

# Heat kernel upper bounds for symmetric Markov semigroups

Zhen-Qing Chen   Panki Kim   Takashi Kumagai   Jian Wang

October 13, 2020

## Abstract

It is well known that Nash-type inequalities for symmetric Dirichlet forms are equivalent to on-diagonal heat kernel upper bounds for the associated symmetric Markov semigroups. In this paper, we show that both imply (and hence are equivalent to) off-diagonal heat kernel upper bounds under some mild assumptions. Our approach is based on a new generalized Davies' method. Our results extend that of [CKS] for Nash-type inequalities with power order considerably and also extend that of [G1] for second order differential operators on a complete non-compact manifold.

**Keywords:** Heat kernel; Nash-type inequality; Dirichlet form; Markov process; Davies' method

**AMS 2020 Mathematics Subject Classification:** Primary 60J35, 35K08; Secondary 31C25, 39B62, 47D07, 60J25.

## 1 Introduction

Analysis on heat kernel has been an active research area for a long time in analysis, geometry and probability. A central topic of the area is to obtain global quantitative bounds on the heat kernel. Nash-type inequalities can effectively characterize upper bounds of the operator norms of semigroups from  $L^1$ -space to  $L^\infty$ -space (i.e., on-diagonal heat kernel upper bounds). For example, let  $(P_t)_{t \geq 0}$  be a symmetric Markov semigroup on  $L^2(E; m)$  associated with a symmetric regular Dirichlet form  $(\mathcal{E}, \mathcal{F})$  on  $L^2(E; m)$ , where  $E$  is a locally compact separable metric space and  $m$  is a  $\sigma$ -finite Radon measure on  $E$  with full support. It is well known that (see [CKS, Theorem 2.1])

$$\|P_t\|_{1 \rightarrow \infty} \leq \varphi(t)e^{\delta t}, \quad t > 0 \tag{1.1}$$

with  $\varphi(t) = ct^{-\nu/2}$  for some constants  $c, \nu > 0$  and  $\delta \geq 0$ , is equivalent to the following Nash-type inequality

$$\|f\|_2^{2+4/\nu} \leq C(\mathcal{E}(f, f) + \delta\|f\|_2^2)\|f\|_1^{4/\nu} \quad \text{for } f \in \mathcal{F}. \tag{1.2}$$

Here and in what follows,  $\|\cdot\|_p$  denotes the  $L^p$ -norm of the space  $L^p(E; m)$  for  $p \in [1, \infty]$ , and  $\|T\|_{p \rightarrow q}$  denotes the operator norm of a bounded operator  $T : L^p(E; m) \rightarrow L^q(E; m)$  for  $p, q \in [1, \infty]$ . For

$\varphi(t) = ct^{-\nu/2}$ , we refer the reader to [N] for the original proof that (1.2) implies (1.1) when  $\delta = 0$ , to [V] for the connection between the estimate (1.1) for  $\nu > 2$  and Sobolev inequalities, and to [BCLS] for the equivalence between Sobolev-type inequalities and Nash-type inequalities. When  $\delta = 0$  and  $\varphi(t)$  is a strictly decreasing differential bijection on  $\mathbb{R}_+ := (0, \infty)$  so that the derivative of  $-\log \varphi(t)$  is of polynomial growth, Coulhon [Co, Theorem II.5] showed that (1.1) is equivalent to the following generalized Nash-type inequality

$$\theta(\|f\|_2^2) \leq \mathcal{E}(f, f) \quad \text{for } f \in \mathcal{F} \text{ with } \|f\|_1 \leq 1, \quad (1.3)$$

where  $\theta(r)$  is a positive function on  $\mathbb{R}_+$  determined by  $\varphi(r)$ ; see Theorem 1.2 or Section 3 for details.

The main goal of this paper is to establish the equivalences among Nash-type inequalities (1.3), on-diagonal heat kernel upper bounds (1.1) and off-diagonal heat kernel upper bounds for a large class of decreasing functions  $\varphi(t)$  and for general symmetric Markov processes which may have discontinuous sample paths. To state our main results, we need to introduce a class of regular functions which will serve as the rate function  $\varphi(t)$  in (1.1).

**Definition 1.1.** A strictly decreasing  $C^1$ -function  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is said to be *regular*, if the following hold:

- (i)  $\varphi(0) := \lim_{r \rightarrow 0} \varphi(r) = \infty$  and  $\varphi(\infty) := \lim_{r \rightarrow \infty} \varphi(r) = 0$ ;
- (ii) there exist a constant  $c \geq 1$  and a decreasing function  $N(t) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that, with  $M(t) := -\log \varphi(t)$ ,  $c^{-1}N(t) \leq M'(t) \leq cN(t)$  on  $\mathbb{R}_+$  and there is a constant  $c_0 > 0$  so that  $M'(t) \leq c_0 M'(at)$  for all  $t > 0$  and  $a \in [1, 2]$ ;
- (iii) for any  $\lambda > 2$  large enough, there exist positive constants  $c(\lambda)$  and  $C(\lambda)$  such that

$$\prod_{k=1}^{\infty} \left( \varphi((\lambda - 1)\lambda^{-(k+1)}t) \right)^{2^{-k}} \leq C(\lambda)\varphi(c(\lambda)t), \quad t > 0.$$

We denote this class of functions by  $\mathcal{R}$ .

Typical examples for regular functions on  $\mathbb{R}_+$  are  $\varphi(r) = r^{-\nu/2}$  with  $\nu > 0$ , or  $\varphi(r) = r^{-\nu/2}$  with  $\nu > 0$  for small  $r > 0$ , and  $\varphi(r) = e^{-r^\alpha}$  with  $\alpha \in (0, 1]$  for large  $r > 0$ . Indeed, for any decreasing function  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $\varphi(0) = \infty$ ,  $\varphi(\infty) = 0$  and that  $1/\varphi$  has the doubling property (i.e., there is a constant  $c_1 \geq 1$  such that  $1/\varphi(2r) \leq c_1/\varphi(r)$  for all  $r > 0$ ), we can find some  $\bar{\varphi} \in \mathcal{R}$  and a constant  $c_2 \geq 1$  so that  $c_2^{-1}\bar{\varphi}(r) \leq \varphi(r) \leq c_2\bar{\varphi}(r)$  for all  $r \in \mathbb{R}_+$ ; see Proposition 2.8.

The following is a special case of the main result, Theorem 4.4, of this paper.

**Theorem 1.2.** *Suppose that  $(\mathcal{E}, \mathcal{F})$  is a symmetric regular Dirichlet form on  $L^2(E; m)$  having no killing inside  $E$  and  $(P_t)_{t \geq 0}$  is its associated semigroup. Let  $\varphi \in \mathcal{R}$  and  $\delta \geq 0$ . The following statements are equivalent.*

(i) The following Nash-type inequality holds:

$$c_1\theta(c_2\|f\|_2^2) \leq \mathcal{E}(f, f) + \delta\|f\|_2^2 \quad \text{for } f \in \mathcal{F} \text{ with } \|f\|_1 \leq 1, \quad (1.4)$$

where  $\theta(r) = -\varphi'(\varphi^{-1}(r))$  and  $c_1, c_2$  are positive constants.

(ii) The semigroup  $(P_t)_{t \geq 0}$  is ultracontractive in the sense that there are constants  $c_3, c_4 > 0$  such that for all  $t > 0$ ,

$$\|P_t\|_{1 \rightarrow \infty} \leq c_3\varphi(c_4t)e^{\delta t}. \quad (1.5)$$

(iii) There are a properly exceptional set  $\mathcal{N} \subset E$  and a heat kernel  $p(t, x, y)$  associated with the semigroup  $(P_t)_{t \geq 0}$  and defined on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$  such that for any  $\varepsilon \in (0, 1)$  there are constants  $C_\varepsilon, c_\varepsilon > 0$  so that for all  $t > 0$  and  $x, y \in E \setminus \mathcal{N}$ ,

$$p(t, x, y) \leq C_\varepsilon\varphi(c_\varepsilon t)e^{\delta t} \exp(-|\psi(y) - \psi(x)| + (1 + \varepsilon)\Lambda(\psi)^2t), \quad (1.6)$$

where

$$\Lambda(\psi)^2 := \max \left\{ \left\| \frac{de^{-2\psi}\Gamma(e^\psi, e^\psi)}{dm} \right\|_\infty, \left\| \frac{de^{2\psi}\Gamma(e^{-\psi}, e^{-\psi})}{dm} \right\|_\infty \right\} < \infty;$$

equivalently,

$$p(t, x, y) \leq C_\varepsilon\varphi(c_\varepsilon t)e^{\delta t} \exp\left(-\frac{d_{\mathcal{E}}(x, y)^2}{4(1 + \varepsilon)t}\right), \quad (1.7)$$

where

$$d_{\mathcal{E}}(x, y) := \sup \{ \psi(x) - \psi(y) : \psi \in \mathcal{F} \cap C_b(E) \text{ with } \Lambda(\psi) \leq 1 \}.$$

In the statement above,  $\Gamma(f, f)$  is the energy measure of  $f$  for  $\mathcal{E}$  when  $f \in \mathcal{F}$ . Precise definitions of energy measure, properly exceptional set, Dirichlet form having no killings inside  $E$ , and heat kernel can be found in Section 2.1.

As mentioned above, the equivalence of (i) and (ii) in Theorem 1.2 is known; see [Co, Theorem II.5] for the case that  $\delta = 0$ . Clearly (iii) implies (ii). So the new contribution of Theorem 1.2 is that (i) implies (iii). When  $\varphi(t) = ct^{-\nu/2}$  for some constants  $c, \nu > 0$  and  $\delta \geq 0$ , that (i) gives (iii) is [CKS, Theorem 3.25], which is the main result of [CKS]. (From it, one can easily extend the result to the case when  $\delta = 0$  and  $\varphi(t) = c(t^{-\nu/2}\mathbb{1}_{\{t \leq 1\}} + t^{-\mu/2}\mathbb{1}_{\{t > 1\}})$  for some constants  $c > 0$  and  $0 < \mu \leq \nu < \infty$ ; see [CKS, Corollary 3.28].) The approach of [CKS] makes use of the condition  $\varphi(t) = ct^{-\nu/2}$  (or equivalently, (1.2)) in an essential way (see, e.g., [CKS, Lemma (3.21)]), which does not seem to work for general  $\varphi \in \mathcal{R}$ . We note that there are quite many cases that the correct on-diagonal heat kernel upper bounds are not of the form  $ct^{-\nu/2}e^{\delta t}$ . For example, on-diagonal estimates for the heat kernel on a non-compact manifold with bounded geometry can be of the form  $c_1t/(\log t)^{1/\alpha}$  for large time, when the volume of the ball grows at least with the rate  $c_2e^{c_3r^\alpha}$  with  $\alpha \in (0, 1]$  for large radius  $r$  (see [BCG] or Example 5.3 below for more details); on-diagonal heat kernel upper bounds for mixtures of symmetric stable-like processes on  $\mathbb{R}^d$  are of the form

$c_4(\Phi^{-1}(t))^d$  for some strictly increasing weighted function  $\Phi$  which satisfies doubling and reverse doubling properties (see [CK1, CK2, CKW1] or Example 5.5 below).

There have been quite many efforts in obtaining on off-diagonal heat kernel estimates under more general settings with more general on-diagonal bounds. Several authors have studied Gaussian off-diagonal bounds for symmetric diffusions by using functional inequalities. Theorem 1.2 is related in spirit to a beautiful result of Grigor'yan [G1, Theorem 1.1] that, under suitable conditions, an isoperimetric-type inequality (that is, a Faber-Krahn type inequality), an on-diagonal upper bound, and an off-diagonal Gaussian upper estimate are equivalent up to some constant multiplies for Brownian motion on a complete non-compact manifold. (Note that these isoperimetric inequalities imply Nash-type inequalities or Sobolev inequalities, but the converse is not true.) The approach of [G1] relies crucially on some constructions that are specific to the setting of second order differential operators, and uses some technical tools like a mean-value type theorem, which do not seem to be easily available in general setting as in our paper.

For strongly local Dirichlet forms, by the Leibniz rule,  $\Lambda(\psi)^2 = d\Gamma(\psi, \psi)/dm$ , and so we have the following statement.

**Corollary 1.3.** *Let  $(\mathcal{E}, \mathcal{F})$  be a strongly local, symmetric regular Dirichlet form on  $L^2(E; m)$  having no killings inside  $E$ , and  $(P_t)_{t \geq 0}$  its associated semigroup. Let  $\varphi \in \mathcal{R}$  and  $\delta \geq 0$ . Then (i) and (ii) in Theorem 1.2 are equivalent to the following*

(iii)' *There are a properly exceptional set  $\mathcal{N} \subset E$  and a heat kernel  $p(t, x, y)$  associated with the semigroup  $(P_t)_{t \geq 0}$  and defined on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$  such that for any  $\varepsilon \in (0, 1)$ , there are positive constants  $c_\varepsilon$  and  $C_\varepsilon$  so that for  $t > 0$  and  $x, y \in E \setminus \mathcal{N}$ ,*

$$p(t, x, y) \leq C_\varepsilon \varphi(c_\varepsilon t) e^{\delta t} \exp\left(-\frac{d_{\mathcal{E}}^*(x, y)^2}{4(1 + \varepsilon)t}\right), \quad (1.8)$$

where  $d_{\mathcal{E}}^*(x, y)$  is the intrinsic distance induced by the Dirichlet form  $(\mathcal{E}, \mathcal{F})$ , i.e.,

$$d_{\mathcal{E}}^*(x, y) := \sup\{\psi(x) - \psi(y) : \psi \in \mathcal{F} \cap C_b(E) \text{ with } d\Gamma(\psi, \psi)/dm \leq 1\}.$$

We note that in general the constant  $\varepsilon$  can not be taken to be zero in the exponent of (1.8). See [Mo, Example 3.1] (especially the last asymptotic estimate there) for a counterexample, and [CS] and the references therein for sharp pointwise Gaussian estimates.

Our approach adopts Davies' idea of obtaining off-diagonal heat kernel upper bounds through establishing  $L^1 \rightarrow L^\infty$ -bounds for some perturbed semigroups. The main novel contributions of this paper are

- We directly apply the generalized Nash-type inequalities to the perturbed semigroups by a localization trick. We do not use the approach in [CKS, Lemma (3.21)] that is based on differential inequalities, which seems to be specific to  $\varphi(t)$  of being  $ct^{-\nu/2}$  considered there.

- We introduce a new iteration argument to obtain the  $L^1 \rightarrow L^\infty$ -bounds for perturbed semi-groups, which works for general regular decay rate functions  $\varphi \in \mathcal{R}$ .

In addition, our approach of establishing that (i) implies (iii) of Theorem 1.2 is much simpler than that of [CKS] even in the special case of  $\varphi(t) = ct^{-\nu/2}$  for some  $c, \nu > 0$ . As mentioned in [G1], there exist many important classes of manifolds on which the heat kernel of Brownian motion decreases faster than polynomially as  $t \rightarrow \infty$ . Since Corollary 1.3 holds for all regular functions, it can be used to deal with the case that the heat kernel has an exponential/subexponential decay for strongly local Dirichlet forms, including a large variety of manifolds mentioned in [G1].

Davies' method goes back to the paper [D1] by E.B. Davies, where explicit constants in the exponential term for Gaussian upper bounds associated with various second order elliptic or hypoelliptic operators on  $\mathbb{R}^d$  were established. The method of [D1] uses logarithmic Sobolev inequality:

$$\int f^2 \log \frac{f^2}{\|f\|_2^2} dm \leq r\mathcal{E}(f, f) + (\log \beta(r))\|f\|_2^2, \quad f \in \mathcal{F}, r > 0.$$

See [D2] for more details on this topic.

Compared to the method using logarithmic Sobolev inequalities, ours have some advantages. For example, our approach based on Nash-type inequalities (1.4) can produce off-diagonal heat kernel upper bounds with the optimal constant  $4 + \varepsilon$  for the  $d_{\mathcal{E}}(x, y)^2$  term in the exponent in (1.7) for general Dirichlet forms, which seems difficult to obtain through logarithmic Sobolev inequalities. (Cf., [Mi, Corollary 5.3], where  $\delta$  should be zero and  $\varphi(t)$  is of the form  $ct^{-\nu/2}\ell_0(t)$  for some positive function  $\ell_0$  on  $(0, \infty)$  that is slowly varying at both 0 and  $\infty$ .) On the other hand, as pointed out in [Co], there are a number of works concerning the relations between Nash-type inequalities and other functional inequalities (for example, isoperimetric inequalities), so it is possible to restate our main result in terms of other functional inequalities instead of Nash-type inequalities. Here are two examples. First, we give the following corollary to Theorem 1.2, which extends the main result of [G1] — [G1, Theorem 1.1] to general symmetric regular Dirichlet forms. Recall that for an open set  $D \subset E$ , the part Dirichlet form  $(\mathcal{E}, \mathcal{F}_D)$  in  $D$  of  $(\mathcal{E}, \mathcal{F})$  is given by

$$\mathcal{F}_D = \{u \in \mathcal{F} : u = 0 \text{ } \mathcal{E}\text{-quasi-everywhere on } D\}.$$

It is known (cf. [CF, FOT]) that  $(\mathcal{E}, \mathcal{F}_D)$  is a regular Dirichlet form on  $L^2(D; m)$  and it corresponds to the part process  $X^D$  of  $X$  associated with  $(\mathcal{E}, \mathcal{F})$  killed upon leaving  $D$ .

**Corollary 1.4.** *Suppose that  $(\mathcal{E}, \mathcal{F})$  is a symmetric regular Dirichlet form on  $L^2(E; m)$  having no killing inside  $E$  and that  $(P_t)_{t \geq 0}$  is its associated semigroup. Let  $\varphi \in \mathcal{R}$  and  $\Theta(t) = N(\varphi^{-1}(1/t))$ , where  $N(t) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a decreasing function given in Definition 1.1(ii). Then, (i), (ii) and (iii) in Theorem 1.2 with  $\delta = 0$  are equivalent to the following isoperimetric-type inequalities:*

- (iv) *There exist constants  $c_5, c_6 > 0$  such that for any pre-compact open set  $D \subset E$ ,*

$$\lambda_1(D) \geq c_5 \Theta(c_6 m(D)), \tag{1.9}$$

where  $\lambda_1(D)$  is the first eigenvalue of the Dirichlet form  $(\mathcal{E}, \mathcal{F}_D)$ , i.e.,

$$\lambda_1(D) = \inf\{\mathcal{E}(f, f) : f \in \mathcal{F} \cap C_c(D) \text{ with } \|f\|_2 = 1\}.$$

(v) There exist constants  $c_7, c_8 > 0$  such that for any pre-compact open set  $D \subset E$  and  $k \geq 1$ ,

$$\lambda_k(D) \geq c_7 \Theta(c_8 m(D)/k),$$

where  $\lambda_k(D)$  is the  $k^{\text{th}}$  eigenvalue of the Dirichlet form  $(\mathcal{E}, \mathcal{F}_D)$ .

Second example is that, though the proof of Theorem 1.2 uses Nash-type inequalities, it can easily be replaced by using the following super-Poincaré inequalities

$$\|f\|_2^2 \leq r\mathcal{E}(f, f) + \beta(r)\|f\|_1^2, \quad f \in \mathcal{F}, r > 0,$$

which obviously are simpler than logarithmic Sobolev inequalities mentioned above. See the proof of off-diagonal Dirichlet heat kernel upper bounds in Subsection 5.1. The reader is referred to [W] for more details on super-Poincaré inequalities.

We now comment on the assumption of  $\Lambda(\psi)^2 < \infty$  in Theorem 1.2. In some contexts (for example, anomalous Markov processes such as diffusions or continuous time random walks on fractals with walk dimension larger than 2), the energy measure  $\Gamma(u, u)$  is singular with respect to the measure  $m$  unless  $u$  is a constant, and so in this case the off-diagonal upper bounds (1.6) in Theorem 1.2 degenerate into the on-diagonal upper bounds (1.5). Indeed, (i) implies (iii) in Theorem 1.2 is a special case of a more general result, Theorem 4.4, of this paper. We emphasize that Theorem 4.4 does not have this absolute continuity assumption and is applicable to the singular contexts mentioned above.

We like to mention another beautiful result due to Grigor'yan [G2] that gives a two-point off-diagonal Gaussian upper bound of  $p(t, x, y)$  for Brownian motions on a smooth connected manifold from the on-diagonal upper bounds at two points  $x$  and  $y$  only rather than at all points. The approach of [G2] is based on an integral maximum principle, which is quite different from the functional-inequalities adopted in the present paper (because the functional-inequalities technique yield a uniform upper bound for heat kernel). The approach of [G2] has been extended to discrete time random walks on graphs in [CGZ] and to continuous time random walks in [F, Ch].

We note that an integral maximum principle was also established in [G1, Lemma 4.3] via the mean-value type theorem. The advantage of this approach based on the integral maximum principle is that it can handle space-inhomogeneous heat kernel estimates; for example, the heat kernel estimate on a manifold  $(M, \rho)$  of a non-negative Ricci curvature has the following upper bound

$$p(t, x, y) \leq \frac{C_\varepsilon}{\sqrt{v(x, \sqrt{t})v(y, \sqrt{t})}} \exp\left(-\frac{\rho(x, y)^2}{(4 + \varepsilon)t}\right), \quad x, y \in M, \varepsilon > 0,$$

where  $v(x, r)$  is the Riemannian volume of the geodesic ball with the centre  $x \in M$  and of the radius  $r$ . See [G1, Section 5] for more details. By slightly modifying the proof of Theorem 1.2), we establish in Theorem 5.1 a possibly space-inhomogeneous off-diagonal upper bound estimate for Dirichlet heat kernel from its possibly space-dependent on-diagonal upper bound. This result will then be used in [CKKW] to study two-sided heat kernel estimates for reflected diffusions with jumps on metric measure spaces that satisfy general volume doubling condition.

The rest of this paper is organized as follows. In Section 2, we briefly describe the setting of our paper and give some definitions for regular properties of functions. In particular, we prove that for any decreasing function  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $1/\varphi$  has the doubling property, there is  $\bar{\varphi} \in \mathcal{R}$  which is comparable to  $\varphi$  by neglecting constants. In Section 3, we mainly recall the results from [Co] about the characterizations of Nash-type inequalities and on-diagonal heat kernel bounds. Section 4 is the main part of our paper, and is devoted to off-diagonal heat kernel bounds. We introduce a new generalized Davies' method and obtain off-diagonal heat kernel bounds from Nash-type inequalities in a very general setting. As an application of our approach, we give in Theorem 5.1 an off-diagonal Dirichlet heat kernel estimates from its possibly space-dependent on-diagonal upper bound. Several examples are given in Subsection 5.2 to illustrate the strength of our main results.

**Notations:** We use “:=” as a way of definition. For  $a, b \in \mathbb{R}$ ,  $a \wedge b := \min\{a, b\}$  and  $a \vee b := \max\{a, b\}$ . We denote  $\mathbb{R}_+ := (0, \infty)$ . For any function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $f(0) := \lim_{r \rightarrow 0} f(r)$  and  $f(\infty) := \lim_{r \rightarrow \infty} f(r)$ . For any open set  $B \subset E$ , denote the set of continuous functions on  $B$  with compact support by  $C_c(B)$ . For a function space  $\mathcal{H}$  on  $E$ , we let  $\mathcal{H}_b := \{f \in \mathcal{H} : f \text{ is bounded}\}$ . The constants  $c, C$  and  $c_i$ , whose exact values are unimportant, are changed in each statement and proof. In particular, we write  $c := c(C_0, C, \eta, \dots)$  which means that the constant  $c$  depends on  $C_0, C, \eta$  and  $\dots$  only.

## 2 Preliminaries

### 2.1 Dirichlet form and heat kernel

In this subsection, we first describe the setting of this paper. Let  $(E, \rho)$  be a locally compact separable metric space, and  $m$  a  $\sigma$ -finite Radon measure on  $E$  with full support. Suppose that  $(\mathcal{E}, \mathcal{F})$  is a regular symmetric Dirichlet form on  $L^2(E; m)$ . This is,

- (i)  $\mathcal{F}$  is a linear subspace of  $L^2(E; m)$ , and  $\mathcal{E}$  is a symmetric bilinear form on  $\mathcal{F} \times \mathcal{F}$  so that  $\mathcal{E}(u, u) \geq 0$  for every  $u \in \mathcal{F}$ ;
- (ii)  $\mathcal{F}$  is a Hilbert space with inner product  $\mathcal{E}_1(u, v) := \mathcal{E}(u, v) + \int_E u(x)v(x) m(dx)$ ;
- (iii) (Markovian property) for any  $u \in \mathcal{F}$ ,  $v := (0 \vee u) \wedge 1 \in \mathcal{F}$  and  $\mathcal{E}(v, v) \leq \mathcal{E}(u, u)$ ;

(iv) (regularity)  $\mathcal{F} \cap C_c(E)$  is dense both in  $(\mathcal{F}, \sqrt{\mathcal{E}_1})$  and in  $(C_c(E), \|\cdot\|_\infty)$ .

We say the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  is strongly local, if for any  $u, v \in \mathcal{F}$  with compact support such that  $u$  is constant on a neighbourhood of the support of  $v$ ,  $\mathcal{E}(u, v) = 0$ .

By the Beurling-Deny formula, any regular symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  admits the following representation for any  $u, v \in \mathcal{F}$ ,

$$\mathcal{E}(u, v) = \mathcal{E}^{(c)}(u, v) + \frac{1}{2} \iint_{E \times E \setminus \text{diag}} (u(x) - u(y))(v(x) - v(y)) J(dx, dy) + \int_E u(x)v(x) \kappa(dx).$$

Here  $\mathcal{E}^{(c)}$  is a symmetric form possessing the strongly local property,  $J(dx, dy)$  is a symmetric Radon measure on  $E \times E \setminus \text{diag}$ , where  $\text{diag}$  denotes the diagonal set, and  $\kappa(dx)$  is a Radon measure on  $E$ . See [FOT, Theorem 4.5.2] or [CF, (4.3.1)]. We call  $J(dx, dy)$  and  $\kappa(dx)$  the jumping measure and the killing measure of  $(\mathcal{E}, \mathcal{F})$ , respectively. *Throughout this paper, we assume the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  admits no killing inside  $E$ ; that is,  $\kappa(dx) = 0$ .*

Let  $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$  be the  $L^2$ -infinitesimal generator of  $(\mathcal{E}, \mathcal{F})$ ; that is,  $u \in \mathcal{D}(\mathcal{L})$ , if and only if  $u \in \mathcal{F}$  and there is some function  $f \in L^2(E; m)$  so that

$$\mathcal{E}(u, v) = - \int_E f(x)v(x) m(dx) \quad \text{for every } v \in \mathcal{F}.$$

In this case, we write  $f$  as  $\mathcal{L}u$ . It is known that there exist a Borel exceptional subset  $\mathcal{N}$  of  $E$  and an  $m$ -symmetric Hunt process  $X = \{X_t, t \geq 0; \mathbb{P}_x, x \in E \setminus \mathcal{N}\}$  on  $E$  associated with the regular symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  on  $L^2(E; m)$  in the sense that the transition semigroup  $(P_t)_{t \geq 0}$  defined by

$$P_t f(x) = \mathbb{E}_x[f(X_t)], \quad x \in E \setminus \mathcal{N}, \quad t \geq 0$$

is a strongly continuous symmetric semigroup on  $L^2(E; m)$  whose infinitesimal generator is the  $L^2$ -infinitesimal generator  $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$  of  $(\mathcal{E}, \mathcal{F})$  on  $L^2(E; m)$ . A set  $\mathcal{N}$  is said to be properly exceptional for the Hunt process  $X$ , if  $m(\mathcal{N}) = 0$  and  $\mathbb{P}_x(X_t \in \mathcal{N} \text{ for some } t > 0) = 0$  for all  $x \in E \setminus \mathcal{N}$ . It is known that a subset has zero  $\mathcal{E}_1$ -capacity, if and only if it is contained in a Borel properly exceptional set. The Hunt process  $X$  is unique up to a set of zero  $\mathcal{E}_1$ -capacity. The transition semigroup  $(P_t)_{t \geq 0}$  can be extended to be a strongly continuous contraction semigroup in  $L^p(E; m)$  for every  $p \in [1, \infty]$ . Every function in  $\mathcal{F}$  admits a quasi-continuous version, and we will always use its quasi-continuous representation. We refer the detailed theory of symmetric Dirichlet forms to [CF, FOT].

**Definition 2.1.** A measurable function  $p(t, x, y) : (0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N}) \rightarrow [0, \infty)$  is called the *heat kernel* associated with the transition semigroup  $(P_t)_{t \geq 0}$ , if  $\mathcal{N}$  is of zero  $\mathcal{E}_1$ -capacity and the



following hold:

$$\begin{aligned} P_t f(x) &= \int p(t, x, y) f(y) m(dy) \quad \text{for all } x \in E \setminus \mathcal{N}, f \in L^1(E; m), \\ p(t, x, y) &= p(t, y, x) \quad \text{for all } t > 0, x, y \in E \setminus \mathcal{N}, \\ p(s+t, x, y) &= \int p(s, x, z) p(t, z, y) m(dz) \quad \text{for all } s, t > 0, x, y \in E \setminus \mathcal{N}. \end{aligned}$$

It was proved in [BBCK, Theorem 3.1] that, if there is a positive left continuous function  $\varphi(t)$  on  $\mathbb{R}_+$  so that

$$\|P_t f\|_\infty \leq \varphi(t) \|f\|_1 \quad \text{for } f \in L^1(E; m) \text{ and } t > 0,$$

then there exists a properly exceptional set  $\mathcal{N}$  such that the semigroup  $(P_t)_{t \geq 0}$  has a heat kernel  $p(t, x, y)$  defined on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$ . Note that we can always extend (without further mentioning)  $p(t, x, y)$  to all  $x, y \in E$  by setting  $p(t, x, y) = 0$  if  $x$  or  $y$  is in  $\mathcal{N}$ . So one can also extend the transition function

$$P(t, x, dy) = p(t, x, y) m(dy) = \mathbb{P}_x(X_t \in dy)$$

of  $X$  into  $(0, \infty) \times E \times \mathcal{B}(E)$ .

Recall that  $\mathcal{F}_b$  is the collection of bounded functions in  $\mathcal{F}$ . It is well known that for any  $u \in \mathcal{F}_b$ , there exists a unique positive Radon measure  $\Gamma(u, u)$  on  $E$  so that

$$\int_E v d\Gamma(u, u) = \mathcal{E}(u, uv) - \frac{1}{2} \mathcal{E}(u^2, v) \quad \text{for every } v \in \mathcal{F}_b. \quad (2.1)$$

Furthermore, for any  $u, v \in \mathcal{F}_b$ ,  $\Gamma(u, v)$  is defined by polarization. The Radon measure  $\Gamma(u, u)$  can be uniquely extended to any  $u \in \mathcal{F}$  as the increasing limit of  $\Gamma(u_n, u_n)$ , where  $u_n := ((-n) \vee u) \wedge n$ . The measure  $\Gamma(u, u)$  is called the *energy measure* of  $u$  for  $(\mathcal{E}, \mathcal{F})$ . See [FOT, Chapter 3.2] or [CF, Chapter 4.3] for more details. When  $\Gamma(u, u)$  is absolutely continuous with respect to the measure  $m$ , its Radon-Nikodym is also called the *square field operator* or the *carré du champ* in some literature. Indeed, we can define  $d\Gamma(u, u)$ , for any  $u \in \mathcal{F}_b$ , by the weak limit of the measure

$$d\Gamma_t(u, u) = \frac{1}{2t} \left( \int_E (u(x) - u(y))^2 P(t, x, dy) \right) m(dx).$$

Then by the fact that

$$(u(x) - u(y))(u(x)v(x) - u(y)v(y)) - u(x)(u(x) - u(y))(v(x) - v(y)) = (u(x) - u(y))^2 v(y)$$

and the symmetry of  $P(t, x, dy) m(dx)$ , we can easily verify that (2.1) holds. Note that, since the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  admits no killings inside  $E$ , for any  $u \in \mathcal{F}_b$ ,

$$\lim_{t \rightarrow 0} \frac{1}{t} \int_E u(x)^2 (1 - P_t 1(x)) m(dx) = 0.$$

We take this opportunity to point out that no killings inside  $E$  condition needs to be imposed in [CKS]; otherwise, [CKS, (1.4) and (3.4)] do not hold. Similar to the remark before [CKS, Theorem (3.9)], we can extend the definition of  $\Gamma(u, u)$  from  $\mathcal{F}_b$  into  $\hat{\mathcal{F}}_b = \{u + c : u \in \mathcal{F}_b, c \in \mathbb{R}\}$  by setting  $\Gamma(u + c, u + c) = \Gamma(u, u)$  for  $u \in \mathcal{F}_b$  and  $c \in \mathbb{R}$ . In particular,  $\Gamma(e^\psi, e^\psi)$  is well defined for every  $\psi \in \mathcal{F}_b$  as  $e^\psi - 1 \in \mathcal{F}_b$ . The reader is referred to [CF, Chapter 4.3] for the construction of the energy measure  $\Gamma(u, u)$  by using a probabilistic approach, and to [FOT, Chapter 3.2] using an analytic approach.

**Example 2.2.** Let  $(M, \rho)$  be a manifold with Riemannian volume measure  $\mu$ , and  $\mathcal{E}(u, u) = \frac{1}{2} \int_M |\nabla u|^2 \sigma \, d\mu$ , where  $\sigma > 0$  is a measurable function on  $M$  with  $\sigma, \sigma^{-1} \in L_{\text{loc}}^\infty(M; \mu)$ . Let  $dm = \sigma \, d\mu$ . Then,  $d\Gamma(u, u) = \frac{1}{2} |\nabla u|^2 \, dm$ .  $\square$

**Example 2.3.** Let  $(\mathcal{E}, \mathcal{F})$  be a symmetric  $\alpha$ -stable-like Dirichlet form on  $L^2(\mathbb{R}^d; dx)$  as follows

$$\begin{aligned} \mathcal{E}(u, v) &= \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \text{diag}} (u(x) - u(y))(v(x) - v(y)) \frac{c(x, y)}{|x - y|^{d+\alpha}} \, dx \, dy, \\ \mathcal{F} &= \{u \in L^2(\mathbb{R}^d; dx) : \mathcal{E}(u, u) < \infty\}, \end{aligned}$$

where  $\alpha \in (0, 2)$  and  $c_0^{-1} \leq c(x, y) = c(y, x) \leq c_0$  for some  $c_0 \geq 1$ . Then,

$$\Gamma(u, u)(dx) = \frac{1}{2} \left( \int_{\{y \in \mathbb{R}^d : y \neq x\}} (u(x) - u(y))^2 \frac{c(x, y)}{|x - y|^{d+\alpha}} \, dy \right) dx.$$

$\square$

## 2.2 Doubling function, regular function and condition (D)

**Definition 2.4.** An increasing function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is said to be a doubling function, if there is a constant  $c > 0$  such that

$$f(2r) \leq cf(r) \quad \text{for } r > 0.$$

The above is equivalent to that there are constants  $C \geq 1$  and  $\eta > 0$  so that

$$\frac{f(R)}{f(r)} \leq C \left( \frac{R}{r} \right)^\eta \quad \text{for any } 0 < r \leq R.$$

In this case, we also say that  $f$  has the doubling property.

The following definition is taken from [Co, p. 515] (see also [G1, Definition 2.1]).

**Definition 2.5.** A differentiable function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is said to satisfy condition (D), if there exists some constant  $\lambda > 0$  such that, with  $M(t) := -\log f(t)$ ,

$$M'(s) \geq \lambda M'(t) \quad \text{for any } t > 0 \text{ and } s \in [t, 2t].$$

For instance,  $C^1$ -functions  $f(t)$  that behave like  $t^{-\delta_1}$  for small  $t$  with  $\delta_1 > 0$ , and  $e^{-c_0 t^{\delta_2}}$  for large  $t$  with  $c_0, \delta_2 > 0$  satisfy condition (D).

We have the following property for the doubling function.

**Lemma 2.6.** *Let  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a decreasing function. Then  $1/\varphi$  satisfies the doubling property, if and only if there exist a strictly decreasing  $C^1$  function  $\bar{\varphi} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , and positive constants  $c_1 \leq c_2$  and  $b_1 \leq b_2$ , such that the following hold for all  $r > 0$ :*

$$c_1\varphi(r) \leq \bar{\varphi}(r) \leq c_2\varphi(r) \quad (2.2)$$

and

$$b_1 \leq -\frac{\bar{\varphi}'(r)}{\bar{\varphi}(r)}r \leq b_2. \quad (2.3)$$

In particular,

(i) *The function  $\bar{\varphi}$  satisfies condition (D). In fact, with  $M(r) := -\log \bar{\varphi}(r)$*

$$\frac{b_1}{r} \leq M'(r) = -\frac{\bar{\varphi}'(r)}{\bar{\varphi}(r)} \leq \frac{b_2}{r}, \quad r > 0.$$

(ii) *If  $\varphi(0) = \infty$  and  $\varphi(\infty) = 0$ , then there is a constant  $c \geq 1$  such that*

$$c^{-1} \frac{r}{\bar{\varphi}^{-1}(r)} \leq -\bar{\varphi}'(\bar{\varphi}^{-1}(r)) \leq c \frac{r}{\bar{\varphi}^{-1}(r)}, \quad r > 0. \quad (2.4)$$

**Proof.** Suppose that (2.3) holds. Then,

$$\log(\bar{\varphi}(2r)/\bar{\varphi}(r)) = \int_r^{2r} \frac{\bar{\varphi}'(u)}{\bar{\varphi}(u)} du \geq -b_2 \int_r^{2r} \frac{1}{u} du = -b_2 \log 2.$$

So  $1/\bar{\varphi}(r) \leq 1/\bar{\varphi}(2r) \leq 2^{b_2}/\bar{\varphi}(r)$ . In particular,  $1/\bar{\varphi}(r)$  is a doubling function. This yields that  $1/\varphi$  satisfies the doubling property as well if (2.2) holds.

Suppose that  $1/\varphi$  satisfies the doubling property. Let  $\varphi_0(2^i) = \varphi(2^i)$  for all  $i \in \mathbb{Z}$  and extend it linearly on  $\mathbb{R}_+$ . Then  $\varphi_0$  is continuous and decreasing on  $\mathbb{R}_+$ . For any  $r > 0$ , choose  $i \in \mathbb{Z}$  such that  $2^i \leq r < 2^{i+1}$ . Then, by the doubling property of the function  $1/\varphi$ , we have

$$\frac{1}{C_1}\varphi(r) \leq \varphi(2^{i+1}) \leq \varphi_0(r) \leq \varphi(2^i) \leq C_1\varphi(r) \quad (2.5)$$

for some constant  $C_1 > 1$  independent of  $i$  and  $r$ . In particular,  $1/\varphi_0$  is a doubling function on  $\mathbb{R}_+$ . Now, we define

$$\bar{\varphi}(r) := \frac{1}{r} \int_r^{2r} \varphi_0(u) du, \quad r > 0.$$

It is clear that  $\varphi_0(2r) \leq \overline{\varphi}(r) \leq \varphi_0(r)$ . This along with (2.5) yields (2.2), and this together with the doubling property of  $1/\varphi_0$  implies the doubling property of  $1/\overline{\varphi}$  as well. Furthermore, noticing that  $\varphi_0$  is continuous and decreasing on  $\mathbb{R}_+$ , for all  $r > 0$ ,

$$\overline{\varphi}'(r) = -\frac{1}{r^2} \int_r^{2r} \varphi_0(u) du + \frac{1}{r}(\varphi_0(2r) - \varphi_0(r)) \geq -\frac{2\varphi_0(r)}{r}$$

and

$$\overline{\varphi}'(r) = -\frac{1}{r^2} \int_r^{2r} \varphi_0(u) du + \frac{1}{r}(\varphi_0(2r) - \varphi_0(r)) \leq -\frac{1}{r^2} \int_r^{2r} \varphi_0(u) du \leq -\frac{\varphi_0(2r)}{r} \leq -C_2 \frac{\varphi_0(r)}{r},$$

where in the last inequality above  $C_2 > 0$  is independent of  $r$  and we used the doubling property of  $1/\varphi_0$ . Combining both inequalities above with (2.5), we know that (2.3) holds; in particular,  $\overline{\varphi}(r)$  is strictly decreasing on  $\mathbb{R}_+$ .

(i) directly follows from (2.3), and (ii) is a consequence of (2.2) and (2.3). The proof is complete.

□

**Lemma 2.7.** *Let  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a decreasing function such that  $1/\varphi$  has the doubling property. Then for any  $\lambda > 2$ , there is a constant  $C(\lambda) > 0$  such that*

$$\prod_{k=1}^{\infty} \left( \varphi((\lambda - 1)\lambda^{-(k+1)}t) \right)^{1/2^k} \leq C(\lambda)\varphi(t), \quad t > 0.$$

**Proof.** Since  $1/\varphi$  is a doubling function on  $\mathbb{R}_+$ , there are constants  $C, \eta > 0$  such that

$$\frac{\varphi(r)}{\varphi(R)} \leq C \left( \frac{R}{r} \right)^\eta, \quad 0 < r \leq R.$$

Then, for any  $\lambda > 2$  and  $t > 0$ ,

$$\begin{aligned} \prod_{k=1}^{\infty} \left( \varphi((\lambda - 1)\lambda^{-(k+1)}t) \right)^{1/2^k} &\leq \prod_{k=1}^{\infty} \left( \varphi((\lambda - 1)\lambda^{-1}t) \cdot C\lambda^{\eta k} \right)^{1/2^k} \\ &= C \left( \varphi((\lambda - 1)\lambda^{-1}t) \right)^{\sum_{k=1}^{\infty} 2^{-k}} \lambda^{\sum_{k=1}^{\infty} \eta k / 2^k} \\ &\leq C_1 \varphi((\lambda - 1)\lambda^{-1}t) \leq C_2 \varphi(t), \end{aligned}$$

where  $C_1, C_2 > 0$  are independent of  $t$  but depend on  $C, \eta$  and  $\lambda$ . □

Combining Lemmas 2.6 and 2.7 with Definition 1.1, we have the following statement.

**Proposition 2.8.** *Let  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a decreasing function such that  $\varphi(0) = \infty$ ,  $\varphi(\infty) = 0$  and that  $1/\varphi$  has the doubling property. Then there exist  $\overline{\varphi} \in \mathcal{R}$  and a constant  $c \geq 1$  such that*

$$c^{-1}\overline{\varphi}(r) \leq \varphi(r) \leq c\overline{\varphi}(r), \quad r \in \mathbb{R}_+.$$

### 3 On diagonal heat kernel upper bounds

In this section, we summarize the relations between Nash-type inequalities and on-diagonal heat kernel upper bounds from [Co]. We emphasize that the results of this section hold for all regular symmetric Dirichlet forms which may have killings inside  $E$ .

**Proposition 3.1.** *Suppose that  $(P_t)_{t \geq 0}$  is a symmetric Markov semigroup on  $L^2(E; m)$  with a regular symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  on  $L^2(E; m)$ . Let  $\delta$  be a non-negative constant.*

(i) *Let  $\theta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be such that  $\int^{+\infty} 1/\theta(s) ds < \infty$ . If*

$$\theta(\|f\|_2^2) \leq \mathcal{E}(f, f) + \delta \|f\|_2^2 \quad \text{for } f \in \mathcal{F} \text{ with } \|f\|_1 \leq 1, \quad (3.1)$$

*then*

$$\|P_t\|_{1 \rightarrow \infty} \leq \varphi(t) e^{\delta t} \quad \text{for } t > 0,$$

*where  $\varphi(t)$  is the inverse function of  $t \mapsto \int_t^\infty 1/\theta(s) ds$ .*

(ii) *Suppose that*

$$\|P_t\|_{1 \rightarrow \infty} \leq \varphi(t) e^{\delta t} \quad \text{for } t > 0,$$

*where  $\varphi(t) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a decreasing function. Then*

$$\tilde{\theta}(\|f\|_2^2) \leq \mathcal{E}(f, f) + \delta \|f\|_2^2 \quad \text{for } f \in \mathcal{F} \text{ with } \|f\|_1 \leq 1, \quad (3.2)$$

*where*

$$\tilde{\theta}(r) := \sup_{t > 0} \frac{r}{t} \log \frac{r}{\varphi(t)}, \quad r > 0. \quad (3.3)$$

**Proof.** For any  $t > 0$  and  $f \in L^2(E; m)$ , define

$$P_t^\delta f(x) = e^{-\delta t} P_t f(x), \quad x \in E.$$

Then,  $(P_t^\delta)_{t \geq 0}$  is a symmetric Markov semigroup on  $L^2(E; m)$  associated with the Dirichlet form  $(\mathcal{E}_\delta, \mathcal{F})$ , where

$$\mathcal{E}_\delta(u, u) := \mathcal{E}(u, u) + \delta \|u\|_2^2, \quad u \in \mathcal{F}.$$

Hence, by applying [Co, Propositions II.1 and II.2] to  $(P_t^\delta)_{t \geq 0}$  and using the fact that  $\|P_t\|_{1 \rightarrow \infty} = e^{\delta t} \|P_t^\delta\|_{1 \rightarrow \infty}$  for all  $t > 0$ , we get the desired conclusion.  $\square$ .

**Remark 3.2.** (i) In [Co, Proposition II.1], the function  $\theta$  is assumed to be continuous but in fact this continuity assumption is not needed; see the proof of [Co, Proposition II.1]. If the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  is conservative; that is, if  $P_t 1 = 1$   $m$ -a.e. for every  $t > 0$ , then (3.1) can be weakened to hold for all  $f \in \mathcal{F}$  with  $\|f\|_1 = 1$ ; see [Co, Remark on p. 513]. Since  $\int^\infty 1/\theta(s) ds < \infty$ , we can well define  $\varphi(0) := \infty$  for the function  $\varphi$  given in Proposition 3.1(i).

(ii) Since  $(P_t)_{t \geq 0}$  is symmetric on  $L^2(E; m)$ ,  $\|P_t\|_{1 \rightarrow 2}^2 = \|P_{2t}\|_{1 \rightarrow \infty}$  for all  $t > 0$ . Note that  $\tilde{\theta}$  defined by (3.3) is convex, because it is a supremum of a class of convex functions and  $\tilde{\theta}(0) = 0$ . Hence, (3.2) is equivalent to

$$\tilde{\theta}(\|f\|_2^2) \leq \mathcal{E}(f, f) + \delta \|f\|_2^2 \quad \text{for } f \in \mathcal{F} \text{ with } \|f\|_1 = 1.$$

We also note that, if  $\varphi(\infty) > 0$ , then  $\tilde{\theta}(r) = 0$  for all  $r \in (0, \varphi(\infty)]$ ; if  $\varphi(0) < \infty$ , then  $\tilde{\theta}(r) = \infty$  for all  $r \in (\varphi(0), \infty)$ .

It is clear that, when  $\theta$  is continuous, the function  $\varphi(t)$  in Proposition 3.1(i) satisfies that

$$-\varphi'(t) = \theta(\varphi(t)), \quad t > 0.$$

In order to study the equivalent characterizations between Nash-type inequalities and on diagonal heat kernel upper bounds, for a given strictly decreasing differentiable bijection  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  (in particular,  $\varphi(0) = \infty$  and  $\varphi(\infty) = 0$ ) we define  $\theta$  by

$$\theta(r) = -\varphi'(\varphi^{-1}(r)), \quad r > 0. \quad (3.4)$$

The following is proved in [Co, Lemma II.3].

**Lemma 3.3.** *Let  $\varphi$  be a strictly decreasing differentiable bijection of  $\mathbb{R}_+$  satisfying (D). Then there is a constant  $c > 0$  such that for all  $r > 0$ ,*

$$\tilde{\theta}(r) \geq c \theta(r),$$

where  $\tilde{\theta}(r)$  and  $\theta(r)$  are defined by (3.3) and (3.4), respectively.

Note that by (3.4),

$$t_2 - t_1 = \int_{\varphi(t_2)}^{\varphi(t_1)} 1/\theta(s) ds \quad \text{for any } 0 < t_1 < t_2 < \infty. \quad (3.5)$$

Since  $\varphi$  is strictly decreasing on  $\mathbb{R}_+$  with  $\varphi(0) = \infty$  and  $\varphi(\infty) = 0$ . we have by (3.5) that  $\int_{\varphi(t)}^{\infty} 1/\theta(s) ds = t$  for  $t > 0$  and  $\int_0^{\infty} 1/\theta(s) ds = \infty$ . Combining Proposition 3.1 with Lemma 3.3, we have the following equivalence between Nash-type inequalities and on diagonal heat kernel upper bounds (see [Co, Theorem II.5]).

**Theorem 3.4.** *Let  $(P_t)_{t \geq 0}$  be a symmetric Markov semigroup associated with a regular symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  on  $L^2(E; m)$ . For a strictly decreasing differentiable bijection  $\varphi$  of  $\mathbb{R}_+$  satisfying condition (D), define  $\theta(x)$  by (3.4). Let  $\delta$  be a non-negative constant. Then the following conditions are equivalent:*

(i) There is a constant  $c_1 > 0$  such that

$$\|P_t\|_{1 \rightarrow \infty} \leq \varphi(c_1 t) e^{\delta t} \quad \text{for } t > 0.$$

(ii) There is a constant  $c_2 > 0$  such that

$$c_2 \theta(\|f\|_2^2) \leq \mathcal{E}(f, f) + \delta \|f\|_2^2 \quad \text{for } f \in \mathcal{F} \text{ with } \|f\|_1 \leq 1.$$

**Remark 3.5.** According to Lemma 2.6 (in particular, by (2.4)), we can replace the function  $\theta$  in (3.4) by

$$\theta_*(r) := \frac{r}{\varphi^{-1}(r)}, \quad r \in \mathbb{R}_+,$$

when  $\varphi$  in Theorem 3.4 has the property that  $1/\varphi(r)$  is a doubling function on  $\mathbb{R}_+$ .

The following table lists some typical regular functions  $\varphi$  and the corresponding  $\theta(r) := -\varphi'(\varphi^{-1}(r))$  (by neglecting constant multiplies). In the table,  $\alpha > 0$ ,  $\beta \in \mathbb{R}$  and  $\gamma \in (0, 1]$ .

$\varphi(r)$	$r^{-\frac{1}{\alpha}}$	$r^{-\frac{1}{\alpha}} \log^\beta(2+r)$	$r^{-\frac{1}{\alpha}} \log^\beta(2+\frac{1}{r})$	$r^{-\frac{1}{\alpha}} e^{-r^\gamma}$	$\log^\beta(1+\frac{1}{r})$
$\theta(r)$	$r^{1+\frac{1}{\alpha}}$	$r^{1+\frac{1}{\alpha}} \log^{-\frac{\beta}{\alpha}}(2+\frac{1}{r})$	$r^{1+\frac{1}{\alpha}} \log^{-\frac{\beta}{\alpha}}(2+r)$	$r^{1+\frac{1}{\alpha}} \vee r \log^{1-\frac{1}{\gamma}}(2+\frac{1}{r})$	$r^{1+\frac{1}{\beta}} \mathbb{1}_{\{0 < r \leq 1\}} + e^{r^{1/\beta}} \mathbb{1}_{\{r > 1\}}$

## 4 Off-diagonal heat kernel upper bounds

In this section, we will derive off-diagonal upper bounds for  $p(t, x, y)$  from Nash-type inequalities. Let  $(P_t)_{t \geq 0}$  be a symmetric Markov semigroup on  $L^2(E; m)$  associated with a regular symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  on  $L^2(E; m)$  that admits no killings inside  $E$ . Following Davies' idea, in order to study off-diagonal heat kernel bounds for  $(P_t)_{t \geq 0}$ , we consider the perturbed semigroup  $(P_t^\psi)_{t \geq 0}$  defined by

$$P_t^\psi f(x) := e^{\psi(x)} P_t(e^{-\psi} f)(x), \quad t \geq 0 \tag{4.1}$$

for some nice function  $\psi \in \mathcal{F}_b$ .

**Definition 4.1.** We say a function  $\psi \in \mathcal{F}_b$  is admissible with respect to  $(\mathcal{E}, \mathcal{F})$ , if there exists a constant  $C_0 > 0$  such that for every  $p \in [1, \infty)$  and  $f \in \mathcal{F}_b$ ,

$$\mathcal{E}(e^\psi f^{2p-1}, e^{-\psi} f) \geq \frac{C_0}{p} \mathcal{E}(f^p, f^p) - \phi(C_0, \psi, p) \|f\|_{2p}^{2p}, \tag{4.2}$$

where  $(\psi, r) \mapsto \phi(C_0, \psi, r)$  is a non-negative function on  $\mathcal{F}_b \times [1, \infty)$  such that

$$\phi(C_0, \psi, r) = \phi(C_0, -\psi, r), \quad \psi \in \mathcal{F}_b, \quad r \geq 1, \tag{4.3}$$

and  $r \mapsto \phi(C_0, \psi, r)$  is increasing and has the doubling property uniformly with respect to  $\psi$  on  $[1, \infty)$ , i.e., there are constants  $C := C(C_0) \geq 1$  and  $\eta := \eta(C_0) > 0$  such that for all  $\psi \in \mathcal{F}_b$ ,

$$\frac{\phi(C_0, \psi, R)}{\phi(C_0, \psi, r)} \leq C \left( \frac{R}{r} \right)^\eta, \quad 1 \leq r \leq R. \quad (4.4)$$

We denote the class of such bounded admissible functions  $\psi$  by  $\mathcal{A}(C_0; \phi, C, \eta)$ .

**Remark 4.2.** Note that if (4.2) holds for  $C_0 > 0$  and  $\phi(C_0, \psi, p)$ , it holds for any  $\tilde{C}_0 \in (0, C_0)$  and  $\phi(\tilde{C}_0, \psi, p) := \phi(C_0, \psi, p)$ .

**Example 4.3.** For a given function  $\psi \in \mathcal{F}_b$  such that

$$\Lambda(\psi)^2 := \max \left\{ \left\| \frac{de^{-2\psi}\Gamma(e^\psi, e^\psi)}{dm} \right\|_\infty, \left\| \frac{de^{2\psi}\Gamma(e^{-\psi}, e^{-\psi})}{dm} \right\|_\infty \right\} < \infty,$$

we have  $\psi \in \mathcal{A}(s; \phi, 5/(1-s), 2)$  for any  $s \in (0, 1)$ , where

$$\phi(s, \psi, r) = \left( 1 + \frac{4r^2 \mathbb{1}_{\{r>1\}}}{1-s} \right) \Lambda(\psi)^2, \quad r \geq 1. \quad (4.5)$$

In particular,  $\phi(s, \psi, 1) = \Lambda(\psi)^2$ . Furthermore, if  $(\mathcal{E}, \mathcal{F})$  is a strongly local Dirichlet form, then  $\Lambda(\psi)^2 = d\Gamma(\psi, \psi)/dm$  and  $\psi \in \mathcal{A}(1; \phi, 1, 1)$  with  $\phi(1, \psi, r) := r\Lambda(\psi)^2$ .

Indeed, it was proved in [CKS, (3.10)] that for all  $f \in \mathcal{F}_b$ ,

$$\mathcal{E}(e^\psi f, e^{-\psi} f) \geq \mathcal{E}(f^p, f^p) - \Lambda(\psi)^2 \|f\|_2^2.$$

This implies that (4.2) holds for  $p = 1$  with  $C_0 = s$  and  $\phi(s, \psi, 1) = \Lambda(\psi)^2$  for any  $s \in (0, 1]$ . On the other hand, we know from the end of the proof for [CKS, Theorem (3.9)] that for all  $p \in [1, \infty)$  and  $f \in \mathcal{F}_b$ ,

$$\mathcal{E}(e^\psi f^{2p-1}, e^{-\psi} f) \geq \frac{2p-1}{p^2} \mathcal{E}(f^p, f^p) - 4\mathcal{E}(f^p, f^p)^{1/2} \Lambda(\psi) \|f\|_{2p}^p - \Lambda(\psi)^2 \|f\|_{2p}^{2p}.$$

Since for any  $a, b \geq 0$ ,  $p \geq 1$  and  $s \in (0, 1]$ ,

$$\frac{2p-1}{p^2} a^2 - 4ab - b^2 \geq \frac{s}{p} a^2 - \left( 1 + \frac{4p^2}{(2-s)p-1} \right) b^2,$$

we have

$$\mathcal{E}(e^\psi f^{2p-1}, e^{-\psi} f) \geq \frac{s}{p} \mathcal{E}(f^p, f^p) - \left( 1 + \frac{4p^2}{(2-s)p-1} \right) \Lambda(\psi)^2 \|f\|_{2p}^{2p} \quad (4.6)$$

$$\geq \frac{s}{p} \mathcal{E}(f^p, f^p) - \left( 1 + \frac{4p^2}{1-s} \right) \Lambda(\psi)^2 \|f\|_{2p}^{2p} \quad (4.7)$$



for all  $p \in [1, \infty)$ ,  $s \in (0, 1)$  and  $f \in \mathcal{F}_b$ . (Note that, by taking  $s = 1/2$  in (4.6) and using  $p \geq 1$ , we get that for all  $p \in [1, \infty)$  and  $f \in \mathcal{F}_b$ ,

$$\mathcal{E}(e^\psi f^{2p-1}, e^{-\psi} f) \geq (2p)^{-1} \mathcal{E}(f^p, f^p) - 9p\Lambda(\psi)^2 \|f\|_{2p}^{2p},$$

which corrects a typo in [CKS, (3.11)]; that is, it should be  $(2p)^{-1}$  not  $p^{-1}$  in front of  $\mathcal{E}(f^p, f^p)$ .) In particular, (4.7) implies that, for  $p > 1$ , (4.2) holds with  $C_0 = s$  and  $\phi(s, \psi, p) = \left(1 + \frac{4p^2}{1-s}\right) \Lambda(\psi)^2$  for any  $s \in (0, 1)$ . Thus for any  $1 < r \leq R$  and  $s \in (0, 1)$ ,

$$\frac{\phi(s, \psi, R)}{\phi(s, \psi, r)} \leq \frac{8R^2/(1-s)}{4r^2/(1-s)} \leq 2(R/r)^2.$$

Recall that, for  $p = 1$ , we have (4.5) that  $\phi(s, \psi, 1) = \Lambda(\psi)^2$  for  $s \in (0, 1]$ . Thus for any  $r > 1$  and  $s \in (0, 1)$ ,

$$\frac{\phi(s, \psi, r)}{\phi(s, \psi, 1)} \leq 1 + \frac{4r^2}{1-s} < \frac{5}{1-s} r^2.$$

We conclude from these two estimates that (4.4) holds with  $C = 5/(1-s)$  and  $\eta = 2$ . In other words,  $\psi \in \mathcal{A}(s; \phi, 5/(1-s), 2)$  with  $\phi(s, \psi, r)$  given by (4.5).

For a strongly local Dirichlet form  $(\mathcal{E}, \mathcal{F})$ , that  $\Lambda(\psi)^2 = d\Gamma(\psi, \psi)/dm$  immediately follows from the Leibniz rule. By the Leibniz rule again and the Cauchy-Schwarz inequality, for all  $p \geq 1$ , we have

$$\mathcal{E}(e^\psi f^{2p-1}, e^{-\psi} f) \geq p^{-1} \mathcal{E}(f^p, f^p) - p\Lambda(\psi)^2 \|f\|_{2p}^{2p};$$

see e.g. [CKS, (3.2)]. This yields that  $\psi \in \mathcal{A}(1; \phi, 1, 1)$ , where  $\phi(1, \psi, r) = r\Lambda(\psi)^2$ .  $\square$

Now, we present the main result of this paper.

**Theorem 4.4.** *Let  $(P_t)_{t \geq 0}$  be a symmetric Markov semigroup on  $L^2(E; m)$  associated with a regular symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  on  $L^2(E; m)$  that has no killings inside  $E$ . Suppose the following Nash-type inequality holds:*

$$\theta(\|f\|_2^2) \leq \mathcal{E}(f, f) + \delta \|f\|_2^2 \quad \text{for } f \in \mathcal{F} \text{ with } \|f\|_1 \leq 1, \quad (4.8)$$

where  $\delta \geq 0$ , and  $\theta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  satisfies that

(i)  $\int_0^\infty 1/\theta(s) ds < \infty$ ;

(ii) there are constant  $c_1, c_2 \geq 1$  and an increasing function  $\theta_*$  on  $\mathbb{R}_+$  so that  $\theta_*(r)/r$  is increasing and

$$c_1^{-1} \theta_*(c_2^{-1} r) \leq \theta(r) \leq c_1 \theta_*(c_2 r) \quad \text{for all } r \in \mathbb{R}_+; \quad (4.9)$$

(iii) for any  $\lambda > 2$  large enough, there exist positive constants  $c(\lambda)$  and  $C(\lambda)$  so that

$$\prod_{k=1}^{\infty} \left( \varphi((\lambda - 1)\lambda^{-(k+1)}t) \right)^{2^{-k}} \leq C(\lambda)\varphi(c(\lambda)t), \quad t > 0, \quad (4.10)$$

where  $\varphi(t)$  is the inverse function of  $t \mapsto \int_t^{\infty} 1/\theta(s) ds$ .

Then there are a properly exceptional set  $\mathcal{N} \subset E$  and a heat kernel  $p(t, x, y)$  associated with the semigroup  $(P_t)_{t \geq 0}$  and defined on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$  so that for any  $\varepsilon > 0$  and  $\psi \in \mathcal{A}(C_0; \phi, C, \eta)$ ,

$$p(t, x, y) \leq C_\varepsilon \varphi(c_\varepsilon t) e^{\delta t} \exp(-|\psi(y) - \psi(x)| + (1 + \varepsilon)\phi(C_0, \psi, 1)t) \quad (4.11)$$

for all  $t > 0$  and  $x, y \in E \setminus \mathcal{N}$ , where  $C_\varepsilon := C_\varepsilon(\theta, C, \eta, \varepsilon) > 0$ ,  $c_\varepsilon := c_\varepsilon(\theta, C, \eta, \varepsilon, C_0) > 0$ , and the dependencies of  $C_\varepsilon$  and  $c_\varepsilon$  on  $\theta$  are only through the constants  $c_1$ ,  $c_2$ ,  $c(\lambda)$  and  $C(\lambda)$  in (4.9) and (4.10).

We note that by the condition (ii) of the function  $\theta$  it holds that  $\int_0^{\infty} 1/\theta(s) ds = \infty$ . In particular, the function  $\varphi$  in (iii) is strictly positive and strictly decreasing on  $\mathbb{R}_+$  with  $\varphi(0) = \infty$  and  $\varphi(\infty) = 0$ .

To prove Theorem 4.4, we need the following estimate for  $\|P_t^\psi\|_{1 \rightarrow \infty}$ , where  $(P_t^\psi)_{t \geq 0}$  is the perturbed semigroup defined by (4.1).

**Proposition 4.5.** *Under the setting of Theorem 4.4, for any  $\varepsilon > 0$  and  $\psi \in \mathcal{A}(C_0; \phi, C, \eta)$ ,*

$$\|P_t^\psi\|_{1 \rightarrow \infty} \leq C_\varepsilon \varphi(c_\varepsilon t) e^{\delta t} \exp((1 + \varepsilon)\phi(C_0, \psi, 1)t) \quad \text{for all } t > 0, \quad (4.12)$$

where  $C_\varepsilon := C_\varepsilon(\theta, C, \eta, \varepsilon) > 0$ ,  $c_\varepsilon := c_\varepsilon(\theta, C, \eta, \varepsilon, C_0) > 0$ , and the dependencies of  $C_\varepsilon$  and  $c_\varepsilon$  on  $\theta$  are only through the constants  $c_1$ ,  $c_2$ ,  $c(\lambda)$  and  $C(\lambda)$  in (4.9) and (4.10).

**Proof.** Without loss of generality, we assume the constant  $C_0$  is less than 1 (see Remark 4.2). We first assume that the function  $r \mapsto \theta(r)/r$  is increasing on  $\mathbb{R}_+$ . For any  $\psi \in \mathcal{A}(C_0; \phi, C, \eta)$ , non-negative  $f \in \mathcal{F}_b$ ,  $t > 0$  and  $p \geq 1$ , we write

$$f_t(x) := P_t^\psi f(x) = e^{\psi(x)} (P_t(e^{-\psi} f))(x) \quad (4.13)$$

and

$$f_{p,t}(x) := \exp(-(\phi(C_0, \psi, p) + \delta p^{-1})t) f_t(x). \quad (4.14)$$

For simplicity, in the following we write  $\phi(\psi, p)$  for  $\phi(C_0, \psi, p)$ . Note that since  $e^\psi - 1 \in \mathcal{F}_b$ ,  $f_t \in \mathcal{F}_b$

and so is  $f_{p,t}$ . According to (4.2) and (4.8), for any  $p \geq 1$  and  $t > 0$ ,

$$\begin{aligned} \frac{d\|f_{p,t}\|_{2p}^{2p}}{dt} &= -(2p\phi(\psi, p) + 2\delta)\|f_{p,t}\|_{2p}^{2p} - 2p\mathcal{E}(e^\psi(f_{p,t})^{2p-1}, e^{-\psi}f_{p,t}) \\ &\leq -2C_0\mathcal{E}(f_{p,t}^p, f_{p,t}^p) - 2\delta\|f_{p,t}\|_{2p}^{2p} \end{aligned} \quad (4.15)$$

$$\begin{aligned} &\leq -2C_0(\mathcal{E}(f_{p,t}^p, f_{p,t}^p) + \delta\|f_{p,t}\|_{2p}^{2p}) \\ &= -2C_0\|f_{p,t}\|_p^{2p}(\mathcal{E}(f_{p,t}^p/\|f_{p,t}\|_p^p, f_{p,t}^p/\|f_{p,t}\|_p^p) + \delta\|f_{p,t}\|_{2p}^{2p}/\|f_{p,t}\|_p^{2p}) \\ &\leq -2C_0\|f_{p,t}\|_p^{2p}\theta(\|f_{p,t}\|_{2p}^{2p}/\|f_{p,t}\|_p^{2p}). \end{aligned} \quad (4.16)$$

From (4.15), we see that  $t \mapsto e^{2\delta t}\|f_{p,t}\|_{2p}^{2p}$  is decreasing on  $\mathbb{R}_+$  for every  $p \geq 1$ . On the other hand, by (4.14),

$$\begin{aligned} \|f_{2p,t}\|_{2p}^{2p} &= \exp(-2p\phi(\psi, 2p)t - \delta t)\|f_t\|_{2p}^{2p} \\ &= \exp(-2p\phi(\psi, 2p)t + 2p\phi(\psi, p)t + \delta t)\|f_{p,t}\|_{2p}^{2p} \\ &= \exp(-[2p(\phi(\psi, 2p) - \phi(\psi, p)) + \delta]t)(e^{2\delta t}\|f_{p,t}\|_{2p}^{2p}). \end{aligned}$$

Since  $\phi(\psi, 2p) \geq \phi(\psi, p)$ , it follows that  $\|f_{2p,t}\|_{2p}^{2p}$  is decreasing in  $t \in \mathbb{R}_+$  for every  $p \geq 1$ . In other words,  $\|f_{p,t}\|_p^p$  is decreasing in  $t \in \mathbb{R}_+$  for every  $p \geq 2$ . Hence we have from (4.16) and the increasing property of the function  $r \mapsto \theta(r)/r$  on  $\mathbb{R}_+$  that for any  $p \geq 2$  and  $t_0 > 0$ ,

$$\frac{d\|f_{p,t}\|_{2p}^{2p}}{dt} \leq -2C_0\|f_{p,t_0}\|_p^{2p}\theta\left(\|f_{p,t}\|_{2p}^{2p}/\|f_{p,t_0}\|_p^{2p}\right) \quad \text{for } t \geq t_0. \quad (4.17)$$

Define

$$\Psi(r) = \frac{1}{2C_0} \int_r^\infty \frac{ds}{\theta(s)}, \quad r > 0,$$

which is finite by our condition (i). We can rewrite (4.17) as

$$\frac{d}{dt}\Psi\left(\frac{\|f_{p,t}\|_{2p}^{2p}}{\|f_{p,t_0}\|_p^{2p}}\right) \geq 1 \quad \text{for } t \geq t_0.$$

In particular, for all  $p \geq 2$  and  $t \geq t_0 > 0$ ,

$$\Psi\left(\frac{\|f_{p,t}\|_{2p}^{2p}}{\|f_{p,t_0}\|_p^{2p}}\right) \geq t - t_0.$$

That is,

$$\|f_{p,t}\|_{2p} \leq (\Psi^{-1}(t - t_0))^{1/(2p)}\|f_{p,t_0}\|_p = (\varphi(2C_0(t - t_0)))^{1/(2p)}\|f_{p,t_0}\|_p, \quad t \geq t_0.$$

Consequently, for all  $p \geq 2$

$$\|f_t\|_{2p} \leq (\varphi(2C_0(t - t_0)))^{1/(2p)} \exp((\phi(\psi, p) + \delta p^{-1})(t - t_0))\|f_{t_0}\|_p, \quad t \geq t_0. \quad (4.18)$$

Fix  $t > 0$ , and, for  $k \geq 1$ , let  $t_k = (\lambda - 1)t \sum_{i=1}^k \lambda^{-i} = t(1 - \lambda^{-k})$ , where  $\lambda > 2^\eta$  is a large constant satisfying (4.10) and independent of  $t$  to be determined later. Applying (4.18) with  $p = 2^k$ ,  $t = t_{k+1}$  and  $t_0 = t_k$ , we get

$$\|f_{t_{k+1}}\|_{2^{k+1}} \leq \left( \varphi(2C_0(\lambda - 1)\lambda^{-(k+1)}t) \right)^{1/2^{k+1}} \exp \left( (\phi(\psi, 2^k) + \delta 2^{-k})(\lambda - 1)\lambda^{-(k+1)}t \right) \|f_{t_k}\|_{2^k}. \quad (4.19)$$

Since  $t \mapsto \|f_{p,t}\|_p$  is decreasing on  $(0, \infty)$  for all  $p \geq 2$ ,  $\|f_{2^k,t}\|_{2^k} \leq \|f_{2^k,t_k}\|_{2^k}$  for  $k \geq 1$ . We have by (4.14) that for all  $k \geq 1$ ,

$$\begin{aligned} \|f_t\|_{2^k} &= \exp \left( (\phi(\psi, 2^k) + \delta 2^{-k})t \right) \|f_{2^k,t}\|_{2^k} \\ &\leq \exp \left( (\phi(\psi, 2^k) + \delta 2^{-k})t \right) \|f_{2^k,t_k}\|_{2^k} \\ &= \|f_{t_k}\|_{2^k} \exp \left( (\phi(\psi, 2^k) + \delta 2^{-k})(t - t_k) \right) \\ &= \|f_{t_k}\|_{2^k} \exp \left( t(\phi(\psi, 2^k) + \delta 2^{-k})\lambda^{-k} \right). \end{aligned} \quad (4.20)$$

As  $\psi \in \mathcal{A}(C_0; \phi, C, \eta)$ , we have

$$\frac{\phi(\psi, R)}{\phi(\psi, r)} \leq C \left( \frac{R}{r} \right)^\eta \quad \text{for } R \geq r \geq 1. \quad (4.21)$$

Thus, using our assumption  $\lambda > 2^\eta$ ,

$$\limsup_{k \rightarrow \infty} (\phi(\psi, 2^k) + \delta 2^{-k})\lambda^{-k} \leq \limsup_{k \rightarrow \infty} (C(2^\eta/\lambda)^k \phi(\psi, 1) + \delta(2\lambda)^{-k}) = 0. \quad (4.22)$$

(4.20) and (4.22) along with the inequality (4.19) yield

$$\begin{aligned} \|f_t\|_\infty &= \lim_{k \rightarrow \infty} \|f_t\|_{2^k} \leq \limsup_{k \rightarrow \infty} \|f_{t_{k+1}}\|_{2^{k+1}} \\ &\leq \left( \prod_{k=1}^{\infty} \left( \varphi(2C_0(\lambda - 1)\lambda^{-(k+1)}t) \right)^{1/2^{k+1}} \right) \\ &\quad \times \exp \left( \lambda^{-1}(\lambda - 1)t \left( \sum_{k=1}^{\infty} \phi(\psi, 2^k)\lambda^{-k} + \delta \sum_{k=1}^{\infty} 2^{-k}\lambda^{-k} \right) \right) \|f_{(\lambda-1)\lambda^{-1}t}\|_2. \end{aligned} \quad (4.23)$$

Furthermore, since  $\|f_{1,s}\|_2$  is decreasing in  $s \in [0, \infty)$ ,  $\|f_{1,s}\|_2 \leq \|f_{1,0}\|_2 = \|f\|_2$ ; that is,

$$\|f_s\|_2 \leq \exp \left( (\phi(\psi, 1) + \delta)s \right) \|f\|_2, \quad s > 0.$$

This together with (4.23) gives us

$$\begin{aligned} \|f_t\|_\infty &\leq \left( \prod_{k=1}^{\infty} \left( \varphi(2C_0(\lambda - 1)\lambda^{-(k+1)}t) \right)^{1/2^{k+1}} \right) \\ &\quad \times \exp \left( \lambda^{-1}(\lambda - 1)t \left( \sum_{k=0}^{\infty} \phi(\psi, 2^k)\lambda^{-k} + \delta \sum_{k=0}^{\infty} 2^{-k}\lambda^{-k} \right) \right) \|f\|_2. \end{aligned} \quad (4.24)$$

Using (4.21), we have that, for  $\lambda > 2^\eta$ ,

$$\sum_{k=0}^{\infty} \phi(\psi, 2^k) \lambda^{-k} \leq \phi(\psi, 1) + C \phi(\psi, 1) \sum_{k=1}^{\infty} 2^{\eta k} \lambda^{-k} = \phi(\psi, 1) \left(1 + \frac{C 2^\eta}{\lambda - 2^\eta}\right).$$

Clearly,

$$\delta \sum_{k=0}^{\infty} 2^{-k} \lambda^{-k} = \frac{\delta}{1 - 2^{-1} \lambda^{-1}}.$$

These together with (4.10) and (4.24) yield that

$$\begin{aligned} \|f_t\|_\infty &\leq (C(\lambda) \varphi(2C_0 c(\lambda)t))^{1/2} \exp\left(t \frac{\lambda-1}{\lambda} \left(\phi(\psi, 1) \left(1 + \frac{C 2^\eta}{\lambda - 2^\eta}\right) + \frac{\delta}{1 - 2^{-1} \lambda^{-1}}\right)\right) \|f\|_2 \\ &\leq (C(\lambda) \varphi(2C_0 c(\lambda)t))^{1/2} \exp\left(t \phi(\psi, 1) \frac{\lambda-1}{\lambda} \left(1 + \frac{C 2^\eta}{\lambda - 2^\eta}\right) + \delta t\right) \|f\|_2. \end{aligned}$$

By the duality and (4.3) (see [CKS, p. 271] for details), we get

$$\|f_t\|_2 \leq (C(\lambda) \varphi(2C_0 c(\lambda)t))^{1/2} \exp\left(t \phi(\psi, 1) \frac{\lambda-1}{\lambda} \left(1 + \frac{C 2^\eta}{\lambda - 2^\eta}\right) + \delta t\right) \|f\|_1.$$

Since  $P_t^\psi f = P_{t/2}^\psi (P_{t/2}^\psi f)$ , it follows that for every non-negative  $f \in \mathcal{F}$ ,

$$\|f_t\|_\infty \leq C(\lambda) \varphi(C_0 c(\lambda)t) \exp\left(t \phi(\psi, 1) \frac{\lambda-1}{\lambda} \left(1 + \frac{C 2^\eta}{\lambda - 2^\eta}\right) + \delta t\right) \|f\|_1. \quad (4.25)$$

For any  $\varepsilon > 0$ , taking  $\lambda := \lambda(\varepsilon, C, \eta) > 2^\eta$  sufficiently large such that

$$\frac{\lambda-1}{\lambda} \left(1 + \frac{C 2^\eta}{\lambda - 2^\eta}\right) < 1 + \varepsilon,$$

we conclude that with  $C_\varepsilon := C(\lambda)$  and  $c_\varepsilon := C_0 c(\lambda)$  (noting that the constant  $\lambda$  depends on  $\varepsilon$ ,  $C$  and  $\eta$  only),

$$\|P_t^\psi f\|_\infty = \|f_t\|_\infty \leq C_\varepsilon \varphi(c_\varepsilon t) e^{\delta t} \exp((1 + \varepsilon) \phi(\psi, 1)t) \|f\|_1 \quad \text{for non-negative } f \in \mathcal{F}_b.$$

The above inequality holds for general  $f \in \mathcal{F}_b$  by applying it to  $f_+$  and  $f_-$  and then using the triangular inequality. Therefore, we get the desired conclusion.

Now we consider the general case of  $\theta$  satisfying conditions (i) and (ii). By condition (ii), there are constants  $c_1, c_2 \geq 1$  and an increasing function  $\theta_*$  so that  $\theta_*(r)/r$  is increasing and

$$c_1^{-1} \theta_*(c_2^{-1} r) \leq \theta(r) \leq c_1 \theta_*(c_2 r) \quad \text{for all } r \in \mathbb{R}_+.$$

Then

$$c_2^{-2} \varphi_*((c_1 c_2)^2 r) \leq \varphi(r) \leq \varphi_*(r) \quad \text{for all } r \in \mathbb{R}_+, \quad (4.26)$$

where  $\varphi_*(r)$  is the inverse function of  $r \mapsto \int_r^\infty c_1/\theta_*(c_2^{-1}s) ds$ . The function  $c_1^{-1}\theta_*(c_2^{-1}r)$  in place of  $\theta(r)$  satisfies conditions (i), (ii) and (iii) in Theorem 4.4. Applying the above proof to  $c_1^{-1}\theta_*(c_2^{-1}r)$  in place of  $\theta(r)$  and noting (4.26), we get the desired assertion.  $\square$

We are now in the position to present the proofs for Theorem 4.4, Theorem 1.2 and Corollary 1.3.

**Proof of Theorem 4.4.** First, by (4.8) and Proposition 3.1(i), we know that

$$\|P_t\|_{1 \rightarrow \infty} \leq \varphi(t)e^{\delta t}, \quad t > 0,$$

which along with [BBCK, Theorem 3.1] yields that there exist a properly exceptional set  $\mathcal{N} \subset E$  and a heat kernel  $p(t, x, y)$  on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$  associated with the semigroup  $(P_t)_{t \geq 0}$  that satisfies

$$p(t, x, y) \leq \varphi(t)e^{\delta t}, \quad t > 0, \quad x, y \in E \setminus \mathcal{N}.$$

Thus we have by Proposition 4.5 that for any  $\varepsilon \in (0, 1]$  and  $\psi \in \mathcal{A}(C_0; \phi, C, \eta)$ ,

$$e^{\psi(x)}p(t, x, y)e^{-\psi(y)} \leq C_\varepsilon\varphi(c_\varepsilon t)e^{\delta t} \exp((1 + \varepsilon)\phi(C_0, \psi, 1)t), \quad t > 0, \quad x, y \in E \setminus \mathcal{N},$$

so the desired estimate follows.  $\square$

**Remark 4.6.** (i) Obviously, (4.8) implies that for any  $\vartheta \in (0, 1]$ ,

$$\vartheta\theta(\|f\|_2^2) \leq \mathcal{E}(f, f) + \vartheta\delta\|f\|_2^2, \quad f \in \mathcal{F} \text{ and } \|f\|_1 \leq 1. \quad (4.27)$$

Then, by Proposition 3.1(i), for any  $t > 0$ ,  $\|P_t\|_{1 \rightarrow \infty} \leq \varphi(\vartheta t)e^{\vartheta\delta t}$ . Thus, the term  $\varphi(c_\varepsilon t)e^{\delta t}$  in (1.7), (1.8), (4.11) and (4.12) can be replaced by  $\varphi(\vartheta c_\varepsilon t)e^{\vartheta\delta t}$ . Note that, [CKS, Theorem 3.25] is stated in that way with  $\vartheta = \varepsilon$ . We stated our results without introducing  $\vartheta \in (0, 1)$  because it can be easily added in whenever needed through the inequality (4.27), and because the bound  $\varphi(\vartheta c_\varepsilon t)e^{\vartheta\delta t}$  does not seem to be better than  $\varphi(c_\varepsilon t)e^{\delta t}$ . Note that for  $\vartheta \in (0, 1)$ , (4.27) is a weaker inequality than (4.8).

(ii) For general  $\varphi$  that satisfies (4.10), by the decreasing property of  $\varphi$  on  $\mathbb{R}_+$  and  $\varphi(0) = \infty$ , we can assume that  $c(\lambda)$  in (4.10) is decreasing with respect to  $\lambda$  and  $C(\lambda)$  is increasing with respect to  $\lambda$  such that  $c(\infty) = 0$  or  $C(\infty) = \infty$ . Thus, according to the proofs of Proposition 4.5 and Theorem 4.4, we know that for the constants  $c_\varepsilon$  and  $C_\varepsilon$  in (4.12) either  $\lim_{\varepsilon \rightarrow 0} c_\varepsilon = 0$  or  $\lim_{\varepsilon \rightarrow 0} C_\varepsilon = \infty$ . In particular, if  $1/\varphi$  has the doubling property, then, by Lemma 2.7, we can take  $c_\varepsilon \equiv c > 0$  and  $\lim_{\varepsilon \rightarrow 0} C_\varepsilon = \infty$ .

**Proofs of Theorem 1.2 and Corollary 1.3.** The equivalence between (i) and (ii) has been proved in Theorem 3.4, and (iii) implies (ii). So we only need to show that (i) implies (iii). Let  $\varphi \in \mathcal{R}$ , and  $\theta(r) = -\varphi'(\varphi^{-1}(r))$  given in (i). Recall that  $t = \int_{\varphi(t)}^{\infty} 1/\theta(s) ds$  for all  $t \in \mathbb{R}_+$  and  $\varphi(\infty) = 0$ . By property (ii) for the regular function, there is a decreasing function  $N(r) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  so that

$$c_1^{-1}N(r) \leq -\frac{\varphi'(r)}{\varphi(r)} \leq c_1N(r) \quad \text{for } r > 0.$$

In other words,

$$c_1^{-1}N(\varphi^{-1}(r)) \leq \frac{\theta(r)}{r} \leq c_1N(\varphi^{-1}(r)) \quad \text{for } r > 0. \quad (4.28)$$

Clearly  $N(\varphi^{-1}(r))$  is a decreasing function. So (4.9) holds with  $\theta_*(r) := rN(\varphi^{-1}(r))$ . Property (iii) for the regular function is just (4.10). Thus by Theorem 4.4, there are a properly exceptional set  $\mathcal{N} \subset E$  and a heat kernel  $p(t, x, y)$  associated with  $(P_t)_{t \geq 0}$  and defined on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$  such that for any  $\varepsilon > 0$  and  $\psi \in \mathcal{A}(C_0; \phi, C, \eta)$ , there are constants  $c_\varepsilon, C_\varepsilon > 0$  so that

$$p(t, x, y) \leq C_\varepsilon \varphi(c_\varepsilon t) e^{\delta t} \exp(-|\psi(y) - \psi(x)| + (1 + \varepsilon)\phi(C_0, \psi, 1)t) \quad (4.29)$$

for all  $t > 0$ , and  $x, y \in E \setminus \mathcal{N}$ .

On the other hand, for any  $\psi \in \mathcal{F}_b$  such that  $\Lambda(\psi)^2 < \infty$ , by Example 4.3, we know that  $\psi \in \mathcal{A}(s; \phi, 5/(1-s), 2)$  for any  $s \in (0, 1)$ , where

$$\phi(s, \psi, r) = \left(1 + \frac{4r^2 \mathbb{1}_{\{r>1\}}}{1-s}\right) \Lambda(\psi)^2.$$

Noting that  $\phi(1/2, \psi, 1) = \Lambda(\psi)^2$ , we have from the above two displays that

$$p(t, x, y) \leq C_\varepsilon \varphi(c_\varepsilon t) e^{\delta t} \exp(-|\psi(y) - \psi(x)| + (1 + \varepsilon)\Lambda(\psi)^2 t)$$

for all  $t > 0$ , and  $x, y \in E \setminus \mathcal{N}$ . This proves (1.6). Furthermore, by optimizing with respect to  $\psi \in \mathcal{F}_b$  with  $\Lambda(\psi)^2 < \infty$  at the right hand side of (1.6), we can get (1.7).

When  $(\mathcal{E}, \mathcal{F})$  is a strongly local Dirichlet form on  $L^2(E; m)$ , we know from Example 4.3 that for any  $\psi \in \mathcal{F}_b$  with  $\Lambda(\psi)^2 < \infty$ ,  $\psi \in \mathcal{A}(1; \phi, 1, 1)$ , where  $\phi(1, \psi, r) = r\Lambda(\psi)^2$ . With this at hand, we deduce from Theorem 1.2 (or from (4.29)) that (i) implies (iii)' in Corollary 1.3.  $\square$

**Remark 4.7.** For anomalous symmetric Markov processes (for example, symmetric diffusions on some fractals that satisfy a sub-Gaussian estimate), the energy measure  $\Gamma(\cdot, \cdot)$  is singular with respect to the underlying reference measure  $m$ ; see [Ku, H, KM]. In particular, in these contexts, the energy measure  $\Gamma(u, u)$  is singular with respect to  $m$  unless  $u$  is a constant, and so the off-diagonal estimates in Theorem 1.2 (iii) degenerate into the on-diagonal estimates. However, motivated by [MS1, MS2], we can check that (4.2) holds true by making use of the so-called cutoff Sobolev

inequality. The reader is referred to [MS1, Proposition 2.4] and [MS2, Lemma 3.5] for details on anomalous diffusions with walk dimension  $d_w > 2$  and on anomalous heavy tailed random walks with walk dimension  $d_w \geq 2$  in the  $d$ -set setting, respectively. Note that, for those two processes, their on-diagonal heat kernels are of polynomials decay (i.e., of the form  $c t^{-d/d_w}$ ). With Theorem 4.4 at hand, we can extend Theorem 1.2 and Corollary 1.3 to these singular settings, and can handle anomalous diffusions, anomalous symmetric jump processes as well as anomalous symmetric diffusions with jumps in general measure spaces with more general scaling functions; see [GT, CKW1, CKW2].

**Proof of Corollary 1.4.** For any pre-compact open set  $D$ , consider the following Dirichlet semigroup associated with  $(P_t)_{t \geq 0}$ :

$$P_t^D f(x) = \mathbb{E}^x(f(X_t) \mathbb{1}_{\{\tau_D > t\}}), \quad f \in L^2(D; m), t > 0,$$

where  $\tau_D$  is the first exit time from the open set  $D$  of the Hunt process  $X = \{X_t, t \geq 0; \mathbb{P}_x, x \in E \setminus \mathcal{N}\}$  associated with the Dirichlet form  $(\mathcal{E}, \mathcal{F})$ .

Suppose that (ii) in Theorem 1.2 holds with  $\delta = 0$ . Then, by [BBCK, Theorem 3.1], we know that the semigroup  $(P_t^D)_{t \geq 0}$  has a heat kernel  $p^D(t, x, y)$  defined on  $(0, \infty) \times (D \setminus \mathcal{N}) \times (D \setminus \mathcal{N})$  for some properly exceptional set  $\mathcal{N}$ . In the following, we extend  $p^D(t, x, y)$  to  $(0, \infty) \times D \times D$  by setting  $p^D(t, x, y) = 0$  when  $x \in \mathcal{N}$  or  $y \in \mathcal{N}$ . According to (ii) in Theorem 1.2 with  $\delta = 0$  again and the fact that  $m(D) < \infty$ , we know that

$$\int_D p^D(t, x, x) m(dx) < \infty$$

for all  $t > 0$ , and so  $P_t^D$  is a Hilbert-Schmidt operator for all  $t > 0$ . Thus, the semigroup  $(P_t^D)_{t \geq 0}$  is compact. By general theory of semigroups for compact operators, there exists an orthonormal basis of eigenfunctions  $\{\phi_n : n \geq 1\} \subset L^2(D; m)$  associated with corresponding eigenvalues  $\{\lambda_n(D) : n \geq 1\}$  satisfying  $0 < \lambda_1(D) < \lambda_2(D) \leq \dots$  and  $\lim_{n \rightarrow \infty} \lambda_n(D) = \infty$ . In particular, for any  $m$ -a.e.  $x, y \in D$  and  $t > 0$ ,

$$p^D(t, x, y) = \sum_{k=1}^{\infty} \exp(-\lambda_k(D)t) \phi_k(x) \phi_k(y).$$

With this expression for  $p^D(t, x, y)$ , that (ii) implies (v) follows from the proof of [G1, Theorem 2.2], also thanks to the property (ii) for the regular function given in Definition 1.1.

It is trivial that (v) implies (iv). Concerning the implication of (iv)  $\Rightarrow$  (ii) we refer to [GH, Lemma 5.4]. Indeed, according to the proof of [GH, Lemma 5.4], (1.9) yields that for any  $f \in \mathcal{F}$  with compact support,

$$c_1 \|f\|_2^2 \Theta(c_2 / \|f\|_2^2) \leq \mathcal{E}(f, f), \quad \|f\|_1 = 1,$$

where  $c_1, c_2$  are positive constants independent of  $f$ . This with (4.28) finishes the proof.  $\square$



## 5 Application and Examples

### 5.1 Application: off-diagonal Dirichlet heat kernel upper bounds

In this subsection, we assume that

- (i)  $V(x, r)$  is a strictly positive function on  $E \times \mathbb{R}_+$  such that  $r \mapsto V(x, r)$  is increasing on  $\mathbb{R}_+$  for any fixed  $x \in E$ , and that there exist constants  $0 < d_1 \leq d_2$ ,  $0 < C_1 \leq C_2$  and  $C_3 \geq 1$  so that

$$C_1 \left(\frac{R}{r}\right)^{d_1} \leq \frac{V(x, R)}{V(x, r)} \leq C_2 \left(\frac{R}{r}\right)^{d_2} \quad \text{for all } x \in E \text{ and } 0 < r \leq R \quad (5.1)$$

and

$$V(x, r) \leq C_3 V(y, r) \quad \text{for all } r > 0 \text{ and } x, y \in E \text{ with } \rho(x, y) < r. \quad (5.2)$$

- (ii) Let  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a strictly increasing continuous function such that  $\phi(0) = 0$  and there exist constants  $C_*, \beta_1 > 0$  so that

$$\frac{\phi(R)}{\phi(r)} \geq C_* \left(\frac{R}{r}\right)^{\beta_1} \quad \text{for all } 0 < r \leq R. \quad (5.3)$$

For any  $x \in E$  and  $R > 0$ , set

$$\beta_{x,R}(r) := \frac{1}{V(x, R)} \max \left\{ \left(\frac{R}{\phi^{-1}(r)}\right)^{d_2}, \left(\frac{R}{\phi^{-1}(r)}\right)^{d_1} \right\}, \quad r > 0. \quad (5.4)$$

Note that  $\beta_{x,R} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is decreasing with  $\beta_{x,R}(0) = \infty$  and  $\beta_{x,R}(\infty) = 0$ . By some elementary calculations, we can show that for all  $x \in E$  and  $R > 0$ ,  $1/\beta_{x,R}$  has the doubling property uniformly such that for any  $0 < s \leq t$ ,

$$\frac{\beta_{x,R}(s)}{\beta_{x,R}(t)} \leq C_*^{-1/\beta_1} (t/s)^{d_2/\beta_1}. \quad (5.5)$$

This is because, when  $0 < s \leq t \leq \phi(R)$ ,

$$\frac{\beta_{x,R}(s)}{\beta_{x,R}(t)} = \frac{\phi^{-1}(t)^{d_2}}{\phi^{-1}(s)^{d_2}} \leq C_*^{-1/\beta_1} (t/s)^{d_2/\beta_1};$$

when  $0 < s < \phi(R) < t$ ,

$$\frac{\beta_{x,R}(s)}{\beta_{x,R}(t)} = R^{d_2-d_1} \frac{\phi^{-1}(t)^{d_1}}{\phi^{-1}(s)^{d_2}} \leq \frac{\phi^{-1}(t)^{d_2}}{\phi^{-1}(s)^{d_2}} \leq C_*^{-1/\beta_1} (t/s)^{d_2/\beta_1};$$

when  $\phi(R) \leq s \leq t$ ,

$$\frac{\beta_{x,R}(s)}{\beta_{x,R}(t)} = \frac{\phi^{-1}(t)^{d_1}}{\phi^{-1}(s)^{d_1}} \leq C_*^{-1/\beta_1} (t/s)^{d_1/\beta_1} \leq C_*^{-1/\beta_1} (t/s)^{d_2/\beta_1}.$$

Define

$$\Psi(r) = \int_r^\infty \frac{\beta_{x,R}^{-1}(s)}{s} ds, \quad r > 0.$$

We claim that there is a constant  $c_1 := c_1(C_*, \beta_1, d_2) \geq 1$  (independent of  $x$  and  $R$ ) such that for all  $r \in \mathbb{R}_+$ ,

$$\Psi(r) \leq c_1 \beta_{x,R}^{-1}(r). \quad (5.6)$$

Consequently, by (5.5) there is a constant  $c_2 := c_2(C_*, \beta_1, d_2) > 1$  (independent of  $x$  and  $R$ ) so that

$$\Psi^{-1}(s) \leq \beta_{x,R}(s/c_1) \leq c_2 \beta_{x,R}(s) \quad \text{for } s > 0. \quad (5.7)$$

Indeed, (5.5) implies that  $0 < s \leq t$ ,

$$\frac{\beta_{x,R}^{-1}(t)}{\beta_{x,R}^{-1}(s)} \leq C_*^{-1/d_2} (s/t)^{\beta_1/d_2}.$$

Thus,

$$\begin{aligned} \Psi(r) &= \sum_{k=0}^{\infty} \int_{2^k r}^{2^{k+1} r} \frac{\beta_{x,R}^{-1}(s)}{s} ds \leq \sum_{k=0}^{\infty} \frac{\beta_{x,R}^{-1}(2^k r)}{2^k r} (2^{k+1} r - 2^k r) \\ &\leq C_*^{-1/d_2} \beta_{x,R}^{-1}(r) \sum_{k=0}^{\infty} 2^{-k\beta_1/d_2} = \frac{C_*^{-1/d_2}}{1 - 2^{-\beta_1/d_2}} \beta_{x,R}^{-1}(r). \end{aligned}$$

This establishes the claim (5.6).

Suppose that  $(\mathcal{E}, \mathcal{F})$  is a symmetric regular Dirichlet form on  $L^2(E; m)$  having no killings inside  $E$ , and that  $(P_t)$  is its associated semigroup. For an open set  $B \subset E$ , let  $(P_t^B)_{t \geq 0}$  be the Dirichlet heat semigroup associated with  $(P_t)_{t \geq 0}$  with the Dirichlet conditions on  $B^c$ , i.e.,

$$P_t^B f(x) = \mathbb{E}_x[f(X_t) \mathbb{1}_{\{t < \tau_B\}}], \quad f \in L^2(B; m), \quad t > 0, \quad x \in B,$$

where  $\tau_B$  is the first exit time from the open set  $B$  by the Hunt process  $X = \{X_t, t \geq 0; \mathbb{P}_x, x \in E \setminus \mathcal{N}\}$  associated with the Dirichlet form  $(\mathcal{E}, \mathcal{F})$ .

**Theorem 5.1.** *Assume that (5.1), (5.2) and (5.3) hold. Suppose that the semigroup  $(P_t)_{t \geq 0}$  has a heat kernel  $p(t, x, y)$  defined on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$  with*

$$p(t, x, x) \leq \frac{c}{V(x, \phi^{-1}(t))}, \quad t > 0, \quad x \in E \setminus \mathcal{N}, \quad (5.8)$$

where  $c > 0$  is independent of  $x$  and  $t$ . Then, for any  $x_0 \in M$  and  $R > 0$ , the Dirichlet heat semigroup  $(P_t^{B(x_0, R)})_{t \geq 0}$  has a heat kernel  $p^{B(x_0, R)}(t, x, y)$  defined on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$  such that

- (i)  $p^{B(x_0, R)}(t, x, y) = 0$  if either  $x$  or  $y$  is in  $B(x_0, R)^c \setminus \mathcal{N}$ ;

(ii) for any  $\varepsilon > 0$  and  $\psi \in \mathcal{A}(C_0; \phi, C, \eta)$ , there is a constant  $C_\varepsilon := C_\varepsilon(C, \eta, d_1, d_2, C_0, C_1, C_2, C_3, C_*, \beta_1, c) > 0$  such that for all  $t > 0$  and  $x, y \in B(x_0, R) \setminus \mathcal{N}$ ,

$$p^{B(x_0, R)}(t, x, y) \leq C_\varepsilon \beta_{x_0, R}(t) \exp(-|\psi(y) - \psi(x)| + (1 + \varepsilon)\phi(C_0, \psi, 1)t),$$

where  $\beta_{x_0, R}(r)$  is defined by (5.4). In particular, for any  $\varepsilon > 0$ , there is a constant  $C_\varepsilon := C_\varepsilon(d_1, d_2, C_1, C_2, C_3, C_*, \beta_1, c) > 0$  so that for all  $t > 0$ ,  $x_0 \in E$ ,  $R > 0$  and  $x, y \in B(x_0, R) \setminus \mathcal{N}$ ,

$$p^{B(x_0, R)}(t, x, y) \leq C_\varepsilon \beta_{x_0, R}(t) \exp(-|\psi(y) - \psi(x)| + (1 + \varepsilon)\Lambda(\psi)^2 t),$$

where

$$\Lambda(\psi)^2 := \max \left\{ \left\| \frac{de^{-2\psi}\Gamma(e^\psi, e^\psi)}{dm} \right\|_\infty, \left\| \frac{de^{2\psi}\Gamma(e^{-\psi}, e^{-\psi})}{dm} \right\|_\infty \right\} < \infty.$$

To prove Theorem 5.1, we need to establish functional inequalities for the Dirichlet form  $(\mathcal{E}, \mathcal{F}_B)$  associated with  $(P_t^B)_{t \geq 0}$ , where  $\mathcal{F}_B = \{u \in \mathcal{F} : u = 0 \text{ } \mathcal{E}\text{-q.e. on } B^c\}$ .

**Lemma 5.2.** *Suppose that (5.1), (5.2) and (5.8) hold. Then, there exists a constant  $\widehat{C} := \widehat{C}(c, C_1, C_2, C_3) > 0$  such that for any  $x_0 \in E$  and  $R > 0$ , the following super-Poincaré inequality for  $(\mathcal{E}, \mathcal{F}_{B(x_0, R)})$  holds:*

$$\|u\|_2^2 \leq r\mathcal{E}(u, u) + \widehat{C}\beta_{x_0, R}(r) \|u\|_1^2, \quad r > 0, u \in \mathcal{F}_{B(x_0, R)}, \quad (5.9)$$

where  $\beta_{x, R}$  is defined by (5.4).

**Proof.** By the Cauchy-Schwarz inequality, for any  $x, y \in E \setminus \mathcal{N}$  and  $t > 0$ ,

$$p(t, x, y) \leq (p(t, x, x)p(t, y, y))^{1/2} \leq \max\{p(t, x, x), p(t, y, y)\}.$$

This along with [Ki, Theorem 2.1] and (5.8) yields that for all  $u \in \mathcal{F} \cap L^1(E; m)$  and  $t > 0$ ,

$$\begin{aligned} \|u\|_2^2 &\leq t\mathcal{E}(u, u) + \|u\|_1^2 \sup_{x \in \text{supp}(u)} p(t, x, x) \\ &\leq t\mathcal{E}(u, u) + \|u\|_1^2 \frac{c}{\inf_{x \in \text{supp}(u)} V(x, \phi^{-1}(t))}. \end{aligned}$$

Consequently, for all  $u \in \mathcal{F} \cap L^1(E; m)$  and  $s > 0$ ,

$$\|u\|_2^2 \leq \phi(s)\mathcal{E}(u, u) + \frac{c \|u\|_1^2}{\inf_{z \in \text{supp}(u)} V(z, s)}. \quad (5.10)$$

For fixed  $x_0 \in E$  and  $R > 0$ , set  $B = B(x_0, R)$ . By (5.1) and (5.2), we have that for any  $z \in B$  and  $s > 0$ ,

$$\frac{V(x_0, R)}{V(z, s)} \leq C_3 \frac{V(z, R)}{V(z, s)} \leq C_3(C_1^{-1} \vee C_2) \max \left\{ \left(\frac{R}{s}\right)^{d_2}, \left(\frac{R}{s}\right)^{d_1} \right\}.$$

Combining this with (5.10), we obtain that for any  $u \in \mathcal{F} \cap C_c(B) \subset \mathcal{F}_B$ , and  $s > 0$ ,

$$\begin{aligned} \|u\|_2^2 &\leq \phi(s)\mathcal{E}(u, u) + \frac{\widehat{C}}{V(x_0, R)} \max \left\{ \left( \frac{R}{s} \right)^{d_2}, \left( \frac{R}{s} \right)^{d_1} \right\} \|u\|_1^2 \\ &= \phi(s)\mathcal{E}(u, u) + \widehat{C}\beta_{x_0, R}(s)\|u\|_1^2, \end{aligned}$$

where  $\widehat{C} = cC_3(C_1^{-1} \vee C_2)$ . As  $\mathcal{F}_b \cap C_c(B)$  is  $\mathcal{E}_1$ -dense in  $\mathcal{F}_B$  and uniformly dense in  $C_c(B)$ , the desired assertion follows.  $\square$

**Proof of Theorem 5.1** Note that the existence of  $p^{B(x_0, R)}(t, x, y)$  as well as the property (i) of  $p^{B(x_0, R)}(t, x, y)$  follow from the existence of  $p(t, x, y)$  and the strong Markov property of the process  $X$ . So it remains to establish property (ii).

In the following, we fix  $x_0 \in E$  and  $R > 0$ , and set  $B = B(x_0, R)$ . For any  $\psi \in \mathcal{A}(C_0; \phi, C, \eta)$ , non-negative  $f \in \mathcal{F}_b$ ,  $t > 0$  and  $p \geq 1$ ,

$$f_t(x) := P_t^{B, \psi} f(x) = e^{\psi(x)} P_t^B(e^{-\psi} f)(x), \quad f_{p,t}(x) := \exp(-\phi(C_0, \psi, p)t) f_t(x), \quad x \in B.$$

Note that  $f_t \in \mathcal{F}_B$  and so  $f_{p,t} \in \mathcal{F}_B$ . According to (4.2) and (5.9), for any  $p \geq 1$ ,  $t > 0$  and  $r > 0$ ,

$$\begin{aligned} \frac{d\|f_{p,t}\|_{2p}^{2p}}{dt} &= -2p\phi(C_0, \psi, p)\|f_{p,t}\|_{2p}^{2p} - 2p\mathcal{E}(e^{\psi} f_{p,t})^{2p-1}, e^{-\psi} f_{p,t}) \leq -2C_0\mathcal{E}(f_{p,t}^p, f_{p,t}^p) \\ &\leq 2C_0 \left( -\frac{1}{r}\|f_{p,t}\|_{2p}^{2p} + \frac{\widehat{C}\beta_{x_0, R}(r)}{r}\|f_{p,t}\|_p^{2p} \right). \end{aligned}$$

Taking

$$r = \beta_{x_0, R}^{-1} \left( \frac{\|f_{p,t}\|_{2p}^{2p}}{2\widehat{C}\|f_{p,t}\|_p^{2p}} \right),$$

we have

$$\frac{d\|f_{p,t}\|_{2p}^{2p}}{dt} \leq -\frac{C_0\|f_{p,t}\|_{2p}^{2p}}{\beta_{x_0, R}^{-1} \left( \frac{\|f_{p,t}\|_{2p}^{2p}}{2\widehat{C}\|f_{p,t}\|_p^{2p}} \right)} < 0 \quad \text{for every } t > 0. \quad (5.11)$$

Define

$$\Psi(r) = \int_r^\infty \frac{\beta_{x_0, R}^{-1}(s)}{C_0 s} ds, \quad r > 0.$$

Since, in view of (5.5),  $t \mapsto 1/\beta_{x_0, R}(t)$  has the doubling property uniformly for all  $x_0 \in E$  and  $R > 0$ , by Lemma 2.7 and its proof, (4.10) hold for  $\beta_{x_0, R}$  with common constants  $c(\lambda) \equiv 1$  and  $C(\lambda)$  for all  $x_0 \in E$  and  $R > 0$ . With this fact, (5.6), (5.7) and (5.11), one can repeat the proof of Proposition 4.5 to see that, for any  $\varepsilon > 0$ , there is a constant  $C_\varepsilon := C_\varepsilon(C, \eta, d_1, d_2, C_0, C_1, C_2, C_3, C_*, \beta_1, c) > 0$  (uniformly on  $x_0$  and  $R$ ) so that

$$\|P_t^{B, \psi}\|_{L^1(B; m) \rightarrow L^\infty(B; m)} \leq C_\varepsilon \beta_{x_0, R}(t) \exp((1 + \varepsilon)\phi(C_0, \psi, 1)t), \quad t > 0.$$

Therefore, the desired property (ii) for  $p^{B(x_0, R)}(t, x, y)$  follows from the proofs of Theorems 4.4 and 1.2.  $\square$

## 5.2 Examples

In this part, we give some examples to illustrate the applications of Theorem 1.2 and Corollary 1.3.

**Example 5.3.** [*Heat kernel for Brownian motion on manifolds*] Let  $(M, \rho)$  be a non-compact manifold with bounded geometry (i.e.,  $M$  has a positive injectivity radius  $r_0$ , and its Ricci curvature is bounded from below) and with the Riemannian volume measure  $m$ . Let  $p(t, x, y)$  be the heat kernel on  $M$ ; that is, it is the smallest positive fundamental solution to the heat equation

$$\frac{\partial u}{\partial t} = \Delta u$$

on  $\mathbb{R}_+ \times M$ . It was proven in [BCG, Theorem 1.1] that, if  $\dim M = d$  and there is a continuous positive strictly increasing function  $v(r)$  on  $[r_0, \infty)$  such that for all  $x \in M$  and  $r \geq r_0$ ,  $m(B(x, r)) \geq v(r)$ , then there are constants  $C, c > 0$  such that for all  $t > 0$  and  $x, y \in M$ ,

$$p(t, x, y) \leq \varphi(t) := \begin{cases} Ct^{-d/2}, & t < r_0^2, \\ C/\gamma(ct), & t \geq r_0^2, \end{cases}$$

where  $\gamma(t)$  is defined by

$$t = r_0^2 + \int_{v(r_0^2)}^{\gamma(t)} v^{-1}(s) ds, \quad t \geq r_0^2.$$

Now suppose that  $\gamma(r)$  is an increasing doubling function on  $[r_0, \infty)$  with  $\gamma(\infty) := \lim_{r \rightarrow \infty} \gamma(r) = \infty$ . By Proposition 2.8, there exist  $\bar{\varphi} \in \mathcal{R}$  and a constant  $c \geq 1$  such that

$$c^{-1}\bar{\varphi}(r) \leq \varphi(r) \leq c\bar{\varphi}(r), \quad r \in \mathbb{R}_+.$$

Then, according to Corollary 1.3, for any  $\varepsilon > 0$ , there is a constant  $C_\varepsilon > 0$  so that for all  $x, y \in M$  and  $t > 0$ ,

$$p(t, x, y) \leq C_\varepsilon \tilde{\varphi}(t) \exp\left(-\frac{\rho(x, y)^2}{4(1 + \varepsilon)t}\right),$$

where

$$\tilde{\varphi}(t) := \begin{cases} t^{-d/2}, & t \leq r_0^2, \\ \gamma(t)^{-1}, & t > r_0^2. \end{cases}$$

Note that, when  $v(r) = c_1 e^{c_2 r^\alpha}$  with  $\alpha \in (0, 1]$  on  $[r_0, \infty)$ , it holds that for large  $t$ ,  $c_3^{-1} \frac{t}{(\log t)^{1/\alpha}} \leq \gamma(t) \leq c_3 \frac{t}{(\log t)^{1/\alpha}}$  with some  $c_3 \geq 1$ ; when  $v(r) = c_4 r^\beta$  with  $\beta \geq 1$  on  $[r_0, \infty)$  (due to the fact that for any non-compact manifold  $M$  with bounded geometry the volume growth is at least linear), it holds that for large  $t$ ,  $c_5 t^{\beta/(1+\beta)} \leq \gamma(t) \leq c_5 t^{\beta/(1+\beta)}$  for some  $c_5 \geq 1$ . In both cases  $\gamma(r)$  are doubling functions such that  $\gamma(\infty) = \infty$ .

We also note that one may apply [G1, Theorem 1.1] to obtain off-diagonal Gaussian estimates for this example as well.  $\square$

**Example 5.4.** [*Heat kernel for Brownian motion on hyperbolic spaces and its subordination*] Suppose that the Nash-type inequality (1.4) holds with  $\delta = 0$  and

$$\theta(r) = \max \left\{ r^{(\alpha+1)/\alpha}, r \log^{(\gamma-1)/\gamma}(2 + r^{-1}) \right\}$$

for some  $\alpha > 0$  and  $\gamma \in (0, 1]$ . Then, by Theorem 1.2 and the table in the end of Section 3, the on-diagonal upper bounds (1.5) and the upper bounds (1.7) hold with  $\varphi(t) = t^{-\alpha} e^{-t^\gamma}$  for the heat kernel of the associated semigroup  $(P_t)_{t \geq 0}$ .

A typical example is Brownian motion  $(B_t)_{t \geq 0}$  in the  $d$ -dimensional hyperbolic space  $\mathbb{H}_k^d$ , which is a simply connected complete  $d$ -dimensional Riemannian manifold with a constant negative sectional curvature  $-k^2$ . In particular, for  $d = 3$ , the assertion above holds with  $\alpha = 3/2$  and  $\gamma = 1$ ; e.g., see [G1, (1.3)]. The reader is also referred to [GN, Theorem 1.1] for the explicit formulas of the heat kernel on  $\mathbb{H}_1^d$  for all  $d \geq 1$ .

Let  $(S_t)_{t \geq 0}$  be a subordinator that is independent of  $(B_t)_{t \geq 0}$  and has the Laplace exponent  $h$ . Consider the subordinated process  $(X_t)_{t \geq 0} := (B_{S_t})_{t \geq 0}$  on the 3-dimensional hyperbolic space  $\mathbb{H}_k^3$ . By [SW, Theorem 1], the Nash-type inequality (1.4) holds with  $\delta = 0$  and  $\theta(r) = \max\{r h(r^{2/3}), r\}$  for the Dirichlet form associated with the subordinated process  $(X_t)_{t \geq 0}$ . In particular, let  $(S_t)_{t \geq 0}$  be a  $\beta$ -stable subordinator whose Laplace exponent is  $h(r) = r^\beta$  for some  $\beta \in (0, 1)$ . Then, (1.4) holds with  $\delta = 0$  and  $\beta(r) = \max\{r^{1+(2\beta/3)}, r\}$ , and so, by Theorem 1.2, the on-diagonal upper bounds (1.5) and the upper bounds (1.7) hold with  $\varphi(t) = t^{-3/(2\beta)} e^{-t}$  for this  $\beta$ -stable subordinated Brownian motion on the 3-dimensional hyperbolic space. This gives us a concrete example of symmetric jump process whose heat kernel decays exponentially for large time.

It is now known that Davies' method presented in [CKS] is very useful to obtain off-diagonal upper bounds for heat kernels of symmetric jump processes, e.g., see [BGK, CKK1, CK1, CK2, CK3]. All the works mentioned above except [CK2, CK3] around Davies' method heavily relied on the assumption that the heat kernel  $p(t, x, y)$  enjoys the polynomial decay in on-diagonal estimate (that is, (1.1) holds with  $\varphi(t) = c t^{-\nu}$  for some  $c, \nu > 0$ ).

In [CK2] (see also [CK3, CKW1]), the correct on-diagonal bound is of the form  $c_0(\Phi^{-1}(t))^{-d}$  for some strictly increasing weighted function  $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  which satisfies doubling and reverse doubling properties; in particular, it is not necessarily of power function type and so one cannot directly apply [CKS, Theorem (3.25)] to get heat kernel off-diagonal upper bound. The approach of [CK2] is first to replace  $c_0(\Phi^{-1}(t))^{-d}$  by a rough upper bound  $c_1 t^{-\theta}$  with some  $c_1, \theta > 0$  for  $t \in (0, 1]$ , and use [CKS, Theorem (3.25)] to obtain a preliminary off-diagonal upper estimate, and then bootstrap it to a precise off-diagonal estimate by a suitable scaling procedure. However, this approach does not work in some other settings such as reflected diffusions with jumps in [CKKW].

Below we give an example of symmetric jump process on  $\mathbb{R}^d$  to show that some estimates in [CK2] can be directly derived by using Theorem 1.2. Note that, the approach of the example below can be easily extended to metric measure spaces setting with uniformly volume doubling and

reversed doubling properties. In particular, with Theorem 1.2 in hand, we can obtain [CK2, Lemma 4.3] directly without using a scaling argument.

**Example 5.5.** [*Symmetric Lévy-like processes with general scaling functions*] Consider a regular symmetric non-local Dirichlet form  $(\mathcal{E}, \mathcal{F})$  on  $L^2(\mathbb{R}^d; dx)$  as follows:

$$\begin{aligned}\mathcal{E}(u, v) &= \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d \setminus \text{diag}} (u(x) - u(y))(v(x) - v(y)) J(x, y) dx dy, \\ \mathcal{F} &= \{u \in L^2(\mathbb{R}^d; dx) : \mathcal{E}(u, u) < \infty\},\end{aligned}$$

where

$$\frac{c^{-1}}{|x - y|^d \phi(|x - y|)} \leq J(x, y) \leq \frac{c}{|x - y|^d \phi(|x - y|)}, \quad x \neq y \quad (5.12)$$

with some  $c \geq 1$  and a strictly increasing function  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  satisfying that

$$\int_0^1 \frac{s}{\phi(s)} ds < \infty$$

and

$$C_L(R/r)^{\beta_1} \leq \phi(R)/\phi(r) \leq C_U(R/r)^{\beta_2}, \quad 0 < r < R < \infty \quad (5.13)$$

for some  $C_L, C_U > 0$ ,  $\beta_1 \in (0, 2)$  and  $\beta_2 \geq \beta_1$ . It was proven in [BKKL, Theorem 3.4] that the Nash-type inequality (1.3) holds with  $\theta(r) = c_0 r / \Phi(r^{-1/d})$ , where

$$\Phi(r) = \frac{r^2}{2 \int_0^r \frac{s}{\phi(s)} ds}.$$

We note that, by [BKKL, Section 2.1],  $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is an increasing function such that

$$\Phi(r) < \phi(r), \quad r \in \mathbb{R}_+ \quad (5.14)$$

and

$$C_L(R/r)^{\beta_1} \leq \Phi(R)/\Phi(r) \leq (R/r)^2 \wedge (C_U(R/r)^{\beta_2}), \quad 0 < r < R < \infty. \quad (5.15)$$

In particular,  $\Phi$  has the doubling property on  $\mathbb{R}_+$  with  $\Phi(0) = 0$  and  $\Phi(\infty) = \infty$ . Therefore, according to Proposition 3.1(i) and [BBCK, Theorem 3.1], there are a properly exceptional set  $\mathcal{N}$  and a heat kernel  $p(t, x, y)$  associated with the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  and defined on  $(0, \infty) \times (\mathbb{R}^d \setminus \mathcal{N}) \times (\mathbb{R}^d \setminus \mathcal{N})$  such that for any  $t > 0$  and  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$ ,

$$p(t, x, y) \leq \frac{c_1}{(\Phi^{-1}(t))^d} \quad (5.16)$$

for some  $c_1 > 0$ .

To consider off-diagonal upper bounds for  $p(t, x, y)$ , we will adopt the truncated argument as in [BGK, CKK1, CK1, CK2]. For each  $\rho > 0$ , we define a bilinear form  $(\mathcal{E}^{(\rho)}, \mathcal{F})$  by

$$\mathcal{E}^{(\rho)}(u, v) = \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} (u(x) - u(y))(v(x) - v(y)) \mathbb{1}_{\{|x-y| \leq \rho\}} J(x, y) dx dy.$$

It follows from (5.12) and (5.13) that there exists a constant  $c_2 > 0$  such that for all  $\rho > 0$  and  $u \in \mathcal{F}$ ,

$$\begin{aligned} \mathcal{E}(u, u) &= \mathcal{E}^{(\rho)}(u, u) + \frac{1}{2} \int_{\{|x-y| \geq \rho\}} (u(x) - u(y))^2 J(x, y) dx dy \\ &\leq \mathcal{E}^{(\rho)}(u, u) + 2 \left( \sup_{x \in \mathbb{R}^d} \int_{\{|x-y| \geq \rho\}} J(x, y) dy \right) \|u\|_2^2 \\ &\leq \mathcal{E}^{(\rho)}(u, u) + \frac{c_2}{\phi(\rho)} \|u\|_2^2. \end{aligned}$$

This implies that  $(\mathcal{E}^{(\rho)}, \mathcal{F})$  is also a regular symmetric Dirichlet form on  $L^2(\mathbb{R}^d; dx)$ , and that the Nash-type inequality (1.4) holds for  $(\mathcal{E}^{(\rho)}, \mathcal{F})$  with  $\theta(r) = c_0 r / \Phi(r^{-1/d})$  and  $\delta = c_2 / \phi(\rho)$ . As mentioned above,  $\Phi$  has the doubling property on  $\mathbb{R}_+$  such that  $\Phi(0) = 0$  and  $\Phi(\infty) = \infty$ , and so  $r \mapsto \frac{1}{[\Phi^{-1}(r)]^d}$  is a regular function, thanks to Proposition 2.8. According to Theorem 1.2, there is a heat kernel  $p^\rho(t, x, y)$  associated with the Dirichlet form  $(\mathcal{E}^{(\rho)}, \mathcal{F})$  and defined on  $(0, \infty) \times (\mathbb{R}^d \setminus \mathcal{N}) \times (\mathbb{R}^d \setminus \mathcal{N})$  such that for any  $t > 0$ ,  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$  and  $\psi \in \mathcal{F}_b$ ,

$$p^\rho(t, x, y) \leq \frac{c_3 e^{c_2 t / \phi(\rho)}}{(\Phi^{-1}(t))^d} \exp\left(-|\psi(x) - \psi(y)| + c_4 t (\|\Gamma_\rho[\psi]\|_\infty \vee \|\Gamma_\rho[-\psi]\|_\infty)\right),$$

where

$$\Gamma_\rho[\psi](x) := \frac{1}{2} \int_{\{|x-y| \leq \rho\}} (e^{\psi(x) - \psi(y)} - 1)^2 J(x, y) dy.$$

Here, for simplicity we take the same properly exceptional set  $\mathcal{N}$  as before.

Next, we fix  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$  and take

$$\psi(z) := \frac{s}{3} (|z - x| \wedge |x - y|) \quad \text{for } z \in \mathbb{R}^d,$$

where  $s > 0$  is a constant to be chosen later. By the facts that  $(1 - e^r)^2 \leq r^2 e^{2|r|}$  for  $r \in \mathbb{R}^1$  and

$$|\psi(z_1) - \psi(z_2)| \leq \frac{s}{3} \left| |z_1 - x| - |z_2 - x| \right| \leq \frac{s}{3} |z_1 - z_2|, \quad z_1, z_2 \in \mathbb{R}^d,$$

it holds for every  $z \in \mathbb{R}^d$  that

$$\begin{aligned} \Gamma_\rho[\psi](z) &= \frac{1}{2} \int_{\{|z-w| \leq \rho\}} (1 - e^{\psi(z) - \psi(w)})^2 J(z, w) dw \\ &\leq \frac{1}{2} \int_{\{|z-w| \leq \rho\}} (\psi(z) - \psi(w))^2 e^{2|\psi(z) - \psi(w)|} J(z, w) dw \\ &\leq \frac{1}{2} \left(\frac{s}{3}\right)^2 e^{2s\rho/3} \int_{\{|z-w| \leq \rho\}} |z - w|^2 J(z, w) dw \\ &\leq c_5 s^2 e^{2s\rho/3} \int_0^\rho \frac{v}{\phi(v)} dv = \frac{c_5}{2} s^2 e^{2s\rho/3} \frac{\rho^2}{\Phi(\rho)} \leq c_6 \frac{e^{s\rho}}{\Phi(\rho)}, \end{aligned}$$



where in the fourth inequality we used (5.12), and the last inequality follows from the inequality  $v^2 e^{2v/3} \leq 9e^v$  for  $v > 0$ . Thus,

$$\|\Gamma_\rho[\psi]\|_\infty \vee \|\Gamma_\rho[-\psi]\|_\infty \leq c_6 \frac{e^{s\rho}}{\Phi(\rho)}.$$

Therefore, we arrive at that for any  $t > 0$ ,  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$ ,

$$p^\rho(t, x, y) \leq \frac{c_3 e^{c_2 t / \psi(\rho)}}{(\Phi^{-1}(t))^d} \exp\left(-\frac{s|x-y|}{3} + c_6 \frac{te^{s\rho}}{\Phi(\rho)}\right). \quad (5.17)$$

We now further assume that  $t \leq \Phi(|x-y|)$ . Define  $\gamma := \frac{\beta_1}{3(d+\beta_1)}$ . Take  $\rho = \gamma|x-y|$  and  $s = \frac{1}{\gamma|x-y|} \log\left(\frac{\Phi(|x-y|)}{t}\right)$ . Then, by (5.15),

$$-\frac{s|x-y|}{3} + c_6 \frac{te^{s\rho}}{\Phi(\rho)} = \frac{1}{3\gamma} \log\left(\frac{t}{\Phi(|x-y|)}\right) + c_6 \frac{\Phi(|x-y|)}{\Phi(\rho)} \leq \frac{1}{3\gamma} \log\left(\frac{t}{\Phi(|x-y|)}\right) + \frac{c_6}{\gamma^2}.$$

On the other hand, due to (5.14) and (5.15),

$$\frac{c_2 t}{\phi(\rho)} \leq \frac{c_2 t}{\Phi(\rho)} \leq \frac{c_2 \Phi(|x-y|)}{\Phi(\rho)} \leq c_7.$$

Thus, combining both estimates above with (5.15) and (5.17), we obtain that for any  $t > 0$  and  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$ ,

$$\begin{aligned} p^\rho(t, x, y) &\leq \frac{c_8}{(\Phi^{-1}(t))^d} \left(\frac{t}{\Phi(|x-y|)}\right)^{1/3\gamma} \\ &= c_8 \left(\frac{\Phi^{-1}(\Phi(|x-y|))}{\Phi^{-1}(t)}\right)^d \left(\frac{t}{\Phi(|x-y|)}\right)^{1+d/\beta_1} \frac{1}{|x-y|^d} \\ &\leq c_9 \left(\frac{\Phi(|x-y|)}{t}\right)^{d/\beta_1} \left(\frac{t}{\Phi(|x-y|)}\right)^{1+d/\beta_1} \frac{1}{|x-y|^d} \\ &= \frac{c_9 t}{|x-y|^d \Phi(|x-y|)}. \end{aligned}$$

This along with [BGK, Lemma 3.1], (5.12), (5.14) and (5.15) yields that for every  $0 < t \leq \Phi(|x-y|)$  and  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$ ,

$$p(t, x, y) \leq \frac{c_9 t}{|x-y|^d \Phi(|x-y|)} + \frac{ct}{\phi(\gamma|x-y|)(\gamma|x-y|)^d} \leq \frac{c_{10} t}{|x-y|^d \Phi(|x-y|)}. \quad (5.18)$$

We conclude from (5.16) and (5.18) that for every  $t > 0$  and  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$ ,

$$p(t, x, y) \leq c_{11} \left( \frac{1}{(\Phi^{-1}(t))^d} \wedge \frac{t}{|x-y|^d \Phi(|x-y|)} \right). \quad (5.19)$$

□

**Remark 5.6.** (i) The proof above simplifies the proof of [BKKL, Theorem 3.8] without using the scaling argument. Moreover, either by following the same arguments in the above proof or by (5.19) with the same arguments as these in the proofs of [CKK2, Theorem 2.4] and [ChKS, Proposition 2.2], one can show that, if we assume  $\sup_{x \in \mathbb{R}^d} \int_{\{|y-x| \geq 1\}} J(x, y) dy < \infty$  and (5.12) holds only for  $0 < |x - y| \leq 1$ , then the upper bound in (5.19) holds true for  $0 < t \leq 1$  and  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$  with  $|x - y| \leq 1$ . Note that, if  $\psi(r) = r^2 \ln^{1+\beta}(2/r)$  for all  $r \in (0, 1]$  with  $\beta > 0$ , then  $c_1 r^2 \ln^\beta(2/r) \leq \Phi(r) \leq c_2 r^2 \ln^\beta(2/r)$  for  $0 < r \leq 1$ . Thus, we recover the main result of [Mi] (see [Mi, Theorem 1.1]) and in fact we get much more.

(ii) We should mention that the heat kernel upper bounds (5.19) are sharp when  $c^{-1}\phi(r) \leq \Phi(r) \leq c\phi(r)$  holds for all  $r > 0$  with some  $c \geq 1$ , see [CK2] for more details. On the other hand, even when the estimate (5.19) is not optimal in general setting, it is the first crucial step to obtain the optimal estimates, see [BKKL, CKW3].

**Acknowledgement.** The authors would thank Alexander Grigor'yan for fruitful discussions and for giving related references on manifolds. They also thank Naotaka Kajino for useful comments on the proof of [D2, Theorem 3.2.4]. The research of Zhen-Qing Chen is partially supported by Simons Foundation Grant 520542. The research of Panki Kim is supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2016R1E1A1A01941893). The research of Takashi Kumagai is supported by JSPS KAKENHI Grant Number JP17H01093 and by the Alexander von Humboldt Foundation. The research of Jian Wang is supported by the National Natural Science Foundation of China (Nos. 11831014 and 12071076), 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).

## References

- [BKKL] 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.
- [BCLS] D. Bakry, T. Coulhon, M. Ledoux and L. Saloff-Coste. Sobolev inequalities in disguise. *Indiana Univ. Math. J.* **44** (1995), 1033–1074.
- [BBCK] M.T. Barlow, R.F. Bass, Z.-Q. Chen and M. Kassmann. Non-local Dirichlet forms and symmetric jump processes. *Trans. Amer. Math. Soc.* **361** (2009), 1963–1999.
- [BCG] M.T. Barlow, T. Coulhon and A. Grigor'yan. Manifolds and graphs with slow heat kernel decay. *Invent. Math.* **144** (2001), 609–649.
- [BGK] M. Barlow, A. Grigor'yan and T. Kumagai. Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.* **626** (2009), 135–157.

- [CKS] E.A. Carlen, S. Kusuoka and D.W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. Heri. Poincaré-Probab. Statist.* **23** (1987), 245–287.
- [Ch] X.X. Chen. Pointwise upper estimates for transition probabilities of continuous time random walks on graphs. *Ann. Inst. Heri. Poincaré-Probab. Statist.* **53** (2017), 27–45.
- [CF] Z.-Q. Chen and M. Fukushima. *Symmetric Markov Processes, Time Change, and Boundary Theory*. Princeton University Press, 2012.
- [CKK1] Z.-Q. Chen, P. Kim and T. Kumagai. Global heat kernel estimates for symmetric jump processes. *Trans. Amer. Math. Soc.* **363** (2011), 5021–5055.
- [CKK2] 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.
- [CKKW] Z.-Q. Chen, P. Kim, T. Kumagai and J. Wang. Heat kernel estimates for reflected diffusions with jumps on metric measure spaces. In preparation.
- [ChKS] Z.-Q. Chen, P. Kim and R. Song. Dirichlet heat kernel estimates for rotationally symmetric Lévy processes. *Proc. Lond. Math. Soc.* **109** (2014), 90–120.
- [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.
- [CK3] Z.-Q. Chen and T. Kumagai. A priori Hölder estimate, parabolic Harnack principle and heat kernel estimates for diffusions with jumps. *Rev. Mat. Iberoam.* **26** (2010), 551–589.
- [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](https://arxiv.org/abs/1604.04035).
- [CKW2] Z.-Q. Chen, T. Kumagai and J. Wang. Heat kernel estimates and parabolic Harnack inequalities for symmetric Dirichlet forms. *Adv. Math.* **374** (2020), paper 107269.
- [CKW3] Z.-Q. Chen, T. Kumagai and J. Wang. Heat kernel estimates for general symmetric pure jump Dirichlet forms. available at [arXiv:1908.07655](https://arxiv.org/abs/1908.07655).
- [Co] T. Coulhon. Ultracontractivity and Nash type inequalities. *J. Funct. Anal.* **141** (1996), 510–539.
- [CGZ] T. Coulhon, A. Grigor’yan and F. Zucca. The discrete integral maximum principle and its applications. *Tohoku Math. J.* **57** (2005), 559–587.
- [CS] T. Coulhon and A. Sikora. Gaussian heat kernel upper bounds via the Phragmén-Lindelöf theorem. *Proc. London Math. Soc.* **96** (2008), 507–544.
- [D1] E.B. Davies. Explicit constants for Gaussian upper bounds on heat kernels. *Amer. J. Math.* **109** (1987), 319–333.
- [D2] E.B. Davies. *Heat Kernels and Spectral Theory*. Cambridge Univ. Press, Cambridge, UK, 1989.
- [F] M. Folz. Gaussian upper bounds for heat kernels of continuous time simple random walks. *Electron. J. Probab.* **16** (2011), 1693–1722.

- [FOT] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet Forms and Symmetric Markov Processes*. Second revised and extended edition. de Gruyter, Berlin, 2011.
- [G1] A. Grigor'yan. Heat kernel upper bounds on a complete non-compact manifold. *Revista Math. Iberoamericana* **10** (1994), 395–452.
- [G2] A. Grigor'yan. Gaussian upper bounds for heat kernel on arbitrary manifolds. *J. Differential Geom.* **45** (1997), 33–52.
- [GH] A. Grigor'yan and J. Hu. Upper bounds of heat kernels on doubling spaces. *Moscow Math. J.* **14** (2014), 505–563.
- [GN] A. Grigor'yan and M. Noguchi. The heat kernel on hyperbolic space. *Bull. London Math. Soc.* **30** (1998), 643–650.
- [GT] A. Grigor'yan and A. Telcs. Two-sided estimates of heat kernels on metric measure spaces. *Ann. Probab.* **40** (2012), 1212–1284.
- [H] M. Hino. On singularity of energy measures on self-similar sets. *Probab. Theory Related Fields* **132** (2005), 265–290.
- [KM] N. Kajino and M. Murugan. On singularity of energy measures for symmetric diffusions with full off-diagonal heat kernel estimates. To appear in *Ann. Probab.*, available at [arXiv:1910.02601](https://arxiv.org/abs/1910.02601).
- [Ki] J. Kigami. Local Nash inequality and inhomogeneity of heat kernels. *Proc. London Math. Soc.* **89** (2004), 525–544.
- [Ku] S. Kusuoka. Dirichlet forms on fractals and products of random matrices. *Publ. Res. Inst. Math. Sci.* **25** (1989), 659–680.
- [Mi] A. Mimica. Heat kernel upper estimates for symmetric jump processes with small jumps of high intensity. *Potential Anal.* **36** (2012), 203–222.
- [Mo] S.A. Molchanov. Diffusion processes and Riemannian geometry. *Uspekhi Mat. Nauk*, **30** (1975), 3–59; (English Trans.) *Russian Math. Surveys* **30** (1975), 1–63.
- [MS1] M. Murugan and L. Saloff-Coste. Davies' method for anomalous diffusions. *Proc. Amer. Math. Soc.* **145** (2017), 1793–1804.
- [MS2] M. Murugan and L. Saloff-Coste. Heat kernel estimates for anomalous heavy-tailed random walks. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **55** (2019), 697–719.
- [N] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.* **80** (1958), 931–954.
- [SW] R.L. Schilling and J. Wang. Functional inequalities and subordination: stability of Nash and Poincaré inequalities. *Math. Z.* **272** (2012), 921–936.
- [W] F.-Y. Wang. *Functional Inequalities, Markov Semigroups and Spectral Theory*. Elsevier Science, 2005.
- [V] N.Th. Varopoulos. Hardy-Littlewood theory for semigroups. *J. Funct. Anal.* **63** (1985), 240–260.

**Zhen-Qing Chen**

Department of Mathematics, University of Washington, Seattle, WA 98195, USA

E-mail: zqchen@uw.edu

**Panki Kim**

Department of Mathematical Sciences and Research Institute of Mathematics, Seoul National University, Seoul 08826, Republic of Korea

E-mail: pkim@snu.ac.kr

**Takashi Kumagai**

Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan

E-mail: kumagai@kurims.kyoto-u.ac.jp

**Jian Wang**

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, Fuzhou 350007, China

Email: jianwang@fjnu.edu.cn