{0}, x, w_{1}\right) \ldots \\
& \ldots q^{\delta, B\left(z_{n_{0}}, \delta_{0}^{1 / 2}\right)}\left(t / n_{0}, w_{n_{0}-1}, y\right) d w_{1} \ldots d w_{n_{0}-1} \geq \widetilde{c_{0}} \geq \widetilde{c_{0}} \delta_{0}^{d / 2} t^{-d / 2} .
\end{aligned}
$$

Similar argument gives (3.6) when $\delta_{0}<t_{0}^{-1}$ and $t \in\left[\delta_{0}, t_{0}^{-1}\right]$.
Fix $\delta \in(0,1)$ and, for simplicity, in this proof we sometimes drop the superscript " $\delta$ " from $Z^{\delta}$ and $q^{\delta}(t, x, y)$. 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 upon leaving the ball $B_{r}$. Then by the proof of Proposition 4.3 in [BBCK], there is a constant $c_{1}=c_{1}(\delta, r)>0$ such that

$$
q^{B_{r}}(t, x, y) \geq c_{1}(r-|x|)^{\beta}(r-|y|)^{\beta} \quad \text { for every } t \in\left[r^{2} / 8, r^{2} / 4\right] \text { and } x, y \in B_{r} .
$$

Define

$$
\varphi_{r}(x)=\left(r^{2}-|x|^{2}\right)^{12 /(2-\beta)} \mathbf{1}_{B_{r}}(x) .
$$

It follows from Lemmas 4.5 and 4.6 of $[\mathrm{BBCK}]$ that for every $t>0$ and $y_{0} \in B_{r}, q^{B_{r}}\left(t, x, y_{0}\right) \in \mathcal{F}^{B_{r}}$ and $\varphi_{r}(\cdot) / q^{B_{r}}\left(t, x, y_{0}\right) \in \mathcal{F}^{B_{r}}$, where $\left(\mathcal{E}, \mathcal{F}^{B_{r}}\right)$ is the Dirichlet form for the killed process $Z^{B_{r}}$.

Note that the Dirichlet form of $\left\{r^{-1} Z_{r^{2} t}, t \geq 0\right\}$ is $\left(\mathcal{E}^{(r)}, \mathcal{F}^{(r)}\right)$, where

$$
\begin{align*}
\mathcal{E}^{(r)}(u, u) & =\int_{\mathbb{R}^{d} \times \mathbb{R}^{d}}(u(x)-u(y))^{2} r^{d+2} J_{\delta}(r x, r y) d x d y  \tag{3.8}\\
\mathcal{F}^{(r)} & =\left\{u \in L^{2}(u, u): \mathcal{E}^{(r)}(u, u)<\infty\right\}=W^{\beta / 2,2}\left(\mathbb{R}^{d}\right) .
\end{align*}
$$

By (2.16) and (3.2), there are constants $c_{2}, c_{3}>0$ independent of $r \geq 1$ and $\delta \in(0,1)$ such that for every $u \in W^{1,2}\left(\mathbb{R}^{n}\right) \subset W^{\beta / 2,2}\left(\mathbb{R}^{d}\right)$,

$$
\begin{equation*}
\mathcal{E}^{(r)}(u, u) \leq c_{2} \int_{\mathbb{R}^{d}} \phi_{r}(\xi)|\widehat{u}(\xi)|^{2} d \xi \leq c_{3} \int_{\mathbb{R}^{d}}|\nabla u(x)|^{2} d x . \tag{3.9}
\end{equation*}
$$

Here $\widehat{u}$ denotes the Fourier transform of $u$.
Define

$$
\begin{equation*}
q_{r}^{B}(t, x, y):=r^{d} q^{B_{r}}\left(r^{2} t, r x, r y\right) . \tag{3.10}
\end{equation*}
$$

It is easy to see $q_{r}^{B}(t, x, y)$ is the transition density function for process $r^{-1} Z_{r^{2} t}^{B_{r}}$. The latter is the subprocess of $\left\{r^{-1} Z_{r^{2} t}, t \geq 0\right\}$ killed upon leaving the unit ball $B(0,1)$, whose Dirichlet form will be denoted as $\left(\mathcal{E}^{(r)}, \mathcal{F}^{(r), B}\right)$. It follows from above there is a constant $c_{4}=c_{4}(\delta, r)>0$ such that

$$
q_{r}^{B}(t, x, y) \geq c_{4}(1-|x|)^{\beta}(1-|y|)^{\beta} \quad \text { for every } t \in[1 / 8,1 / 4] \text { and } x, y \in B(0,1) .
$$

Recall that

$$
\phi(x)=c_{5}\left(1-|x|^{2}\right)^{12 /(2-\beta)} \mathbf{1}_{B(0,1)}(x),
$$

where $c_{5}$ is a normalizing constant so that $\int_{\mathbb{R}^{d}} \phi(x) d x=1$. Let $x_{0} \in B(0,1)$ and Define

$$
\begin{aligned}
u(t, x) & :=q_{r}^{B}\left(t, x, x_{0}\right), \quad v(t, x):=q_{r}^{B}\left(t, x, x_{0}\right) / \phi(x)^{1 / 2}, \\
H(t) & :=\int_{B(0,1)} \phi(y) \log u(t, y) d y, \\
G(t) & :=\int_{B(0,1)} \phi(y) \log v(t, y) d y=\int_{B(0,1)} \phi(y) \log u(t, y) d y-\frac{1}{2} \int_{B(0,1)} \phi(x) \log \phi(x) d x \\
& =H(t)-c_{6} .
\end{aligned}
$$

By Lemma 4.7 of [BBCK],

$$
\begin{equation*}
G^{\prime}(t)=-\mathcal{E}^{(r)}\left(u(t, \cdot), \frac{\phi}{u(t, \cdot)}\right) \tag{3.11}
\end{equation*}
$$

(The reason we work with $J_{\delta}$ rather than $J$ is so that we can use [BBCK, Lemma 4.7] to obtain above (3.11). The remainder of the argument does not use the condition on $J_{\delta}$, and in particular the constants can be taken to be independent of $\delta \in(0,1)$.)

Write $J^{(r)}(x, y):=r^{d+2} J_{\delta}(r x, r y)$ and $\kappa_{B}^{(r)}(x):=2 \int_{\mathbb{R}^{d} \backslash B(0,1)} J^{(r)}(x, y) d y$ for $x \in B:=B(0,1)$. Then we have from (3.8) and (3.11),

$$
\begin{aligned}
G^{\prime}(t)=- & \int_{B} \int_{B} \frac{[u(t, y)-u(t, x)]}{u(t, x) u(t, y)}[u(t, x) \phi(y)-\phi(x) u(t, y)] J^{(r)}(x, y) d y d x \\
& -\int_{B} \phi(x) \kappa_{B}^{(r)}(x) d x
\end{aligned}
$$

The main step is to show that for all $t$ in $(0,1]$ one has

$$
\begin{equation*}
G^{\prime}(t) \geq-c_{7}+c_{8} \int_{B}(\log u(t, y)-H(t))^{2} \phi(y) d y \tag{3.12}
\end{equation*}
$$

for positive constants $c_{7}, c_{8}$.
Setting $a=u(t, y) / u(t, x)$ and $b=\phi(y) / \phi(x)$, we see that

$$
\begin{align*}
\frac{[u(t, y)-u(t, x)]}{u(t, x) u(t, y)} & {[u(t, x) \phi(y)-\phi(x) u(t, y)] } \\
& =\phi(x)\left(b-\frac{b}{a}-a+1\right) \\
& =\phi(x)\left[\left(\left(1-b^{1 / 2}\right)^{2}-b^{1 / 2}\left(\frac{a}{b^{1 / 2}}+\frac{b^{1 / 2}}{a}-2\right)\right]\right. \tag{3.13}
\end{align*}
$$

Using the inequality

$$
A+\frac{1}{A}-2 \geq(\log A)^{2}, \quad A>0
$$

with $A=a / \sqrt{b}$, the right hand side of (3.13) is bounded above by

$$
\left(\phi(x)^{1 / 2}-\phi(y)^{1 / 2}\right)^{2}-\sqrt{\phi(x) \phi(y)}(\log v(t, y)-\log v(t, x))^{2} .
$$

Substituting in the formula for $G^{\prime}(t)$ and using Proposition 3.2,

$$
\begin{aligned}
H^{\prime}(t)=G^{\prime}(t) & \geq-c_{9}+\int_{B} \int_{B}(\log v(t, y)-\log v(t, x))^{2} \sqrt{\phi(x) \phi(y)} J^{(r)}(x, y) d x d y \\
& \geq-c_{9}+c_{10} \int_{B}(\log v(t, y)-G(t))^{2} \phi(y) d y \\
& \geq-c_{11}+c_{12} \int_{B}(\log u(t, y)-H(t))^{2} \phi(y) d y
\end{aligned}
$$

which gives (3.12). Note that in the first inequality we used the fact that

$$
\int_{B} \int_{B}\left(\phi(x)^{1 / 2}-\phi(y)^{1 / 2}\right)^{2} J^{(r)}(x, y) d x d y+\int_{B} \phi(x) \kappa_{B}^{(r)}(x) d x=\mathcal{E}^{(r)}\left(\phi^{1 / 2}, \phi^{1 / 2}\right)<\infty
$$

which follows from (3.9) and in the last inequality we used the fact that

$$
\begin{aligned}
& \int_{B}(\log u(t, y)-H(t))^{2} \phi(y) d y \\
= & \int_{B}\left(\log v(t, y)-G(t)+\frac{1}{2} \log \phi(y)-c_{6}\right)^{2} \phi(y) d y \\
\leq & 2 \int_{B}(\log v(t, y)-G(t))^{2} \phi(y) d y+2 \int_{B}\left(\frac{1}{2} \log \phi(y)-c_{6}\right)^{2} \phi(y) d y \\
= & 2 \int_{B}(\log v(t, y)-G(t))^{2} \phi(y) d y+c_{13} .
\end{aligned}
$$

Let $q_{r}(t, x, y):=r^{d} q\left(r^{2} t, r x, r y\right)$, which is the transition density function with respect to the Lebesgue measure on $\mathbb{R}^{d}$ for the process $Z_{t}^{(r)}:=r^{-1} Z_{r^{2} t}$, whose non-local Dirichlet form is given by the jumping intensity measure $r^{d+2} J(r x, r y)$. Using Theorem 2.4 and the fact that $R_{1} \leq \frac{1}{4} \leq \frac{r}{4}$ where $R_{1}$ is given in Theorem 2.4, for $r^{2} t \geq 1$,

$$
\begin{aligned}
& \mathbb{P}_{x}\left(Z_{t}^{(r)} \notin B(x, 1 / 4)\right) \\
= & \int_{B(x, 1 / 4)^{c}} r^{d} q\left(r^{2} t, r x, r y\right) d y \\
= & \int_{B(r x, r / 4)^{c}} q\left(r^{2} t, r x, z\right) d z \\
\leq & c_{14} \int_{\left\{z \in \mathbb{R}^{d}: C_{1}|z-r x| \geq \max \left\{C_{1} r / 4, r^{2} t\right\}\right\}} e^{-c_{15}|z-r x|} d z \\
& +c_{16} \int_{\left\{z \in \mathbb{R}^{d}: r^{2} t \geq C_{1}|z-r z| \geq C_{1} r / 4\right\}} r^{-d} t^{-d / 2} \exp \left(-\frac{c_{17}|z-r z|^{2}}{r^{2} t}\right) d z \\
\leq & c_{18} \int_{\left\{w \in \mathbb{R}^{d}:|w| \geq r / 4\right\}} e^{-c_{15}|w|} d w+c_{19} \int_{r / 4}^{r^{2} t / C_{1}} r^{-d} t^{-d / 2} \exp \left(-\frac{c_{17} s^{2}}{r^{2} t}\right) s^{d-1} d s \\
\leq & c_{18} \int_{\left\{w \in \mathbb{R}^{d}:|w| \geq r / 4\right\}} e^{-c_{15}|w|} d w+c_{19} \int_{1 /(4 \sqrt{t})}^{\infty} \exp \left(-c_{17} u^{2}\right) u^{d-1} d u .
\end{aligned}
$$

Let $t_{0} \in(0,1 / 2)$ be small so that

$$
c_{19} \int_{1 /\left(4 \sqrt{t_{0}}\right)}^{\infty} \exp \left(-c_{17} u^{2}\right) u^{d-1} d u<1 / 16
$$

and

$$
c_{18} \int_{\left\{w \in \mathbb{R}^{d}:|w| \geq 1 /\left(4 \sqrt{t_{0}}\right)\right\}} e^{-c_{15}|w|} d w<1 / 16 .
$$

We then have

$$
\mathbb{P}_{x}\left(Z_{t}^{(r)} \notin B(x, 1 / 4)\right)<1 / 16+1 / 16=1 / 8 \quad \text { for every } r \geq t_{0}^{-1 / 2} \text { and } 0<t \leq t_{0}
$$

By Lemma 3.8 of [BBCK], we have for every $r \geq t_{0}^{-1 / 2}$,

$$
\mathbb{P}_{x}\left(\sup _{s \in\left[0, t_{0}\right]}\left|Z_{s}^{(r)}-Z_{0}^{(r)}\right|>1 / 4\right) \leq 1 / 4 .
$$

Therefore, with $r \geq t_{0}^{-1 / 2}$, for every $t \leq t_{0}$,

$$
\begin{aligned}
\int_{B(0,1 / 4)} u(t, x) d x & \geq \mathbb{P}_{0}\left(\sup _{s \in\left[0, t_{0}\right]}\left|Z_{s}^{(r)}-Z_{0}^{(r)}\right|<1 / 4\right) \\
& =1-\mathbb{P}_{0}\left(\sup _{s \in\left[0, t_{0}\right]}\left|Z_{s}^{(r)}-Z_{0}^{(r)}\right| \geq 1 / 4\right) \geq \frac{3}{4} .
\end{aligned}
$$

Here the conservativeness of $Z_{t}^{(r)}$ is used in the first equality.
Choose $K$ such that $\mu_{d}(B(0,1 / 4)) e^{-K}=\frac{1}{4}$ and define

$$
D_{t}:=\left\{x \in B(0,1 / 4): u(t, x) \geq e^{-K}\right\} .
$$

By Proposition 2.2, if $t \leq t_{0}$

$$
\begin{aligned}
\frac{3}{4} \leq \int_{B(0,1 / 4)} u(t, x) d x & =\int_{D_{t}} u(t, x) d x+\int_{B(0,1 / 4) \backslash D_{t}} u(t, x) d x \\
& \leq c_{20} t^{-d / \alpha} \mu_{d}\left(D_{t}\right)+\mu_{d}(B(0,1 / 4)) e^{-K}=c_{20} t^{-d / \alpha} \mu_{d}\left(D_{t}\right)+\frac{1}{4}
\end{aligned}
$$

Therefore

$$
\mu_{d}\left(D_{t}\right) \geq \frac{t^{d / \alpha}}{c_{21}} \geq c_{22}>0 \quad \text { if } t \in\left[\varepsilon / 4, t_{0}\right] .
$$

Note that the positive constant $c_{22}=c_{22}(\varepsilon)$ can be chosen to be independent of $r \geq t_{0}^{-1 / 2}$ and $x_{0} \in B(0,1 / 2)$.

Jensen's inequality tells us that if $t \leq t_{0}$

$$
H(t)=\int_{B}(\log u(t, x)) \phi(x) d x \leq \log \int_{B} u(t, x) \phi(x) d x \leq \log \|\phi\|_{\infty}:=\bar{H} .
$$

On $D_{t}, \log u(t, x) \geq-K$ so there are only four possible cases:
(a) If $\log u(t, x)>0$ and $H(t) \leq 0$, then $(\log u(t, x)-H(t))^{2} \geq H(t)^{2}$.
(b) If $\log u(t, x)>0$ and $0<H(t) \leq \bar{H}$, then

$$
(\log u(t, x)-H(t))^{2} \geq 0 \geq H(t)^{2}-\bar{H}^{2} .
$$

(c) If $-K \leq \log u(t, x) \leq 0$ and $|H(t)| \geq 2 K$, then $(\log u(t, x)-H(t))^{2} \geq \frac{1}{4} H(t)^{2}$.
(d) If $-K \leq \log u(t, x) \leq 0$ and $|H(t)|<2 K$, then

$$
(\log u(t, x)-H(t))^{2} \geq 0 \geq \frac{1}{4} H(t)^{2}-K^{2} .
$$

Thus we conclude

$$
(\log u(t, x)-H(t))^{2} \geq \frac{1}{4} H(t)^{2}-(\bar{H} \vee K)^{2} \quad \text { on } D_{t} .
$$

Since $\phi$ is bounded below by $c_{23}>0$ on $B(0,1 / 4)$, then

$$
\begin{aligned}
-c_{11}+c_{12} \int_{B}(\log u(t, x)-H(t))^{2} \phi(x) d x & \geq c_{12} \int_{D_{t}}(\log u(t, x)-H(t))^{2} \phi(x) d x-c_{11} \\
& \geq c_{24} \mu_{d}\left(D_{t}\right)\left(\frac{1}{4} H(t)^{2}-(\bar{H} \vee K)^{2}\right)-c_{11}
\end{aligned}
$$

We therefore have

$$
H^{\prime}(t) \geq F H(t)^{2}-E, \quad t \in\left[\varepsilon / 4, t_{0}\right]
$$

for some positive constants $E$ and $F$ that are independent of $r \geq t_{0}^{-1 / 2}$.
Now we do some calculus. Let $t_{2} \in\left[\varepsilon / 2, t_{0} \wedge 2\right]$ and let $Q:=\max \left(16 E,(16 E / F)^{1 / 2}\right)$. Suppose $H\left(t_{2}\right) \leq-Q$. Since $H^{\prime}(t) \geq-E$ and $t_{2}-t<t_{0} \wedge 2 \leq 2$,

$$
\begin{equation*}
H\left(t_{2}\right)-H(t) \geq-2 E \quad \text { for } t \in\left[\varepsilon / 4, t_{2}\right] . \tag{3.14}
\end{equation*}
$$

This implies $H(t) \leq-Q / 2$. Since $F Q^{2} / 4 \geq 4 E, E<\frac{F}{2} H(t)^{2}$ and hence

$$
H^{\prime}(t) \geq \frac{F}{2} H(t)^{2} .
$$

Integrating $H^{\prime} / H^{2} \geq F / 2$ over $\left[\frac{\varepsilon}{4}, t_{2}\right]$ yields

$$
\frac{1}{H\left(t_{2}\right)}-\frac{1}{H(\varepsilon / 4)} \leq-\frac{F}{2}\left(t_{2}-\varepsilon / 4\right) \leq-\frac{F \varepsilon}{8} .
$$

Since $H(\varepsilon / 4) \leq-Q / 2<0$, we have $1 / H\left(t_{2}\right) \leq-F \varepsilon / 16$, that is,

$$
H\left(t_{2}\right) \geq-\frac{16}{F \varepsilon} .
$$

This proves that either $H\left(t_{2}\right) \geq-Q$ or $H\left(t_{2}\right) \geq-16 /(F \varepsilon)$. Thus in either case, $H\left(t_{2}\right) \geq-U$, where $U=U(\varepsilon):=\max \{Q, 16 /(F \varepsilon)\}>0$, and so $G\left(t_{2}\right)=H\left(t_{2}\right)-c_{6} \geq-U-c_{6}$.

Now for every $x_{0}, x_{1} \in B(0,1 / 2)$, applying the above first with $x_{0}$ and then with $x_{0}$ replaced by $x_{1}$, we have

$$
\begin{aligned}
\log q_{r}^{B}\left(2 t_{2}, x_{0}, x_{1}\right) & =\log \int_{B} q_{r}^{B}\left(t_{2}, x_{0}, z\right) q_{r}^{B}\left(t_{2}, x_{1}, z\right) d z \\
& \geq \log \int_{B} q_{r}^{B}\left(t_{2}, x_{0}, z\right) q_{r}^{B}\left(t_{2}, x_{1}, z\right) \phi(z) d z-\log \|\phi\|_{\infty} \\
& \geq \int_{B} \log \left(q_{r}^{B}\left(t_{2}, x_{0}, z\right) q_{r}^{B}\left(t_{2}, x_{1}, z\right)\right) \phi(z) d z-\log \|\phi\|_{\infty} \\
& =\int_{B} \log q_{r}^{B}\left(t_{2}, x_{0}, z\right) \phi(z) d z+\int_{B} \log q_{r}^{B}\left(t_{2}, x_{1}, z\right) \phi(z) d z-\log \|\phi\|_{\infty} \\
& \geq-2\left(U+c_{26}\right),
\end{aligned}
$$

that is, $q_{r}^{B}\left(2 t_{2}, x_{0}, x_{1}\right) \geq e^{-2\left(U+c_{26}\right)}$. A repeated use of the semigroup property (but at most $2 / t_{2}$ more times) then shows $q_{r}^{B}\left(t, x_{0}, x_{1}\right) \geq c_{27}(\varepsilon)$ for every $t \in[\varepsilon / 2,2]$ and $x_{0}, x_{1} \in B(0,1 / 2)$. Taking $\varepsilon=1 / 4$, we have for every $r \geq t_{0}^{-1 / 2}, x, y \in B(0,1 / 2)$ and $t \in[1 / 4,2]$,

$$
r^{d} q^{B_{r}}\left(r^{2} t, r x, r y\right)=q_{r}^{B}(t, x, y) \geq c_{28},
$$

in particular,

$$
q^{B_{r}}\left(r^{2}, r x, r y\right) \geq c_{28} r^{-d}
$$

Thus we have

$$
q^{B(0, \sqrt{t})}(t, x, y) \geq c_{28} t^{-d / 2} \quad \text { for } t \geq t_{0}^{-1} \text { and } x, y \in B(0, \sqrt{t} / 2) .
$$

Clearly the above inequality holds with $B(0, \sqrt{t})$ and $B(0, \sqrt{t} / 2)$ being replaced by any other ball $B\left(x_{0}, \sqrt{t}\right)$ and $B\left(x_{0}, \sqrt{t} / 2\right)$ of the same radius, respectively. Consequently,

$$
q(t, x, y) \geq q^{B\left(x_{0}, \sqrt{t}\right)}(t, x, y) \geq c_{28} t^{-d / 2} \quad \text { for } t \geq t_{0}^{-1} \text { and }|x-y|^{2} \leq t .
$$

This proves the theorem.

For any ball $B \subset \mathbb{R}^{d}$, let $\left(\mathcal{E}^{\delta, B}, \mathcal{F}^{\delta, B}\right)$ denote the Dirichlet form of the subprocess $Z^{\delta, B}$ of $Z^{\delta}$ killed upon leaving the ball $B$. It is shown in [BBCK, Theorem 1.5 and Theorem 2.6] that $\left(\mathcal{E}^{\delta}, \mathcal{F}^{\delta}\right)$ and $\left(\mathcal{E}^{\delta, B}, \mathcal{F}^{\delta, B}\right)$ converge as $\delta \rightarrow 0$ to $(\mathcal{E}, \mathcal{F})$ and $\left(\mathcal{E}^{B}, \mathcal{F}^{B}\right)$, respectively in the sense of Mosco, where $B$ is a ball in $\mathbb{R}^{d}$. Therefore the semigroup of $Z^{\delta}$ and $Z^{\delta, B}$ converge in $L^{2}$ to that of $Z$ and $Z^{B}$, respectively. By the same proof as that for [BBCK, Theorem 1.3], we deduce from Theorem 3.4 the following lower bound estimate for the heat kernel of $Z$, which extends Theorem 1.3 in [BBCK].

Theorem 3.5 Suppose the Dirichlet form $(\mathcal{E}, \mathcal{F})$ is given by (2.14)-(2.15) with the jumping kernel $J$ satisfying the condition (2.16). For each $t_{0}>0$, there exists $c_{1}=c_{1}\left(t_{0}\right)>0$, such that for every $x_{0} \in \mathbb{R}^{d}, t \geq t_{0}$,

$$
q^{B\left(x_{0}, t^{1 / 2}\right)}(t, x, y) \geq c_{1} t^{-d / 2} \quad \text { for q.e. } x, y \in B\left(x_{0}, \sqrt{t} / 2\right)
$$

and

$$
q(t, x, y) \geq c_{1} t^{-d / 2} \quad \text { for q.e. } x, y \text { with }|x-y|^{2} \leq t .
$$

Now we return to the case for the Dirichlet form ( $\mathcal{Q}, \mathcal{D}$ ) given by (1.2)-(1.3).
Theorem 3.6 There exist $c_{0}, c_{1}, c_{2}, c_{3}, c_{4}>0$ such that

$$
p(t, x, y) \geq \begin{cases}c_{0} t^{-d / 2} & \text { when } t \geq R_{*}^{\alpha},|x-y|^{2} \leq t  \tag{3.15}\\ c_{1}\left(\frac{t}{|x-y|}\right)^{c_{2}|x-y|} & \text { when }|x-y| \geq \max \left\{t / C_{*}, R_{*}\right\} \\ c_{3} t^{-d / 2} \exp \left(-\frac{c_{4}|x-y|^{2}}{t}\right) & \text { when } C_{*}|x-y| \leq t \leq|x-y|^{2}\end{cases}
$$

where $R_{*}$ and $C_{*}$ are the constants given in Proposition 2.1 and in Theorem 2.3, respectively.

Proof. By Theorem 3.5, we only need to show the second and third inequalities in (3.15). We first prove the second inequality in (3.15). Let $R:=|x-y|$ and $c_{+}:=\left(4 / R_{*}\right) \vee\left(C_{*} / T_{*}\right) \geq 1$. Let $l \geq 2$ be a positive integer such that $c_{+} R<l \leq c_{+} R+1$ and let $x=x_{0}, x_{1}, \cdots, x_{l}=y$ be such that $\left|x_{i}-x_{i+1}\right| \leq 2 R / l \leq 2 / c_{+}$for $i=1, \cdots, l-1$. (Here we used the fact that $\mathbb{R}^{d}$ is a geodesic space.) Since $t / l \leq C_{*} R / l \leq C_{*} / c_{+} \leq T_{*}$ and $2 R / l \leq 2 / c_{*} \leq R_{*} / 2$, by Proposition 2.1(ii), we have for all $\left(y_{i}, y_{i+1}\right) \in B\left(x_{i}, R_{*} / 4\right) \times B\left(x_{i+1}, R_{*} / 4\right)$

$$
\begin{equation*}
p\left(t / l, y_{i}, y_{i+1}\right) \geq c_{0}\left((t / l)^{-d / \alpha} \wedge \frac{(t / l)}{(R / l)^{d+\alpha}}\right) \geq c_{1}\left((t / l)^{-d / \alpha} \wedge(t / l)\right)=c_{1} t / l \tag{3.16}
\end{equation*}
$$

since $t / l \leq T_{*} \leq 1$. Let $B_{i}=B\left(x_{i}, R_{*} / 4\right)$. Using (3.16), we have

$$
\begin{aligned}
p(t, x, y) & \geq \int_{B_{1}} \ldots \int_{B_{l-1}} p\left(t / l, x, y_{1}\right) \ldots p\left(t / l, y_{l-1}, y\right) d y_{1} \ldots d y_{l-1} \\
& \geq c_{1}(t / l) \Pi_{i=1}^{l-1} c_{2}(t / l)=\left(c_{3} t / l\right)^{l} \asymp\left(c_{4} t / R\right)^{c_{+} R+1} \geq c_{5}(t / R)^{c_{6} R}
\end{aligned}
$$

and the proof is completed.
We next prove the third inequality in (3.15). Take maximum $l \in \mathbb{N}$ such that $t / l \leq(R / l)^{2}$; then $R^{2} / t-1<l \leq R^{2} / t$. Since $t \geq C_{*} R$, we can take $t / l \geq C_{*}^{2}$. Let $x=x_{0}, x_{1}, \cdots, x_{l}=y$ be such that $\left|x_{i}-x_{i+1}\right| \asymp R / l$ for $i=1, \cdots, l-1$. Since $(R / l)^{2} \asymp t / l \geq C_{*}^{2}$, by Theorem 3.5 , we have

$$
\begin{equation*}
p\left(t / l, x_{i}, x_{i+1}\right) \geq c_{1}(t / l)^{-d / 2} \tag{3.17}
\end{equation*}
$$

Using (3.17), we have

$$
\begin{aligned}
p(t, x, y) & \geq \int_{B_{1}} \ldots \int_{B_{l-1}} p\left(t / l, x, y_{1}\right) \ldots p\left(t / l, y_{l-1}, y\right) d y_{1} \ldots d y_{l-1} \\
& \geq c_{1}(t / l)^{-d / 2} \Pi_{i=1}^{l-1}\left(c_{2}(t / l)^{-d / 2}(R / l)^{d}\right) \\
& \geq c_{1}(t / l)^{-d / 2} c_{2}^{l-1} \\
& \geq c_{1}(t / l)^{-d / 2} \exp \left(-c_{3} l\right) \\
& \geq c_{4} t^{-d / 2} \exp \left(-\frac{c_{5}|x-y|^{2}}{t}\right)
\end{aligned}
$$

This completes the proof.

Remark 3.7 In [CK2], the following two-sided transition density function estimate was obtained for the relativistic $\alpha$-stable-like process where $J(x, y) \asymp|x-y|^{-d-\alpha} e^{-|x-y|}$ : for $t \leq 1$ :

$$
c_{1}\left(t^{-d / \alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\right) e^{-c_{2}|x-y|} \leq p(t, x, y) \leq c_{3}\left(t^{-d / \alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}\right) e^{-c_{4}|x-y|}
$$

Theorem 2.3 and 3.6 show that when $|x-y| \rightarrow \infty$, the behavior of the heat kernel for finite range $\alpha$-stable-like process is different from that of relativistic $\alpha$-stable-like process.

We will use the following near diagonal lower bound for the killed process in the next section. Recall that $R_{*} \in(0,1)$ is the constant given in Proposition 2.1(ii).

Proposition 3.8 For every $c_{1} \in(0,1), c_{2}, c_{3}>0$, there is a constant $c_{4}>0$ such that for every $x_{0} \in \mathbb{R}^{d}$ and $r \leq R_{*}$,

$$
\begin{equation*}
p^{B\left(x_{0}, r\right)}(t, x, y) \geq c_{4} t^{-d / \alpha} \quad \text { for q.e. } x, y \in B\left(x_{0}, c_{1} r\right) \text { and } t \in\left[c_{2} r^{\alpha}, c_{3} r^{\alpha}\right] . \tag{3.18}
\end{equation*}
$$

Proof. Let $\kappa:=c_{2} /\left(2 c_{3}\right)$ and $B_{r}:=B\left(x_{0}, r\right)$. We first show that there is a constant $c_{5} \in(0,1)$ so that (3.18) holds for every $r \leq R_{*}$, quasi-every $x, y \in B\left(x_{0}, c_{1} r\right)$ and $t \in\left[\kappa c_{5} r^{\alpha}, c_{5} r^{\alpha}\right]$. We will use the following Dynkin-Hunt formula, which is easy to establish using the strong Markov property, since we know the existence of the heat kernels:

$$
\begin{equation*}
p^{B_{r}}(t, x, y)=p(t, x, y)-\mathbb{E}^{x}\left[1_{\left\{\tau_{B_{r}} \leq t\right\}} p\left(t-\tau_{B_{r}}, X_{\tau_{B_{r}}}, y\right)\right] . \tag{3.19}
\end{equation*}
$$

For $r \leq R_{*}$ and $t \in\left[\kappa c_{5} r^{\alpha}, c_{5} r^{\alpha}\right]$, and $x, y \in B\left(x_{0}, c_{1} r\right)$, by (3.19) and Proposition 2.1(i) and (ii) $\left(|x-y| \leq 2 c_{1} r \leq 2 c_{1}\left(\kappa c_{5}\right)^{-1 / \alpha} t^{-1 / \alpha}\right)$, we have

$$
\begin{equation*}
p^{B_{r}}(t, x, y) \geq c_{6} c_{5}^{1+d / \alpha} t^{-d / \alpha}-c_{7} \mathbb{E}^{x}\left[1_{\left\{\tau_{B_{r}} \leq t\right\}}\left(\left(t-\tau_{B_{r}}\right)^{-d / \alpha} \wedge \frac{t-\tau_{B_{r}}}{\left|X_{\tau_{B_{r}}}-y\right|^{d+\alpha}}\right)\right], \tag{3.20}
\end{equation*}
$$

where constants $c_{6}, c_{7}$ are independent of $c_{5} \in(0,1]$. Observe that

$$
\left|X_{\tau_{B_{r}}}-y\right| \geq\left(1-c_{1}\right) r, \quad t-\tau_{B_{r}} \leq t \leq c_{5} r^{\alpha}
$$

and so

$$
\begin{equation*}
\frac{t-\tau_{B_{r}}}{\left|X_{\tau_{B_{r}}}-y\right|^{d+\alpha}} \leq \frac{t-\tau_{B_{r}}}{\left.\left(\left(1-c_{1}\right) r\right)\right)^{d+\alpha}} \leq \frac{c_{5}^{1+d / \alpha}}{\left(1-c_{1}\right)^{d+\alpha}} t^{-d / \alpha} . \tag{3.21}
\end{equation*}
$$

Note that if $c_{5}<\left(\left(1-c_{1}\right) / 2\right)^{\alpha}$, by Proposition 2.1 (i), for $t \leq c_{5} r^{\alpha}$

$$
\begin{aligned}
& \mathbb{P}_{x}\left(X_{t} \notin B\left(x,\left(1-c_{1}\right) r / 2\right)\right)=\int_{B\left(x,\left(1-c_{1}\right) r / 2\right)^{c}} p(t, x, y) d y \\
& \leq c_{7} \int_{B\left(x,\left(1-c_{1}\right) r / 2\right)^{c}} \frac{t}{|x-y|^{d+\alpha}} d z \leq c_{8} \frac{t}{r^{\alpha}} \leq c_{8} c_{5}
\end{aligned}
$$

where $c_{8}$ is independent of $c_{5}$. Now applying [BBCK, Lemma 3.8], we have

$$
\begin{equation*}
\mathbb{P}_{x}\left(\tau_{B\left(x,\left(1-c_{1}\right) r\right)} \leq t\right) \leq 2 c_{8} c_{5} . \tag{3.22}
\end{equation*}
$$

Consequently, we have from (3.20), (3.21) and (3.22)

$$
\begin{aligned}
p^{B_{r}}(t, x, y) & \geq\left(c_{6} c_{5}^{1+d / \alpha}-c_{7} \frac{c_{5}^{1+d / \alpha}}{\left(1-c_{1}\right)^{d+\alpha}} \mathbb{P}_{x}\left(\tau_{B_{r}} \leq t\right)\right) t^{-d / \alpha} \\
& \geq\left(c_{6} c_{5}^{1+d / \alpha}-c_{7} \frac{c_{5}^{1+d / \alpha}}{\left(1-c_{1}\right)^{d+\alpha}} \mathbb{P}_{x}\left(\tau_{B\left(x,\left(1-c_{1}\right) r\right)} \leq t\right)\right) t^{-d / \alpha} \\
& \geq c_{5}^{1+d / \alpha}\left(c_{6}-2 c_{8} c_{7} \frac{c_{5}}{\left(1-c_{1}\right)^{d+\alpha}}\right) t^{-d / \alpha} .
\end{aligned}
$$

Clearly we can choose $c_{5}<\left(\left(1-c_{1}\right) / 2\right)^{\alpha}$ small so that $p^{B_{r}}(t, x, y) \geq c_{9} t^{-d / \alpha}$. This establishes (3.18) for any $x_{0} \in \mathbb{R}^{d}, r \leq R_{*}$ and $t \in\left[\kappa c_{5} r^{\alpha}, c_{5} r^{\alpha}\right]$.

Now for $r \leq R_{*}$ and $t \in\left[c_{2} r^{\alpha}, c_{3} r^{\alpha}\right]$, define $k_{0}=\left[2 c_{3} / c_{5}\right]+1$. Here for $a \geq 1,[a]$ denotes the largest integer that does not exceed $a$. Then $t / k_{0} \in\left[\kappa c_{5} r^{\alpha}, c_{5} r^{\alpha}\right]$. Using semigroup $k_{0}$ times, we conclude that for q.e. $x, y \in B\left(x_{0}, c_{1} r\right)$ and $t \in\left[c_{2} r^{\alpha}, c_{3} r^{\alpha}\right]$,

$$
\begin{aligned}
& p^{B\left(x_{0}, r\right)}(t, x, y) \\
= & \int_{B\left(x_{0}, r\right)} \cdots \int_{B\left(x_{0}, r\right)} p^{B\left(x_{0}, r\right)}\left(t / k_{0}, x, w_{1}\right) \ldots p^{B\left(x_{0}, r\right)}\left(t / k_{0}, w_{n-1}, y\right) d w_{1} \ldots d w_{n-1} \\
\geq & \int_{B\left(x_{0},\left(t / k_{0}\right)^{1 / \alpha / 2)}\right.} \cdots \int_{B\left(x_{0},\left(t / k_{0}\right)^{1 / \alpha} / 2\right)} p^{B\left(x_{0}, r\right)}\left(t / k_{0}, x, w_{1}\right) \ldots p^{B\left(x_{0}, r\right)}\left(t / k_{0}, w_{n-1}, y\right) d w_{1} \ldots d w_{n-1} \\
\geq & c_{9}\left(t / k_{0}\right)^{-d / \alpha}\left(c_{9}\left(t / k_{0}\right)^{-d / \alpha} c_{1} r^{d}\right)^{k_{0}-1} \\
\geq & c_{10} t^{-d / \alpha},
\end{aligned}
$$

where $c_{10}:=c_{9}^{k_{0}} k_{0}^{d / \alpha}\left(c_{1} c_{9} c_{5}^{-d / \alpha}\right)^{k_{0}-1}$. The proof of (3.18) is now completed.

## 4 Applications of Heat Kernel Estimates

### 4.1 Parabolic Harnack Inequality

We first introduce a space-time process $Z_{s}:=\left(V_{s}, X_{s}\right)$, where $V_{s}=V_{0}-s$. The filtration generated by $Z$ satisfying the usual condition will be denoted as $\left\{\widetilde{\mathcal{F}}_{s} ; s \geq 0\right\}$. The law of the space-time process $s \mapsto Z_{s}$ starting from $(t, x)$ will be denoted as $\mathbb{P}^{(t, x)}$.

We say that a non-negative Borel measurable function $h(t, x)$ on $[0, \infty) \times \mathbb{R}^{d}$ is parabolic (or caloric) on $D=(a, b) \times B\left(x_{0}, r\right)$ if for every relatively compact open subset $D_{1}$ of $D, h(t, x)=$ $\mathbb{E}^{(t, x)}\left[h\left(Z_{\tau_{D_{1}}}\right)\right]$ for every $(t, x) \in D_{1} \cap\left([0, \infty) \times \mathbb{R}^{d}\right)$, where $\tau_{D_{1}}=\inf \left\{s>0: Z_{s} \notin D_{1}\right\}$.

For each $r>0$, we define

$$
\psi(r):=r^{\alpha} \vee r^{2} .
$$

Theorem 4.1 For every $\delta \in(0,1)$, there exists $c=c(\alpha, \delta)>0$ such that for every $x_{0} \in \mathbb{R}^{d}, t_{0} \geq 0$, $R>0$ and every non-negative function $u$ on $[0, \infty) \times \mathbb{R}^{d}$ that is parabolic on $\left(t_{0}, t_{0}+6 \delta \psi(R)\right) \times$ $B\left(x_{0}, 4 R\right)$,

$$
\begin{equation*}
\sup _{\left(t_{1}, y_{1}\right) \in Q_{-}} u\left(t_{1}, y_{1}\right) \leq c \inf _{\left(t_{2}, y_{2}\right) \in Q_{+}} u\left(t_{2}, y_{2}\right), \tag{4.1}
\end{equation*}
$$

where $Q_{-}=\left(t_{0}+\delta \psi(R), t_{0}+2 \delta \psi(R)\right) \times B\left(x_{0}, R\right)$ and $Q_{+}=\left(t_{0}+3 \delta \psi(R), t_{0}+4 \delta \psi(R)\right) \times B\left(x_{0}, R\right)$.
To prove the theorem, we need one notion and one lemma. According to [BBK], we say (UJS) holds if

$$
\begin{equation*}
J(x, y) \leq \frac{c}{r^{d}} \int_{B(x, r)} J\left(x^{\prime}, y\right) d x^{\prime} \quad \text { whenever } r \leq \frac{1}{2}|x-y|, x, y \in \mathbb{R}^{d} . \tag{UJS}
\end{equation*}
$$

For $R>0$, we say (UJS) $\leq R$ holds if the above holds for all $x, y \in \mathbb{R}^{d}$ and $r \leq \frac{|x-y|}{2} \wedge R$.

It is easy to check that finite range jump process satisfies $(\mathrm{UJS})_{\leq 1}$.
The following lemma corresponds to [CK1, Lemma 4.9] (also [CK2, Lemmas 6.1]). The statement is changed (in the sense that the size of two space-time balls are different and the initial points are also different) and the proof requires major changes from the original ones.

Lemma 4.2 Let $R \leq R_{*}$ and $\delta<1$. $Q_{1}=\left[t_{0}+2 \delta R^{\alpha} / 3, t_{0}+5 \delta R^{\alpha}\right] \times B\left(x_{0}, 3 R / 2\right), Q_{2}=\left[t_{0}+\right.$ $\left.\delta R^{\alpha} / 3, t_{0}+6 \delta R^{\alpha}\right] \times B\left(x_{0}, 2 R\right)$ and define $Q_{-}$and $Q_{+}$as in Theorem 4.1. Let $h:[0, \infty) \times \mathbb{R}^{d} \rightarrow \mathbb{R}_{+}$ be bounded and supported in $[0, \infty) \times B\left(x_{0}, 3 R\right)^{c}$. Then there exists $C_{1}=C_{1}(\delta)>0$ such that the following holds:

$$
\mathbb{E}^{\left(t_{1}, y_{1}\right)}\left[h\left(Z_{\tau_{Q_{1}}}\right)\right] \leq C_{1} \mathbb{E}^{\left(t_{2}, y_{2}\right)}\left[h\left(Z_{\tau_{2}}\right)\right] \quad \text { for }\left(t_{1}, y_{1}\right) \in Q_{-} \text {and }\left(t_{2}, y_{2}\right) \in Q_{+}
$$

Proof. Without loss of generality, assume that $t_{0}=0$. Denote $B_{c R}=B\left(x_{0}, c R\right)$. Using the Lévy system formula,

$$
\begin{align*}
\mathbb{E}^{\left(t_{2}, y_{2}\right)}\left[h\left(Z_{\tau_{Q_{2}}}\right)\right] & =\mathbb{E}^{\left(t_{2}, y_{2}\right)}\left[h\left(t_{2}-\left(\tau_{B_{2 R}} \wedge\left(t_{2}-\delta R^{\alpha} / 3\right)\right), X_{\tau_{B_{2 R}} \wedge\left(t_{2}-\delta R^{\alpha} / 3\right)}\right)\right] \\
& =\mathbb{E}^{\left(t_{2}, y_{2}\right)}\left[\int_{0}^{t_{2}-\delta R^{\alpha} / 3} \mathbf{1}_{\left\{t \leq \tau_{B_{2 R}}\right\}} d t \int_{B_{3 R}^{c}} h\left(t_{2}-t, v\right) J\left(X_{t}, v\right) d v\right] \\
& =\int_{0}^{t_{2}-\delta R^{\alpha} / 3} h\left(t_{2}-t, v\right) d t \int_{B_{3 R}^{c}} \mathbb{E}^{\left(t_{2}, y_{2}\right)}\left[\mathbf{1}_{\left\{t \leq \tau_{B_{2 R}}\right\}} J\left(X_{t}, v\right)\right] d v \\
& =\int_{\delta R^{\alpha} / 3}^{t_{2}} h(s, v) d s \int_{B_{3 R}^{c}} \mathbb{E}^{\left(0, y_{2}\right)}\left[\mathbf{1}_{\left\{t_{2}-s \leq \tau_{B_{2 R}}\right\}} J\left(X_{t_{2}-s}, v\right)\right] d v \\
& =\int_{\delta R^{\alpha} / 3}^{t_{2}} d s \int_{B_{3 R}^{c}} h(s, v) d v \int_{B_{2 R}} p^{B_{2 R}}\left(t_{2}-s, y_{2}, z\right) J(z, v) d z  \tag{4.2}\\
& \geq \int_{\delta R^{\alpha} / 3}^{t_{1}} d s \int_{B_{3 R}^{c}} h(s, v) d v \int_{B_{2 R}} p^{B_{2 R}}\left(t_{2}-s, y_{2}, z\right) J(z, v) d z \\
& \geq \int_{\delta R^{\alpha} / 3}^{t_{1}} d s \int_{B_{3 R}^{c}} h(s, v) d v \int_{B_{3 R / 2}} p^{B_{2 R}\left(t_{2}-s, y_{2}, z\right) J(z, v) d z} \tag{4.3}
\end{align*}
$$

Since $6 \delta R^{\alpha} \geq t_{2}-s \geq t_{2}-t_{1} \geq \delta R^{\alpha}$ for $s \in\left[\delta R^{\alpha} / 3, t_{1}\right]$, by Proposition 3.8, we have that the right hand side of (4.3) is greater than or equal to

$$
\frac{c_{1}}{R^{d}} \int_{\delta R^{\alpha} / 3}^{t_{1}} d s \int_{B_{3 R}^{c}} h(s, v) d v \int_{B_{3 R / 2}} J(z, v) d z
$$

So, the proof is complete once we obtain

$$
\begin{equation*}
\mathbb{E}^{\left(t_{1}, y_{1}\right)}\left[h\left(Z_{\tau_{Q_{1}}}\right)\right] \leq \frac{c_{2}}{R^{d}} \int_{\delta R^{\alpha} / 3}^{t_{1}} d s \int_{B_{3 R}^{c}} h(s, v) d v \int_{B_{3 R / 2}} J(z, v) d z \tag{4.4}
\end{equation*}
$$

Analogous to (4.2), we have by using the Lévy system,

$$
\mathbb{E}^{\left(t_{1}, y_{1}\right)}\left[h\left(Z_{\tau_{Q_{1}}}\right)\right]=\int_{2 \delta R^{\alpha} / 3}^{t_{1}} d s \int_{B_{3 R}^{c}} h(s, v) d v \int_{B_{3 R / 2}} p^{B_{3 R / 2}\left(t_{1}-s, y_{1}, z\right) J(z, v) d z . . . . ~}
$$

Since

$$
\begin{aligned}
& \int_{B_{3 R / 2}} p^{B_{3 R / 2}\left(t_{1}-s, y_{1}, z\right)} \int_{B_{3 R}^{c}} J(z, v) h(s, v) d v d z \\
= & \int_{B_{5 R / 4}} p^{B_{3 R / 2}}\left(t_{1}-s, y_{1}, z\right) \int_{B_{3 R}^{c}} J(z, v) h(s, v) d v d z \\
& +\int_{B_{3 R / 2} \backslash B_{5 R / 4}} p^{B_{3 R / 2}\left(t_{1}-s, y_{1}, z\right) \int_{B_{3 R}^{c}} J(z, v) h(s, v) d v d z=I_{1}+I_{2} .}
\end{aligned}
$$

When $z \in B_{3 R / 2} \backslash B_{5 R / 4}$, we have $\left|y_{1}-z\right| \geq R / 4$, so by Proposition $2.1(\mathrm{i}), p_{s}^{B_{3 R / 2}}\left(y_{1}, z\right) \leq c_{3} R^{-d}$ for some constant $c_{3}>0$ and $\int_{0}^{t_{1}} I_{2} d s$ is less than or equal to the right hand side of (4.4). For $z \in B_{5 R / 4}$ by $\mathbf{U J S}_{\leq 1}$,

$$
\begin{aligned}
\int_{B_{3 R}^{c}} J(z, v) h(s, v) d v & \leq \frac{c_{4}}{R^{d}} \int_{B(z, R / 6)} \int_{B_{3 R}^{c}} J\left(z^{\prime}, v\right) h(s, v) d v d z^{\prime} \\
& \leq \frac{c_{4}}{R^{d}} \int_{B_{3 R / 2}} \int_{B_{3 R}^{c}} J\left(z^{\prime}, v\right) h(s, v) d v d z^{\prime}
\end{aligned}
$$

since $B(z, R / 6) \subset B_{3 R / 2}$. Since the right hand side of the above inequality does not depend on $z$ anymore, multiplying both sides by $p^{B_{3 R / 2}}\left(t_{1}-s, y_{1}, z\right)$ and integrating over $z \in B_{5 R / 4}$ (and further integrating over $\int_{2 \delta R^{\alpha} / 3}^{t_{1}} d s$ ), we obtain $\int_{0}^{t_{1}} I_{1} d s$ is less than or equal to the right hand side of (4.4). This proves the lemma.

Proof of Theorem 4.1. Let $R_{*}$ denote the constant given in Proposition 2.1. We first consider the case that $u$ is non-negative and bounded on $[0, \infty) \times \mathbb{R}^{d}$.
(1) Suppose $R \leq R_{*} / 2$. When $t \in\left(0, R_{*}^{\alpha}\right]$ and $|x-y| \leq R_{*}$, one can prove [CK1, Lemmas 4.11] (also see [CK2, Lemmas 6.2]) from our heat kernel estimates in Proposition 2.1. Given our Lemma 4.2 and the lemmas corresponding to [CK1, Lemmas 4.11, and 4.13], the proof of the parabolic Harnack inequality is similar to those in [CK1, CK2] with some modification. We skip the details here. Interested reader can find its full proof in [CKK].
(2) Suppose $R \in\left(R_{*} / 2,1\right]$ and let $\left(t_{1}, x_{1}\right) \in Q_{-}$and $\left(t_{2}, x_{2}\right) \in Q_{+}$. Without loss of generality, we may assume $x_{0}=0$ and $t_{0}=0$. We further assume that $\left|x_{1}-x_{2}\right| \leq R_{*} / 8$. If not, we just repeat the argument below at most $16\left[R / R^{*}\right]$ times.

For notational convenience, denote $R_{*} / 2$ by $r_{*}$ and let $B^{1}=B\left(x_{1}, r_{*}\right), B^{2}=B\left(x_{1}, r_{*} / 2\right)$. Define

$$
Q_{1}=\left(t_{1}+\frac{\delta}{2} \psi\left(r_{*}\right), t_{1}+\frac{3 \delta}{4} \psi\left(r_{*}\right)\right) \times\left(B^{1} \backslash B^{2}\right) \quad \text { and } \quad Q_{2}=\left[0, t_{2}\right] \times B^{2}
$$

Since $u$ is parabolic, by the case (1) but with

$$
\left(t_{1}-\frac{\delta}{4} \psi\left(r_{*}\right), t_{1}+\frac{\delta}{4} \psi\left(r_{*}\right)\right) \times B^{1} \quad \text { and } \quad\left(t_{1}+\frac{\delta}{2} \psi\left(r_{*}\right), t_{1}+\frac{3 \delta}{4} \psi\left(r_{*}\right)\right) \times B^{1}
$$

in place of $Q_{-}$and $Q_{+}$respective, we have

$$
\begin{aligned}
u\left(t_{2}, x_{2}\right) & =\mathbb{E}^{\left(t_{2}, x_{2}\right)}\left[u\left(Z_{\tau_{Q_{2}}}\right)\right] \\
& \geq \mathbb{E}^{\left(t_{2}, x_{2}\right)}\left[u\left(Z_{\tau_{Q_{2}}}\right): Z_{\tau_{Q_{2}}} \in Q_{1}\right] \geq c_{1} u\left(t_{1}, x_{1}\right) \mathbb{P}^{\left(t_{2}, x_{2}\right)}\left(Z \tau_{Q_{2}} \in Q_{1}\right)
\end{aligned}
$$

Since $|y-z|<2 r_{*}=R_{*}<1$ for every $(y, z) \in B^{2} \times B^{1}$, we have by the Lévy system formula for $X$ that

$$
\begin{aligned}
\mathbb{P}^{\left(t_{2}, x_{2}\right)}\left(Z_{\tau_{Q_{2}}} \in Q_{1}\right) & =\mathbb{P}^{x_{2}}\left(X_{\tau_{B^{2}}} \in B^{1}, t_{2}-t_{1}-\frac{3 \delta}{4} \psi\left(r_{*}\right)<\tau_{B^{2}}<t_{2}-t_{1}-\frac{\delta}{2} \psi\left(r_{*}\right)\right) \\
& \geq c_{2} \int_{t_{2}-t_{1}-\frac{3 \delta}{4} \psi\left(r_{*}\right)}^{t_{2}-t_{1}-\frac{\delta}{2} \psi\left(r_{*}\right)} \int_{B^{2}}\left(\int_{B^{1} \backslash \overline{B^{2}}} \frac{p^{B^{2}}\left(s, x_{2}, y\right)}{|y-z|^{d+\alpha}} d z\right) d y d s \\
& \geq c_{3} \int_{t_{2}-t_{1}-\frac{3 \delta}{4} \psi\left(r_{*}\right)}^{t_{2}-t_{1}-\frac{\delta}{2} \psi\left(r_{*}\right)} \int_{B^{2}} p^{B^{2}}\left(s, x_{2}, y\right) d y d s
\end{aligned}
$$

for some positive constants $c_{2}=c_{2}(\alpha, d)$ and $c_{3}=c_{3}\left(\alpha, d, R_{*}\right)$. Note that

$$
\frac{\delta}{4} \psi\left(r_{*}\right) \leq t_{2}-t_{1}-\frac{3 \delta}{4} \psi\left(r_{*}\right) \leq t_{2}-t_{1}-\frac{\delta}{2} \psi\left(r_{*}\right) \leq 3 \delta \psi\left(2 r_{*}\right)
$$

Applying Proposition 3.8 to $p^{B^{2}}\left(s, x_{2}, y\right)$, we have

$$
\mathbb{P}^{\left(t_{2}, x_{2}\right)}\left(Z_{\tau_{Q_{2}}} \in Q_{1}\right) \geq c_{4} \int_{t_{2}-t_{1}-\frac{3 \delta}{4} \psi\left(r_{*}\right)}^{t_{2}-t_{1}-\frac{\delta}{2} \psi\left(r_{*}\right)} \int_{B\left(x_{1}, \psi\left(r_{*} / 8\right)\right)} s^{-d / \alpha} \mu_{d}(d y) d s \geq c_{5} \frac{\delta}{4} \psi\left(r_{*}\right)>0
$$

This proves that $u\left(t_{2}, x_{2}\right) \geq c_{6} u\left(t_{1}, x_{1}\right)$ for some positive constant $c_{6}=c_{6}\left(d, \alpha, R_{*}, \delta\right)$.
(3) Now let's consider the case $R \geq 1$. We will use balayage; see [BG, Chapter VI] for details, and see [BBCK, Theorem 1.7] and [BBK, Proposition 3.3] for similar arguments. Without loss of generality, we may assume $x_{0}=0$ and $t_{0}=0$. Let $B=B(0,4 R), B^{\prime}=B(0,3 R), E=$ $(0,6 \delta \psi(R)) \times \overline{B^{\prime}}, Q=(0,6 \delta \psi(R)) \times B$. As in the proof of [BBCK, Theorem 1.7], we define $u_{E}$, the réduite of $u$ with respect to $E$ by

$$
u_{E}(s, x)=\mathbb{E}^{(s, x)}\left[u\left(V_{T_{E}}, X_{T_{E}}\right): T_{E}<\tau_{Q}\right]
$$

where $T_{E}=\inf \left\{s \geq 0: Z_{s} \in E\right\}$; then $u=u_{E}$ on $E$. By the balayage formula, there exists a measure $\nu_{E}$ supported on $\bar{E}$ such that

$$
\begin{equation*}
u_{E}(t, x)=\int_{E} p^{B}(t-r, x, z) \nu_{E}(d r, d z) \quad \text { for all } \quad(t, x) \in Q \tag{4.5}
\end{equation*}
$$

where $p^{B}(s, x, y)=0$ if $s<0$.
Let $\left(t_{1}, x_{1}\right) \in Q_{-}$and $\left(t_{2}, x_{2}\right) \in Q_{+}$and observe that

$$
3 \delta \psi(R) \geq t_{2}-r \geq t_{2}-t_{1} \geq \delta \psi(R) \quad \text { for every } r \in\left[0, t_{1}\right]
$$

It follows from Theorem 3.5 and semigroup property that

$$
p^{B}\left(t_{2}-r, y, z\right) \geq c_{1} R^{-d}, \quad \text { for all } y, z \in \overline{B^{\prime}}, r \in\left[0, t_{1}\right]
$$

The above gives us that

$$
u_{E}\left(t_{2}, x_{2}\right) \geq \int_{\left[0, t_{1}\right] \times \overline{B^{\prime}}} p^{B}(t-r, x, z) \nu_{E}(d r, d z) \geq \frac{c_{1}}{R^{d}} \nu_{E}\left(\left[0, t_{1}\right] \times \overline{B^{\prime}}\right)
$$

Thus in order to prove the parabolic Harnack inequality, it suffices to show the following for each $\left(t_{1}, x_{1}\right) \in Q_{-}$;

$$
\begin{equation*}
u_{E}\left(t_{1}, x_{1}\right)=\int_{\left[0, t_{1}\right] \times \overline{B^{\prime}}} p^{B}\left(t_{1}-r, x_{1}, z\right) \nu_{E}(d r, d z) \leq \frac{c_{2}}{R^{d}} \nu_{E}\left(\left[0, t_{1}\right] \times \overline{B^{\prime}}\right) . \tag{4.6}
\end{equation*}
$$

Since the jumps of the process $X$ are bounded by 1 and $R \geq 1, u_{E}$ is parabolic (caloric) on $(0,6 \delta \psi(R)) \times B(0,2 R)$. It follows that the support of $\nu_{E}$ is contained in $\bar{E} \backslash(0,6 \delta \psi(R)) \times B(0,2 R)$. Thus, we can write

$$
u_{E}\left(t_{1}, x_{1}\right)=\int_{F_{1}\left(t_{1}\right) \cup F_{2}} p^{B}\left(t_{1}-r, x_{1}, z\right) \nu_{E}(d r, d z),
$$

where $F_{1}(t):=[0, t] \times\left(\overline{B^{\prime}} \backslash B(0, R)\right), F_{2}=\{0\} \times \overline{B^{\prime}}$. If $(r, z) \in F_{1}\left(t_{1}\right)$, then $\left|x_{1}-z\right| \geq R$, so by (2.3) when $t_{1}-r \geq \delta \psi(R)$ and by Proposition 2.1 (i) and Theorem 2.3 otherwise, we have $p^{B}\left(t_{1}-r, x_{1}, z\right) \leq c_{3} R^{-d}$. If $(r, z) \in F_{2}$, then $t_{1}-r \geq \delta \psi(R)$ and by (2.3) again, we have $p^{B}\left(t_{1}-r, x_{1}, z\right) \leq c_{3} R^{-d}$. Thus,

$$
u_{E}\left(t_{1}, x_{1}\right) \leq \frac{c_{3}}{R^{d}} \nu_{E}\left(F_{1}\left(t_{1}\right) \cup F_{2}\right) \leq \frac{c_{2}}{R^{d}} \nu_{E}\left(\left[0, t_{1}\right] \times \overline{B^{\prime}}\right)
$$

and (4.6) is established.
Finally, we will prove (4.1) when $u$ is not necessarily bounded on $[0, \infty) \times \mathbb{R}^{d}$. Let $U$ be a bounded domain such that $\overline{Q_{-} \cup Q_{+}} \subset U \subset \bar{U} \subset\left(t_{0}, t_{0}+6 \delta \psi(R)\right) \times B\left(x_{0}, 4 R\right)$. For any $n \in \mathbb{N}$, define $u_{n}(t, x)=\mathbb{E}^{(t, x)}\left[(u \wedge n)\left(Z_{\tau_{U}}\right)\right]$. Then $u_{n}$ is non-negative and bounded on $[0, \infty) \times \mathbb{R}^{d}$, parabolic on $U$ and $\lim _{n \rightarrow \infty} u_{n}(t, x)=u(t, x)$ for $x \in[0, \infty) \times \mathbb{R}^{d}$. From the above arguments, we see that (4.1) holds for $u_{n}$ with the constant $c$ independent of $n$. Letting $n \rightarrow \infty$, we obtain (4.1) for $u$.

By the same proof as that for [CK1, Theorem 4.14] or [CK2, Proposition 4.14], we have the following Hölder continuity for parabolic functions.

Theorem 4.3 For every $R_{0} \in(0,1]$, there are constants $c=c\left(R_{0}\right)>0$ and $\kappa>0$ such that for every $0<R \leq R_{0}$ and every bounded parabolic function $h$ in $Q\left(0, x_{0}, 2 R\right):=\left(0,(2 R)^{\alpha}\right) \times B\left(x_{0}, 2 R\right)$,

$$
\begin{equation*}
|h(s, x)-h(t, y)| \leq c\|h\|_{\infty, R} R^{-\kappa}\left(|t-s|^{1 / \alpha}+|x-y|\right)^{\kappa} \tag{4.7}
\end{equation*}
$$

holds for $(s, x),(t, y) \in Q\left(0, x_{0}, R\right)$, where $\|h\|_{\infty, R}:=\sup _{(t, y) \in\left[0,(2 R)^{\alpha}\right] \times \mathbb{R}^{d}}|h(t, y)|$. In particular, for the transition density function $p(t, x, y)$ of $X$, for any $t_{0} \in(0,1)$, there are constants $c=c\left(t_{0}\right)>0$ and $\kappa>0$ such that for any $t, s \in\left[t_{0}, 1\right]$ and $\left(x_{i}, y_{i}\right) \in \mathbb{R}^{d} \times \mathbb{R}^{d}$ with $i=1,2$,

$$
\begin{equation*}
\left|p\left(s, x_{1}, y_{1}\right)-p\left(t, x_{2}, y_{2}\right)\right| \leq c t_{0}^{(-d+\kappa) / \alpha}\left(|t-s|^{1 / \alpha}+\left|x_{1}-x_{2}\right|+\left|y_{1}-y_{2}\right|\right)^{\kappa} \tag{4.8}
\end{equation*}
$$

Remark 4.4 (i) Since the heat kernel $p(t, x, y)$ is Hölder continuous, the estimates derived in previous sections for $p(t, x, y)$ hold for every $x, y \in \mathbb{R}^{d}$.
(ii) Note that the proof of Theorem 4.3 needs only the short time heat kernel estimates in Proposition 2.1 on $p(t, x, y)$ for $t \in\left(0, T^{*}\right]$ and for q.e. $x, y \in \mathbb{R}^{d}$ having $|x-y| \leq R_{*}$ for some $T_{*} \in(0,1)$ and $R_{*} \in(0,1]$. Therefore as long as a pure jump symmetric strong Markov process $Y$ has a transition density function $p(t, x, y)$ that has the two sided short time finite range estimates as that in Theorem 4.3 for $t \in\left(0, t_{0}\right)$ and $(x, y) \in \mathbb{R}^{d} \times \mathbb{R}^{d}$ with $|x-y| \leq r_{0}$ for some $t_{0}$ and $r_{0}>0$, it can be established directly from these heat kernel estimate that every bounded parabolic functions of $Y$ is Hölder continuous. If in addition we have (UJS) $\leq 1$ and the lower bound on the heat kernel $p^{B}(t, x, y)$ as in Proposition 3.8, then the parabolic Harnack inequality (Theorem 4.1) can be proved for $R \leq R_{*} / 2$.
(iii) In fact, $(\mathrm{UJS})_{\leq 1}$ is necessary for the parabolic Harnack inequality for $R \leq 1$. This is proved in [BBK, Proposition 4.7] for the discrete space setting, and the proof for the continuous space case can be found in [CKK].
(iv) There is a minor gap in the proof of [CK2, Lemma 6.1]. Condition (UJS) $\leq 1$ should be imposed on the jumping kernel $J$ for this lemma and consequently for the main results (such as Theorems 1.2 and 4.12) of [CK2]. Note that $(U J S)_{\leq 1}$ is automatically satisfied if $\psi \equiv 1$ in (1.12) of [CK2] (corresponds to the case $\gamma_{1}=\gamma_{2}=0$ ). A sufficient condition for $J$ to satisfy condition $(\mathrm{UJS})_{\leq 1}$ is that the function $\psi$ in (1.12) of [CK2] has the property that $\psi(r+1) \leq c_{0} \psi(r)$ for every $r \geq 1$.

Suppose that $Y$ is the Hunt process associated with Dirichlet form $(\mathcal{Q}, \mathcal{D})$ given by (1.2)-(1.3) whose jumping intensity kernel $J(x, y)$ has the property that $J(x, y) \mathbf{1}_{\{(x, y): d(x, y)>\kappa\}}$ is bounded and

$$
\begin{equation*}
J(x, y)=\frac{c(x, y)}{|x-y|^{d+\alpha}} \quad \text { for }|x-y| \leq 1 \quad \text { and } \quad \sup _{x \in \mathbb{R}^{d}} \int_{\left\{y \in \mathbb{R}^{d}:|y-x|>1\right\}} J(x, y) d y<\infty \tag{4.9}
\end{equation*}
$$

where $c(x, y)$ is a function that is bounded between two positive constants and is symmetric in $x$ and $y$. Then by the Meyer's construction method (see [BBCK, Lemma 3.6] and [BGK, Lemma $3.1(\mathrm{c})$ and (3.18)]), the process $Y$ can be constructed from the finite range $\alpha$-stable-like process $X$ having jump intensity kernel $\frac{c(x, y)}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y| \leq 1\}}$ and so $Y$ has a transition density function $q(t, x, y)$ with respect to $\mu_{d}$. Moreover, for any ball $B \subset \mathbb{R}^{d}$,

$$
\begin{equation*}
q(t, x, y) \geq e^{-t\|\mathcal{J}\|_{\infty}} p(t, x, y) \quad \text { and } \quad q^{B}(t, x, y) \geq e^{-t\|\mathcal{J}\|_{\infty}} p^{B}(t, x, y) \tag{4.10}
\end{equation*}
$$

and

$$
\begin{equation*}
q(t, x, y) \leq p(t, x, y)+t\left\|J_{1}\right\|_{\infty} \tag{4.11}
\end{equation*}
$$

where $p(t, x, y)$ is the transition density function of $X$,

$$
J_{1}(x, y):=J(x, y) \mathbf{1}_{\{d(x, y)>1\}} \quad \text { and } \quad \mathcal{J}(x):=\int_{\mathbb{R}^{d}} J_{1}(x, y) \mu_{d}(d y)
$$

Thus using the heat kernel estimate for $p(t, x, y)$ in Proposition 2.1 and Proposition 3.8, by the same line of argument as that in the Remark 4.4(ii) we have the following.

Theorem 4.5 The Hölder continuity estimate (4.7) holds for bounded parabolic functions of $Y$. In particular, all these applies to the transition density function $q(t, x, y)$ of $Y$. Moreover, if we assume $(U J S)_{\leq R_{1}}$ in addition, then the parabolic Harnack inequality (4.1) holds for non-negative parabolic functions of $Y$ with $R \leq R_{1} / 2$.

The full detail of the above theorem will be given in more general format in [CKK].
Remark 4.6 Very recently, Kassmann [Ka, Theorem 1.1] proved by a quite different analytic method the Hölder continuity for bounded harmonic functions of symmetric pure jump processes whose jumping intensity kernel $J$ satisfies the condition

$$
J(x, y)=c(x, y)|x-y|^{-d-\alpha} \quad \text { for }|x-y| \leq 1 \quad \text { and } \quad J(x, y) \leq c|x-y|^{-d-\eta} \quad \text { for }|x-y|>1,
$$

where $\eta>0$ and $c>0$ are two positive constants. Clearly, such type of jumping kernel is a special case of those given by (4.9). Since every harmonic function is parabolic, our Theorem 4.5 recovers and extends the main result of [Ka]. See $[\mathrm{S}]$ for some related work on the Hölder continuity of bounded harmonic functions for a class of non-local operators.

### 4.2 Two-sided Green Function Estimates

When $d=1,2$, the finite range $\alpha$-stable-like processes are all recurrent. So in this subsection, we assume $d \geq 3$ and give two-sided sharp estimates the Green function for $G(x, y)$ of finite range stable-like process $X$ in $\mathbb{R}^{d}$ where

$$
G(x, y):=\int_{0}^{\infty} p(t, x, y) d t, \quad x, y \in \mathbb{R}^{d} .
$$

Theorem 4.7 There exists $c=c(\alpha, d)>1$ such that

$$
c^{-1}\left(\frac{1}{|x-y|^{d-\alpha}}+\frac{1}{|x-y|^{d-2}}\right) \leq G(x, y) \leq c\left(\frac{1}{|x-y|^{d-\alpha}}+\frac{1}{|x-y|^{d-2}}\right), \quad x, y \in \mathbb{R}^{d} .
$$

Proof. We first note that for every $T, M \in[0, \infty)$

$$
\begin{equation*}
\int_{T}^{\infty} t^{-\frac{d}{2}} e^{-\frac{M|x-y|^{2}}{2 t}} d t=\frac{1}{|x-y|^{d-2}} \int_{0}^{\frac{|x-y|^{2}}{T}} u^{\frac{d-4}{2}} e^{-\frac{1}{2} M u} d u \tag{4.12}
\end{equation*}
$$

Recall that $R_{*}<1$ and $T_{*}=R_{*}^{\alpha}$ are the constants from Proposition 2.1(ii). Using (4.12), it is easy to see that, if $|x-y| \leq R_{*}$, by Proposition 2.1 (i) and Theorem 2.3

$$
\begin{aligned}
G(x, y) & \leq c_{1} \int_{0}^{|x-y|^{\alpha}} \frac{t}{|x-y|^{d+\alpha}} d t+c_{2} \int_{|x-y|^{\alpha}}^{T_{*}} t^{-d / \alpha} d t+c_{3} \int_{T_{*}}^{\infty} t^{-\frac{d}{2}} e^{-\frac{c_{4}|x-y|^{2}}{2 t}} d t \\
& \leq \frac{c_{5}}{|x-y|^{d-\alpha}}+\frac{c_{3}}{|x-y|^{d-2}} \int_{0}^{\frac{|x-y|^{2}}{T_{*}}} u^{\frac{d-4}{2}} e^{-\frac{1}{2} c_{4} u} d u \leq \frac{c_{6}}{|x-y|^{d-\alpha}} .
\end{aligned}
$$

On the other hand if $|x-y|>R_{*}$, by Theorem 2.3 and (4.12)

$$
\begin{aligned}
G(x, y) & \leq c_{7} \int_{0}^{C_{*}|x-y|} \exp \left(-c_{8}|x-y| \log \frac{|x-y|}{t}\right) d t+c_{9} \int_{C_{*}|x-y|}^{\infty} t^{-\frac{d}{2}} e^{-\frac{c_{4}|x-y|^{2}}{2 t}} d t \\
& \leq c_{7} \int_{0}^{C_{*}|x-y|} \exp \left(-c_{10}|x-y|\right) d t+\frac{c_{9}}{|x-y|^{d-2}} \int_{0}^{\frac{|x-y|}{C_{*}}} u^{\frac{d-4}{2}} e^{-\frac{1}{2} c_{4} u} d u \\
& \leq c_{7} C_{*}|x-y| \exp \left(-c_{10}|x-y|\right)+\frac{c_{11}}{|x-y|^{d-2}} \leq \frac{c_{12}}{|x-y|^{d-2}}
\end{aligned}
$$

where $C_{*}<1$ is given in Theorem 2.3.
The lower bounded is easier. If $|x-y| \leq R_{*}$, by Proposition 2.1(ii)

$$
G(x, y) \geq c_{13} \int_{0}^{|x-y|^{\alpha}} \frac{t}{|x-y|^{d+\alpha}} d t=\frac{c_{13}}{2|x-y|^{d-\alpha}}
$$

If $|x-y| \geq R_{*}$, by Theorem 3.6 and (4.12)

$$
G(x, y) \geq c_{14} \int_{|x-y|^{2}}^{\infty} t^{-\frac{d}{2}} d t=\frac{c_{14}}{|x-y|^{d-2}} \int_{0}^{1} u^{\frac{d-4}{2}} d u
$$

Remark 4.8 Under some mild assumptions on bounded open set $D$, when $c(x, y)$ is a constant, Green function $G_{D}(x, y)$ for $X$ in $D$ is comparable to the one for isotropically symmetric stable process in $D$ (see [GR, KS2]). Theorem 4.7 shows that, unlike bounded open sets, the behavior of the Green function for $X$ in $\mathbb{R}^{d}$ is different from the behavior of the Green function for isotropically symmetric stable process in $\mathbb{R}^{d}$.

Now let's consider the more general non-local Dirichlet form $(\mathcal{E}, \mathcal{F})$ in $(2.14)-(2.15)$ with the jumping kernel $J$ satisfying the condition (2.16). Recall that $q(t, x, y)$ is the transition density function for the associated Hunt process $Z$ with respect to the Lebesgue measure on $\mathbb{R}^{d}$. For $d \geq 3$, let

$$
V(x, y):=\int_{0}^{\infty} q(t, x, y) d t, \quad x, y \in \mathbb{R}^{d}
$$

Using Theorems 2.4 and 3.5 instead of Theorems 2.3 and 3.6 respectively in the proof of Theorem 4.7, we get the Green function estimate for the process $Z$ for $|x-y| \geq 1$.

Theorem 4.9 There exists $c=c(\alpha, d)>1$ such that

$$
c^{-1} \frac{1}{|x-y|^{d-2}} \leq V(x, y) \leq c \frac{1}{|x-y|^{d-2}} \quad \text { for }|x-y| \geq 1
$$

### 4.3 Differentiability of Spectral Functions

In [TT, TT2, Ts], the differentiability of spectral functions for symmetric stable processes are studied.

Recall that $X$ is a finite range stable-like process considered in this paper whose Dirichlet form $(\mathcal{Q}, \mathcal{D})$ is given by (1.2)-(1.3) whose jumping intensity kernel $J(x, y)=\frac{c(x, y)}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y| \leq 1\}}$. Let $\mu$ be a signed measure in Kato class $\mathbf{K}_{\infty}(X)$ as introduced in [C]. The associated spectral function $C(\lambda)$ is defined to be

$$
C(\lambda)=-\inf \left\{\mathcal{Q}(u, u)+\lambda \int_{\mathbb{R}^{d}} u(x)^{2} \mu(d x): u \in \mathcal{D} \text { with } \int_{\mathbb{R}^{d}} u(x)^{2} d x=1\right\}
$$

Using the heat kernel estimates established in this paper, by an almost same argument as that in [TT, TT2] and [Ts], it can be shown that if $d \leq 4$ and if the extended Dirichlet space $\left(\mathcal{Q}, \mathcal{D}_{e}\right)$ is compactly embedded into $\left.L^{2}\left(\mathbb{R}^{d},|\mu|\right)\right)$, then $\lambda \mapsto C(\lambda)$ is differentiable on $\mathbb{R}$. But we will not go into details about it.

## 5 Weighted Poincaré Inequality of Fractional Order

Throughout this section, $r \geq 1, \sigma \in(0, \infty)$ and $\alpha \in(0,2)$. Recall that $\mu_{d}$ denotes the Lebesgue measure in $\mathbb{R}^{d}$. In this section, the exact values of the constants $c$ 's are always independent of $r$ and they might change from one appearance to another. Let $\mathcal{M}(\sigma)$ be the set of all non-increasing function $\Psi$ from $[0,1]$ to $[0,1]$ such that $\Psi(s)>\Psi(1)=0$ for every $s \in[0,1)$ and

$$
\begin{equation*}
\Psi\left(s+\frac{1}{2}\left((1-s) \wedge \frac{1}{2}\right)\right) \geq \sigma \Psi(s), \quad s \in(0,1) \tag{5.1}
\end{equation*}
$$

We will use $\mathcal{N}(\sigma)$ to denote all the functions $\Phi$ of the form $c \Psi(|x|)$ for some $\Psi \in \mathcal{M}(\sigma)$ having $\int_{\mathbb{R}^{d}} \Phi(x) d x=1$. Note that, when $\beta \in(0,2), c\left(1-|x|^{2}\right)^{12 /(2-\beta)} \mathbf{1}_{B(0,1)}(x)$ is in $\mathcal{N}\left((1 / 8)^{12 /(2-\beta)}\right)$. Condition (5.1) says that for each $\Phi \in \mathcal{N}(\sigma)$, values of $\Phi$ at points with comparable distance from the unit sphere $\partial B(0,1)$ are comparable. This implies that values of $\Phi$ in balls in Whitney-type covering, which will be discussed below, are universally comparable to each other. This property will be used in many places below.

For $\Phi \in \mathcal{N}(\sigma)$, define

$$
u_{\Phi}:=\int_{B(0,1)} u(x) \Phi(x) d x
$$

This section is devoted to prove the following form of weighted Poincaré inequality.
Theorem 5.1 For every $d \geq 1,0<\alpha<2$ and $\sigma \in(0, \infty)$, there exists a positive constant $c_{1}=c_{1}(d, \alpha, \sigma)$ independent of $r \geq 1$, such that for every $\Phi \in \mathcal{N}(\sigma)$ and $u \in L^{1}(B(0,1), \Phi(x) d x)$,

$$
\begin{aligned}
& \int_{B(0,1)}\left(u(x)-u_{\Phi}\right)^{2} \Phi(x) d x \\
\leq & c_{1} \int_{B(0,1) \times B(0,1)}(u(x)-u(y))^{2} \frac{r^{2-\alpha}}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y| \leq 1 / r\}}(\Phi(x) \wedge \Phi(y)) d x d y
\end{aligned}
$$

Moreover, the constant $c_{1}$ stays bounded for $\alpha \in(0,2)$.

The exponent $2-\alpha$ of $r$ in the integral above is quite delicate to get. We will prove the above theorem through several lemmas. For the remainder of this section, we fix $\sigma \in(0, \infty)$ and $\Phi \in \mathcal{N}(\sigma)$.

We first prove the following simple lemma. Let

$$
u_{B(x, s)}:=\frac{1}{\mu_{d}(B(x, s))} \int_{B(x, s)} u(y) d y
$$

Lemma 5.2 For every $B(z, s) \subset B(0,1)$ and every $u \in L^{1}(B(z, s), d x)$,

$$
\int_{B(z, s)}\left(u(x)-u_{B(z, s)}\right)^{2} d x \leq \frac{1}{\mu_{d}(B(z, s))} \int_{B(z, s)} \int_{B(z, s)}(u(x)-u(y))^{2} d x d y
$$

Proof. By Cauchy-Schwartz inequality,

$$
\begin{aligned}
\int_{B(z, s)}\left(u(x)-u_{B(z, s)}\right)^{2}(x) d x & =\int_{B(z, s)}\left(\frac{1}{\mu_{d}(B(z, s))} \int_{B(z, s)}(u(x)-u(y)) d y\right)^{2} d x \\
& \leq \frac{1}{\mu_{d}(B(z, s))} \int_{B(z, s)} \int_{B(z, s)}(u(x)-u(y))^{2} d x d y
\end{aligned}
$$

Recall Whitney-type coverings (see [SC, Section 5.3.3] for details): We first let

$$
\overline{\mathcal{W}}:=\left\{B: \text { the center of the ball } B \text { is in } B(0,1) \text { and } r(B)=\frac{1}{10^{3}} \rho(B)\right\}
$$

where $r(B)$ is the radius of the ball $B$ and $\rho(B)$ denotes the Euclidean distance between the ball $B$ and $B(0,1)^{c}$. In the sequel, for $\lambda>0$ and a ball $B=B(x, r)$ centered at $x$ with radius $r$, we denote $\lambda B$ the concentric ball $B(x, \lambda r)$ with radius $\lambda r$.

Start $\mathcal{W}$ by picking a ball $B^{0} \in \overline{\mathcal{W}}$ with the largest possible radius. Pick the next ball $B^{1}$ to be a ball in $\overline{\mathcal{W}}$ which does not intersect $B^{0}$ and has maximal radius. Assuming that $k$ balls $B^{0}, \cdots, B^{k-1}$ have already been picked, pick the next ball $B^{k}$ to be a ball in $\overline{\mathcal{W}}$ which does not intersect $\cup_{j=0}^{k-1} B^{j}$ and has maximal radius. Though this procedure, we get a sequence of disjoint balls $\mathcal{W}:=\left\{B^{0}, \cdots, B^{k-1}, B^{k}, \cdots\right\}$ from $\overline{\mathcal{W}}$. Moreover, the Whitney-type decomposition of the unit ball $B(0,1)$ has the following properties (see, for example, page 135 of [SC]).

$$
\begin{equation*}
B(0,1)=\bigcup_{B \in \mathcal{W}} 2 B \tag{1}
\end{equation*}
$$

(2) There exists a positive constant $K$ such that

$$
\begin{equation*}
\sup _{y \in B(0,1)} \#\left\{B \in \mathcal{W}: y \in 10^{2} B\right\} \leq K \tag{5.2}
\end{equation*}
$$

where $\# S$ is the number of elements in the set $S$.

There exists a ball $B(0) \in \mathcal{W}$ such that $0 \in 2 B(0)$. We pick an fix such a ball $B(0)$ and call it the central ball of $\mathcal{W}$. For any $B \in \mathcal{W}$, let $\gamma_{B}$ be the straight line segment between the center of $B$ and the origin. Let

$$
\overline{\mathcal{W}}(B):=\left\{A \in \mathcal{W}: 2 A \cap \gamma_{B} \neq \emptyset\right\} .
$$

Now we define the chain $\mathcal{W}(B):=\left(B_{0}, B_{1}, \cdots, B_{l(B)-1}\right)$ with $B_{0}=B(0)$ and $B_{l(B)-1}=B$ as follows; Starting from the origin, let $y_{0}$ be the first point along $\gamma_{B}$ which does not belong to $2 B_{0}$. Define $B_{1}$ to be (any) one of balls in $\overline{\mathcal{W}}(B)$ such that $y_{0} \in 2 B_{1}$. Inductively, having $B_{0}, B_{1}, \cdots, B_{k}$ constructed, let $y_{k}$ be the first point along $\gamma_{B}$ which does not belong to $\cup_{j=0}^{k} 2 B_{j}$. Define $B_{k+1}$ to be (any) one of balls in $\overline{\mathcal{W}}(B)$ such that $y_{k} \in 2 B_{k+1}$. When the last chosen is not $B$, we simply add $B$ as the last ball in $\mathcal{W}(B)$.

Using Lemma 5.2, the next lemma can be proved easily.

Lemma 5.3 There exists a positive constant $c=c(d)$ such that for every $B \in \mathcal{W}, B_{i}, B_{i+1} \in \mathcal{W}(B)$ and for every $u \in L^{1}(B(0,1), \Phi d x)$,

$$
\left|u_{4 B_{i}}-u_{4 B_{i+1}}\right| \leq \sum_{j=0}^{1} \frac{c}{\mu_{d}\left(B_{i+j}\right)}\left(\int_{4 B_{i+j}} \int_{4 B_{i+j}}(u(x)-u(y))^{2} d x d y\right)^{1 / 2}
$$

Proof. Note that

$$
\begin{aligned}
& \left(\mu_{d}\left(4 B_{i} \cap 4 B_{i+1}\right)\right)^{1 / 2}\left|u_{4 B_{i}}-u_{4 B_{i+1}}\right| \\
= & \left(\int_{4 B_{i} \cap 4 B_{i+1}}\left|u_{4 B_{i}}-u_{4 B_{i+1}}\right|^{2} \mu_{d}(d x)\right)^{1 / 2} \\
\leq & \left(\int_{4 B_{i}}\left|u(x)-u_{4 B_{i}}\right|^{2} \mu_{d}(d x)\right)^{1 / 2}+\left(\int_{4 B_{i+1}}\left|u(x)-u_{4 B_{i+1}}\right|^{2} \mu_{d}(d x)\right)^{1 / 2} .
\end{aligned}
$$

Now the lemma follows from our Lemma 5.2 and the fact that

$$
\mu_{d}\left(4 B_{i} \cap 4 B_{i+1}\right) \geq c \max \left\{\mu_{d}\left(B_{i}\right), \mu_{d}\left(B_{i+1}\right)\right\}
$$

(see Lemma 5.3.7 in [SC]).

Lemma 5.4 There exists a positive constant $c=c(d, \sigma)$ such that for every $B \in \mathcal{W}, B_{i}, B_{i+1} \in$ $\mathcal{W}(B)$ and for every $u \in L^{1}(B(0,1), \Phi d x)$,

$$
\sqrt{\Phi_{B}}\left|u_{4 B_{i}}-u_{4 B_{i+1}}\right| \leq \sum_{j=0}^{1} \frac{c}{\mu_{d}\left(B_{i+j}\right)}\left(\int_{4 B_{i+j}} \int_{4 B_{i+j}}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)^{1 / 2}
$$

Proof. Since the values of $\Phi$ are universally comparable to each other on $4 B$ for every $B \in \mathcal{W}$, we have from Lemma 5.3

$$
\begin{align*}
& \left|u_{4 B_{i}}-u_{4 B_{i+1}}\right|  \tag{5.3}\\
\leq & \sum_{j=0}^{1} \frac{c}{\left(\mu_{d}\left(B_{i+j}\right)\right)^{1 / 2}\left(\int_{B_{i+j}} \Phi(y) d y\right)^{1 / 2}}\left(\int_{4 B_{i+j}} \int_{4 B_{i+j}}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)^{1 / 2}
\end{align*}
$$

Note that

$$
\begin{equation*}
\rho(A)=10^{3} r(A) \geq \frac{10^{3}}{4} r(B)=\frac{1}{4} \rho(B) \quad \text { for every } A \in \mathcal{W}(B) \tag{5.4}
\end{equation*}
$$

(See Lemma 5.3.6 in [SC].) Using (5.1), (5.4) and the fact that $\Psi$ is non-increasing, there exists a positive constant $c$ independent of $B$ such that

$$
\max _{y \in B} \Phi(y) \leq c \min _{y \in A} \Phi(y) \quad \text { for every } A \in \mathcal{W}(B)
$$

Thus we have

$$
\begin{equation*}
\Phi_{B}=\frac{1}{\mu_{d}(B)} \int_{B} \Phi(y) d y \leq c \frac{1}{\mu_{d}\left(B_{i}\right)} \int_{B_{i}} \Phi(y) d y \quad \text { for every } B_{i} \in \mathcal{W}(B) \tag{5.5}
\end{equation*}
$$

The lemma follows from (5.3) and (5.5).
The proof of the next lemma is similar to that of Theorem 5.3 .4 on page 141-143 of [SC]. For reader's convenience, we nevertheless spell out the details of the proof here.

Lemma 5.5 There exists a positive constant $c=c(d, \sigma)$ such that for every $u \in L^{1}(B(0,1), \Phi d x)$,

$$
\int_{B(0,1)}\left(u(x)-u_{\Phi}\right)^{2} \Phi(x) d x \leq c \sum_{A \in \mathcal{W}} \frac{1}{\mu_{d}(A)} \int_{4 A \times 4 A}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y
$$

Proof. Note that

$$
\begin{aligned}
& \int_{B(0,1)}\left(u(x)-u_{\Phi}\right)^{2} \Phi(x) d x \\
\leq & 2 \int_{B(0,1)}\left(u(x)-u_{4 B(0)}\right)^{2} \Phi(x) d x+2\left(\int_{B(0,1)} \Phi(x) d x\right)\left(u_{\Phi}-u_{4 B(0)}\right)^{2} \\
\leq & 2 \int_{B(0,1)}\left(u(x)-u_{4 B(0)}\right)^{2} \Phi(x) d x+2 \int_{B(0,1)}\left(u(x)-u_{4 B(0)}\right)^{2} \Phi(x) d x \\
\leq & 4 \sum_{B \in \mathcal{W}} \int_{4 B}\left(u(x)-u_{4 B(0)}\right)^{2} \Phi(x) d x \\
\leq & 8 \sum_{B \in \mathcal{W}} \int_{4 B}\left(u(x)-u_{4 B}\right)^{2} \Phi(x) d x+8 \sum_{B \in \mathcal{W}}\left(u_{4 B}-u_{4 B(0)}\right)^{2} \int_{4 B} \Phi(x) d x \\
\leq & c \sum_{B \in \mathcal{W}} \frac{1}{\mu_{d}(B)} \int_{4 B \times 4 B}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y \\
& +c \sum_{B \in \mathcal{W}} \int \mathbf{1}_{B}(z)\left(\left|u_{4 B}-u_{4 B(0)}\right|\left(\Phi_{B}\right)^{1 / 2}\right)^{2} d z,
\end{aligned}
$$

where in the last inequality, we used the fact that the values of $\Phi$ are universally comparable to each other on $4 B$ for every $B \in \mathcal{W}$. To establish the lemma, it suffices to deal with the second summation above.

By Lemma 5.4, we get

$$
\begin{aligned}
& \left|u_{4 B}-u_{4 B(0)}\right|\left(\Phi_{B}\right)^{1 / 2} \mathbf{1}_{B}(z) \\
\leq & \sum_{i=0}^{l(B)-2}\left|u_{4 B_{i}}-u_{4 B_{i+1}}\right|\left(\Phi_{B}\right)^{1 / 2} \mathbf{1}_{B}(z) \\
\leq & c \sum_{i=0}^{l(B)-1} \frac{1}{\mu_{d}\left(B_{i}\right)}\left(\int_{4 B_{i}} \int_{4 B_{i}}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)^{1 / 2} \mathbf{1}_{B}(z) \\
= & c \sum_{i=0}^{l(B)-1} \frac{1}{\mu_{d}\left(B_{i}\right)}\left(\int_{4 B_{i}} \int_{4 B_{i}}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)^{1 / 2} \mathbf{1}_{10^{4} B_{i}}(z) \mathbf{1}_{B}(z) \\
\leq & c \sum_{A \in \mathcal{W}} \frac{1}{\mu_{d}(A)}\left(\int_{4 A} \int_{4 A}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)^{1 / 2} \mathbf{1}_{10^{4} A}(z) \mathbf{1}_{B}(z)
\end{aligned}
$$

In the first equality above, we have used the fact that $B \subset 10^{4} B_{i}$ (Lemma 5.3 .8 in [SC]). Since the balls in $\mathcal{W}$ are disjoint, summing both sides over $B \in \mathcal{W}$ and taking the square, we get

$$
\begin{aligned}
& \sum_{B \in \mathcal{W}} \mathbf{1}_{B}(z)\left(\left|u_{4 B}-u_{4 B(0)}\right|\left(\Phi_{B}\right)^{1 / 2}\right)^{2} \\
\leq & c\left(\sum_{A \in \mathcal{W}} \frac{1}{\mu_{d}(A)}\left(\int_{4 A} \int_{4 A}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)^{1 / 2} \mathbf{1}_{10^{4} A}(z)\right)^{2}
\end{aligned}
$$

Integrating over $z \in B(0,1)$, and using Lemma 5.3.12 in $[\mathrm{SC}]$ and the fact the balls in $\mathcal{W}$ are disjoint, we have

$$
\begin{aligned}
& \sum_{B \in \mathcal{W}} \int \mathbf{1}_{B}(z)\left(\left|u_{4 B}-u_{4 B(0)}\right| \Phi_{B}^{1 / 2}\right)^{2} d z \\
\leq & c \int\left(\sum_{A \in \mathcal{W}} \frac{1}{\mu_{d}(A)}\left(\int_{4 A} \int_{4 A}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)^{1 / 2} \mathbf{1}_{10^{4} A}(z)\right)^{2} d z \\
\leq & c \int\left(\sum_{A \in \mathcal{W}} \frac{1}{\mu_{d}(A)}\left(\int_{4 A} \int_{4 A}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)^{1 / 2} \mathbf{1}_{A}(z)\right)^{2} d z \\
\leq & c \int \sum_{A \in \mathcal{W}} \frac{1}{\left(\mu_{d}(A)\right)^{2}}\left(\int_{4 A} \int_{4 A}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right) \mathbf{1}_{A}(z) d z \\
\leq & c \sum_{A \in \mathcal{W}} \frac{1}{\mu_{d}(A)}\left(\int_{4 A} \int_{4 A}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y\right)
\end{aligned}
$$

This completes the proof for the lemma.

Lemma 5.6 There exists a positive constant $c=c(d, \sigma)$ such that for every $u \in L^{1}(B(0,1), \Phi d x)$,

$$
\begin{aligned}
& \int_{B(0,1)}\left(u(x)-u_{\Phi}\right)^{2} \Phi(x) d x \\
\leq & \frac{c}{10^{2 \alpha}} \int_{B(0,1) \times B(0,1)} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{10^{2}}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y .
\end{aligned}
$$

Proof. Since $|x-y| \leq 8 r(A) \leq \frac{1}{10^{2}}$ if $x, y \in 4 A$, we have for every $A \in \mathcal{W}$

$$
\begin{aligned}
& \frac{1}{\mu_{d}(A)} \int_{4 A \times 4 A}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & \frac{c}{(r(A))^{d}} \int_{4 A \times 4 A} \frac{(u(x)-u(y))^{2}|x-y|^{d+\alpha}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{10^{2}}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & \frac{c}{10^{2 \alpha}} \int_{4 A \times 4 A} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{\left.10^{2}\right\}}\right.}(\Phi(x) \wedge \Phi(y)) d x d y .
\end{aligned}
$$

It then follows from Lemma 5.5 and (5.2) that

$$
\begin{aligned}
& \int_{B(0,1)}\left(u(x)-u_{\Phi}\right)^{2} \Phi(x) d x \\
\leq & \frac{c}{10^{2 \alpha}} \sum_{A \in \mathcal{W}} \int_{4 A \times 4 A} \frac{(u(x)-u(y))^{2}}{\left.|x-y|\right|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{10^{2}}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & \frac{c}{10^{2 \alpha}} \int_{B(0,1) \times B(0,1)} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{\left.10^{2}\right\}}\right.}(\Phi(x) \wedge \Phi(y)) d x d y .
\end{aligned}
$$

Due to Lemma 5.6, we have Theorem 5.1 for $1 \leq r \leq 10^{2}$. So, from now we may assume $r>10^{2}$.
Lemma 5.7 There exists a positive constant $c=c(d, \sigma)$ such that for every $r>10^{2}$ for every $u \in L^{1}(B(0,1), \Phi d x)$,

$$
\begin{aligned}
& \int_{B(0,1)}\left(u(x)-u_{\Phi}\right)^{2} \Phi(x) d x \\
\leq & c \int_{B(0,1) \times B(0,1)}(u(x)-u(y))^{2} \frac{r^{-\alpha}}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y|<1 / r\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
& +c \int_{B\left(0,1-\frac{10}{r}\right) \times B\left(0,1-\frac{10}{r}\right)} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y|<\frac{1}{10^{2}}\right.}(\Phi(x) \wedge \Phi(y)) d x d y .
\end{aligned}
$$

Proof. By Lemma 5.5, we have

$$
\begin{aligned}
& \int_{B(0,1)}\left(u(x)-u_{\Phi}\right)^{2} \Phi(x) d x \leq c \sum_{A \in \mathcal{W}} \int_{4 A \times 4 A} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}}\left(\frac{|x-y|}{r(A)}\right)^{d}(\Phi(x) \wedge \Phi(y)) d x d y \\
& \leq c\left(\sum_{A \in \mathcal{W}: r(A) \leq \frac{1}{10 r}}+\sum_{A \in \mathcal{W}: r(A)>\frac{1}{10 r}}\right) \int_{4 A \times 4 A} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}}(\Phi(x) \wedge \Phi(y)) d x d y \\
= & I+I I .
\end{aligned}
$$

If $A \in \mathcal{W}$ and $r(A) \leq \frac{1}{10 r}$, then $|x-y| \leq 8 r(A)<\frac{1}{r}$ for every $x, y \in 4 A$. So using (5.2), we have

$$
\begin{aligned}
I & \leq c \sum_{A \in \mathcal{W}: r(A) \leq \frac{1}{10 r}} \int_{4 A \times 4 A}(u(x)-u(y))^{2} \frac{r^{-\alpha}}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y| \leq 1 / r\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
& \leq c \int_{B(0,1) \times B(0,1)}(u(x)-u(y))^{2} \frac{r^{-\alpha}}{|x-y|^{d+\alpha}} \mathbf{1}_{\{|x-y|<1 / r\}}(\Phi(x) \wedge \Phi(y)) d x d y
\end{aligned}
$$

On the other hands, if $A \in \mathcal{W}$ and $r(A)>\frac{1}{10 r}$, then for every pair of points $x, y$ in $4 A$, we have $|x-y| \leq 8 r(A)<\frac{1}{10^{2}}$ and

$$
\operatorname{dist}(x, \partial B(0,1)) \geq \rho(A)-4 r(A)>10^{2} r(A) \geq \frac{10}{r}
$$

Therefore, using (5.2) we have

$$
\begin{aligned}
I I & \leq c \sum_{A \in \mathcal{W}: r(A)>\frac{1}{10 r}} \int_{4 A \times 4 A} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{10^{2}}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
& \leq c \int_{B\left(0,1-\frac{10}{r}\right) \times B\left(0,1-\frac{10}{r}\right)} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y|<\frac{1}{\left.10^{2}\right\}}\right.}(\Phi(x) \wedge \Phi(y)) d x d y
\end{aligned}
$$

For our purpose, we need to construct another covering; For each $r>10^{2}$, we let $\mathcal{V}=\mathcal{V}_{r}:=$ $\left\{B^{1}, \cdots, B^{k(r)}\right\}$ be a maximum sequence of disjoint balls with radius $\frac{1}{400 r}$ that we can put inside $B\left(0,1-\frac{10}{r}\right)$. Note that

$$
B\left(0,1-\frac{10}{r}\right) \subset \bigcup_{B \in \mathcal{V}} 2 B \subset \bigcup_{B \in \mathcal{V}} 10^{2} B \subset B\left(0,1-\frac{9}{r}\right)
$$

For every $y \in B(0,1)$, since $\bigcup_{B \in \mathcal{V}: y \in 2 B} B \subset B\left(y, \frac{3}{400 r}\right)$,

$$
\#\{B \in \mathcal{V}: y \in 2 B\} \cdot \mu_{d}\left(B\left(0, \frac{1}{400 r}\right)\right) \leq \mu_{d}\left(B\left(y, \frac{3}{400 r}\right)\right)
$$

Therefore we have

$$
\begin{equation*}
\sup _{y \in B(0,1)} \#\{B \in \mathcal{V}: y \in 2 B\} \leq 3^{d} \tag{5.6}
\end{equation*}
$$

Recall that $\rho(B)$ denotes the Euclidean distance between the ball $B$ and $B(0,1)^{c}$. For balls $A$ and $B$ in $\mathcal{V}$ with $\operatorname{dist}(A, B)>\frac{1}{40 r}$ and $\rho(B) \geq \rho(A)$, we construct the path $\gamma_{A, B}$ starting from $A$ in the following way. Let $x_{A}$ be the center of $A$ and $x_{B}$ be the center of $B$. If $\left|x_{B}\right| \geq 1 /(400 r)$, then let $y_{B}:=\frac{\left|x_{A}\right|}{\left|x_{B}\right|} x_{B}$ so that $x_{B}$ is in the straight line segment from $y_{B}$ to 0 . Let $\gamma_{A, B}^{2}$ be the straight line segment from $y_{B}$ to $x_{B}$. We also let $\gamma_{A, B}^{1}$ be the shortest path from $x_{A}$ to $y_{B}$ with $\gamma_{A, B}^{1} \subset \partial B\left(0,\left|x_{A}\right|\right)$. In this case, $\gamma_{A, B}$ is the union of $\gamma_{A, B}^{1}$ and $\gamma_{A, B}^{2}$ starting from $x_{A}$ and ending at $x_{B}$ via $y_{B}$. If $\left|x_{B}\right|<1 /(400 r)$, let $\gamma_{A, B}$ be simply a straight line segment between 0 and $x_{A}$.

For $A, B \in \mathcal{V}$ with $\rho(B) \geq \rho(A)$, let

$$
\overline{\mathcal{V}}(A, B):=\left\{C \in \mathcal{V}: 2 C \cap \gamma_{A, B} \neq \emptyset\right\}
$$

and define the chain $\mathcal{V}(A, B):=\left(C_{0}, C_{1}, \cdots, C_{l(A, B)-1}\right)$ with $C_{0}=A$ and $C_{l(A, B)-1}=B$ similar to the chain in the Whitney-type coverings; Starting from the center of $A$, let $y_{0}$ be the first point along $\gamma_{A, B}$ which does not belong to $2 C_{0}$. Define $C_{1}$ to be one of balls in $\overline{\mathcal{V}}(A, B)$ such that $y_{0} \in 2 C_{1}$. Inductively, having $C_{0}, C_{1}, \cdots, C_{k}$ constructed, let $y_{k}$ be the first point along $\gamma_{A, B}$ which does not belong to $\cup_{j=0}^{k} 2 C_{j}$. Define $C_{k+1}$ to be one of balls in $\overline{\mathcal{V}}(A, B)$ such that $y_{k} \in 2 C_{k+1}$. When the last chosen is not $B$, we add $B$ as the last ball in $\mathcal{V}(A, B)$.

In the sequel, for every path $\gamma$ in $\mathbb{R}^{d}$ we denote by $|\gamma|$ the length of $\gamma$.
Lemma 5.8 There exists a positive constant $c=c(d)$ such that for every $r>10^{2}$ and every $A, B \in \mathcal{V}$ with $\rho(B) \geq \rho(A),\left|\gamma_{A, B}\right|>\frac{1}{4 r}$ and $\operatorname{dist}(A, B) \leq \frac{1}{50}$,

$$
\begin{equation*}
|x-y| \geq \frac{c}{r} \# \overline{\mathcal{V}}(A, B) \geq \frac{c}{r} \# \mathcal{V}(A, B) \geq\left|\gamma_{A, B}\right|, \quad \text { for every }(x, y) \in 2 A \times 2 B \tag{5.7}
\end{equation*}
$$

In particular,

$$
\begin{equation*}
\# \mathcal{V}(A, B) \leq \# \overline{\mathcal{V}}(A, B) \leq c r \tag{5.8}
\end{equation*}
$$

Proof. It is easy to see that the length of $\gamma_{A, B}$ is less than or equal to $4|x-y|$ for every $(x, y) \in A \times B$. Thus by using the fact that balls $C^{\prime}$ 's in $\overline{\mathcal{V}}(A, B)$ are disjoint and that $\cup_{C \in \overline{\mathcal{V}}(A, B)} C$ is within the $\frac{1}{100 r}$-neighborhood of $\gamma_{A B}$, we have

$$
\# \overline{\mathcal{V}}(A, B) \cdot\left(\frac{1}{400 r}\right)^{d}=c \sum_{C \in \overline{\mathcal{V}}(A, B)} \mu_{d}(C) \leq c|x-y| r^{1-d}
$$

and so $\# \overline{\mathcal{V}}(A, B) \leq c r|x-y|$.
On the other hand, since $2 C^{\prime}$ s in $\mathcal{V}(A, B)$ covers $\gamma_{A, B}$, it is easy to see that

$$
E:=\left\{x \in B(0,1): \operatorname{dist}\left(x, \gamma_{A, B}\right)<\frac{1}{400 r}\right\} \subset \bigcup_{C \in \mathcal{V}(A, B)} 3 C
$$

and that

$$
\mu_{d}(E) \geq c\left|\gamma_{A, B}\right|\left(\frac{1}{r}\right)^{d-1}
$$

Thus

$$
c\left|\gamma_{A, B}\right| r^{1-d} \leq \mu_{d}(E) \leq \sum_{C \in \overline{\mathcal{V}}(A, B)} \mu_{d}(3 C)=\# \mathcal{V}(A, B) \cdot\left(\frac{3}{400 r}\right)^{d}
$$

and so $\left|\gamma_{A, B}\right| \leq \frac{c}{r} \# \mathcal{V}(A, B)$. The lemma is proved.
The proof of the next lemma is similar to the one of Lemma 5.3. So we skip its proof.

Lemma 5.9 Let $A, B \in \mathcal{V}$ with $\rho(B) \geq \rho(A)$. There exists a positive constant $c=c(d)$ such that for every $C_{i}, C_{i+1} \in \mathcal{V}(A, B)$ and for every $u \in L^{1}(B(0,1), \Phi d x)$,

$$
\left|u_{2 C_{i}}-u_{2 C_{i+1}}\right|^{2} \leq \sum_{j=0}^{1} \frac{c}{\left(\mu_{d}\left(2 C_{i+j}\right)\right)^{2}} \int_{2 C_{i+j}} \int_{2 C_{i+j}}(u(x)-u(y))^{2} d x d y
$$

Lemma 5.10 There exist positive constants $c=c(d, \sigma)$ and $c_{1}=c_{1}(d)$ such that for every $r>10^{2}$ and every $A, B \in \mathcal{V}$ with $\rho(B) \geq \rho(A)$ and $\left|\gamma_{A, B}\right| \geq \frac{1}{4 r}$,

$$
\begin{aligned}
& \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y|<\frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & c c_{1}^{\alpha}(\# \mathcal{V}(A, B))^{1-d-\alpha} \sum_{C \in \mathcal{V}(A, B)} \int_{2 C} \int_{2 C} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}}(\Phi(x) \wedge \Phi(y)) d x d y
\end{aligned}
$$

Proof. Let $l:=\# \mathcal{V}(A, B) \geq 2$. For every $y \in A$ and $x \in B$,

$$
\begin{aligned}
& (u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) \\
\leq & (l+2)(\Phi(x) \wedge \Phi(y))\left(\left|u(x)-u_{2 A}\right|^{2}+\left|u(x)-u_{2 B}\right|^{2}+\sum_{i=0}^{l-1}\left|u_{2 C_{i}}-u_{2 C_{i+1}}\right|^{2}\right) \\
\leq & 2 l\left((\Phi(x) \wedge \Phi(y))\left|u(y)-u_{2 A}\right|^{2}+(\Phi(x) \wedge \Phi(y))\left|u(x)-u_{2 B}\right|^{2}\right. \\
& \left.+\sum_{i=0}^{l-1}(\Phi(x) \wedge \Phi(y))\left|u_{2 C_{i}}-u_{2 C_{i+1}}\right|^{2}\right) .
\end{aligned}
$$

Note that from the construction of the chain $\mathcal{V}(A, B)$, it is easy to see that there exists a constant $c$ independent of $r$ such that for every $A, B \in \mathcal{V}$ and $C \in \mathcal{V}(A, B)$,

$$
\int_{2 A} \int_{2 B}(\Phi(x) \wedge \Phi(y)) d x d y \leq c \int_{2 C_{i}} \int_{2 C_{i+1}}(\Phi(x) \wedge \Phi(y)) d x d y
$$

Obviously

$$
\int_{2 A} \int_{2 B}\left|u(x)-u_{2 B}\right|^{2}(\Phi(x) \wedge \Phi(y)) d x d y \leq \mu_{d}(2 B) \int_{2 B}\left|u(x)-u_{2 B}\right|^{2} \Phi(x) d x
$$

and

$$
\int_{2 A} \int_{2 B}\left|u(y)-u_{2 A}\right|^{2}(\Phi(x) \wedge \Phi(y)) d x d y \leq \mu_{d}(2 A) \int_{2 A}\left|u(y)-u_{2 A}\right|^{2} \Phi(y) d y
$$

Thus we have, for every $y \in A$ and $x \in B$,

$$
\begin{aligned}
& \int_{2 A} \int_{2 B}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & 2 l\left(\int_{2 A} \int_{2 B}(\Phi(x) \wedge \Phi(y))\left|u(y)-u_{2 A}\right|^{2} d x d y+\int_{2 A} \int_{2 B}(\Phi(x) \wedge \Phi(y))\left|u(x)-u_{2 B}\right|^{2} d x d y\right. \\
& \left.+\sum_{i=0}^{l-1} \int_{2 A} \int_{2 B}(\Phi(x) \wedge \Phi(y))\left|u_{2 C_{i}}-u_{2 C_{i+1}}\right|^{2} d x d y\right) \\
\leq \quad & c l\left(\mu_{d}(2 A)\left|\int_{2 A}\right| u(y)-\left.u_{2 A}\right|^{2} \Phi(y) d y+\mu_{d}(2 B) \int_{2 B}\left|u(x)-u_{2 B}\right|^{2} \Phi(x) d x\right. \\
& \left.+\sum_{i=0}^{l-1}\left|u_{2 C_{i}}-u_{2 C_{i+1}}\right|^{2} \int_{2 C_{i}} \int_{2 C_{i+1}}(\Phi(x) \wedge \Phi(y)) d x d y\right)
\end{aligned}
$$

We apply Lemma 5.2 to the first two integrals in the above and apply Lemma 5.9 to the integrals in the summation above. Then using the fact that the values of $\Phi$ are universally comparable on each $A, B, C_{i}$, we get that

$$
\begin{equation*}
\int_{2 A} \int_{2 B}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y \leq c l \sum_{C \in \mathcal{V}(A, B)} \int_{2 C} \int_{2 C}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) d x d y \tag{5.9}
\end{equation*}
$$

Note that, using (5.7), we have that for $x \in B$ and $y \in A$ with $|x-y|<\frac{1}{100}$

$$
\begin{equation*}
\frac{1}{100} \geq|x-y| \geq c \frac{l}{r} \geq c l|z-w|, \quad \forall z, w \in C \in \mathcal{V}(A, B) \tag{5.10}
\end{equation*}
$$

Therefore, from (5.9)-(5.10), we conclude that

$$
\begin{aligned}
& \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}}(\Phi(x) \wedge \Phi(y)) \mathbf{1}_{\left\{|x-y|<\frac{1}{100}\right\}} d x d y \\
\leq & \frac{1}{10^{2 \alpha}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}}(\Phi(x) \wedge \Phi(y)) \mathbf{1}_{\left\{|x-y|<\frac{1}{100}\right\}} d x d y \\
= & \frac{1}{10^{2 \alpha}}\left(c \frac{r}{l}\right)^{d+\alpha} \int_{2 A} \int_{2 B}(u(x)-u(y))^{2}(\Phi(x) \wedge \Phi(y)) \mathbf{1}_{\left\{|x-y|<\frac{1}{100}\right\}} d x d y \\
\leq & c c_{1}^{\alpha} l^{1-d-\alpha} \sum_{C \in \mathcal{V}(A, B)} \int_{2 C} \int_{2 C} \frac{(u(z)-u(w)))^{2}}{|z-w|^{d+\alpha}}(\Phi(z) \wedge \Phi(w)) d z d w
\end{aligned}
$$

Recall that $[a]$ denote the largest integer which is no larger than $a$ and define for $C \in \mathcal{V}$

$$
C(\mathcal{V}):=\{(A, B): A, B \in \mathcal{V} \text { with } \rho(B) \geq \rho(A) \text { and } C \in \mathcal{V}(A, B)\}
$$

The following is a key lemma to count the number of chains containing each $C \in \mathcal{V}$.

Lemma 5.11 There exists a positive constant $c=c(d)$ such that for every $r>10^{2}, 30 \leq l \leq[16 r]$ and $C \in \mathcal{V}$,

$$
\begin{equation*}
\#\left\{(A, B) \in C(\mathcal{V}): \frac{100+l}{400 r}<\left|\gamma_{A, B}\right| \leq \frac{101+l}{400 r}\right\} \leq c l^{d} . \tag{5.11}
\end{equation*}
$$

Proof. Without loss of generality, we assume $d \geq 2$. (The case of $d=1$ is easier.) Fix $r>10^{2}$, $30 \leq l \leq[16 r]$ and $C \in \mathcal{V}$. We will order $(A, B) \in C(\mathcal{V})$ so that $\rho(B) \geq \rho(A)$. Let $x_{C}$ be the center of the ball $C$. If $\left|x_{C}\right| \leq 4 /(400 r)$, then $\left|x_{B}\right| \leq 6 /(400 r)$, so the number of possible choice for $B$ is less than $c 2^{d}$. Since $(100+l) /(400 r) \leq\left|\gamma_{A, B}\right| \leq(101+l) /(400 r)$, the number of possible choice for $A$ is $c^{d-1}$, so (5.11) holds in this case. We thus assume $\left|x_{C}\right|>4 /(400 r)$. Define $H_{x_{C}}:=B\left(0,\left|x_{C}\right|+2 /(400 r)\right) \backslash B\left(0,\left|x_{C}\right|-2 /(400 r)\right)$. Since $2 C \cap \gamma_{A, B} \neq \emptyset, H_{x_{C}} \cap \gamma_{A, B} \neq \emptyset$. Let $y_{B}^{\prime}$ be the first point along $\gamma_{A, B}$ (starting from $x_{B}$ ) which belongs to $H_{x_{C}} \cap \gamma_{A, B}$. Also, let $z_{A, B}$ be the first point along $\gamma_{A, B}$ (starting from $x_{B}$ ) which belongs to $2 C$, and let $\gamma_{B}$ be the sub-path of $\gamma_{A, B}$ starting from $z_{A, B}$ ending at $x_{B}$.

Let $m /(400 r) \leq\left|\gamma_{B}\right|<(m+1) /(400 r)$ where $0 \leq m \leq l+100$ and consider the following two cases:

$$
\text { Case (i) }\left|y_{B}^{\prime}-z_{A, B}\right| \leq \frac{5}{400 r}, \quad \text { Case (ii) }\left|y_{B}^{\prime}-z_{A, B}\right|>\frac{5}{400 r} \text {. }
$$

For Case (i), the number of possible choices for $y_{B}^{\prime}$ and $B$ is less than $c 2^{d}$ when $C$ is given and $m$ is fixed. Once $y_{B}^{\prime}$ is fixed, the number of possible choice for $A$ is $c(l-m+106)^{d-1}$, since the arclength between $z_{A, B}$ and $x_{A}$ along the curve $\gamma_{A, B}$ is at most $\frac{101+l-m}{400 r}$ and $\left|y_{B}^{\prime}-z_{A, B}\right| \leq 5 /(400 r)$. Summing over $m$, the number of possible choices for $A$ and $B$ is less than

$$
c^{\prime} \sum_{m=0}^{l+100}(l-m+106)^{d-1} \leq c^{\prime \prime} l^{d} .
$$

For Case (ii), let $i \leq m$ be such that $i /(400 r) \leq\left|z_{A, B}-y_{B}\right|<(i+1) /(400 r)$ where $y_{B}:=\frac{\left|x_{A}\right|}{\left|x_{B}\right|} x_{B}$. In this case, $\left|y_{B}-y_{B}^{\prime}\right| \leq 4 /(400 r)$ and $i \geq 1$. Since $y_{B} \in \partial B\left(0,\left|x_{A}\right|\right) \subset H_{x_{C}}$, given $C$, the number of possible choices for $y_{B}$ and $B$ is less than $c i^{d-2}$ when $m$ and $i$ are fixed. Observe that given $C$ and $B, y_{B}^{\prime}$ and $x_{B}$ are determined. Since $x_{A} \in \partial B\left(0,\left|x_{A}\right|\right) \subset H_{x_{C}}$, given $C$ and $B$, the number of possible choice for $x_{A}$ is less than $c((l-m+i+101) / i)^{d-2}$ when $m$ and $i$ are fixed. Summing over $m$ and $i$, the number of possible choices for $A$ and $B$ is less than

$$
c^{\prime} \sum_{m=1}^{l+100} \sum_{i=1}^{m} i^{d-2}\left(\frac{l-m+i+101}{i}\right)^{d-2}=c^{\prime} \sum_{m=1}^{l+100} \sum_{i=1}^{m}(l-m+i+101)^{d-2} \leq c^{\prime \prime} l^{d} .
$$

We thus obtain (5.11).
Lemma 5.12 There exist positive constants $c=c(d, \sigma)$ and $c_{1}=c_{1}(d)$ such that for every $r \geq 10^{2}$

$$
\begin{aligned}
& \sum_{\substack{A, B \in \mathcal{V} \\
\text { dist }(A, B)>\frac{1}{3 r}}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & c c_{1}^{\alpha} \int_{B(0,1) \times B(0,1)}(u(x)-u(y))^{2} \frac{r^{2-\alpha}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{r}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y .
\end{aligned}
$$

Proof. For $(x, y) \in 2 A \times 2 B$ with $|x-y| \leq \frac{1}{100}$, it is elementary to check that $\left|\gamma_{A, B}\right|<\frac{1}{25}$. Thus, by Lemma 5.10 , we have

$$
\begin{aligned}
& \sum_{\substack{A, B \in \mathcal{V} \\
\operatorname{dist}(A, B)>\frac{1}{3 r}}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & c c_{1}^{\alpha} \sum_{\substack{A, B \in \mathcal{V}: \rho(B) \geq \rho(A) \\
\frac{130}{400 r}<\left|\gamma_{A, B}\right|<\frac{1}{25}}}(\# \mathcal{V}(A, B))^{1-d-\alpha} \sum_{C \in \mathcal{V}(A, B)} \int_{2 C} \int_{2 C} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & c c_{1}^{\alpha} \sum_{C \in \mathcal{V}}\left(\sum_{l=30}^{[16 r]} \sum_{\substack{(A, B) \in C(\mathcal{V}) \\
\frac{100+l}{400 r<\left|\gamma_{A, B}\right| \leq \frac{100+l+1}{400 r}}}}(\# \mathcal{V}(A, B))^{1-d-\alpha}\right) \\
& \times \int_{2 C} \int_{2 C} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{r}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y .
\end{aligned}
$$

Applying (5.7), we see that

$$
\begin{aligned}
& \sum_{\substack{A, B \in \mathcal{V} \\
\operatorname{dist}(A, B)>\frac{1}{3 r}}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
& \leq c c_{1}^{\alpha} \sum_{C \in \mathcal{V}}\left(\sum_{l=30}^{[16 r]} l^{1-d-\alpha} \cdot \#\left\{(A, B) \in C(\mathcal{V}): \frac{100+l}{400 r}<\left|\gamma_{A, B}\right| \leq \frac{101+l}{400 r}\right\}\right) \\
& \times \int_{2 C} \int_{2 C} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{r}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y
\end{aligned}
$$

By Lemma 5.11,

$$
\sum_{l=30}^{[16 r]} l^{1-d-\alpha} \cdot \#\left\{(A, B) \in C(\mathcal{V}): \frac{100+l}{400 r}<\left|\gamma_{A, B}\right| \leq \frac{101+l}{400 r}\right\} \leq c \sum_{l=30}^{[16 r]} l^{1-\alpha} \leq c r^{2-\alpha}
$$

Thus we conclude that

$$
\begin{aligned}
& \sum_{\substack{A, B \in \mathcal{V} \\
\text { dist }(A, B)>\frac{1}{3 r}}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
& \leq c c_{1}^{\alpha} r^{2-\alpha} \sum_{C \in \mathcal{V}} \int_{2 C} \int_{2 C} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{r}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
& \leq c c_{1}^{\alpha} \int_{B(0,1) \times B(0,1)}(u(x)-u(y))^{2} \frac{r^{2-\alpha}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{r}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y .
\end{aligned}
$$

In the last inequality above, we have used (5.6).

Proof of Theorem 5.1: By Lemma 5.7, it is enough to show the following claim; there exist constants $c=c(d, \sigma)>0$ and $c_{1}(d)>0$ such that for every $r>10^{2}$ and $u \in L^{1}(B(0,1), \Phi d x)$

$$
\begin{align*}
& \int_{B\left(0,1-\frac{10}{r}\right) \times B\left(0,1-\frac{10}{r}\right)} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & c c_{1}^{\alpha} \int_{B(0,1) \times B(0,1)}(u(x)-u(y))^{2} \frac{r^{2-\alpha}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{r}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y . \tag{5.12}
\end{align*}
$$

Note that

$$
\begin{aligned}
& \int_{B\left(0,1-\frac{10}{r}\right) \times B\left(0,1-\frac{10}{r}\right)} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & \sum_{A, B \in \mathcal{V}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & \sum_{\substack{A, B \in \mathcal{V} \\
\operatorname{dist}(A, B) \leq \frac{1}{3 r}}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{r}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
& +\sum_{\substack{A, B \in \mathcal{V} \\
\operatorname{dist}(A, B)>\frac{1}{3 r}}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
\leq & c r^{2-\alpha} \int_{B(0,1) \times B(0,1)} \frac{(u(x)-u(y))^{2}}{|x-y|^{d+\alpha}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{r}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y \\
& +\sum_{\substack{A, B \in \mathcal{V} \\
\operatorname{dist}(A, B)>\frac{1}{3 r}}} \int_{2 A} \int_{2 B} \frac{(u(x)-u(y))^{2}}{|x-y|^{d}} \mathbf{1}_{\left\{|x-y| \leq \frac{1}{100}\right\}}(\Phi(x) \wedge \Phi(y)) d x d y .
\end{aligned}
$$

In the last inequality above, we have used (5.6) and the fact $r^{2-\alpha} \geq 1$. Thus (5.12) follows from Lemma 5.12.

Acknowledgment: We thank Zoran Vondraček for pointing out an error in a preliminary version of this paper.

## References

[BBCK] M.T. Barlow, R.F. Bass, Z.-Q. Chen and M. Kassmann. Non-local Dirichlet forms and symmetric jump processes. To appear in Trans. Amer. Math. Soc..
[BBK] M.T. Barlow, R.F. Bass and T. Kumagai. Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps. To appear in Math. Z..
[BGK] M.T. Barlow, A. Grigor'yan and T. Kumagai. Heat kernel upper bounds for jump processes. To appear in J. Reine Angew. Math..
[BL] R. F. Bass and D. A. Levin. Transition probabilities for symmetric jump processes. Trans. Amer. Math. Soc. 354 (2002), 2933-2953.
[B] J. Bertoin. Lévy Processes. Cambridge University Press, 1996.
[BG] R.M. Blumenthal and R.K. Getoor. Markov processes and potential theory. Academic Press, Reading, MA, 1968.
[CSS] L. A. Caffarelli, S. Salsa and Luis Silvestre. Regularity estimates for the solution and the free boundary to the obstacle problem for the fractional Laplacian. Invent. Math. 171(1) (2008) 425-461.
[CKS] E.A. Carlen, S. Kusuoka and D.W. Stroock. Upper bounds for symmetric Markov transition functions. Ann. Inst. H. Poincaré-Probab. Statist. 23 (1987), 245-287.
[C] Z.-Q. Chen. Gaugeability and conditional gaugeability. Trans. Ameri. Math. Soc. 354 (2002), 4639-4679.
[CKK] Z.-Q. Chen, P. Kim and T. Kumagai. Notes on heat kernel estimates and parabolic Harnack inequality for jump processes. In preparation.
[CK1] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for stable-like processes on $d$-sets. Stochastic Process Appl. 108 (2003), 27-62.
[CK2] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for jump processes of mixed types on metric measure spaces. Probab. Theory Relat. Fields, 140 (2008), 277-317.
[CR] Z.-Q. Chen and S. Rohde. SLE driven by symmetric stable processes. Preprint 2007.
[FS] E. B. Fabes and D. W. Stroock. A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. Arch. Rational Mech. Anal., 96(4) (1986), 327-338.
[FOT] M. Fukushima, Y. Oshima and M. Takeda. Dirichlet forms and symmetric Markov processes. de Gruyter, Berlin, 1994.
[GT] D. Gilbarg and N. S. Trudinger. Elliptic Partial Differential Equations of Second Order. Second Edition. Springer-Verlag, 1983.
[GR] T. Grzywny and M. Ryznar. Estimates of Green function for some perturbations of fractional Laplacian. To appear in Illinois J. Math..
[HK] J. Hu and T. Kumagai. Nash-type inequalities and heat kernels for non-local Dirichlet forms. Kyushu J. Math. 60 (2006), 245-265.
[HPR] S.R. Hurst, E. Platen and S.T. Rachev. Option pricing for a logstable asset price model. Math. Comput. Modelling 29 (1999), 105-119.
[IW] N. Ikeda and S. Watanabe. Stochastic Differential Equations and Diffusion Processes. Second Edition. North-Holland Publishing Co., 1989.
[JW] A. Janicki and A. Weron. Simulation and Chaotic Behavior of $\alpha$-Stable Processes. Dekker, 1994.
[J] D. Jerison. The Poincaré inequality for vector fields satisfying Hörmander's condition. Duke Math. J. 53(2) (1986), 503-523.
[Ka] M. Kassmann. A priori estimates for integro-differential operators with measurable kernels. To appear in Calc. Var. Partial Differential Equations.
[KS1] P. Kim and R. Song. Potential theory of truncated stable processes. Math. Z. 256(1) (2007), 139-173.
[KS2] P. Kim and R. Song. Boundary behavior of harmonic functions for truncated stable processes. J. Theoret. Probab. 21 (2008), 287-321.
[KSZ] J. Klafter, M. F. Shlesinger and G. Zumofen. Beyond Brownian motion. Physics Today, 49 (1996), 33-39.
[K] V. Kolokoltsov. Symmetric stable laws and stable-like jump-diffusions. Proc. London Math. Soc. 80 (2000), 725-768.
[M] A. Matacz. Financial modeling and option theory with the truncated Lévy process. Int. J. Theor. Appl. Finance 3(1) (2000), 143-160.
[SC] L. Saloff-Coste. Aspects of Sobolev-type inequalities. Cambridge University Press, Cambridge, 2002.
[SS] L. Saloff-Coste and D. W. Stroock. Opérateurs uniformément sous-elliptiques sur les groupes de Lie. J. Funct. Anal. 98(1) (1991), 97-121.
[S] L. Silvestre. Hölder estimates for solutions of integro differential equations like the fractional Laplace. Indiana Univ. Math. J. 55 (2006), 1155-1174.
[SV] R. Song and Z. Vondracek. Parabolic Harnack inequality for the mixture of Brownian motion and stable process. Tohoku Math. J. 59(1) (2007), 1-19.
[TT] M. Takeda and K. Tsuchida. Criticality of generalized Schrödinger operators and differentiability of spectral functions. Adv. Studies Pure Math. 41 (2004), 333-350.
[TT2] M. Takeda and K. Tsuchida. Differentiability of spectral functions for symmetric $\alpha$-stable processes. Trans. Amer. Math. Soc. 359 (2007), 4031-4054.
[Ts] K. Tsuchida. Differentiability of spectral functions for relativistic $\alpha$-stable processes with application to large deviations. Potential Anal. 28(1) (2008), 17-33.

## Zhen-Qing Chen

Department of Mathematics, University of Washington, Seattle, WA 98195, USA
E-mail: zchen@math.washington.edu

## Panki Kim

Department of Mathematics, Seoul National University, Seoul 151-742, South Korea
E-mail: pkim@snu.ac.kr

## Takashi Kumagai

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


[^0]:    *Research partially supported by NSF Grant DMS-06000206.
    ${ }^{\dagger}$ Research partially supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2007-331-C00037).
    ${ }^{\ddagger}$ Research partially supported by the Grant-in-Aid for Scientific Research (B) 18340027.

