Stability of heat kernel estimates and parabolic Harnack inequalities for general symmetric pure jump processes

Zhen-Qing Chen, Takashi Kumagai, and Jian Wang

Abstract. In this paper, we survey recent work on heat kernel estimates for general symmetric pure jump processes on metric measure space (M, ρ, μ) generated by the following type of non-local Dirichlet forms

$$\mathcal{E}(f, g) = \int_{M \times M} (f(x) - f(y))(g(x) - g(y))\, J(dx, dy),$$

where $J(dx, dy)$ is a symmetric Radon measure on $M \times M \setminus \mathrm{diag}$ that may have different growth behaviors for small and large jumps. Under general volume doubling condition on (M, ρ, μ) and some mild quantitative assumptions on $J(dx, dy)$ that are allowed to have light tails of polynomial decay at infinity, we present stability results for two-sided heat kernel estimates and heat kernel upper bounds as well as the corresponding parabolic Harnack inequalities. The results extend considerably those for mixture of symmetric stable-like jump processes in metric measure spaces, and more interestingly, they have connections to these for symmetric diffusions with jumps.

1. Preliminaries

In this section, we first recall the history on heat kernel estimates for mixture of symmetric stable-like processes on metric measure space, and then present an interesting example, which is not of mixture of stable-like type and exhibits some new phenomena. In particular, this example will lead us to consider general symmetric pure jump Markov processes whose jumping kernel can have light tails at infinity, which is (a spacial but important case of) the subject of this paper.

1.1. Mixture of symmetric stable-like processes. Two-sided heat kernel estimates and parabolic Harnack inequalities have a long history in the theory of

2010 *Mathematics Subject Classification.* Primary 60J35, 35K08, 60J75; Secondary 31C25, 60J25, 60J45.

Key words and phrases. Symmetric pure jump process, non-local Dirichlet form, heat kernel estimate, parabolic Harnack inequality, jumping kernel.

The first author was supported in part by Simons Foundation Grant 520542 and a Victor Klee Faculty Fellowship at UW.

The second author was supported in part by JSPS KAKENHI Grant Number JP17H01093 and by the Alexander von Humboldt Foundation.

The third author was supported in part by the National Natural Science Foundation of China (No. 11831014), the Program for Probability and Statistics: Theory and Application (No. IRTL1704) and the Program for Innovative Research Team in Science and Technology in Fujian Province University (IRTSTFJ).

partial differential equations and diffusion processes (which are strong Markov processes with continuous sample paths). There are many beautiful results including the De Giorgi-Nash-Moser theory and Aronson's Gaussian estimates in these areas. Studies of transition density functions for Markov processes with discontinuous sample paths, or equivalently, heat kernels for non-local operators, are relatively recent.

We start with a (rotationally) symmetric α-stable process $Z = \{Z_t, t \ge 0; \mathbb{P}_x, x \in \mathbb{R}^d\}$ on $\mathbb{R}^d$ with $\alpha \in (0,2)$, which is a Lévy process such that

$$\mathbb{E}_x e^{i\langle Z_t - x, \xi\rangle} = e^{-t|\xi|^\alpha} \quad \text{for every } x,\, \xi \in \mathbb{R}^d.$$

The infinitesimal generator of a symmetric α-stable process Z on $\mathbb{R}^d$ is the fractional Laplacian $\Delta^{\alpha/2} := -(-\Delta)^{\alpha/2}$, which is a prototype of non-local operators. The fractional Laplacian can be written in the form

$$\Delta^{\alpha/2} u(x) = c \lim_{\varepsilon \to 0} \int_{\{|x-y|>\varepsilon\}} \frac{u(y) - u(x)}{|y-x|^{d+\alpha}}\, dy$$

for some constant $c = c(d,\alpha) > 0$. It is well-known (see e.g. [**BG**]) that the transition density function $p(t,x,y)$ of the symmetric α-stable process has the following two sided estimates:

$$p(t,x,y) \simeq t^{-d/\alpha} \wedge \frac{t}{|x-y|^{d+\alpha}}, \quad x, y \in \mathbb{R}^d, t > 0.$$

In this paper, we write $f \simeq g$, if there exist constants $c_1, c_2 > 0$ such that $c_1 g(x) \le f(x) \le c_2 g(x)$ for some range of x. For $a, b \in \mathbb{R}$, set $a \vee b := \max\{a,b\}$ and $a \wedge b := \min\{a,b\}$. Note that, in contrast to the standard Gaussian tail for Brownian motions, heat kernel of symmetric α-stable processes has polynomial decay when the distance $|x-y| \to \infty$. This is called the heavy tail phenomenon for stable processes. Due to this property, many systems in physics and economies can and have been modeled by non-Gaussian stable processes.

Let (M, ρ, μ) be an Ahlfors d-regular set; that is, $\mu(B_\rho(x,r)) \simeq r^d$ for all $r > 0$ and $x \in M$, where $B_\rho(x,r) := \{y \in M : \rho(x,y) < r\}$. Let $J(x,y)$ be a positive symmetric function on $M \times M$ such that

$$J(x,y) \simeq \rho(x,y)^{-(d+\alpha)}, \quad x, y \in M \tag{1.1}$$

for some $0 < \alpha < 2$. Consider a non-local regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(M;\mu)$ given by

$$\mathcal{E}(f,g) = \int_{M \times M \setminus \mathrm{diag}} (f(x) - f(y)(g(x) - g(y))J(x,y)\, \mu(dx)\, \mu(dy), \quad f, g \in \mathcal{F},$$

where diag denotes the diagonal set $\{(x,x) : x \in M\}$. The reader is referred to [**CF, FOT**] for terminology and theory of symmetric Dirichlet forms and their associated symmetric Markov processes. It was established in Chen and Kumagai [**CK**] that the symmetric strong Markov process X associated with $(\mathcal{E}, \mathcal{F})$ on $L^2(M;\mu)$ has infinite lifetime, and has a jointly Hölder continuous transition density function $p(t,x,y)$ with respect to the measure μ, which enjoys the following two-sided estimates

$$p(t,x,y) \simeq t^{-d/\alpha} \wedge \frac{t}{\rho(x,y)^{d+\alpha}} \qquad \text{for any } (t,x,y) \in (0,\infty) \times M \times M. \tag{1.2}$$

We call the above Hunt process X a symmetric α-stable-like process on M. Note that when $M = \mathbb{R}^d$ and $J(x,y) = c(x,y)|x-y|^{-(d+\alpha)}$ with $0 < c_1 \le c(x,y) \le$

$c_2 < \infty$ for some constants $\alpha \in (0,2)$ and $c_1, c_2 > 0$, X is a symmetric α-stable-like process on $\mathbb{R}^d$, and the associated infinitesimal generator can be viewed as the analog to a divergence form operator for fractional Laplacians. Since $J(x,y)$ is the weak limit of $p(t,x,y)/t$ as $t \to 0$, heat kernel estimate (1.2) implies (1.1). Therefore, the results from [**CK**] give a stable characterization for α-stable-like heat kernel estimates when $\alpha \in (0,2)$ and the metric measure space (M,ρ,μ) is a d-set for some constant $d > 0$. This result has later been extended in [**CK2**] to mixture of stable-like processes on more general metric measure spaces, under some growth condition on the weighted function ϕ such as

$$\int_0^r \frac{s}{\phi(s)}\, ds \le c\frac{r^2}{\phi(r)} \quad \text{for } r > 0. \tag{1.3}$$

For α-stable-like processes where $\phi(r) = r^\alpha$, condition (1.3) corresponds exactly to $0 < \alpha < 2$. Some of the key methods used in [**CK**] were inspired by a previous work [**BL**] on symmetric random walks with stable-like long range jumps on integer lattice $\mathbb{Z}^d$. See [**BBK2, MS1**] for related works on long range random walks on graphs.

The notion of d-set arises in the theory of function spaces and in fractal geometry. Geometrically, self-similar sets are typical examples of d-sets. There are many self-similar fractals on which there exist sub-diffusive processes with walk dimension $d_w > 2$ (that is, diffusion processes with scaling relation *time* $\approx$ *space*^{d_w}). For example, the walk dimension of Brownian motions on the Sierpinski gasket in $\mathbb{R}^n$ $(n \ge 2)$ is $\log(n+3)/\log 2$; see [**B**]. A direct calculation shows that the β-subordination of the sub-diffusive processes on these fractals are jump processes whose Dirichlet forms $(\mathcal{E},\mathcal{F})$ are of the form given above with $\alpha = \beta d_w$ in (1.1), and their transition density functions have two-sided estimates (1.2). Note that as $\beta \in (0,1)$, $\alpha \in (0, d_w)$ so α can be larger than 2. When $\alpha > 2$, the approach in [**CK**] ceases to work as it is hopeless to construct good cut-off functions a priori in this case. A long standing open problem in the field was whether the estimate (1.2) holds for generic symmetric jump processes having jumping kernel of the form (1.1) for any $\alpha \in (0, d_w)$. A related open question was to characterize the heat kernel estimate (1.2) by conditions that are stable under "rough isometries".

These open problems have recently been solved affirmatively in [**CKW1**]. Actually, in [**CKW1**] we obtained stability of two-sided heat kernel estimates for mixture of symmetric stable-like jump processes on general metric measure spaces that satisfy the volume doubling condition and the reverse volume doubling condition.

In details, let (M,ρ,μ) be a locally compact separable metric space, and μ a positive Radon measure on M with full support. In what follows, we will refer to such a triple (M,ρ,μ) as a *metric measure space*. We assume $\mu(M) = \infty$. Denote the open ball centered at x with radius r by $B(x,r)$ and $\mu(B(x,r))$ by $V(x,r)$.

Definition 1.1. (i) We say that (M,ρ,μ) satisfies the *volume doubling property* (VD), if there exists a constant $C_\mu \ge 1$ such that

$$V(x,2r) \le C_\mu V(x,r) \quad \text{for all } x \in M \text{ and } r > 0.$$

(ii) We say that (M,ρ,μ) satisfies the *reverse volume doubling property* (RVD), if there exist constants $l_\mu, c_\mu > 1$ such that

$$V(x,l_\mu r) \ge c_\mu V(x,r) \quad \text{for all } x \in M \text{ and } r > 0.$$

We remark that under RVD, $\mu(M)=\infty$ if and only if M has infinite diameter. We also note that when M is connected and unbounded, VD implies RVD.

We consider the following regular non-local *Dirichlet form* $(\mathcal{E},\mathcal{F})$ on $L^2(M;\mu)$:

$$\mathcal{E}(f,g)=\int_{M\times M\setminus \mathrm{diag}} (f(x)-f(y)(g(x)-g(y))\, J(dx,dy), \quad f,g\in \mathcal{F}, \tag{1.4}$$

where $J(dx,dy)$ is a symmetric Radon measure on $M\times M\setminus \mathrm{diag}$. Associated with the regular Dirichlet form $(\mathcal{E},\mathcal{F})$ on $L^2(M;\mu)$ is a μ-symmetric *Hunt process* $X=\{X_t, t\ge 0; \mathbb{P}^x, x\in M\setminus\mathcal{N}\}$. Here $\mathcal{N}\subset M$ is a properly exceptional set for $(\mathcal{E},\mathcal{F})$ in the sense that $\mathcal{N}$ is nearly Borel, $\mu(\mathcal{N})=0$ and $M_\partial\setminus\mathcal{N}$ is X-invariant. This Hunt process is unique up to a properly exceptional set; see [**FOT**, Theorem 4.2.8]. Since $(\mathcal{E},\mathcal{F})$ only has non-local part, the associated Hunt process X is of the pure jump type. We fix X and $\mathcal{N}$, and write $M_0=M\setminus\mathcal{N}$. Let $\{P_t\}_{t\ge 0}$ be the *transition semigroup* of the Markov process X (or of the Dirichlet form $(\mathcal{E},\mathcal{F})$ on $L^2(M;\mu)$). The *transition density function* for the Markov process X is a measurable function $p(t,x,y): M_0\times M_0\to(0,\infty)$ for every $t>0$, such that

$$\mathbb{E}^x[f(X_t)]=P_tf(x)=\int p(t,x,y)f(y)\,\mu(dy) \quad \text{for all } x\in M_0, f\in L^\infty(M;\mu), \tag{1.5}$$

$$p(t,x,y)=p(t,y,x) \quad \text{for all } t>0,\ x,y\in M_0,$$

$$p(s+t,x,z)=\int p(s,x,y)p(t,y,z)\,\mu(dy) \quad \text{for all } s,t>0,\ x,z\in M_0.$$

In literature, $p(t,x,y)$ is also called the *heat kernel* of the process X or of the Dirichlet form $(\mathcal{E},\mathcal{F})$ on $L^2(M;\mu)$.

Let $\mathbb{R}_+:=[0,\infty)$, and $\phi:\mathbb{R}_+\to\mathbb{R}_+$ be a strictly increasing continuous function with $\phi(0)=0$ and $\phi(1)=1$ so that that there exist constants $c_1,c_2>0$ and $\beta_2\ge\beta_1>0$ such that

$$c_1\Big(\frac{R}{r}\Big)^{\beta_1}\le \frac{\phi(R)}{\phi(r)} \ \le\ c_2\Big(\frac{R}{r}\Big)^{\beta_2} \quad \text{for all } 0<r\le R.$$

In [**CKW1**, Theorem 1.13], the following are shown to be equivalent under the VD and RVD condition:

(i) there exists a heat kernel $p(t,x,y)$ associated with $(\mathcal{E},\mathcal{F})$, which has the following estimates for all $t>0$ and all $x,y\in M$,

$$p(t,x,y)\simeq \frac{1}{V(x,\phi^{-1}(t))}\wedge\frac{t}{V(x,\rho(x,y))\phi(\rho(x,y))}. \tag{1.6}$$

Here and in what follows, $\phi^{-1}(t)$ is the inverse function of the strictly increasing function $t\mapsto\phi(t)$.

(ii) J_ϕ and a cut-off Sobolev inequality $\mathrm{CSJ}(\phi)$ hold, where J_ϕ means that there exists a non-negative symmetric function $J(x,y)$ so that for $\mu\times\mu$-almost all $x,y\in M$,

$$J(dx,dy)=J(x,y)\,\mu(dx)\,\mu(dy), \tag{1.7}$$

and

$$J(x,y)\simeq\frac{1}{V(x,\rho(x,y))\phi(\rho(x,y))}. \tag{1.8}$$

See Definition 2.3 below for the definition of $\mathrm{CSJ}(\phi)$.

(iii) J_ϕ and E_ϕ hold, where E_ϕ means that

$$\mathbb{E}^x \tau_{B(x,r)} \simeq \phi(r) \quad \text{for all } x \in M_0 \text{ and } r > 0, \tag{1.9}$$

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

We emphasize that in the above result from [**CKW1**, Theorem 1.13], the underlying metric measure space (M, ρ, μ) is only assumed to satisfy the general VD and RVD. Neither uniform VD nor uniform RVD property is assumed. We do not assume M to be connected nor (M, ρ) to be geodesic. See related works for symmetric stable-like processes with metric measure spaces satisfying Ahlfors d-set condition [**GHH**], and for random walks with α-stable-like long range jumps on connected locally finite infinite graphs satisfying global Ahlfors d-set condition [**MS2**].

We point out that, for symmetric jump processes associated with non-local Dirichlet forms $(\mathcal{E}, \mathcal{F})$ above, parabolic Harnack inequalities are strictly weaker than the two-sided heat kernel estimates. It is established in [**CKW2**, Corollary 1.21] that the parabolic Harnack inequalities together with a suitable lower bound for the jumping kernel is equivalent to the two-sided heat kernel estimates. (See [**BBK2**, Theorem 1.4] for the corresponding result for random walks with α-stable-like long range jumps with $0 < \alpha < 2$ on graphs satisfying the global Ahlfors d-set condition.) This is in contrast to symmetric diffusion processes, where parabolic Harnack inequalities are equivalent to two-sided heat kernel estimates. The reader is referred to [**CKW2**] for the stability results of parabolic Harnack inequalities for symmetric pure jump Dirichlet forms, and to [**CKW3**] for various characterizations of elliptic Harnack inequalities for symmetric pure jump processes.

1.2. Motivating example: beyond mixture of stable-like processes.

Consider the following example.

Example 1.2. Let (M, ρ, μ) be a metric measure space such that the volume doubling (VD) condition holds. Suppose $X := \{X_t, t \geq 0; \mathbb{P}^x, x \in M\}$ is a conservative symmetric diffusion process on M that has a transition density function $q(t, x, y)$ with respect to μ such that

$$q(t, x, y) \asymp \frac{1}{V(x, t^{1/\beta})} \exp\left(-c\left(\frac{\rho(x, y)^\beta}{t}\right)^{1/(1-\beta)}\right), \quad t > 0, x, y \in M \tag{1.10}$$

for some $\beta \geq 2$. Here and in what follows, for two positive functions $f(t, z)$ and $g(t, z)$, notation $f \asymp g$ means that there exist positive constants c_i $(i = 1, \cdots, 4)$ such that

$$c_1 f(c_2 t, z) \leq g(t, z) \leq c_3 f(c_4 t, z).$$

For example, a celebrate result due to Aronson [**Ar**] says that a symmetric diffusion process X on $\mathbb{R}^d$ associated with the uniformly elliptic operator $\mathcal{L} = \sum_{i,j=1}^d \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial}{\partial x_j}\right)$ has the two-sided estimates (1.10) with $\beta = 2$, $V(x, r) \simeq r^d$ and $\rho(x, y) = |x - y|$ being the Euclidean metric. This is also the case for Brownian motions on the Sierpinski gasket with $\beta = \log(n + 3)/\log 2$ in (1.10), or on the Sierpinski carpet in $\mathbb{R}^n$ $(n \geq 2)$ with $\beta > 2$ in (1.10); see [**B, Ku**].

Let $S := (S_t)_{t \geq 0}$ be a subordinator with $S_0 = 0$ that is independent of X and has the Laplace exponent

$$f(r) = \int_0^\infty (1 - e^{-rs})\nu(s)\, ds, \quad r > 0,$$

where

$$\nu(s)=\frac{1}{s^{1+\gamma_1}}\mathbf{1}_{\{0<s\le 1\}}+\frac{1}{s^{1+\gamma_2}}\mathbf{1}_{\{s>1\}}$$

with $\gamma_1\in(0,1)$ and $\gamma_2\in(1,\infty)$. Let $Y:=(Y_t)_{t\ge 0}$ be the subordinate process defined by $Y_t:=X_{S_t}$ for all $t>0$. For a set $A\subset M$, define the exit time $\tau_A^Y=\inf\{t>0:Y_t\notin A\}$. The subordinate process Y has the following properties which are established in [**CKW5**, Example 1.1].

Assertion 1. The subordinated process Y is a symmetric jump process such that

(i) its jumping kernel $J(dx,dy)$ has a density with respect to the product measure $\mu\times\mu$ on $M\times M\setminus\mathrm{diag}$ given by

$$J(x,y)\simeq\begin{cases}\frac{1}{V(x,\rho(x,y))\rho(x,y)^{\alpha_1}}, & \rho(x,y)\le 1,\\ \frac{1}{V(x,\rho(x,y))\rho(x,y)^{\alpha_2}}, & \rho(x,y)>1,\end{cases}\tag{1.11}$$

where $\alpha_1=\gamma_1\beta$ and $\alpha_2=\gamma_2\beta$, and diag stands for the diagonal of $M\times M$.

(ii) for any $x_0\in M$ and $r>0$,

$$\mathbb{E}^x\left[\tau^Y_{B(x_0,r)}\right]\simeq r^{\alpha_1}\vee r^{\beta}.\tag{1.12}$$

Assertion 2. The process Y has a jointly continuous transition density function $p(t,x,y)$ with respect to the measure μ on M so that

$$p(t,x,y)\simeq\frac{1}{V(x,t^{1/\alpha_1})}\wedge(tJ(x,y))\quad\text{for } t\le 1$$

and

$$\begin{aligned}&c_1\left(\frac{1}{V(x,t^{1/\beta})}\mathbf{1}_{\{\rho(x,y)\le t^{1/\beta}\}}+\frac{t}{V(x,\rho(x,y))\rho(x,y)^{\alpha_2}}\mathbf{1}_{\{\rho(x,y)>t^{1/\beta}\}}\right)\\ &\le p(t,x,y)\le c_2\left(\frac{1}{V(x,t^{1/\beta})}\wedge(tJ(x,y)+q(c_3t,x,y))\right)\quad\text{for } t>1,\end{aligned}$$

where $q(t,x,y)$ is the transition density function for the diffusion process X of the form (1.10). If, in addition, (M,ρ,μ) is connected and satisfies the chain condition, then

$$p(t,x,y)\asymp\begin{cases}\frac{1}{V(x,t^{1/\alpha_1})}\wedge(tJ(x,y)), & t\le 1,\\ \frac{1}{V(x,t^{1/\beta})}\wedge(tJ(x,y)+q(t,x,y)), & t>1.\end{cases}\tag{1.13}$$

The first property of Assertion 1 says that the scaling function of jumping kernel for the process Y is $\phi_j(r):=r^{\alpha_1}\vee r^{\alpha_2}$ with $\alpha_1<\beta<\alpha_2$, while the second property of Assertion 1 indicates that the scaling function of the process Y is $r^{\alpha_1}\vee r^{\beta}$. The scaling function of the process Y is different from the associated jumping kernel at large scale (that is, when $r>1$). Thus, comparing (1.11)-(1.12) with (1.8)-(1.9), we can see that the behavior of the symmetric jump process Y in Example 1.2 is different from that of symmetric α_1-stable-like or mixed stable-like processes studied in [**CK, CKW1**], where the scale functions for large jumps are all assumed to be less than that for the diffusion processes if there is one (for example, $r\mapsto r^2$ in the Euclidean space case). Due to these differences, the process Y above has two-sided heat kernel estimates (1.13), which is of a different shape from (1.6). In particular, this may appear surprising at the first glance that in (1.13) there is the diffusive scaling $\phi_c(r):=r^{\beta}$ when $r>1$ involved. But it becomes quite reasonable if one thinks more about it as the jumping kernel $J(dx,dy)$ of Y has finite second moment in the case of $\beta=2$. Note that in this example, $\phi_j(r)\geq \phi_c(r)$ for all $r>0$.

The purpose of this article is to summarize our recent work in [**CKW1**, **CKW2**, **CKW5**] on stable characterizations of heat kernel estimates and parabolic Harnack inequalities for general symmetric pure jump Dirichlet forms, which include a large class of symmetric pure jump Dirichlet forms that have light jumping kernel at infinity and thus possibly exhibit diffusive behaviors.

2. Heat kernel for symmetric pure jump processes

In this section, we survey some recent developments on heat kernel estimates for general symmetric pure jump processes. The reader is referred to [**CKW5**] for further details.

2.1. Two scaling functions. Let (M,ρ,μ) be a *metric measure space*. We assume that all balls are relatively compact and assume for simplicity that $\mu(M)=\infty$. We do not assume M to be connected nor (M,ρ) to be geodesic. Throughout this paper, we always assume that VD holds, and occasionally we also assume RVD holds. We are concerned with regular non-local Dirichlet forms $(\mathcal{E},\mathcal{F})$ on $L^2(M;\mu)$ given by (1.4). Let $X=\{X_t,t\geq 0;\mathbb{P}^x,x\in M\setminus\mathcal{N}\}$ be the associated Hunt process.

In order to consider heat kernel estimates for the Dirichlet form $(\mathcal{E},\mathcal{F})$ above, that may have light jumping kernel at infinity, we need to introduce two different scaling functions. Set $\mathbb{R}_+:=[0,\infty)$. Let $\phi_j:\mathbb{R}_+\to\mathbb{R}_+$ and $\phi_c:[1,\infty)\to\mathbb{R}_+$ be strictly increasing continuous functions with $\phi_j(0)=\phi_c(0)=0$, $\phi_j(1)=\phi_c(1)=1$ and satisfying that there exist constants $c_{1,\phi_j},c_{2,\phi_j},c_{1,\phi_c},c_{2,\phi_c}>0$, $\beta_{2,\phi_j}\geq\beta_{1,\phi_j}>0$ and $\beta_{2,\phi_c}\geq\beta_{1,\phi_c}>1$ such that

$$
\begin{aligned}
c_{1,\phi_j}\Big(\frac{R}{r}\Big)^{\beta_{1,\phi_j}} &\leq \frac{\phi_j(R)}{\phi_j(r)} \leq c_{2,\phi_j}\Big(\frac{R}{r}\Big)^{\beta_{2,\phi_j}} \quad \text{for all } 0<r\leq R,\\
c_{1,\phi_c}\Big(\frac{R}{r}\Big)^{\beta_{1,\phi_c}} &\leq \frac{\phi_c(R)}{\phi_c(r)} \leq c_{2,\phi_c}\Big(\frac{R}{r}\Big)^{\beta_{2,\phi_c}} \quad \text{for all } 0<r\leq R.
\end{aligned}
\tag{2.1}
$$

Since $\beta_{1,\phi_c}>1$, we know from [**BGT**, Definition, p. 65; Definition, p. 66; Theorem 2.2.4 and its remark, p. 73] that there exists a strictly increasing function $\bar\phi_c:\mathbb{R}_+\to\mathbb{R}_+$ such that there is a constant $c_1\geq 1$ with

$$
c_1^{-1}\phi_c(r)/r\leq\bar\phi_c(r)\leq c_1\phi_c(r)/r \quad \text{for all } r>0. \tag{2.2}
$$

Clearly, by (2.1) and (2.2), there exist constants $c_{1,\bar\phi_c},c_{2,\bar\phi_c}>0$ such that

$$
c_{1,\bar\phi_c}\Big(\frac{R}{r}\Big)^{\beta_{1,\phi_c}-1} \leq \frac{\bar\phi_c(R)}{\bar\phi_c(r)} \leq c_{2,\bar\phi_c}\Big(\frac{R}{r}\Big)^{\beta_{2,\phi_c}-1} \quad \text{for all } 0<r\leq R.
$$

Define

$$
\beta_*:=\sup\Big\{\beta>0: \text{there is a constant } c_*>0 \text{ so that } \phi_j(R)/\phi_j(r)\geq c_*(R/r)^\beta \text{ for } 0<r<R\leq 1\Big\},
$$

and

$$\beta^* := \sup\Big\{\beta>0: \text{there is a contant } c^*>0 \text{ so that } \phi_j(R)/\phi_j(r)\ge c^*(R/r)^\beta \\ \text{for all } R>r\ge 1\Big\}.$$

Throughout this paper, we assume that there is a constant $c_0\ge 1$ so that

$$\phi_c(r)\le c_0\phi_j(r) \text{ on } [0,1] \text{ if } \beta_*>1 \text{ and } \phi_c(r)\le c_0\phi_j(r) \text{ on } (1,\infty) \text{ if } \beta^*>1. \quad (2.3)$$

We point out that, by (2.1) with $\beta_{1,\phi_c}>1$, ϕ_c is not comparable to ϕ_j on $[0,1]$ when $\beta_*\le 1$, and ϕ_c is not comparable to ϕ_j on $[1,\infty)$ when $\beta^*\le 1$.

Roughly speaking, the function ϕ_j plays the role of the scaling function in the jumping kernel; while ϕ_c is a scale function that should be intrinsically determined by ϕ_j and the metric measure space (M,ρ,μ) which will possibly appear in the expression of heat kernel estimates. However, we do not have a universal formula for ϕ_c. For example, when the state space M (such as $\mathbb{R}^d$ or a nice fractal) has a nice diffusion process, ϕ_c can be the scaling function of the diffusion in some cases but can also be a different scale function in some other cases; see Examples 1.2 and 4.1. In some cases, ϕ_c can just be ϕ_j on a part or the whole of $[0,\infty)$. To cover a wide spectrum of scenarios, in the formulation and characterization we allow ϕ_c to be any function that satisfies conditions (2.1) and (2.3).

Given ϕ_c and ϕ_j as above, we set

$$\phi(r):=\begin{cases}\phi_j(r)\mathbf{1}_{\{\beta_*\le 1\}}+\phi_c(r)\mathbf{1}_{\{\beta_*>1\}} & \text{for } 0<r\le 1,\\ \phi_j(r)\mathbf{1}_{\{\beta^*\le 1\}}+\phi_c(r)\mathbf{1}_{\{\beta^*>1\}} & \text{for } r>1.\end{cases} \quad (2.4)$$

In view of the assumptions above, ϕ is strictly increasing on $\mathbb{R}_+$ such that $\phi(0)=0$, $\phi(1)=1$ and there exist constants $c_{1,\phi},c_{2,\phi}>0$ so that

$$c_{1,\phi}\Big(\frac{R}{r}\Big)^{\beta_{1,\phi}}\le\frac{\phi(R)}{\phi(r)}\ \le\ c_{2,\phi}\Big(\frac{R}{r}\Big)^{\beta_{2,\phi}} \quad \text{for all } 0<r\le R,$$

where $\beta_{1,\phi}=\beta_{1,\phi_c}\wedge\beta_{1,\phi_j}$ and $\beta_{2,\phi}=\beta_{2,\phi_c}\vee\beta_{2,\phi_j}$. Clearly, we have by (2.3) that

$$\phi(r)\le c_0\phi_j(r) \quad \text{for every } r\ge 0.$$

Denote by $\phi_j^{-1}(t)$, $\phi_c^{-1}(t)$ and $\bar\phi_c^{-1}(t)$ the inverse functions of the strictly increasing functions $t\mapsto\phi_j(t)$, $t\mapsto\phi_c(t)$ and $t\mapsto\bar\phi_c(t)$, respectively. Then the inverse function ϕ^{-1} of ϕ is given by

$$\phi^{-1}(r):=\begin{cases}\phi_j^{-1}(r)\mathbf{1}_{\{\beta_*\le 1\}}+\phi_c^{-1}(r)\mathbf{1}_{\{\beta_*>1\}} & \text{for } 0<r\le 1,\\ \phi_j^{-1}(r)\mathbf{1}_{\{\beta^*\le 1\}}+\phi_c^{-1}(r)\mathbf{1}_{\{\beta^*>1\}} & \text{for } r>1.\end{cases}$$

Throughout this paper, we will fix the notations for these functions ϕ_c, ϕ_j, ϕ and $\bar\phi_c$. In particular, as we will see from the results below, ϕ is the "true" scaling function for the process X. For example, $\phi(r)=r^{\alpha_1}\vee r^\beta$ for the process Y in Example 1.2, and the scaling function for its jumping kernel is $\phi_j(r)=r^{\alpha_1}\vee r^{\alpha_2}$, where $\alpha_1<\beta<\alpha_2$.

2.2. Formulas for heat kernel estimates. In this subsection, we present formulas of heat kernel estimates for general symmetric pure jump processes. As we will see, the processes will enjoy heat kernel estimates with different t forms that the scaling functions ϕ_c, ϕ_j and $\bar\phi_c$ are fully or partly involved, according to different ranges of the indexes β_* and β^*.

Recall that the function $\bar\phi_c(r)$ is a strictly increasing function satisfying (2.2). For any $t>0$ and $x,y\in M_0$, set

$$p^{(j)}(t,x,y):=\frac{1}{V(x,\phi_j^{-1}(t))}\wedge\frac{t}{V(x,\rho(x,y))\phi_j(\rho(x,y))} \tag{2.5}$$

and

$$p^{(c)}(t,x,y):=\frac{1}{V(x,\phi_c^{-1}(t))}\exp\left(-\frac{\rho(x,y)}{\bar\phi_c^{-1}(t/\rho(x,y))}\right). \tag{2.6}$$

Here, $p^{(j)}(t,x,y)$ follows from two-sided heat kernel estimates for mixture of symmetric stable-like processes on metric measure space; see (1.6); while $p^{(c)}(t,x,y)$ is partly motivated by two-sided heat kernel estimates for strongly local Dirichlet forms, see [**GT, HK**].

DEFINITION 2.1. (i) We say that $\mathrm{HK}(\phi_j,\phi_c)$ holds if there exists a heat kernel $p(t,x,y)$ of the semigroup $\{P_t\}$ associated with $(\mathcal{E},\mathcal{F})$, which has the following estimates for all $x,y\in M_0$,

$$p(t,x,y)\asymp\begin{cases}\begin{cases}p^{(j)}(t,x,y) & \text{if } \beta_*\le 1\\ \frac{1}{V(x,\phi_c^{-1}(t))}\wedge\left(p^{(c)}(t,x,y)+p^{(j)}(t,x,y)\right) & \text{if } \beta_*>1\end{cases} & \text{for } t\le 1,\\ \begin{cases}p^{(j)}(t,x,y) & \text{if } \beta^*\le 1\\ \frac{1}{V(x,\phi_c^{-1}(t))}\wedge\left(p^{(c)}(t,x,y)+p^{(j)}(t,x,y)\right) & \text{if } \beta^*>1\end{cases} & \text{for } t>1.\end{cases} \tag{2.7}$$

(ii) We say $\mathrm{HK}_-(\phi_j,\phi_c)$ holds if the upper bound in (2.7) holds but the lower bound is replaced by the following statement: there are constants $c_1,c_2>0$ so that

$$p(t,x,y)\ge c_1\begin{cases}\frac{1}{V(x,\phi^{-1}(t))}, & \rho(x,y)\le c_2\phi^{-1}(t),\\ \frac{t}{V(x,\rho(x,y))\phi_j(\rho(x,y))}, & \rho(x,y)>c_2\phi^{-1}(t).\end{cases}$$

(iii) We say $\mathrm{UHK}(\phi_j,\phi_c)$ holds if the upper bound in (2.7) holds.

The scale function ϕ_c plays a role in the definitions of $\mathrm{HK}(\phi_j,\phi_c)$ and $\mathrm{HK}_-(\phi_j,\phi_c)$ only for $t\le 1$ when $\beta_*>1$ and for $t>1$ when $\beta^*>1$. The cut-off time 1 here is not important – it can be replaced by any fixed constant $T>0$. Furthermore, it follows from Theorem 3.3 below that $\mathrm{HK}_-(\phi_j,\phi_c)$ (and so $\mathrm{HK}(\phi_j,\phi_c)$) is stronger than $\mathrm{PHI}(\phi)$, which in turn yields the Hölder regularity of parabolic functions; see the proof of [**CKW4**, Theorem 1.17]. In particular, this implies that if $\mathrm{HK}_-(\phi_j,\phi_c)$ (respectively, $\mathrm{HK}(\phi_j,\phi_c)$) holds, then it can be strengthened to hold for all $x,y\in M$, and consequently the Hunt process X can be refined to start from every point in M.

We note that the expression of $\mathrm{HK}(\phi_j,\phi_c)$ takes different form depending on whether $\beta_*\le 1$ (respectively, $\beta^*\le 1$) or not. This is because when $\beta_*>1$ (respectively, $\beta^*>1$), the heat kernel estimate may involve another function ϕ_c that is intrinsically determined by ϕ_j but we do not have a generic formula for it under our general setting. However it does not necessarily mean that the heat kernel estimates for $p(t,x,y)$ has a phase transition exactly at $\beta_*=1$ or $\beta^*=1$. For example, suppose (M,ρ,μ) is a connected metric measure space satisfying the chain condition on which there is a symmetric diffusion process enjoying the heat kernel estimate (1.10) with $\beta>1$ as in Example 1.2. Consider a symmetric pure jump process Y on this space whose jumping density has bounds (1.11) with $\alpha_1<\beta$ and $\alpha_2>0$. Clearly, $\beta_*=\alpha_1$ and $\beta^*=\alpha_2$. When $\alpha_2<\beta$, the transition density function $p(t,x,y)$ of Y has estimates (2.5) with $\phi_j(r)=r^{\alpha_1}\mathbf{1}_{\{0<r\le 1\}}+r^{\alpha_2}\mathbf{1}_{\{r>1\}}$ as in [**CKW1**], whereas for $\alpha_2>\beta$, $p(t,x,y)$ satisfies (1.13) as in Example 1.2. Hence for this example, phase transition for the expression of heat kernel estimates for $p(t,x,y)$ occurs at $\alpha_2=\beta$.

2.3. Jumping kernel and functional inequalities. To state our stable characterizations of heat kernel estimates for general symmetric pure jump processes, we need four definitions.

Definition 2.2. Let $\psi:\mathbb{R}_+\to\mathbb{R}_+$. We say J_ψ holds, if there exists a non-negative symmetric function $J(x,y)$ so that (1.7) is satisfied, and (1.8) holds with ψ in place of ϕ for all $x,y\in M$; that is,

$$\frac{c_1}{V(x,\rho(x,y))\psi(\rho(x,y))}\le J(x,y)\le\frac{c_2}{V(x,\rho(x,y))\psi(\rho(x,y))}. \tag{2.8}$$

We say that $\mathrm{J}_{\psi,\le}$ (resp. $\mathrm{J}_{\psi,\ge}$) if (1.7) holds and the upper bound (resp. lower bound) in (2.8) holds.

Note that, since $\phi(r)\le c_0\phi_j(r)$ for all $r>0$, $\mathrm{J}_{\phi_j,\le}$ implies $\mathrm{J}_{\phi,\le}$; that is, $\mathrm{J}_{\phi,\le}$ is weaker than $\mathrm{J}_{\phi_j,\le}$.

Let $U\subset V$ be open sets of M with $U\subset\overline{U}\subset V$. We say a non-negative bounded measurable function φ is a *cut-off function for* $U\subset V$, if $\varphi=1$ on U, $\varphi=0$ on V^c and $0\le\varphi\le 1$ on M. For $f,g\in\mathcal{F}$, we define the carré du-Champ operator $\Gamma(f,g)$ for the symmetric pure jump Dirichlet form $(\mathcal{E},\mathcal{F})$ by

$$\Gamma(f,g)(dx)=\int_{y\in M}(f(x)-f(y))(g(x)-g(y))\,J(dx,dy).$$

Clearly $\mathcal{E}(f,g)=\Gamma(f,g)(M)$. We now introduce the following cut-off Sobolev inequality $\mathrm{CSJ}(\phi)$ that controls the energy of cut-off functions.

Definition 2.3. Let $\mathcal{F}_b=\mathcal{F}\cap L^\infty(M,\mu)$. We say that condition $\mathrm{CSJ}(\phi)$ holds, if there exist constants $C_0\in(0,1]$ and $C_1,C_2>0$ such that for every $0<r\le R$, almost all $x_0\in M$ and any $f\in\mathcal{F}$, there exists a cut-off function $\varphi\in\mathcal{F}_b$ for $B(x_0,R)\subset B(x_0,R+r)$ such that

$$\begin{aligned}&\int_{B(x_0,R+(1+C_0)r)}f^2\,d\Gamma(\varphi,\varphi)\\&\le C_1\int_{U\times U^*}(f(x)-f(y))^2\,J(dx,dy)+\frac{C_2}{\phi(r)}\int_{B(x_0,R+(1+C_0)r)}f^2\,d\mu,\end{aligned} \tag{2.9}$$

where $U=B(x_0,R+r)\setminus B(x_0,R)$ and $U^*=B(x_0,R+(1+C_0)r)\setminus B(x_0,R-C_0r)$.

Remark 2.4. (i) $\mathrm{CSJ}(\phi)$ for symmetric pure jump Dirichlet forms is first introduced in [**CKW1**] as a counterpart of $\mathrm{CSA}(\phi)$ for strongly local

Dirichlet forms (see [**AB, BB, BBK1**]). A similar condition is called condition (AB) in [**GHH**] for the case $\phi(r)=r^\alpha$. As pointed out in [**CKW1**, Remark 1.6(ii)], the main difference between CSJ(ϕ) and CSA(ϕ) is that the integrals in the left hand side and in the second term of the right hand side of the inequality (2.9) are over $B(x_0,R+(1+C_0)r)$ instead of over $B(x_0,R+r)$ in [**AB**]. Note that the integral over $B(x_0,R+r)^c$ is zero in the left hand side of (2.9) for the case of strongly local Dirichlet forms. As we see from the approach of [**CKW1**] in the study of stability of heat kernel estimates for symmetric mixed stable-like processes, it is important to enlarge the ball $B(x_0,R+r)$ and integrate over $B(x_0,R+(1+C_0)r)$ rather than over $B(x_0,R+r)$.

(ii) Denote by $\mathcal{F}_{loc}$ the space of functions locally in $\mathcal{F}$; that is, $f\in\mathcal{F}_{loc}$ if and only if for any relatively compact open set $U\subset M$ there exists $g\in\mathcal{F}$ such that $f=g$ μ-a.e. on U. Since each ball is relatively compact and (2.9) uses the property of f on $B(x_0,R+(1+C_0)r)$ only, CSJ(ϕ) also holds for any $f\in\mathcal{F}_{loc}$.

We next introduce the Faber-Krahn inequality and the (weak) Poincaré inequality.

Definition 2.5. We say the MMD space $(M,\rho,\mu,\mathcal{E})$ satisfies the *Faber-Krahn inequality* FK(ϕ) if there exist positive constants C and p such that for any ball $B(x,r)$ and any open set $D\subset B(x,r)$,

$$\lambda_1(D)\ge \frac{C}{\phi(r)}(V(x,r)/\mu(D))^p,$$

where $\lambda_1(D)=\inf\{\mathcal{E}(f,f):\, f\in\mathcal{F}_D \text{ with } \|f\|_2=1\}$ and $\mathcal{F}_D$ is defined to be the $\sqrt{\mathcal{E}(\cdot,\cdot)+\|\cdot\|_2^2}$-closure in $\mathcal{F}$ of $\mathcal{F}\cap C_c(D)$.

Definition 2.6. We say that the (*weak*) *Poincaré inequality* PI(ϕ) holds if there exist constants $C>0$ and $\kappa\ge 1$ such that for any ball $B_r=B(x,r)$ with $x\in M$ and for any $f\in\mathcal{F}_b$,

$$\int_{B_r}(f-\overline{f}_{B_r})^2\,d\mu\le C\phi(r)\int_{B_{\kappa r}\times B_{\kappa r}}(f(y)-f(x))^2\,J(dx,dy),$$

where $\overline{f}_{B_r}=\frac{1}{\mu(B_r)}\int_{B_r} f\,d\mu$ is the average value of f on B_r.

2.4. Stable characterizations of two-sided heat kernel estimates. With the notations above, we can now state the following stable characterizations of two-sided heat kernel estimates and upper bounds of heat kernel for general symmetric pure jump process from [**CKW5**].

Theorem 2.7. *Assume that the metric measure space (M,ρ,μ) satisfies* VD *and* RVD*, and the functions ϕ_c and ϕ_j satisfy* (2.1) *and* (2.3)*. Let $\phi(r)$ be defined by* (2.4)*. The following are equivalent.*

(1) HK$_-(\phi_j,\phi_c)$.
(2) PI(ϕ), J$_{\phi_j}$ *and* CSJ(ϕ).

If, in additional, (M,ρ,μ) is connected and satisfies the chain condition, then each assertion above is equivalent to

(3) HK(ϕ_j,ϕ_c).

We refer the reader to [**CKW5**, Theorem 1.11] for more equivalent characterizations of $\mathrm{HK}_-(\phi_j,\phi_c)$ and $\mathrm{HK}(\phi_j,\phi_c)$. We emphasize again that the connectedness and the chain condition of the underlying metric measure space (M,ρ,μ) are only used to derive optimal lower bounds off-diagonal estimates for heat kernel when the time is small (i.e., from $\mathrm{HK}_-(\phi_j,\phi_c)$ to $\mathrm{HK}(\phi_j,\phi_c)$), while for the equivalence between (1) and (2) in the result above, the metric measure space (M,ρ,μ) is only assumed to satisfy the general VD and RVD; that is, we do not assume M to be connected nor (M,ρ) to be geodesic. (In fact, (2) $\Longrightarrow$ (1) in Theorem 2.7 holds true under VD and (2.1), without assuming RVD). Furthermore, we do not assume the uniform comparability of volume of balls; that is, we do not assume the existence of a non-decreasing function V on $[0,\infty)$ with $V(0)=0$ so that $\mu(B(x,r))\asymp V(r)$ for all $x\in M$ and $r>0$.

We also have the following characterizations for $\mathrm{UHK}(\phi_j,\phi_c)$ (see [**CKW5**, Theorem 1.12] for more equivalent characterizations). In the following, we say $(\mathcal{E},\mathcal{F})$ *on* $L^2(M;\mu)$ *is conservative* if its associated Hunt process X has infinite lifetime. This is equivalent to $P_t1=1$ on M_0 for every $t>0$. It follows from [**CKW1**, Proposition 3.1] that any equivalent statement of Theorem 2.7 implies that the process X is conservative. We also point out that $\mathrm{UHK}(\phi_j,\phi_c)$ alone does not imply the conservativeness of the associated Dirichlet form $(\mathcal{E},\mathcal{F})$ (see [**CKW1**, Proposition 3.1 and Remark 3.2] for more details).

THEOREM 2.8. *Assume that the metric measure space (M,ρ,μ) satisfies* VD *and* RVD*, and that the functions ϕ_c and ϕ_j satisfy* (2.1) *and* (2.3). *Let $\phi(r)$ be defined by* (2.4). *Then the following are equivalent:*

(1) $\mathrm{UHK}(\phi_j,\phi_c)$ *and* $(\mathcal{E},\mathcal{F})$ *on* $L^2(M;\mu)$ *is conservative.*

(2) $\mathrm{FK}(\phi)$, $\mathrm{J}_{\phi_j,\le}$ *and* $\mathrm{CSJ}(\phi)$.

We emphasize that the above two theorems are equivalent characterizations and stability results. It is possible that none of the statements hold with a bad selection of ϕ_c.

2.5. Further remarks on Theorems 2.7 and 2.8. In this subsection, we make further comments on the formulations of $\mathrm{HK}_-(\phi_j,\phi_c)$ and $\mathrm{HK}(\phi_j,\phi_c)$, and discuss relations of the main results above (Theorems 2.7 and 2.8) to those in the literature. Recall that the function ϕ is defined by (2.4) and $p^{(j)}(t,x,y),p^{(c)}(t,x,y)$ are defined as in (2.5) and (2.6).

REMARK 2.9. (i) By simple calculations, we have

$$p^{(c)}(t,x,y)\preceq \frac{1}{V(x,\phi_c^{-1}(t))}\wedge\frac{t}{V(x,\rho(x,y))\phi_c(\rho(x,y))}\quad \text{on } (0,\infty)\times M_0\times M_0,$$

where $f\preceq g$ means that there exists a constant $c>0$ such that $f(x)\le cg(x)$ for the specified range of x. Hence, when $\beta^*>1$ and $\phi_c=\phi_j$ on $[1,\infty)$, it holds that

$$p^{(c)}(t,x,y)\preceq \frac{1}{V(x,\phi_j^{-1}(t))}\wedge\frac{t}{V(x,\rho(x,y))\phi_j(\rho(x,y))}\quad \text{on } (1,\infty)\times M_0\times M_0.$$

Consequently, in this case we have,

$$\frac{1}{V(x,\phi_c^{-1}(t))}\wedge\big(p^{(c)}(t,x,y)+p^{(j)}(t,x,y)\big)\asymp p^{(j)}(t,x,y)\quad \text{on } (1,\infty)\times M_0\times M_0.$$

Similarly, one can check that in the case of $\beta_* > 1$ and $\phi_c = \phi_j$ on $[0,1]$,

$$\frac{1}{V(x,\phi_c^{-1}(t))} \wedge \left(p^{(c)}(t,x,y) + p^{(j)}(t,x,y)\right) \asymp p^{(j)}(t,x,y) \quad \text{on } (0,1] \times M_0 \times M_0.$$

Therefore, when $\phi = \phi_j$ on $[0,\infty)$ (that is, when $\phi_c = \phi_j$ on $[0,1]$ if $\beta_* > 1$ and $\phi_c = \phi_j$ on $(1,\infty)$ if $\beta^* > 1$), $\mathrm{HK}(\phi_j,\phi_c)$ is just the heat kernel estimate $p^{(j)}(t,x,y)$. Thus, in this case, Theorems 2.7 and 2.8 are essentially the main results (Theorems 1.13 and 1.15) of [**CKW1**]. We note that by the proof of [**CKW1**, Lemma 4.1], $\mathrm{J}_{\phi_j,\geq}$ implies $\mathrm{PI}(\phi_j)$ so in the case of $\phi = \phi_j$ on $[0,\infty)$, we can drop condition $\mathrm{PI}(\phi)$ from the statement of Theorem 2.7.

(ii) When $\phi(r) = \phi_c(r)$ on $[0,\infty)$ (that is, when $\beta_* \wedge \beta^* > 1$, and $\phi_c(r)$ is a strictly increasing function satisfying (2.1) and (2.3)), $\mathrm{HK}_-(\phi_j,\phi_c)$ and $\mathrm{HK}(\phi_j,\phi_c)$ are just $\mathrm{HK}(\Phi,\psi)$ and $\mathrm{SHK}(\Phi,\psi)$ in [**BKKL2**, Definition 2.8] with

$$\psi = \phi_j \quad \text{and} \quad \Phi = \phi_c.$$

In this case, our Theorems 2.7 and 2.8 have also been independently obtained in [**BKKL2**, Theorem 2.14, Corollary 2.15, Theorem 2.17 and Corollary 2.18].

(iii) When $\phi(r) = \phi_c(r)\mathbf{1}_{[0,1]} + \phi_j(r)\mathbf{1}_{(1,\infty)}$ (that is, when $\beta_* > 1$ and either $\beta^* \leq 1$ or $\beta^* > 1$ with $\phi_c(r) = \phi_j(r)$ for all $r \in [1,\infty)$), $\mathrm{HK}(\phi_j,\phi_c)$ is reduced into

$$p(t,x,y) \asymp \begin{cases} \frac{1}{V(x,\phi_c^{-1}(t))} \wedge \left(p^{(c)}(t,x,y) + p^{(j)}(t,x,y)\right), & 0 < t \leq 1, \\ p^{(j)}(t,x,y), & t > 1. \end{cases} \tag{2.10}$$

In this case $\mathrm{HK}(\phi_j,\phi_c)$ is of the same form as that of $\mathrm{HK}(\phi_c,\phi_j)$ in [**CKW4**] for symmetric diffusions with jumps, or equivalent, for symmetric Dirichlet forms that contain both the strongly local part and the pure jump part; see [**CKW4**, Definition 1.11 and Remark 1.12] for details. However, there are differences between them. In [**CKW4**] the function ϕ_c in $\mathrm{HK}(\phi_c,\phi_j)$ is the scaling function of the diffusion (i.e. the strongly local part of Dirichlet forms), while in the present paper the function ϕ_c in $\mathrm{HK}(\phi_j,\phi_c)$ is determined by ϕ_j and the underlying metric measure space (M,ρ,μ).

In Remark 2.9, we have discussed the form of $\mathrm{HK}(\phi_j,\phi_c)$ for the cases of $\phi(r) = \phi_j(r)$ on $[0,\infty)$, $\phi(r) = \phi_c(r)$ on $[0,\infty)$, and $\phi(r) = \phi_c(r)\mathbf{1}_{[0,1]} + \phi_j(r)\mathbf{1}_{(1,\infty)}$, respectively. We now discuss the remaining case of $\phi(r)$.

Remark 2.10. This remark is concerned with the case that $\phi(r) = \phi_j(r)\mathbf{1}_{[0,1]} + \phi_c(r)\mathbf{1}_{(1,\infty)}$. It consists of two subcases of $\beta^* > 1$ where either $\beta_* \leq 1$ or $\beta_* > 1$ with $\phi_c(r) = \phi_j(r)$ for all $r \in (0,1]$.

(i) In this case, we can rewrite the expression of $\mathrm{HK}(\phi_j,\phi_c)$ in the following way. For $0 < t \leq 1$,

$$p(t,x,y) \simeq p^{(j)}(t,x,y) \simeq \begin{cases} \frac{1}{V(x,\phi_j^{-1}(t))}, & \rho(x,y) \leq c_1\phi_j^{-1}(t), \\ \frac{t}{V(x,\rho(x,y))\phi_j(\rho(x,y))}, & \rho(x,y) > c_1\phi_j^{-1}(t). \end{cases}$$

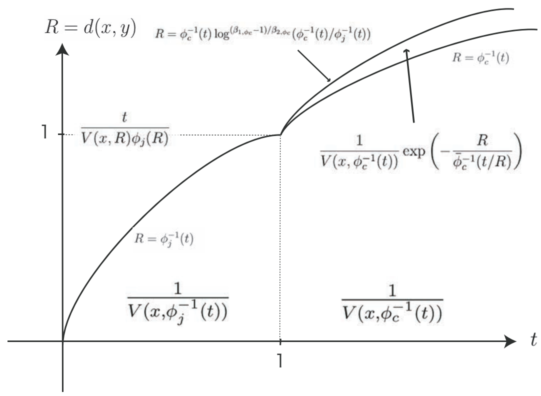

FIGURE 1. Dominant term in the heat kernel estimates $\mathrm{HK}(\phi_j, \phi_c)$ for $p(t,x,y)$ for the case in Remark 2.10.

For $t > 1$,

$$p(t,x,y) \asymp \frac{1}{V(x,\phi_c^{-1}(t))} \wedge \left(p^{(c)}(t,x,y) + p^{(j)}(t,x,y)\right)$$

$$\asymp \begin{cases} \frac{1}{V(x,\phi_c^{-1}(t))}, & \rho(x,y) \le c_2\phi_c^{-1}(t), \\ \frac{t}{V(x,\rho(x,y))\phi_j(\rho(x,y))} + \frac{1}{V(x,\phi_c^{-1}(t))} \exp\left(-\frac{\rho(x,y)}{\bar{\phi}_c^{-1}(t/\rho(x,y))}\right), & \rho(x,y) > c_2\phi_c^{-1}(t). \end{cases}$$

In particular, for $t \in (0,1]$, the heat kernel estimates $\mathrm{HK}(\phi_j, \phi_c)$ are completely dominated by the jumping kernel for the associated Dirichlet form $(\mathcal{E}, \mathcal{F})$. For $t > 1$, we have the following more explicit expression of $\mathrm{HK}(\phi_j, \phi_c)$:

$$p(t,x,y) \asymp \begin{cases} \frac{1}{V(x,\phi_c^{-1}(t))}, & \rho(x,y) \le c_2\phi_c^{-1}(t), \\ \frac{1}{V(x,\phi_c^{-1}(t))} \exp\left(-\frac{\rho(x,y)}{\bar{\phi}_c^{-1}(t/\rho(x,y))}\right), & c_2\phi_c^{-1}(t) < \rho(x,y) < t_*, \\ \frac{t}{V(x,\rho(x,y))\phi_j(\rho(x,y))}, & \rho(x,y) \ge t_*, \end{cases}$$

where t_* satisfies

$$c_3\phi_c^{-1}(t) \log^{(\beta_{1,\phi_c}-1)/\beta_{2,\phi_c}}(\phi_c^{-1}(t)/\phi_j^{-1}(t)) \le t_* \le c_4\phi_c^{-1}(t) \log^{(\beta_{2,\phi_c}-1)/\beta_{1,\phi_c}}(\phi_c^{-1}(t)/\phi_j^{-1}(t)),$$

and β_{1,ϕ_c} and β_{2,ϕ_c} are given in (2.1). In this case, one can check that only the information of ϕ_c on $[1,\infty)$ is actually needed for the expression of $\mathrm{HK}(\phi_j, \phi_c)$ because $t/\rho(x,y) \ge c_3 > 0$ holds when $t > 1$ and $c_2\phi_c^{-1}(t) \le \rho(x,y) \le t_*$. Figure 1 indicates in this case which term is the dominant one for the estimate of $p(t,x,y)$ in each region.

(ii) The definition of $\mathrm{HK}(\phi_j, \phi_c)$ in this case is different from $\mathrm{HK}(\phi_c, \phi_j)$ in [**CKW4**] for symmetric diffusions with jumps. Denote by the heat kernel by $\bar{p}(t,x,y)$ in the case of symmetric diffusion with jumps. We say $\mathrm{HK}(\phi_c, \phi_j)$ holds if (2.10) holds. Thus the expressions for $\mathrm{HK}(\phi_j, \phi_c)$ for symmetric jump processes with lighter jumping tails and $\mathrm{HK}(\phi_c, \phi_j)$ for symmetric diffusions with jumps are exactly switched over the time interval $(0,1]$ and $(1,\infty)$. The reason for this difference is as follows. For diffusions with jumps, due to the heavy tail property of jumps, the behaviors of the processes for large time are dominated by the pure jump part, and the behaviors for small time enjoy the continuous nature from diffusions as well as some interactions with jumps. For symmetric pure jump processes with lighter tails considered in this case, the small time behavior of the processes are controlled by the jumping kernels; however, since large jumps of the associated process is so light (as a typical example, one can consider symmetric jump processes on $\mathbb{R}^d$ whose associated jumping measure has finite second moments), it yields similar long time estimates as those for diffusions.

(iii) The lower bound in $\mathrm{HK}_-(\phi_j, \phi_c)$ can be expressed more explicitly; namely, there are constants $c_1, c_2 > 0$ so that

$$p(t,x,y) \ge c_1 \begin{cases} p^{(j)}(t,x,y), & 0 < t \le 1, \\ \frac{1}{V(x,\phi_c^{-1}(t))}, & t > 1,\ \rho(x,y) \le c_2\phi_c^{-1}(t), \\ \frac{t}{V(x,\rho(x,y))\phi_j(\rho(x,y))}, & t > 1,\ \rho(x,y) > c_2\phi_c^{-1}(t). \end{cases}$$

The next remark is on the formulation and proofs of Theorems 2.7 and 2.8.

REMARK 2.11. Motivated by Example 1.2, we start with two scaling functions ϕ_j and ϕ_c, which is quite natural. This allows us to treat the cases of $0 < r \le 1$ and $r > 1$ separably with possibly different scaling indices. It also allows us to incorporate the heat kernel formulation considered in [**BKKL1, BKKL2**]. The significance of this viewpoint is further illustrated by Example 4.1, where the local lower scaling index of the scale function ϕ_j on $(1, \infty)$ is strictly larger than 1, while the local lower scaling index of the scale function ϕ_j on $(0, 1)$ can take any value in $(0, 2)$. Both Example 1.2 and Example 4.1 are outside the main settings of [**BKKL1, BKKL2**].

The most difficult case in the proof of Theorems 2.7 and 2.8 is when $\phi(r) = \phi_j(r)\mathbf{1}_{[0,1]} + \phi_c(r)\mathbf{1}_{(1,\infty)}$; that is, the case discussed in Remark 2.10. On the other hand, as mentioned in Remark 2.10(ii), the expressions of heat kernel estimates for the process studied in this case are the same as these for diffusions with jumps studied in [**CKW4**] but with time interval $t \le 1$ and $t > 1$ switched. It turns out that some ideas and strategies from [**CKW4**] on diffusions with jumps can be adapted in this case but there are also new ingredients needed to deal with symmetric pure jump processes having light tails in [**CKW5**]. Our approach can yield more concise and explicit heat kernel estimate $\mathrm{HK}(\phi_j, \phi_c)$ as mentioned in Remark 2.10(i). Furthermore, our approach also gives us the stable characterizations of parabolic Harnack inequalities (see Theorem 3.3 below), which are useful as indicated in Example 4.2.

3. Stability of parabolic Harnack inequalities

Let $Z := \{V_s, X_s\}_{s\ge 0}$ be the space-time process corresponding to X, where $V_s = V_0 - s$. Denote the law of the space-time process Z starting from (t, x) by $\mathbb{P}^{(t,x)}$. For every open subset D of $[0, \infty) \times M$, define $\tau_D = \inf\{s > 0 : Z_s \notin D\}$.

DEFINITION 3.1. (i) We say that a Borel measurable function $u(t, x)$ on $[0, \infty) \times M$ is *parabolic* (or *caloric*) on $D = (a, b) \times B(x_0, r)$ for the process X if there is a properly exceptional set $\mathcal{N}_u$ associated with the process

X so that for every relatively compact open subset U of D, $u(t,x) = \mathbb{E}^{(t,x)} u(Z_{\tau_U})$ for every $(t,x) \in U \cap ([0,\infty) \times (M \setminus \mathcal{N}_u))$.

(ii) We say that the *parabolic Harnack inequality* (PHI(ϕ)) holds for the process X, if there exist constants $0 < C_1 < C_2 < C_3 < C_4$, $C_5 > 1$ and $C_6 > 0$ such that for every $x_0 \in M$, $t_0 \ge 0$, $R > 0$ and for every non-negative function $u = u(t,x)$ on $[0,\infty) \times M$ that is parabolic on cylinder $Q(t_0, x_0, \phi(C_4 R), C_5 R) := (t_0, t_0 + \phi(C_4 R)) \times B(x_0, C_5 R)$,

$$\text{ess sup}_{Q_-} u \le C_6 \, \text{ess inf}_{Q_+} u, \tag{3.1}$$

where $Q_- := (t_0 + \phi(C_1 R), t_0 + \phi(C_2 R)) \times B(x_0, R)$ and $Q_+ := (t_0 + \phi(C_3 R), t_0 + \phi(C_4 R)) \times B(x_0, R)$.

The next definition was introduced in [**BBK2**] in the setting of graphs, and then extended in [**CKK1**] to the general setting of metric measure spaces.

Definition 3.2. We say that UJS holds if there is a symmetric function $J(x,y)$ so that (1.7) holds, and there is a constant $c > 0$ such that for any $x, y \in M$ and $0 < r \le \rho(x,y)/2$,

$$J(x,y) \le \frac{c}{V(x,r)} \int_{B(x,r)} J(z,y) \, \mu(dz).$$

The following result is established in [**CKW5**, Theorem 1.15], which extends the corresponding result in [**CKW2**, Theorem 1.20 and Corollary 1.21] from jump kernels of mixture stable types to more general jumping kernels. It gives the stable characterization of parabolic Harnack inequalities, as well as the relation between parabolic Harnack inequalities and two-sided heat kernel estimates.

Theorem 3.3. *Suppose that the metric measure space (M, ρ, μ) satisfies* VD *and* RVD*, and that the functions ϕ_c and ϕ_j satisfy* (2.1) *and* (2.3)*. Let $\phi(r)$ be defined by* (2.4)*. Then*

$$\mathrm{PHI}(\phi) \Longleftrightarrow \mathrm{PI}(\phi) + \mathrm{J}_{\phi,\le} + \mathrm{CSJ}(\phi) + \mathrm{UJS}.$$

Consequently,

$$\mathrm{HK}_-(\phi_j, \phi_c) \Longleftrightarrow \mathrm{PHI}(\phi) + \mathrm{J}_{\phi_j}.$$

If, additionally, the metric measure space (M, ρ, μ) is connected and satisfies the chain condition, then

$$\mathrm{HK}(\phi_j, \phi_c) \Longleftrightarrow \mathrm{PHI}(\phi) + \mathrm{J}_{\phi_j}.$$

Like [**CKW2**, Theorem 1.20], we can obtain more equivalent statements for PHI(ϕ). But we will not go into details here for the sake of space consideration. We also note that similar to the setting of [**CKW2**], the proof of Theorem 3.3 is proved under an additional assumption that the jumping measure $J(dx,dy)$ in the non-local Dirichlet form $(\mathcal{E}, \mathcal{F})$ in (1.4) is of the form $J(dx,dy) = J(x,dy)\,\mu(dx)$ on $M \times M \setminus \text{diag}$. Under this assumption, it is shown in [**CKW2**, Proposition 3.3] that PHI(ϕ) implies that $J(x,dy)$ is absolutely continuous with respect to the measure $\mu(dy)$ for μ-a.e. $x \in M$ and so $J(dx,dy)$ is absolutely continuous with respect to $\mu(dx)\,\mu(dy)$. Recently it is observed in [**LM**] that the proof of [**CKW2**, Proposition 3.3] can be refined to show that PHI(ϕ) always implies that $J(dx,dy)$ is absolutely continuous with respect to the product measure $\mu(dx)\,\mu(dy)$ on $M \times M \setminus \text{diag}$.

4. Examples

To illustrate the main results presented above, we give two more examples in this section. They are taken from Examples 5.2 and 5.3 from [**CKW5**], where the reader can find details of the proofs for the assertions below.

The first example shows that for Theorem 2.7, the scale function $\phi_c(r)$ does not need to be the scaling function corresponding to diffusion processes (e.g. see Example 1.2) even in the case of the Euclidean space.

EXAMPLE 4.1. Let $M = \mathbb{R}^d$ and $\mu(dx) = dx$. Consider the following jumping kernel

$$J(x, y) \simeq \frac{1}{|x-y|^d \phi_j(|x-y|)},$$

where ϕ_j is a strictly increasing continuous function with $\phi_j(0) = 0$ and $\phi_j(1) = 1$ so that

(i) there are constants $c_{1,\phi_j}, c_{2,\phi_j} > 0$ and $0 < \beta_{1,\phi_j} \le \beta_{2,\phi_j} < 2$ such that

$$c_{1,\phi_j}\Big(\frac{R}{r}\Big)^{\beta_{1,\phi_j}} \le \frac{\phi_j(R)}{\phi_j(r)} \ \le\ c_{2,\phi_j}\Big(\frac{R}{r}\Big)^{\beta_{2,\phi_j}} \quad \text{for all } 0 < r \le R \le 1; \tag{4.1}$$

(ii) there are constants $c^*_{1,\phi_j}, c^*_{2,\phi_j} > 0$ and $1 < \beta^*_{1,\phi_j} \le \beta^*_{2,\phi_j} < \infty$ such that

$$c^*_{1,\phi_j}\Big(\frac{R}{r}\Big)^{\beta^*_{1,\phi_j}} \le \frac{\phi_j(R)}{\phi_j(r)} \ \le\ c^*_{2,\phi_j}\Big(\frac{R}{r}\Big)^{\beta^*_{2,\phi_j}} \quad \text{for all } 1 \le r \le R. \tag{4.2}$$

Let $(\mathcal{E}, \mathcal{F})$ be the regular symmetric pure jump Dirichlet form on $L^2(\mathbb{R}^d; dx)$ having above $J(x, y)$ as its jumping kernel, where $\mathcal{F} = \{f \in L^2(\mathbb{R}^d; dx) : \mathcal{E}(f, f) < \infty\}$. Define

$$\phi(r) := \mathbf{1}_{[0,1]}(r)\phi_j(r) + \mathbf{1}_{(1,\infty)}(r)\phi_c(r),$$

where

$$\phi_c(r) = \begin{cases} r^2, & r \in (0, 1], \\ \frac{\Phi(r)}{\Phi(1)}, & r \in [1, \infty) \end{cases}$$

and

$$\Phi(r) = \frac{r^2}{2\int_0^r s/\phi_j(s)\, ds}, \quad r > 0.$$

It is shown in [**CKW5**, Example 5.2] that $\mathrm{HK}(\phi_j, \phi_c)$ holds for $(\mathcal{E}, \mathcal{F})$ by Theorem 2.7.

The above assertion improves [**BKKL1**, Theorem 1.4 and Corollary 1.5], in which an extra condition that $\beta_{1,\phi_j} > 1$ is required. We also note that in this example, $\phi_c(r)$ does not need to be comparable to the quadratic function $r \mapsto r^2$ for all $r > 0$. Nevertheless, we can treat symmetric pure jump forms with the growth order of $\phi_j(r)$ not necessarily strictly less than 2, as in [**BKKL1**]. By [**BGT**, Corollaries 2.6.2 and 2.6.4], $\phi_c(r)$ and $\phi_j(r)$ are comparable on $[1, \infty)$ if and only if $\beta^*_{2,\beta_j} < 2$, in which case, the heat kernel estimate $\mathrm{HK}(\phi_j, \phi_c)$ is reduced to $\mathrm{HK}(\phi_j)$ in [**CKW1**] (see Remark 2.9(i)). Therefore, the "diffusive scaling" appears in $\mathrm{HK}(\phi_j, \phi_c)$ only when $\beta^*_{2,\phi_j} \ge 2$. For example, when $\phi_j(r) = r^\alpha \vee r^2$ for all $r > 0$ with $\alpha \in (0, 2)$, we can take

$$\phi_c(r) := r^2\mathbf{1}_{\{0\le r\le 1\}} + \frac{r^2}{\log(e-1+r)}\mathbf{1}_{\{r>1\}}$$

and so

$$\phi(r) = r^\alpha \mathbf{1}_{\{0\le r\le 1\}} + \frac{r^2}{\log(e-1+r)}\mathbf{1}_{\{r>1\}}.$$

It holds that for any $t>1$ and $x,y\in\mathbb{R}^d$,

$$p^{(c)}(t,x,y) \asymp \frac{1}{V(x,(t\log(1+t))^{1/2})}\exp\left(-\frac{\rho(x,y)^2}{t\log(1+t/\rho(x,y))}\right).$$

This does not belong to the so-called (sub)-Gaussian estimates. Indeed long time behavior of the process is super-diffusive, and its heat kernel estimates are "super-Gaussian".

Our second example is to illustrate our stable characterization of parabolic Harnack inequalities. The assertion of the example below is a counterpart of [**CKK1**, Theorem 1.4], which is concerned with a local version of parabolic Harnack inequalities. One can use the assertion below to recover the parabolic Harnack inequalities for a large class of symmetric jump processes studied in [**CKK2**].

Example 4.2. Let $M=\mathbb{R}^d$ and $\mu(dx)=dx$. Consider a non-negative symmetric function $J(x,y)$ on $\mathbb{R}^d\times\mathbb{R}^d$ such that

$$J(x,y) \simeq \frac{1}{|x-y|^d\phi_j(|x-y|)}, \qquad |x-y|\le 1,$$
$$J(x,y) \preceq \frac{1}{|x-y|^{d+2}}, \qquad |x-y|>1$$

and

$$\sup_{x\in\mathbb{R}^d}\int_{\{|x-y|\ge 1\}}|x-y|^2 J(x,y)\,dy<\infty. \tag{4.3}$$

Here, ϕ_j is a strictly increasing continuous function with $\phi_j(0)=0$ and $\phi_j(1)=1$ so that (4.1) holds with $0<\beta_{1,\phi_j}\le\beta_{2,\phi_j}<2$.

Let $(\mathcal{E},\mathcal{F})$ be the regular symmetric pure jump Dirichlet form on $L^2(\mathbb{R}^d;dx)$ having the above $J(x,y)$ as its jumping kernel, where $\mathcal{F}=\{f\in L^2(\mathbb{R}^d;dx): \mathcal{E}(f,f)<\infty\}$. It can be shown that that PHI(ϕ) holds with $\phi(r)=\phi_j(r)\mathbf{1}_{\{0\le r\le 1\}}+r^2\mathbf{1}_{\{r>1\}}$, if and only if UJS holds for the jumping kernel $J(x,y)$ given above.

References

[AB] S. Andres and M.T. Barlow. Energy inequalities for cutoff-functions and some applications. *J. Reine Angew. Math.* **699** (2015), 183–215.

[Ar] D.G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.* **73** (1967), 890–896.

[BKKL1] J. Bae, J. Kang, P. Kim and J. Lee. Heat kernel estimates for symmetric jump processes with mixed polynomial growths. *Ann. Probab.* **47** (2019), 2830–2868.

[BKKL2] J. Bae, J. Kang, P. Kim and J. Lee. Heat kernel estimates and their stabilities for symmetric jump processes with general mixed polynomial growths on metric measure spaces, `arXiv:1904.10189`.

[B] M.T. Barlow. Diffusions on fractals. In: *Lectures on Probability Theory and Statistics, Ecole d'Éte de Probabilités de Saint-Flour XXV - 1995*, 1–121. Lect. Notes Math. **1690**, Springer 1998.

[BB] M.T. Barlow and R.F. Bass. Stability of parabolic Harnack inequalities. *Trans. Amer. Math. Soc.* **356** (2003), 1501–1533.

[BBK1] M.T. Barlow, R.F. Bass and T. Kumagai. Stability of parabolic Harnack inequalities on metric measure spaces. *J. Math. Soc. Japan* **58** (2006), 485–519.

[BBK2] M.T. Barlow, R. F. Bass and T. Kumagai. Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps. *Math. Z.* **261** (2009), 297–320.

[BL] R.F. Bass and D. A. Levin. Transition probabilities for symmetric jump processes. *Trans. Amer. Math. Soc.* **354** (2002), 2933–2953.

[BGT] N.H. Bingham, C.M. Goldie and J.L. Teugels. *Regular Variation*, Encyclopedia of Mathematics and Its Applications **27**, Cambridge Univ. Press, Cambridge, 1987.

[BG] R.M. Blumenthal and R.K. Getoor. Some theorems on stable processes. *Trans. Amer. Math. Soc.* **95**, (1960), 263–273.

[CF] Z.-Q. Chen and M. Fukushima. *Symmetric Markov Processes, Time Change, and Boundary Theory*. Princeton Univ. Press, Princeton and Oxford, 2012.

[CKK1] Z.-Q. Chen, P. Kim and T. Kumagai. On heat kernel estimates and parabolic Harnack inequality for jump processes on metric measure spaces. *Acta Math. Sin. (Engl. Ser.)* **25** (2009), 1067–1086.

[CKK2] Z.-Q. Chen, P. Kim and T. Kumagai. Global heat kernel estimates for symmetric jump processes. *Trans. Amer. Math. Soc.* **363** (2011), 5021–5055.

[CK] 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.

[CKW1] Z.-Q. Chen, T. Kumagai and J. Wang. Stability of heat kernel estimates for symmetric non-local Dirichlet forms. To appear in *Memoirs Amer. Math. Soc.*, available at `arXiv:1604.04035`.

[CKW2] Z.-Q. Chen, T. Kumagai and J. Wang. Stability of parabolic Harnack inequalities for symmetric non-local Dirichlet forms. To appear in *J. European Math. Soc.*, available at `arXiv:1609.07594`.

[CKW3] Z.-Q. Chen, T. Kumagai and J. Wang. Elliptic Harnack inequalities for symmetric non-local Dirichlet forms. *J. Math. Pures Appl.* **125** (2019), 1–42.

[CKW4] Z.-Q. Chen, T. Kumagai and J. Wang. Heat kernel and parabolic Harnack inequalities for symmetric Dirichlet forms. To appear in *Advances in Math.*, available at `arXiv:1908.07650`.

[CKW5] Z.-Q. Chen, T. Kumagai and J. Wang. Heat kernel estimates for general symmetric pure jump Dirichlet forms. Preprint 2019, available at `arXiv:1908.07655`.

[FOT] M. Fukushima, Y. Oshima and M. Takeda. *Dirichlet Forms and Symmetric Markov Processes*. de Gruyter, Berlin, 2nd rev. and ext. ed., 2011.

[GHH] A. Grigor'yan, E. Hu and J. Hu. Two-sided estimates of heat kernels of jump type Dirichlet forms. *Advances in Math.* **330** (2018), 433–515.

[GT] A. Grigor'yan and A. Telcs. Two-sided estimates of heat kernels on metric measure spaces. *Ann. Probab.* **40** (2012), 1212–1284.

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

[Ku] T. Kumagai. Random walks on disordered media and their scaling limits. Lect. Notes Math. **2101**. École d'Été de probabilités de Saint-Flour. Springer 2014.

[LM] G. Liu and M. Murugan. Parabolic Harnack inequality implies the existence of jump kernel, Preprint 2020, available at `arXiv:2005.12450`

[MS1] M. Murugan and L. Saloff-Coste. Transition probability estimates for long range random walks. *New York Journal of Mathematics* **21** (2015), 723–757.

[MS2] M. Murugan and L. Saloff-Coste. Heat kernel estimates for anomalous heavy-tailed random walks. *Ann. Inst. Henri Poincaré-Probab. Stat.* **55** (2019), 697–719.

Department of Mathematics, University of Washington, Seattle, WA 98195, USA
E-mail address: `zqchen@uw.edu`

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

College of Mathematics and Informatics & Fujian Key Laboratory of Mathematical Analysis and Applications (FJKLMAA) & Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University, 350007, Fuzhou, P.R. China
E-mail address: `jianwang@fjnu.edu.cn`