

# Non-local Dirichlet Forms and Symmetric Jump Processes

Martin T. Barlow\*, Richard F. Bass†, Zhen-Qing Chen‡  
and Moritz Kassmann§

(May 1, 2007)

## Abstract

We consider the non-local symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  given by

$$\mathcal{E}(f, f) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (f(y) - f(x))^2 J(x, y) dx dy$$

with  $\mathcal{F}$  the closure with respect to  $\mathcal{E}_1$  of the set of  $C^1$  functions on  $\mathbb{R}^d$  with compact support, where  $\mathcal{E}_1(f, f) := \mathcal{E}(f, f) + \int_{\mathbb{R}^d} f(x)^2 dx$ , and where the jump kernel  $J$  satisfies

$$\kappa_1 |y - x|^{-d-\alpha} \leq J(x, y) \leq \kappa_2 |y - x|^{-d-\beta}$$

for  $0 < \alpha < \beta < 2$ ,  $|x - y| < 1$ . This assumption allows the corresponding jump process to have jump intensities whose sizes depend on the position of the process and the direction of the jump. We prove upper and lower estimates on the heat kernel. We construct a strong Markov process corresponding to  $(\mathcal{E}, \mathcal{F})$ . We prove a parabolic Harnack inequality for nonnegative functions that solve the heat equation with respect to  $\mathcal{E}$ . Finally we construct an example where the corresponding harmonic functions need not be continuous.

*Keywords:* Jump processes, symmetric processes, integro-differential operators, Harnack inequality, Dirichlet forms, heat kernel, harmonic, parabolic.

*Subject Classification:* Primary 60J35, Secondary 60J75, 45K05, 31B05

## 1 Introduction

In this paper we introduce a class of symmetric Markov processes of pure jump type. Our assumptions allow the jump intensities to depend on both the position of the process

---

\*Research partially supported by NSERC (Canada)

†Research partially supported by NSF grant DMS-0601783.

‡Research partially supported by NSF grant DMS-0600206.

§Research partially supported by DFG (Germany) through Sonderforschungsbereich 611.

and the direction of the jump. Thus our processes can be highly anisotropic. Although very little regularity is assumed, nevertheless we are able to obtain a number of results concerning these processes.

We begin by considering non-local symmetric Dirichlet forms. Set

$$\mathcal{E}(f, f) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (f(y) - f(x))^2 J(x, y) dx dy, \quad (1.1)$$

$$\mathcal{F} = \overline{C_c^1(\mathbb{R}^d)}^{\mathcal{E}_1}, \quad (1.2)$$

where the jump kernel  $J(x, y)$  is a function of  $x$  and  $y$  satisfying the following conditions

(A1)  $J(x, y) = J(y, x)$  for all  $x$  and  $y$ ;

(A2)  $J(x, y) = 0$  for  $|x - y| \geq 1$ ;

(A3) There exist  $\alpha, \beta \in (0, 2)$ ,  $\beta > \alpha$  and positive  $\kappa_1, \kappa_2$  such that

$$\kappa_1 |y - x|^{-d-\alpha} \leq J(x, y) \leq \kappa_2 |y - x|^{-d-\beta} \quad \text{for } |y - x| < 1.$$

Here  $\mathcal{E}_1(f, f) := \mathcal{E}(f, f) + \|f\|_2^2$ ,  $C_c^1(\mathbb{R}^d)$  denotes the space of  $C^1$  functions on  $\mathbb{R}^d$  with compact support, and  $\mathcal{F}$  is the closure of  $C_c^1(\mathbb{R}^d)$  with respect to the metric  $\mathcal{E}_1(f, f)^{1/2}$ .

We obtain the following four main results in this paper. We emphasize that we make no continuity assumptions whatsoever on the jump kernel  $J$ .

- (i) We show that there is a Hunt  $X$  process associated to  $(\mathcal{E}, \mathcal{F})$ , and  $X$  has a symmetric transition density function  $p(t, x, y)$ , with respect to Lebesgue measure on  $\mathbb{R}^d$ . We call this transition density function the heat kernel corresponding to  $(\mathcal{E}, \mathcal{F})$  and derive upper and lower bounds.
- (ii) We show that the strong Markov process  $X$  corresponding to  $(\mathcal{E}, \mathcal{F})$  is conservative and can be realized as a weak limit of certain more regular jump processes.
- (iii) We establish a parabolic Harnack inequality for nonnegative functions that solve the heat equation with respect to  $(\mathcal{E}, \mathcal{F})$ .
- (iv) We construct a counterexample to show that harmonic functions with respect to  $X$  need not be continuous on  $\mathbb{R}^d$ .

We now discuss each of these points in turn.

We first show there exists a Hunt process associated to  $(\mathcal{E}, \mathcal{F})$ .

**Theorem 1.1** *Suppose (A1)–(A3) hold.*

- (a) *There exists  $\mathcal{N} \subset \mathbb{R}^d$  having zero capacity with respect to the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  and there is a Hunt process  $(X, \mathbb{P}^x)$  with state space  $\mathbb{R}^d \setminus \mathcal{N}$  such that for every  $f \in L^2(\mathbb{R}^d, dx)$  and  $t > 0$ ,  $x \mapsto \mathbb{E}^x[f(X_t)]$  is a quasi-continuous version of  $T_t f$ , where  $\{T_t, t \geq 0\}$  is the  $L^2$ -semigroup associated with the closed form  $(\mathcal{E}, \mathcal{F})$ .*

(b) *The process  $X$  is conservative.*

Note that the Hunt process  $X$  can start from any point in  $\mathbb{R}^d \setminus \mathcal{N}$  and that once it starts from  $\mathbb{R}^d \setminus \mathcal{N}$  the process  $X$  together with its left hand limits takes values in  $\mathbb{R}^d \setminus \mathcal{N}$  up to and strictly before its lifetime  $\zeta$ . The set  $\mathcal{N}$  is called the properly exceptional set of  $X$  (or, equivalently, of  $(\mathcal{E}, \mathcal{F})$ ) and it has zero Lebesgue measure. For simplicity, sometimes we just say that  $X$  is a Hunt process associated with  $(\mathcal{E}, \mathcal{F})$  starting from quasi-everywhere in  $\mathbb{R}^d$ . For more on terminology and properties of Dirichlet forms, we refer the reader to [FOT94].

Let  $P(t, x, dy)$  be the transition probability for the Hunt process  $X$  associated with  $(\mathcal{E}, \mathcal{F})$ .

**Theorem 1.2** *Assume conditions (A1)–(A3) hold. There is a properly exceptional set  $\mathcal{N} \subset \mathbb{R}^d$  of  $X$ , a positive symmetric kernel  $p(t, x, y)$  defined on  $(0, \infty) \times (\mathbb{R}^d \setminus \mathcal{N}) \times (\mathbb{R}^d \setminus \mathcal{N})$ , and positive constants  $C_1$  and  $C_2$  (depending on the constants in (A1)–(A3)) such that  $P(t, x, dy) = p(t, x, y)dy$ , and*

$$p(t, x, y) \leq C_1 t^{-d/\alpha} e^{C_2 t} \quad \text{for every } t > 0 \text{ and } x, y \in \mathbb{R}^d \setminus \mathcal{N}. \quad (1.3)$$

Moreover, for every  $t > 0$  and  $y \in \mathbb{R}^d \setminus \mathcal{N}$ ,  $x \mapsto p(t, x, y)$  is quasi-continuous on  $\mathbb{R}^d$ .

We also obtain lower bounds on the heat kernel. Let  $B(x, r)$  denote the open ball of radius  $r$  centered at  $x$ . Let  $p^D(t, x, y)$  be the transition densities for the subprocess  $X^D$  of  $X$  killed upon exiting the ball  $D$ .

**Theorem 1.3** *Assume conditions (A1)–(A3) hold. Let  $y_0 \in \mathbb{R}^d$ ,  $T > 1/2$ , and  $\delta \in (0, 1/2)$ . Let  $R > 0$  and  $B = B(y_0, R)$ . There exists a properly exceptional set  $\mathcal{N}$  and a positive constant  $C$  that depends on  $R, T, \alpha, \beta, \kappa_1, \kappa_2$ , and  $\delta$ , but not on  $y_0$  such that for all  $t \in [\delta, T]$*

$$p^B(t, x, y) \geq C \quad (1.4)$$

for every  $(x, y) \in (B(y_0, 3R/4) \setminus \mathcal{N}) \times (B(y_0, 3R/4) \setminus \mathcal{N})$ .

**Remark 1.4** The jump kernel  $J$  does not have any scaling properties, and so one should not expect the results in Theorems 1.2 and 1.3 to be scale invariant. In particular, the constant  $C$  in Theorem 1.3 depends on  $R$ .

One of the difficulties in working with the process associated to  $(\mathcal{E}, \mathcal{F})$  is that we do not know that  $\mathcal{F}$  defined above is equal to

$$\mathcal{F}_\infty = \{f \in L^2(\mathbb{R}^d, dx) : \mathcal{E}(f, f) < \infty\}.$$

(A similar problem arises when studying the minimal Brownian motion on a manifold.) To circumvent this and other difficulties, we will approximate  $J$  by jump kernels  $J_\xi$  which

have regular behavior for  $|x - y| \leq \xi$ . We introduce the following condition, which we will assume from time to time. Let  $\xi > 0$ .

(A4)( $\xi$ ) The jump kernel  $J$  satisfies

$$J(x, y) = \kappa_2 |y - x|^{-d-\beta} \quad \text{when } |y - x| < \xi.$$

If  $J$  satisfies (A1)–(A3), define

$$J_\xi(x, y) = J(x, y) \mathbb{1}_{(|x-y|>\xi)} + \kappa_2 |x - y|^{-d-\beta} \mathbb{1}_{(|x-y|\leq\xi)}. \quad (1.5)$$

Let  $(\mathcal{E}^{(\xi)}, \mathcal{F}^{(\xi)})$  be the regular Dirichlet form on  $\mathbb{R}^d$  defined by (1.1)–(1.2) with  $J_\xi$  in place of  $J$ . We can then prove (see Lemma 2.5 below) that

$$\mathcal{F}^{(\xi)} = \{f \in L^2(\mathbb{R}^d, dx) : \mathcal{E}^{(\xi)}(f, f) < \infty\}.$$

We also have

**Theorem 1.5** *As  $\xi \rightarrow 0+$ ,  $(\mathcal{E}^{(\xi)}, \mathcal{F}^{(\xi)})$  converges to  $(\mathcal{E}, \mathcal{F})$  in the sense of Mosco.*

**Remark 1.6** See Definition 2.1 for the meaning of Mosco convergence.

- (i) Mosco convergence (see [Mo94]) implies that the semigroups of the processes  $X^{(\xi)}$  associated to  $\mathcal{E}^{(\xi)}$  converge in  $L^2(\mathbb{R}^d, dx)$  to the semigroup of the process  $X$ . We establish Theorem 1.3 by first proving the result for  $X^{(\xi)}$ , with constants independent of  $\xi$ ; taking a limit then gives the result for  $X$ .
- (ii) Studying  $X$  by first assuming (A4)( $\xi$ ) and then taking limits is analogous to a common procedure in the study of elliptic operators in divergence form. There one often first assumes the coefficients are smooth and obtains estimates that do not depend on the smoothness, and then uses a limiting procedure.
- (iii) We prove Theorem 1.5 by first establishing a simple sufficient criterion for Mosco convergence to hold. This is of independent interest.
- (iv) It seems to be difficult to establish a similar result if we approximate  $J(x, y)$  from below when  $|x - y|$  is small.
- (v) As remarked above, we do not know in general that  $\mathcal{F} = \mathcal{F}_\infty$ . However if the jump kernel  $J$  is “translation equivalent” near the diagonal, that is, if there exist constants  $c > 1$  and  $\delta > 0$  such that

$$c^{-1}J(x, y) \leq J(x - z, y - z) \leq cJ(x, y)$$

for a.e.  $x, y, z \in \mathbb{R}^d$  with  $0 < |x - y| < \delta$ , then it is not difficult to show that  $\mathcal{F} = \mathcal{F}_\infty$ . Since this result is not used in this paper, we omit its proof.

Let  $X$  be the Hunt process associated with the regular Dirichlet form  $(\mathcal{E}, \mathcal{F})$  on  $\mathbb{R}^d$ , which has an exceptional set  $\mathcal{N}$ . We say a function  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  is *harmonic* on a ball  $B(y, r)$  if  $h(X_{t \wedge \tau_{B(y, r_1)}})$  is a  $\mathbb{P}^x$ -martingale with right continuous paths for every  $r_1 \in (0, r)$  and every  $x \in B(y, r_1) \setminus \mathcal{N}$ . Here  $\tau_{B(y, r_1)} = \inf\{t : X_t \notin B(y, r_1)\}$ .

Set  $V_t = V_0 - t$  and let  $\mathbb{P}^{(s, x)}$  be the law of  $(V_t, X_t)$  started at  $(s, x)$ . We say a function  $u : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$  is *caloric* on  $Q = (a, b) \times B(x_0, r)$  with respect to  $\mathcal{E}$  if  $u(V_{t \wedge \tau_{Q_1}}, X_{t \wedge \tau_{Q_1}})$  is a  $\mathbb{P}^{(s, x)}$ -martingale with right continuous paths for every open subset  $Q_1$  of  $Q$  with  $\overline{Q_1} \subset Q$  and for every  $(s, x) \in Q_1 \cap (\mathbb{R}_+ \times (\mathbb{R}^d \setminus \mathcal{N}))$ . Here  $\tau_{Q_1} = \inf\{t : (V_t, X_t) \notin Q_1\}$ .

We prove that nonnegative functions that are caloric with respect to  $\mathcal{E}$  satisfy a parabolic Harnack inequality.

**Theorem 1.7** *Suppose the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  is given by (1.1)–(1.2) with  $J(x, y)$  satisfying (A1)–(A3). Let  $t_0 \geq 0, R \geq 1$ , and  $T > 0$ . There exists a positive real  $C = C(\alpha, \beta, \kappa_1, \kappa_2, d, R, T)$  such that if  $x_0 \in \mathbb{R}^d$  and  $u$  is nonnegative and bounded in  $(t_0, t_0 + 5T) \times \mathbb{R}^d$  and is caloric on  $Q = (t_0, t_0 + 5T) \times B(x_0, 4R)$  with respect to  $\mathcal{E}$ , then*

$$\text{ess sup}_{Q^-} u \leq C \text{ess inf}_{Q^+} u,$$

where  $Q^- = [t_0 + T, t_0 + 2T] \times B(x_0, R)$  and  $Q^+ = [t_0 + 3T, t_0 + 4T] \times B(x_0, R)$ .

**Remark 1.8** Concerning the hypotheses and statement of Theorem 1.7, we make the following remarks.

- (i) We assume  $u(t, \cdot)$  is bounded in  $\mathbb{R}^d$  in order to ensure that the random variable  $u(V_{t \wedge \tau_{Q_1}}, X_{t \wedge \tau_{Q_1}})$  is integrable. However the constant  $C$  does not depend on this bound.
- (ii) Harmonic functions are caloric, so the parabolic Harnack inequality implies that an elliptic Harnack inequality also holds.
- (iii) Assumption (A3) does not satisfy any type of scaling property. As a result, one cannot expect the parabolic Harnack inequality to be scale invariant, i.e., that the constant  $C$  could be chosen independently of  $R$  or  $T$ . Since an example in [BK05a] shows that scale invariance can fail for the elliptic Harnack inequality, scale invariance can also fail for the parabolic Harnack inequality. This phenomenon is well-known in the theory of degenerate partial differential equations, see [CW86], [GW90].
- (iv) We shall see in Theorem 1.9 below that it may not be possible to extend a harmonic function  $h$  to a continuous function on  $\mathbb{R}^d$ . Thus we have to use the essential supremum and essential infimum in Theorem 1.7.
- (v) Assumption (A2) rules out any jumps of size larger than 1. An example in [BK05a] shows that the large jumps, although in many ways less interesting, can cause the Harnack inequality to fail. For similar reasons we cannot replace the ball of radius 1 by arbitrarily small balls in Theorem 1.7.

- (vi) We allow  $0 < \alpha < \beta < 2$  with no other restriction on  $\alpha$  and  $\beta$ . This should be contrasted with the situation in [BK05a], which considers non-local operators that are non-symmetric, and where in addition it was required that  $\beta - \alpha < 1$ .

Following [FS86] many papers have used heat kernel estimates to prove Harnack inequalities. The usual procedure is to obtain an oscillation inequality, and from this one obtains a Harnack inequality. We cannot use this approach here, since our counterexample shows that the constant in the oscillation inequality can blow up as the radius  $r_n$  of the ball approaches 0. Instead, the proof of Theorem 1.7 uses a balayage argument; this approach is new and is of independent interest.

As we mentioned above, harmonic functions need not be continuous. We prove

**Theorem 1.9** *There exists a Dirichlet form  $(\mathcal{E}, \mathcal{F})$  given by (1.1)–(1.2) with the jump kernel  $J$  satisfying (A1)–(A3), but where there also exists a bounded harmonic function that cannot be extended to be a continuous function on  $B(0, 1)$ .*

- Remark 1.10**
- (i) We will also show that continuity can fail for  $P_t f$ , even when  $f$  is smooth.
  - (ii) Our construction also gives an example of a martingale problem for which uniqueness fails. The example also shows that the existence of a properly exceptional set  $\mathcal{N}$  in Theorem 1.1 and Theorem 1.2 is essential and cannot be dropped in general.
  - (iii) The harmonic function we construct may be continuous outside a set  $\mathcal{N}$  of capacity 0.

Heat kernel estimates and Harnack inequalities have a long history in the theory of partial differential equations. After path breaking work by DeGiorgi [DG57] and Nash [Nas58] on regularity, Moser [Mos61] proved a scale invariant Harnack inequality for functions that are harmonic with respect to second order elliptic operators in divergence form. This was extended in Moser [Mos64] to solutions to the heat equation, i.e., the parabolic case; see also [Mos71]. A quite different proof of this was given in Fabes-Stroock [FS86]. The Harnack inequality for operators in nondivergence form was established by Krylov-Safonov [KS80]. However the corresponding theory of Harnack inequalities for jump processes is still largely unknown.

Non-local operators such as those considered in this paper arise in the study of models of financial markets (see [SS06] and the references therein). They also arise in the study of the Dirichlet-to-Neumann map, particularly for subelliptic operators or in rough domains.

Harnack inequalities for non-local operators have been considered in [BL02a] and [SV04] for fixed order, non-symmetric operators, [BL02b] and [CK03] for fixed order, symmetric operators. A scale dependent Harnack inequality has been established in [BK05a] for variable order, non-symmetric operators. Additionally, regularity of harmonic functions is considered for variable order, non-symmetric operators in [BK05b], [HKa05]. For heat kernel estimates and parabolic Harnack principle for symmetric non-local Dirichlet forms

on  $d$ -sets, see [CK03], [HKu05] for fixed order and [CK06] for variable order. See [SU05] for related results for processes given in terms of pseudo-differential operators.

The paper is organized as follows. The Mosco convergence results are established in the next section. In Section 3, we obtain some upper bounds for the fundamental solution of the operator corresponding to  $\mathcal{E}$  and prove that the corresponding process  $X$  is conservative. In Section 4, we consider lower bound estimates for the heat kernel. The parabolic Harnack inequality is established in Section 5. The counterexample is constructed in Section 6. We use  $c_i, c$  or  $C$  to denote finite positive constants that depend only on  $\alpha, \beta, \kappa_i$  or  $d$  and whose exact value is not important and may change from line to line. Further dependencies are mentioned explicitly. We denote the Lebesgue measure of a Borel set  $A$  by  $|A|$ . If  $A$  is a Borel set and  $Y$  a right continuous process, we use the notation

$$T_A^Y = T_A = \inf\{t > 0 : Y_t \in A\}, \quad \tau_A^Y = \tau_A = \inf\{t > 0 : Y_t \notin A\}. \quad (1.6)$$

For processes  $Y$  with paths that are right continuous with left limits, we let  $Y_{t-}$  be the left hand limit at time  $t$  and  $\Delta Y_t := Y_t - Y_{t-}$  the jump at time  $t$ .

## 2 Mosco convergence

Let us first recall the definition of Mosco convergence and its properties. Let  $E$  be a locally compact separable metric space and  $m$  a Radon measure on  $E$  with full support. Given a densely defined quadratic form  $(\mathcal{E}, \mathcal{F})$  in  $L^2(E; m)$ , we can extend its domain of definition to  $L^2(E; m)$  by setting  $\mathcal{E}(u, u) = \infty$  for  $u \in L^2(E; m) \setminus \mathcal{F}$ . Throughout this section we will use this extension and, unless otherwise specified, all the quadratic forms encountered will be assumed to be densely defined in  $L^2(E; m)$ . Recall that given  $(\mathcal{E}, \mathcal{F})$  we set  $\mathcal{E}_1(f, f) = \mathcal{E}(f, f) + \|f\|_2^2$ .

**Definition 2.1** *A sequence of closed quadratic forms  $\{(\mathcal{E}^n, \mathcal{F}^n)\}$  on  $L^2(E; m)$  is said to be convergent to a closed quadratic form  $(\mathcal{E}, \mathcal{F})$  on  $L^2(E; m)$  in the sense of Mosco (cf. [Mo94]) if*

- (a) *For every sequence  $\{u_n, n \geq 1\}$  in  $L^2(E; m)$  that converges weakly to  $u$  in  $L^2(E; m)$ ,*

$$\liminf_{n \rightarrow \infty} \mathcal{E}^n(u_n, u_n) \geq \mathcal{E}(u, u),$$

- (b) *For every  $u \in L^2(E; m)$ , there is a sequence  $\{u_n, n \geq 1\}$  in  $L^2(E; m)$  converging strongly to  $u$  in  $L^2(E; m)$  such that*

$$\limsup_{n \rightarrow \infty} \mathcal{E}^n(u_n, u_n) \leq \mathcal{E}(u, u).$$

Let  $\{P_t, t \geq 0\}$  and  $\{P_t^n, t \geq 0\}$  be the semigroups of  $(\mathcal{E}, \mathcal{F})$  and  $(\mathcal{E}^n, \mathcal{F}^n)$ , respectively, and  $\{G_\alpha, \alpha > 0\}$  and  $\{G_\alpha^n, \alpha > 0\}$  their corresponding resolvents, respectively. The following result is known (see Theorem 2.4.1 and Corollary 2.6.1 of [Mo94]).

**Proposition 2.2** Let  $(\mathcal{E}, \mathcal{F})$  and  $\{(\mathcal{E}^n, \mathcal{F}^n), n \geq 1\}$  be closed quadratic forms on  $L^2(E; m)$ . The following are equivalent:

- (a)  $(\mathcal{E}^n, \mathcal{F}^n)$  converges to  $(\mathcal{E}, \mathcal{F})$  in the sense of Mosco;
- (b) For every  $\alpha > 0$  and  $f \in L^2(E; m)$ ,  $G_\alpha^n f$  converges to  $G_\alpha f$  in  $L^2(E; m)$ ;
- (c) For every  $t > 0$  and  $f \in L^2(E; m)$ ,  $P_t^n f$  converges to  $P_t f$  in  $L^2(E; m)$ .

Here is a criterion for Mosco convergence to hold.

**Theorem 2.3** Suppose

- (i)  $\mathcal{F}^n \subset \mathcal{F}$  for every  $n \geq 1$  and  $\mathcal{E}^n(u, u) \geq \mathcal{E}(u, u)$  for every  $u \in \mathcal{F}^n$ .
- (ii) There is a common core  $\mathcal{C}$  for the Dirichlet forms  $(\mathcal{E}^n, \mathcal{F}^n)$  and  $(\mathcal{E}, \mathcal{F})$  such that

$$\lim_{n \rightarrow \infty} \mathcal{E}^n(u, u) = \mathcal{E}(u, u) \quad \text{for every } u \in \mathcal{C}.$$

Then  $(\mathcal{E}^n, \mathcal{F}^n)$  converges to  $(\mathcal{E}, \mathcal{F})$  in the sense of Mosco.

**Proof.** Let  $\{u_k, k \geq 1\}$  be a sequence in  $L^2(E; m)$  that converges weakly to  $u$  in  $L^2(E; m)$ . Without loss of generality, we may assume that  $\lim_{k \rightarrow \infty} \mathcal{E}^k(u_k, u_k)$  exists and is finite. This in particular implies that  $u_k \in \mathcal{F}^k \subset \mathcal{F}$  for every  $k \geq 1$  and  $\sup_{k \geq 1} \mathcal{E}(u_k, u_k) < \infty$ . Since  $\{u_k, k \geq 1\}$  is bounded in  $L^2(E; m)$ , taking a subsequence if necessary, we may assume that the Cesaro mean of  $\{u_k, k \geq 1\}$  converges in  $(\mathcal{E}_1, \mathcal{F})$  to some function  $v$  (see page 14 of [Si74]). As  $u_k$  converges weakly to  $u$ , we must have  $u = v \in \mathcal{F}$ . Therefore,

$$\liminf_{k \rightarrow \infty} \mathcal{E}^k(u_k, u_k) \geq \limsup_{k \rightarrow \infty} \mathcal{E}(u_k, u_k) \geq \limsup_{k \rightarrow \infty} \mathcal{E} \left( \frac{1}{k} \sum_{j=1}^k u_j, \frac{1}{k} \sum_{j=1}^k u_j \right) \geq \mathcal{E}(u, u).$$

The second inequality follows since the triangle inequality tells us that

$$\mathcal{E} \left( \frac{1}{k} \sum_{j=1}^k u_j, \frac{1}{k} \sum_{j=1}^k u_j \right)^{1/2} \leq \frac{1}{k} \sum_{j=1}^k \mathcal{E}(u_j, u_j)^{1/2}.$$

This shows that the condition (a) of Definition 2.1 is satisfied.

For any  $u \in \mathcal{F}$ , there exists a sequence  $\{v_j\} \subset \mathcal{C}$  converging strongly to  $u$  in  $L^2(E; m)$  such that

$$\lim_{j \rightarrow \infty} \mathcal{E}(v_j, v_j) = \mathcal{E}(u, u).$$

Since for each  $j \geq 1$ ,

$$\lim_{n \rightarrow \infty} \mathcal{E}^n(v_j, v_j) = \mathcal{E}(v_j, v_j),$$

using induction we can find an increasing subsequence  $\{n_j\}$  such that

$$|\mathcal{E}^{n_j}(v_j, v_j) - \mathcal{E}(v_j, v_j)| \leq 2^{-j} \quad \text{for } n \geq n_j.$$



Put  $u_1 = \cdots = u_{n_1-1} = 0$  and  $u_{n_j} = \cdots = u_{n_{j+1}-1} = v_j$  for  $j \geq 1$ . It is easy to see that  $\{u_k, k \geq 1\}$  is a sequence in  $\mathcal{C}$  converging strongly to  $u$  in  $L^2(E; m)$  such that

$$\lim_{k \rightarrow \infty} \mathcal{E}^k(u_k, u_k) = \mathcal{E}(u, u).$$

For  $u \in L^2(E, m) \setminus \mathcal{F}$ , since  $\mathcal{E}(u, u) = \infty$ , it trivially holds that

$$\limsup_{k \rightarrow \infty} \mathcal{E}^k(u_k, u_k) \leq \mathcal{E}(u, u).$$

This shows that the condition (b) in Definition 2.1 is satisfied. Hence we have shown that  $(\mathcal{E}^k, \mathcal{F}^k)$  is Mosco-convergent to  $(\mathcal{E}, \mathcal{F})$ .  $\square$

Let  $X^k$  be the Hunt process associated with  $(\mathcal{E}^k, \mathcal{F}^k)$  and  $X^{k,B}$  be the subprocess of  $X^k$  killed upon exiting an open set  $B$ . It is known (see [FOT94]) that the Dirichlet form  $(\mathcal{E}^k, \mathcal{F}^{k,B})$  of  $X^k$  is given by

$$\mathcal{F}^{k,B} = \{u \in \mathcal{F}^k : u = 0 \text{ } \mathcal{E}^k\text{-q.e. on } B^c\}. \quad (2.1)$$

**Theorem 2.4** *Suppose  $B$  is an open set and the following hold.*

- (i)  $\mathcal{F}^k \subset \mathcal{F}$  and  $\mathcal{E}^k(u, u) \geq \mathcal{E}(u, u)$  for every  $u \in \mathcal{F}^k$  and every  $k \geq 1$ .
- (ii) There is a common core  $\mathcal{C}$  for the Dirichlet forms  $(\mathcal{E}^k, \mathcal{F}^k)$  and  $(\mathcal{E}, \mathcal{F})$  such that

$$\lim_{k \rightarrow \infty} \mathcal{E}^k(u, u) = \mathcal{E}(u, u) \quad \text{for every } u \in \mathcal{C}.$$

Furthermore, there is a common core  $\mathcal{C}_B \subset \mathcal{C}$  for the Dirichlet forms  $(\mathcal{E}^k, \mathcal{F}^{k,B})$  and  $(\mathcal{E}, \mathcal{F}^B)$ .

Then  $(\mathcal{E}^k, \mathcal{F}^{k,B})$  converges to  $(\mathcal{E}, \mathcal{F}^B)$  in the sense of Mosco.

**Proof.** To emphasize the domain of definition, for this proof only, we write  $\mathcal{E}^B$  and  $\mathcal{E}^{k,B}$  for  $(\mathcal{E}, \mathcal{F}^B)$  and  $(\mathcal{E}^k, \mathcal{F}^{k,B})$ , respectively. With this notation,  $\mathcal{E}^B(u, u) = \infty$  when  $u \notin \mathcal{F}^B$  and  $\mathcal{E}^{k,B}(u, u) = \infty$  when  $u \notin \mathcal{F}^{k,B}$ .

First note that, since  $\mathcal{E}_1^k(u, u) \geq \mathcal{E}_1(u, u)$ ,

$$\mathcal{F}^k \subset \mathcal{F} \quad \text{and} \quad \mathcal{F}^{k,B} \subset \mathcal{F}^B.$$

For any  $v_k$  that converges weakly to  $v$  in  $L^2(B; dx)$ , we claim that

$$\liminf_{k \rightarrow \infty} \mathcal{E}^{k,B}(v_k, v_k) \geq \mathcal{E}(v, v). \quad (2.2)$$

Suppose that the left hand side of (2.2) is finite. Then there is a subsequence  $\{n_k\}$  such that

$$\lim_{k \rightarrow \infty} \mathcal{E}^{n_k, B}(v_{n_k}, v_{n_k}) = \liminf_{k \rightarrow \infty} \mathcal{E}^{k, B}(v_k, v_k) \quad \text{and} \quad \sup_{k \geq 1} \mathcal{E}^{n_k, B}(v_{n_k}, v_{n_k}) < \infty.$$

In particular, this implies that  $v_{n_k} \in \mathcal{F}^{n_k, B}$  and

$$\sup_{k \geq 1} \mathcal{E}_1^B(v_{n_k}, v_{n_k}) \leq \sup_{k \geq 1} \mathcal{E}_1^{n_k, B}(v_{n_k}, v_{n_k}) < \infty.$$

By taking a subsequence if necessary, the Cesaro mean of  $\{v_{n_k}, k \geq 1\}$  converges in  $\mathcal{F}^B$  with respect to the Hilbert norm  $\sqrt{\mathcal{E}_1^B}$  to a function  $w$ , which has to be  $v$ . This implies that  $v \in \mathcal{F}^B$ . By extending  $v_{n_k}$  and  $v$  to take the value zero off  $B$ , we have  $v_{n_k} \in \mathcal{F}^{n_k}$  and  $v \in \mathcal{F}$ . By Theorem 2.3,  $(\mathcal{E}^k, \mathcal{F}^k)$  converges to  $(\mathcal{E}, \mathcal{F})$  in the sense of Mosco, and we have in particular that

$$\liminf_{k \rightarrow \infty} \mathcal{E}^k(v_k, v_k) \geq \mathcal{E}(v, v).$$

As  $\mathcal{E}^{k, B}$  and  $\mathcal{E}^B$  agree with  $\mathcal{E}^k$  and  $\mathcal{E}$  on  $\mathcal{F}^{k, B}$  and  $\mathcal{F}^B$ , respectively, this proves (2.2).

Noting that  $\mathcal{C}_B$  is a common core for  $(\mathcal{E}, \mathcal{F}^B)$  and  $(\mathcal{E}^k, \mathcal{F}^{k, B})$ , it can be shown that the second condition (b) in Definition 2.1 holds for  $(\mathcal{E}^{k, B}, \mathcal{F}^{k, B})$  and  $(\mathcal{E}^B, \mathcal{F}^B)$  in much the same way as in the proof of Theorem 2.3.  $\square$

The following lemma gives a concrete characterization of the domain of the Dirichlet form  $(\mathcal{E}, \mathcal{F})$  given by (1.1)-(1.2).

**Lemma 2.5** *Let  $J(x, y)$  satisfy (A1)–(A3), and (A4)( $\xi$ ) for some  $\xi \in (0, 1)$ . Then*

$$\mathcal{F} = \{f \in L^2(\mathbb{R}^d, dx) : \mathcal{E}(f, f) < \infty\} = \mathbb{H}^{\beta/2}(\mathbb{R}^d).$$

*The corresponding Hunt process  $X$  is a conservative strong Markov process which can start from any point in  $\mathbb{R}^d$ .*

**Proof.** For two bilinear forms  $\mathcal{E}$  and  $\mathcal{C}$  having a common core  $C_c^1(\mathbb{R}^d)$ , let us write  $\mathcal{E}(f, f) \approx \mathcal{C}(f, f)$  if there is a finite constant  $c_1 > 0$  such that  $c_1^{-1}\mathcal{E}(f, f) \leq \mathcal{C}(f, f) \leq c_1\mathcal{E}(f, f)$  for every  $f \in C_c^1(\mathbb{R}^d)$ . As  $(\mathcal{E}, \mathcal{F})$  satisfies (A1)–(A3) and (A4)( $\xi$ ), we have for  $f \in C_c^1(\mathbb{R}^d)$ ,

$$\begin{aligned} \mathcal{E}_1(f, f) &\approx \int_{|x-y| \leq \xi} \frac{(f(x) - f(y))^2}{|x-y|^{d+\beta}} dx dy + \|f\|_2^2 \\ &\approx \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{(f(x) - f(y))^2}{|x-y|^{d+\beta}} dx dy + \|f\|_2^2 \\ &:= \mathcal{C}_1(f, f). \end{aligned}$$

So  $\mathcal{F} = \overline{C_c^1(\mathbb{R}^d)^{\mathcal{E}_1}} = \overline{C_c^1(\mathbb{R}^d)^{\mathcal{C}_1}} = \mathbb{H}^{\beta/2}(\mathbb{R}^d) = \{f : \mathcal{C}_1(f, f) < \infty\}$ , which is the same as  $\{f : \mathcal{E}_1(f, f) < \infty\}$ .

Note that the process  $X$  can be constructed from the Lévy process  $Z$  on  $\mathbb{R}^d$  whose Lévy measure is

$$J_0(h) dh = |h|^{-d-\beta} \mathbb{1}_{(|h| \leq \xi)} dh \quad (2.3)$$

using Remark 3.4. As  $Z$  is conservative and can start from every point in  $\mathbb{R}^d$ , the same is true of  $X$ .  $\square$

**Proof of Theorem 1.5:** Let  $\delta_k$  be a sequence of positive numbers decreasing to 0. Set

$$J_k(x, y) = \begin{cases} J(x, y) & \text{for } |x - y| \geq \delta_k; \\ \kappa_2 |y - x|^{-d-\beta} & \text{for } |x - y| < \delta_k, \end{cases} \quad (2.4)$$

and define  $(\mathcal{E}^k, \mathcal{F}^k)$  in the same way as we defined  $(\mathcal{E}, \mathcal{F})$  in (1.1)–(1.2). Note that  $\mathcal{E}^k$  satisfies A4( $\delta_k$ ). Take  $\mathcal{E}^k(f, f) = +\infty$  if  $f \in L^2(\mathbb{R}^d, dx) \setminus \mathcal{F}^k$ .

It is clear that  $J_k(x, y)$  decreases to  $J(x, y)$  as  $k \uparrow \infty$ , and so  $\mathcal{F}^k \subset \mathcal{F}$  and  $\mathcal{E}^k(u, u) \geq \mathcal{E}(u, u)$  on  $\mathcal{F}^k$  for every  $k \geq 1$ . By Lemma 2.5,

$$\mathcal{F}^k = \mathbb{H}^{\beta/2}(\mathbb{R}^d).$$

and thus in particular  $\mathcal{F}^k$  is independent of  $k$ . Note that  $C_1^1(\mathbb{R}^d)$  is the common core of  $(\mathcal{E}^k, \mathcal{F}^k)$  for  $k \geq 1$  and for  $(\mathcal{E}, \mathcal{F})$ , and that

$$\lim_{k \rightarrow \infty} \mathcal{E}^k(u, u) = \mathcal{E}(u, u) \quad \text{for every } u \in C_c^1(\mathbb{R}^d).$$

Theorem 1.5 now follows from Theorem 2.3.  $\square$

**Theorem 2.6** *Let  $B$  be a ball and define  $\mathcal{F}^{k,B}$  by (2.1), where  $(\mathcal{E}^k, \mathcal{F}^k)$  is as in the proof of Theorem 1.5. Then  $(\mathcal{E}^k, \mathcal{F}^{k,B})$  converges in the sense of Mosco to  $(\mathcal{E}, \mathcal{F}^B)$ .*

**Proof.** Note that  $C_c^1(B) \subset C_c^1(\mathbb{R}^d)$  is a common core for  $(\mathcal{E}, \mathcal{F}^B)$  and  $(\mathcal{E}^k, \mathcal{F}^{k,B})$ . The conclusion of the corollary follows directly from Theorem 2.4 and the proof of Theorem 1.5.  $\square$

### 3 Upper bounds for the heat kernel

Throughout this section we will assume that the jump kernel  $J$  satisfies (A1)–(A3). We begin with the proof of Theorem 1.1(a), which is easy. Theorem 1.1(b) will be proved at the end of this section.

**Proof of Theorem 1.1(a):** Let  $C_\infty(\mathbb{R}^d)$  denote the space of continuous functions on  $\mathbb{R}^d$  that vanish at infinity and let  $\|\cdot\|_\infty$  denote the supremum norm in  $C_\infty(\mathbb{R}^d)$ . It is easy to check by using Fatou's lemma that the bilinear form  $(\mathcal{E}, \mathcal{F})$  is a closed form (cf. [FOT94, Example 1.2.4]). As  $C_c^1(\mathbb{R}^d)$  is dense both in  $(\mathcal{F}, \mathcal{E}_1)$  and in  $(C_\infty, \|\cdot\|_\infty)$ ,  $(\mathcal{E}, \mathcal{F})$  is a regular Dirichlet form on  $\mathbb{R}^d$ . Our result now follows from [FOT94, Chapter 7].  $\square$

It is well known that Nash's inequality implies the operator norm estimate for the transition semigroup  $P_t$  from  $L^1(\mathbb{R}^d)$  to  $L^\infty(\mathbb{R}^d)$ . However this only implies for every  $t > 0$  the existence of  $p(t, x, y)$  almost everywhere on  $\mathbb{R}^d \times \mathbb{R}^d$  such that for every  $f \geq 0$  on  $\mathbb{R}^d$ ,

$$P_t f(x) = \int_{\mathbb{R}^d} p(t, x, y) f(y) dy \quad \text{for a.e. } x \in \mathbb{R}^d.$$

We need something stronger.

Since the following result has independent interest, we state and prove it in a more general context. For the next two theorems only, let  $E$  be a locally compact separable metric space and  $m$  a Radon measure on  $E$  whose support is all of  $E$ . A symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  in  $L^2(E, m)$  is said to be regular if  $C_c(E) \cap \mathcal{F}$  is dense both in  $(\mathcal{F}, \mathcal{E}_1)$  and in  $(C_c(E), \|\cdot\|_\infty)$ . It is well known (cf. [FOT94]) that a regular symmetric Dirichlet form  $(\mathcal{E}, \mathcal{F})$  has associated with it a symmetric Hunt process  $X$  that can start from every point outside a properly exceptional set  $\mathcal{N}$  (cf. Theorem 1.1). For  $x \in E \setminus \mathcal{N}$ , we use  $\{P(t, x, dy), t \geq 0\}$  to denote the transition probability of  $X$ . The transition semigroup  $\{P_t, t \geq 0\}$  of  $X$  is defined for  $x \in E \setminus \mathcal{N}$  by

$$P_t f(x) := \mathbb{E}^x [f(X_t)] \quad \text{for } f \geq 0 \text{ on } E \text{ and } t > 0.$$

**Theorem 3.1** *Let  $E$ ,  $m$ , and  $P_t$  be as above. Assume that there is a positive left continuous function  $M(t)$  on  $(0, \infty)$  such that*

$$\|P_t f\|_\infty \leq M(t) \|f\|_1 \quad \text{for every } f \in L^1(E, m) \text{ and } t > 0. \quad (3.1)$$

*Then there is a properly exceptional set  $\mathcal{N} \subset E$  of  $X$  and a positive symmetric kernel  $p(t, x, y)$  defined on  $(0, \infty) \times (E \setminus \mathcal{N}) \times (E \setminus \mathcal{N})$  such that  $P(t, x, dy) = p(t, x, y) m(dy)$ ,*

$$p(t+s, x, y) = \int p(t, x, z) p(s, z, y) dz \quad \text{for every } x, y \in E \setminus \mathcal{N} \text{ and } t, s > 0,$$

and

$$p(t, x, y) \leq M(t) \quad \text{for every } t > 0 \text{ and } x, y \in E \setminus \mathcal{N}. \quad (3.2)$$

*Moreover, there is an  $\mathcal{E}$ -nest  $\{F_k, k \geq 1\}$  of compact sets so that  $\mathcal{N} = E \setminus \bigcup_{k=1}^\infty F_k$  and that for every  $t > 0$  and  $y \in E \setminus \mathcal{N}$ ,  $x \mapsto p(t, x, y)$  is continuous on each  $F_k$ .*

**Proof.** Let  $\mathcal{N}$  be a properly exceptional set of  $X$ . Recall that the transition semigroup  $\{P_t, t \geq 0\}$  of  $X$  is defined for  $x \in E \setminus \mathcal{N}$  by

$$P_t f(x) := \mathbb{E}^x [f(X_t)] \quad \text{for } f \geq 0 \text{ on } E \text{ and } t > 0.$$

Let  $\{f_k, k \geq 1\} \subset C_c(E) \cap \mathcal{F}$  be dense in both  $L^2(E, m)$  and  $L^1(E, m)$ . For each fixed  $t > 0$  and  $k \geq 0$ ,  $P_t f_k$  is quasi-continuous on  $E$ . Thus for each  $t > 0$ , there is an  $\mathcal{E}$ -nest  $\{F_n^{(t)}, n \geq 1\}$  consisting of an increasing sequence of compact sets such that  $P_t f_k$  is continuous on each  $F_n^{(t)}$  for every  $k \geq 1$  (cf. [FOT94, Theorem 2.1.2]). Let  $\mathcal{N}_t := E \setminus \bigcup_{n=1}^\infty F_n^{(t)}$ , which is  $\mathcal{E}$ -polar and in particular has zero  $m$ -measure.

Inequality (3.1) yields that for every  $n \geq 1$ ,

$$\sup_{x \in F_n^{(t)}} |P_t f_j(x) - P_t f_k(x)| \leq M(t) \|f_j - f_k\|_1.$$

Since  $\{f_k, k \geq 1\} \subset C_c(E) \cap \mathcal{F}$  is dense in  $L^1(E, m)$ , it follows that  $P_t f$  is continuous on each  $F_n^{(t)}$  and

$$\sup_{x \in E \setminus \mathcal{N}_t} |P_t f(x)| \leq M(t) \|f\|_1$$

for every  $f \in L^1(E, m)$ . Therefore for every  $t > 0$  and  $x \in E \setminus \mathcal{N}_t$ , there is an integrable kernel  $y \mapsto p_0(t, x, y)$  defined  $m$ -a.e. on  $E$  such that

$$\mathbb{E}^x [f(X_t)] = P_t f(x) = \int_E p_0(t, x, y) f(y) dy \quad \text{for every } f \in L^1(E, m) \quad (3.3)$$

and

$$p_0(t, x, y) \leq M(t) \quad \text{for } m\text{-a.e. } y \in E \quad (3.4)$$

From the semigroup property  $P_{t+s} = P_t P_s$ , we have for every  $t, s > 0$  and  $x \in E \setminus (\mathcal{N}_{t+s} \cup \mathcal{N}_t)$ ,

$$p_0(t+s, x, y) = \int_E p_0(t, x, z) p_0(s, z, y) m(dy) \quad \text{for } m\text{-a.e. } y \in E. \quad (3.5)$$

Note that since  $P_t$  is symmetric, we have for each fixed  $t > 0$ ,

$$p_0(t, x, y) = p_0(t, y, x) \quad \text{for } m\text{-a.e. } (x, y) \in E \times E.$$

By enlarging the properly exceptional set  $\mathcal{N}$  if necessary, we may and do assume that  $\mathcal{N} \supset \cup_{t \in \mathbb{Q}_+} \mathcal{N}_t$ . For every  $t > 0$  and  $x, y \in E \setminus \mathcal{N}$ , let  $s \in \mathbb{Q}_+$  be less than  $t/3$  and define

$$p(t, x, y) := \int_E p_0(s, x, w) \left( \int_E p_0(t-2s, w, z) p_0(s, y, z) m(dz) \right) m(dw). \quad (3.6)$$

By (3.5) the above definition is independent of the choice of  $s \in \mathbb{Q}_+ \cap (0, t/3)$ . Clearly,  $p(t, x, y) = p(t, y, x)$  for every  $x, y \in E \setminus \mathcal{N}$ . By the semigroup property and (3.3), we have for  $\varphi \geq 0$  on  $E$  and  $x \in E \setminus \mathcal{N}$ ,

$$\begin{aligned} & \mathbb{E}^x [\varphi(X_t)] \\ &= \int_E \left( \int_E p_0(s, x, w) \left( \int_E p_0(t-2s, w, z) p_0(s, y, z) m(dz) \right) m(dw) \right) \varphi(y) m(dy) \\ &= \int_E \left( \int_E p_0(s, x, w) \left( \int_E p_0(t-2s, w, z) p_0(s, y, z) m(dz) \right) m(dw) \right) \varphi(y) m(dy) \\ &= \int_E p(t, x, y) \varphi(y) m(dy). \end{aligned}$$

Thus  $p(t, x, y)$  coincides with  $p_0(t, x, y)$   $m$ -a.e. on  $E \times E$ . Note that it follows from (3.4) and (3.6) that for every  $t > 0$  and  $x, y \in E \setminus \mathcal{N}$ ,

$$p(t, x, y) \leq M(t-2s) \quad \text{for every } s \in \mathbb{Q}_+ \text{ and } s < t/3.$$

Taking  $s \downarrow 0$  yields

$$p(t, x, y) \leq M(t) \quad \text{for every } t > 0 \text{ and } x, y \in E \setminus \mathcal{N}.$$

For  $t, s > 0$  and  $x, y \in E \setminus \mathcal{N}$ , take  $s_0 \in \mathbb{Q}_+ \cap (0, (t \wedge s)/3)$ , and we have by (3.5)-(3.6)

$$\begin{aligned}
& p(t+s, x, y) \\
&= \int_E p_0(s_0, x, w) \left( \int_E p_0(t+s-2s_0, w, z) p_0(s_0, y, z) m(dz) \right) m(dw) \\
&= \int_{E^5} p_0(s_0, x, w) p_0(t-2s_0, w, u_1) p_0(s_0, u_1, u_2) p_0(s_0, u_2, v) p_0(s-2s_0, v, z) \\
&\quad p_0(s_0, y, z) m(dw) m(du_1) m(du_2) m(dz) m(dv) \\
&= \int_E p(t, x, v) p(s, v, y) m(dv).
\end{aligned}$$

We may assume that there is an  $\mathcal{E}$ -nest  $\{F_n, n \geq 1\}$  of compact sets such that  $\mathcal{N} = E \setminus (\cup_{n=1}^\infty F_n)$  and  $P_t f_k$  is continuous on  $F_n$  for each  $k \geq 1$ , each  $t$  rational, and each  $n$ . It follows from inequality (3.1),

$$\sup_{x \in F_n} |P_t f_j(x) - P_t f_k(x)| \leq M(t) \|f_j - f_k\|_1$$

for every  $t \in \mathbb{Q}_+$  and  $n, k \geq 1$ . Since  $\{f_j, j \geq 1\}$  is dense in  $L^1(E, m)$ , we conclude that  $P_t f$  is continuous on each  $F_n$  whenever  $f \in L^1(E, m)$ . By (3.4), the function  $w \mapsto \int_E p_0(t-2s, w, z) p_0(s, y, z) dz$  is  $L^1$ -integrable on  $E$ , and so as a function of  $x$ ,  $p(t, x, y)$  is continuous on each  $F_n$  for every real  $t > 0$  and  $y \in E \setminus \mathcal{N}$ . This proves the theorem.  $\square$

In order to get off-diagonal estimates for  $p(t, x, y)$  from the on-diagonal estimate (3.2), we need the following.

**Theorem 3.2** *Let the heat kernel  $p(t, x, y)$  and the properly exceptional set  $\mathcal{N}$  be as in Theorem 3.1. Suppose that  $\psi \in C_c(E)$  and that there is a positive left continuous function  $M_\psi(t)$  on  $(0, \infty)$  such that*

$$\|P_t^\psi f\|_\infty \leq M_\psi(t) \|f\|_1 \quad \text{for every } f \in L^1(E, m) \text{ and } t > 0, \quad (3.7)$$

where  $\{P_t^\psi, t \geq 0\}$  is the semigroup defined by  $P^\psi f(x) := e^{\psi(x)} P_t(e^{-\psi} f)(x)$ . Then

$$p(t, x, y) \leq e^{-\psi(x)+\psi(y)} M_\psi(t) \quad \text{for every } t > 0 \text{ and } x, y \in E \setminus \mathcal{N}.$$

**Proof.** Clearly by Theorem 3.1,  $\{P_t^\psi, t \geq 0\}$  admits a heat kernel

$$p^\psi(t, x, y) := e^{\psi(x)} p(t, x, y) e^{-\psi(y)} \quad \text{for } t > 0 \text{ and } x, y \in E \setminus \mathcal{N}.$$

Since  $\psi \in C_c(E)$ , for every  $x \in E \setminus \mathcal{N}$  and  $t > 0$ ,  $y \mapsto p^\psi(t, x, y)$  is  $L^1$ -integrable. Recall the  $\mathcal{E}$ -nest  $\{F_k, k \geq 1\}$  from Theorem 3.1. Since

$$p^\psi(t, x, y) = \int_E p^\psi(t-s, x, z) p^\psi(s, z, y) m(dz),$$

for each fixed  $y \in E \setminus \mathcal{N}$ , it follows from Theorem 3.1 that  $x \mapsto p^\psi(t, x, y)$  is continuous on each  $F_k$ . Thus by (3.7), for every  $s \in (0, t)$  and  $x, y \in E \setminus \mathcal{N}$ ,

$$\begin{aligned} p^\psi(t, x, y) &\leq M_\psi(t-s) \int_E p^\psi(s, z, y) dz \\ &= M_\psi(t-s) e^{\psi(x)} \mathbb{E}_x [e^{-\psi(X_s)}]. \end{aligned}$$

Since  $M_\psi(t)$  is left continuous and  $\psi \in C_c(E)$ , letting  $s \downarrow 0$ , we have by the bounded convergence theorem that

$$p^\psi(t, x) \leq M_\psi(t) e^{\psi(x)} e^{-\psi(x)} = M_\psi(t),$$

and the conclusion of the theorem follows.  $\square$

We are now ready to prove Theorem 1.2. For  $0 < s < 1$ , we use  $\mathbb{H}^s(\mathbb{R}^d)$  to denote the usual Sobolev space of fractional order:

$$\mathbb{H}^s(\mathbb{R}^d) := \left\{ v \in L^2(\mathbb{R}^d, dx) : \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{|v(x) - v(y)|^2}{|x - y|^{d+2s}} dy dx < \infty \right\}. \quad (3.8)$$

**Proof of Theorem 1.2:** We begin with the following inequality of Nash form: for all functions  $u \in \mathbb{H}^{\frac{\alpha}{2}}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$

$$\left( \int_{\mathbb{R}^d} u^2 dx \right)^{1+\frac{\alpha}{d}} \leq c_1 \left( \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|u(x) - u(y)|^2}{|x - y|^{d+\alpha}} dy dx \right) \left( \int_{\mathbb{R}^d} |u(x)| dx \right)^{\frac{2\alpha}{d}}, \quad (3.9)$$

where  $c_1$  is a positive constant depending only on the space dimension  $d$ . This may be proved using the continuous Sobolev embedding  $\mathbb{H}^{\frac{\alpha}{2}}(\mathbb{R}^d) \hookrightarrow L^{2d/(d-\frac{\alpha}{2})}(\mathbb{R}^d)$  and interpolation in  $L^p(\mathbb{R}^d)$  spaces. An alternative way of proving this is to recall that the transition densities for a symmetric stable process of order  $\alpha$  are bounded by  $ct^{-d/\alpha}$  and then to apply Theorem 3.25 of [CKS87].

We then deduce from (A3)

$$\begin{aligned} \|u\|_2^{2+(2\alpha/d)} &\leq c_1 \left( \int_{|x-y|<1} \frac{|u(x) - u(y)|^2}{|x - y|^{d+\alpha}} dy dx + c_2 \int_{\mathbb{R}^d} u(x)^2 dx \right) \|u\|_1^{2\alpha/d} \\ &\leq c_1 (\kappa_1^{-1} \mathcal{E}(u, u) + c_2 \|u\|_2^2) \|u\|_1^{2\alpha/d}. \end{aligned} \quad (3.10)$$

Let  $\{P_t, t \geq 0\}$  denote the transition semigroup of  $X$ ; that is,

$$P_t f(x) := \mathbb{E}^x [f(X_t)] \quad \text{for } f \geq 0 \text{ on } \mathbb{R}^d \text{ and } t > 0.$$

It follows from Theorem 2.1 of [CKS87] that

$$e^{-c_2 \kappa_1 t} \|P_t f\|_\infty \leq c_3 t^{-d/\alpha} \|f\|_1. \quad (3.11)$$

Noting that  $(\mathcal{E}, \mathcal{F})$  is a regular Dirichlet form on  $\mathbb{R}^d$ , the conclusion of this theorem now follows immediately from Theorem 3.1.  $\square$

Let  $B \subset \mathbb{R}^d$  be a ball. Denote by  $X^B$  be the subprocess of  $X$  killed upon leaving  $B$ . Let  $\{P^B(t, x, dy), t > 0\}$  be the transition probability of  $X^B$ . We will need the existence and regularity of the transition density of  $X^B$ .

**Theorem 3.3** *Assume conditions (A1)–(A3) hold. Let  $\mathcal{N}$  be the properly exceptional set of  $X$  in Theorem 1.2. There exist a positive symmetric kernel  $p^B(t, x, y)$  defined on  $(0, \infty) \times (B \setminus \mathcal{N}) \times (B \setminus \mathcal{N})$  such that  $P^B(t, x, dy) = p^B(t, x, y)dy$ , and*

$$p^B(t, x, y) \leq C_1 t^{-d/\alpha} e^{C_2 t} \quad \text{for every } t > 0 \text{ and } x, y \in B \setminus \mathcal{N}, \quad (3.12)$$

where the constants  $C_i$  depends on  $\alpha, \beta, \kappa_1, d$ . Moreover, for every  $t > 0$  and  $y \in B \setminus \mathcal{N}$ ,  $x \mapsto p^B(t, x, y)$  is quasi-continuous on  $B$ .

**Proof.** Let  $\mathcal{N}$  and  $p(t, x, y)$  be the properly exceptional set and the transition density function, resp., in Theorem 1.2. Define  $\tau_B = \inf\{t > 0 : X_t \notin B\}$ . Then

$$p^B(t, x, y) := p(t, x, y) - \mathbb{E}^x [p(t - \tau_B, X_{\tau_B}, y); \tau_B < t], \quad x, y \in B \setminus \mathcal{N},$$

is the transition density function for  $X^B$ . It is easy to see that  $p^B(t, x, y)$  is symmetric and  $y \mapsto p(t, x, y)$  is quasi-continuous. The bound (3.12) is immediate from (1.3).  $\square$

Note that the Dirichlet form for  $X^B$  is  $(\mathcal{E}, \mathcal{F}^B)$ , where

$$\mathcal{F}^B = \{u \in \mathcal{F} : u = 0 \text{ } \mathcal{E}\text{-q.e. on } B^c\}. \quad (3.13)$$

So for  $u \in \mathcal{F}^B$ ,

$$\mathcal{E}(u, u) = \int_{B \times B} (u(x) - u(y))^2 J(x, y) dx dy + \int_B u(x)^2 \kappa_B(x) dx, \quad (3.14)$$

where

$$\kappa_B(x) = 2 \int_{B^c} J(x, y) dy. \quad (3.15)$$

**Remark 3.4** We will use several times the following construction of Meyer [Mey75]. Suppose we have a jump kernel  $J(x, y)$  and another jump kernel  $J_0(x, y) \leq J(x, y)$  such that

$$\mathcal{J}(x) := \int_{\mathbb{R}^d} (J(x, y) - J_0(x, y)) dy \leq c_1 \quad \text{for all } x. \quad (3.16)$$

Let

$$q(x, y) = \frac{J(x, y) - J_0(x, y)}{\mathcal{J}(x)}. \quad (3.17)$$



Let  $Z^{(0)} = \{Z_t^{(0)}, t \geq 0\}$  be the process corresponding to the jump kernel  $J_0$ . Then we can construct a process  $Z$  corresponding to the jump kernel  $J$  as follows. Let  $S_1$  be an exponential random variable of parameter 1 independent of  $Z^0$ , let  $C_t = \int_0^t \mathcal{J}(Z_s^{(0)}) ds$ , and let  $U_1$  be the first time that  $C_t$  exceeds  $S_1$ . We let  $Z_s = Z_s^{(0)}$  for  $0 \leq s \leq U_1$ .

At time  $U_1$  we introduce a jump from  $Z_{U_1-}$  to  $Y_1$ , where  $Y_1$  is chosen at random according to the distribution  $q(Z_{U_1-}, y) dy$ . We set  $Z_{U_1} = Y_1$ , and repeat, using an independent exponential  $S_2$ , etc. Since  $\mathcal{J}(x)$  is bounded, only finitely many new jumps are introduced in any bounded time interval. In [Mey75] it is proved that the resulting process corresponds to the kernel  $J$ . See also [INW66].

Note that if  $\mathcal{N}_0$  is the null set corresponding to  $Z^{(0)}$  then this construction yields that  $\mathcal{N} \subset \mathcal{N}_0$ .

**Remark 3.5** Conversely, we can also remove a finite number of jumps from a process  $Z$  to obtain a new process  $Z^{(0)}$ . Suppose  $J(x, y) = J_0(x, y) + J_1(x, y)$ , where we have  $\int J_1(x, y) dy \leq c_1$  for all  $x$  and for simplicity we also assume that  $J_0(x, y)J_1(x, y) = 0$ . One starts with the process  $Z$  (associated with  $J$ ), runs it until the stopping time  $S_1 = \inf\{t : J_1(Z_{t-}, Z_t) > 0\}$ , and at that time restarts  $Z$  at the point  $Z_{S_1-}$ . One then repeats this procedure. Meyer [Mey75] proves that the resulting process  $Z^{(0)}$  will correspond to the jump kernel  $J_0$ . In this case we have  $\mathcal{N}_0 \subset \mathcal{N}$ .

We will need the following bound, which arises easily from Remark 3.4.

**Lemma 3.6** *Let  $Z^{(0)}, Z$  be as in Remark 3.4, and  $F \in \sigma(Z_t^{(0)}, 0 \leq t < \infty)$ . Then*

$$\mathbb{P}^x(\{Z_s = Z_s^{(0)} \text{ for all } 0 \leq s \leq t\} \cap F) \geq e^{-t\|\mathcal{J}\|_\infty} \mathbb{P}^x(F).$$

*In particular, if  $D \subset \mathbb{R}^d$  and  $A \subset D$ ,*

$$\mathbb{P}^x(Z_t \in A, \tau_D^Z > t) \geq e^{-t\|\mathcal{J}\|_\infty} \mathbb{P}^x(Z_t^{(0)} \in A, \tau_D^{Z^{(0)}} > t). \quad (3.18)$$

**Proof.** We have, writing  $c_1 = \|\mathcal{J}\|_\infty$  and  $G = \{Z_s = Z_s^{(0)}, 0 \leq s \leq t\}$ ,

$$\mathbb{P}^x(G \cap F) = \mathbb{P}^x(U_1 > t, F) = \mathbb{P}^x(C_t < S_1, F) \geq \mathbb{P}^x(S_1 > c_1 t, F) = \mathbb{P}^x(F) e^{-c_1 t}.$$

The last equality holds because  $S_1$  is independent of the process  $Z^{(0)}$ .

For the second assertion, let  $F = \{Z_t^{(0)} \in A, \tau_D^{Z^{(0)}} > t\}$ , and  $G = \{Z_s = Z_s^{(0)} \text{ for all } 0 \leq s \leq t\}$ . Then,

$$\mathbb{P}^x(Z_t \in A, \tau_D^Z > t) \geq \mathbb{P}^x(Z_t \in A, \tau_D^Z > t, G) = \mathbb{P}^x(F \cap G) \geq e^{-t\|\mathcal{J}\|_\infty} \mathbb{P}^x(F).$$

□

Let  $J(x, y) = J_0(x, y) + J_1(x, y)$ , and let  $\mathcal{J}, q$  be defined by (3.16), (3.17). Write  $p^{(0)}(t, x, y)$  for the transition density of the process  $Z^{(0)}$  associated with  $J_0$ . Let

$$r(t, x, y) = \int q(x, z) p(t, z, y) dz. \quad (3.19)$$

The following lemma, which follows quite easily from Meyer's construction, is proved in [BGK, Lemma 3.1].

**Lemma 3.7** (a) For any Borel set  $B$

$$\mathbb{P}^x(Z_t \in B) = \mathbb{P}^x(Z_t^{(0)} \in B, S_1 > t) + \mathbb{E}^x \int_0^t \int_B r(t-s, Z_s^{(0)}, z) \mathcal{J}(Z_s^{(0)}) dz ds. \quad (3.20)$$

(b) If  $\|J_1\|_\infty < \infty$  then

$$p(t, x, y) \leq p^{(0)}(t, x, y) + t\|J_1\|_\infty. \quad (3.21)$$

We need the following standard result.

**Lemma 3.8** Let  $X$  be conservative. Suppose that there exist positive constants  $r$ ,  $t$  and  $p$  such that

$$\mathbb{P}^x(|X_s - x| > r) \leq p \quad \text{for all } x \in \mathbb{R}^d \setminus \mathcal{N} \text{ and } 0 \leq s \leq t. \quad (3.22)$$

Then

$$\mathbb{P}^x\left(\sup_{0 \leq s \leq t} |X_s - X_0| > 2r\right) \leq 2p \quad \text{for all } x \in \mathbb{R}^d \setminus \mathcal{N}.$$

**Proof.** Let  $S = \inf\{t : |X_t - X_0| > 2r\}$ . Then using the strong Markov property of  $X$  and (3.22)

$$\begin{aligned} \mathbb{P}^x\left(\sup_{s \leq t} |X_s - x| > 2r\right) &= \mathbb{P}^x(S \leq t) \\ &\leq \mathbb{P}^x(|X_t - x| > r) + \mathbb{P}^x(S \leq t \text{ and } |X_t - x| \leq r) \\ &\leq p + \mathbb{P}^x(S \leq t \text{ and } |X_t - X_S| > r) \\ &= p + \mathbb{E}^x\left(\mathbb{1}_{\{S \leq t\}} \mathbb{P}^{X_S}(|X_{t-S} - X_0| > r)\right) \leq 2p. \end{aligned}$$

□

We now use Lemma 3.7 to obtain off-diagonal upper bounds on  $p(t, x, y)$ ; this technique was introduced in [BGK].

**Proposition 3.9** Let  $\mathcal{N}$  be the properly exceptional set of  $X$  in Theorem 1.2. There exist constants  $t_0 > 0$ ,  $c_1, c_2, c_3 > 0$  such that for every  $x, y \in \mathbb{R}^d \setminus \mathcal{N}$ ,

$$p(t, x, y) \leq c_2 t e^{-c_3|x-y|} \quad \text{for } t \in (0, 2], |x - y| \geq \frac{1}{16}. \quad (3.23)$$

**Proof.** For  $\delta \in (0, e)$  let  $J^{(\delta)}(x, y) = J(x, y)\mathbb{1}_{(|x-y| < \delta)}$ , and let  $\mathcal{E}^{(\delta)}$  be defined by (1.1) with  $J^{(\delta)}$  in place of  $J$ . (Of course, if  $\delta \geq 1$ , then  $J^{(\delta)} = J$  and  $\mathcal{E}^{(\delta)} = \mathcal{E}$ .) Clearly  $(\mathcal{E}^{(\delta)}, \mathcal{F})$  is a regular Dirichlet form on  $\mathbb{R}^d$ . Thus there is a Hunt process  $X^{(\delta)}$  associated with  $(\mathcal{E}^{(\delta)}, \mathcal{F})$  and a properly exceptional set  $\mathcal{N}(\delta)$  so that  $X$  starts from every point in  $\mathbb{R}^d \setminus \mathcal{N}(\delta)$ . Since using Meyer's procedure we can construct  $X^{(\delta)}$  from  $X$  and vice versa, we can take  $\mathcal{N}(\delta) = \mathcal{N}$ . Let  $p^{(\delta)}(t, x, y)$  be the transition density of  $X^{(\delta)}$ .

By (A1)–(A3) we have

$$\begin{aligned} \int_{B(x,\delta)^c} J(x,y) dy &= \int (J(x,y) - J^{(\delta)}(x,y)) dy \leq c_4 \delta^{-\beta} \mathbf{1}_{\{\delta < 1\}} \leq c_4 \delta^{-\beta}, \\ \int_{B(x,\delta)} |x-y|^2 J(x,y) dy &= \int |x-y|^2 J^{(\delta)}(x,y) dy \leq c_5 \delta^{2-\beta}. \\ \|J - J^{(\delta)}\|_\infty &= \sup_{|x-y| \geq \delta} J(x,y) \leq c_6 \delta^{-d-\beta} \mathbf{1}_{\{\delta < 1\}} \leq c_6 \delta^{-d-\beta}. \end{aligned}$$

Starting from (3.10) we have

$$\begin{aligned} \|u\|_2^{2+(2\alpha/d)} &\leq c_7 \left( \int \int_{|x-y| < (\delta \wedge 1)} \frac{|u(x) - u(y)|^2}{|x-y|^{d+\alpha}} dy dx + (\delta \wedge 1)^{-\alpha} \|u\|_2^2 \right) \|u\|_1^{2\alpha/d} \\ &\leq c_8 (\kappa_1^{-1} \mathcal{E}^{(\delta)}(u, u) + (\delta \wedge 1)^{-\alpha} \|u\|_2^2) \|u\|_1^{2\alpha/d}. \end{aligned}$$

Theorem 3.2 in Section 3 and Theorem 3.25 of [CKS87] now give

$$p^{(\delta)}(t, x, y) \leq c_9 t^{-d/\alpha} e^{c_{10} t (\delta \wedge 1)^{-\alpha}} e^{-E_\delta(2t, x, y)} \quad \text{for } x, y \in \mathbb{R}^d \setminus \mathcal{N}(\delta), \quad (3.24)$$

where

$$\begin{aligned} \Gamma_\delta(f, f)(x) &= \int (f(x) - f(y))^2 J^{(\delta)}(x, y) dy, \\ \Lambda_\delta(\psi)^2 &= \|e^{-2\psi} \Gamma_\delta(e^\psi, e^\psi)\|_\infty \vee \|e^{2\psi} \Gamma_\delta(e^{-\psi}, e^{-\psi})\|_\infty, \\ E_\delta(t, x, y) &= \sup \{ |\psi(x) - \psi(y)| - t \Lambda_\delta(\psi)^2 : \psi \in \mathcal{F} \cap C_b(\mathbb{R}^d) \text{ with } \Lambda_\delta(\psi) < \infty \}. \end{aligned}$$

Let  $t > 0$  and  $x_0, y_0 \in \mathbb{R}^d \setminus \mathcal{N}$ ; write  $R = |x_0 - y_0|$ . Note that the set  $\mathcal{C}$  of compactly supported Lipschitz continuous functions is a core for  $\mathcal{E}^{(\delta)}$ . Let  $\lambda > 0$  and

$$\psi(x) = \lambda(R - |x_0 - x|)^+.$$

So  $|\psi(x) - \psi(y)| \leq \lambda|x - y|$ . Noting that  $|e^t - 1|^2 \leq t^2 e^{2|t|}$ , we obtain

$$\begin{aligned} e^{-2\psi(x)} \Gamma_\delta(e^\psi, e^\psi)(x) &= \int (e^{\psi(x)-\psi(y)} - 1)^2 J^{(\delta)}(x, y) dy \\ &\leq e^{2\lambda\delta} \lambda^2 \int |x-y|^2 J^{(\delta)}(x, y) dy \\ &\leq c_{11} (\lambda\delta)^2 e^{2\lambda\delta} \delta^{-\beta} \leq c_{12} e^{3\lambda\delta} \delta^{-\beta}. \end{aligned}$$

Since the same bound holds for  $e^{2\psi(x)} \Gamma_\delta(e^{-\psi}, e^{-\psi})(x)$  we have

$$-E_\delta(2t, x_0, y_0) \leq -\lambda R + c_{12} t \delta^{-\beta} e^{3\lambda\delta}. \quad (3.25)$$

In what follows we will always choose  $\delta$  so that  $\delta \leq e$ , and  $t$  such that  $t < \delta^\beta$ . Hence  $t \leq \delta^\beta \leq c_{13} \delta^\alpha$ , so that the term  $e^{c_{10} t (\delta \wedge 1)^{-\alpha}}$  in (3.24) is bounded by  $c_{14}$ . We take

$$\lambda = \frac{1}{3\delta} \log \left( \frac{\delta^\beta}{t} \right) > 0. \quad (3.26)$$

So

$$-E_\delta(2t, x_0, y_0) \leq -\frac{R}{3\delta} \log\left(\frac{\delta^\beta}{t}\right) + c_{15}t\delta^{-\beta}\left(\frac{\delta^\beta}{t}\right) = \frac{R}{3\delta} \log\left(\frac{t}{\delta^\beta}\right) + c_{16}. \quad (3.27)$$

Hence by (3.24),

$$p^{(\delta)}(t, x_0, y_0) \leq c_{17}t^{-d/\alpha}\left(\frac{t}{\delta^\beta}\right)^{R/(3\delta)}. \quad (3.28)$$

We need to consider two cases. Set  $R_0 = 3e(1+d/\alpha)$ . First, suppose  $R \geq R_0$  and  $t \leq 1$ . Set  $\delta = e$ . Then since there are no jumps larger than 1, we have  $p(t, x_0, y_0) = p^{(\delta)}(t, x_0, y_0)$ , and by (3.28)

$$p(t, x_0, y_0) \leq c_{17}t^{-d/\alpha}\left(\frac{t}{\delta^\beta}\right)^{R/(3\delta)} \leq c_{18}te^{-\beta R/3e}. \quad (3.29)$$

Next, suppose  $0 < R \leq R_0$ , and set  $\delta = R\alpha/(3(\alpha+d))$ . We assume that  $t < \delta^\beta = c_{19}R^\beta$ . Then since  $R/(3\delta) = 1 + d/\alpha$ , (3.21) and (3.28) give

$$p(t, x_0, y_0) \leq c_{20}\frac{t}{\delta^{\beta+\beta d/\alpha}} + c_{21}\frac{t}{\delta^{\beta+d}} \leq c_{22}\frac{t}{R^{\beta+\beta d/\alpha}} + c_{23}\frac{t}{R^{\beta+d}} \quad (3.30)$$

Combining (3.29) and (3.30), we deduce there exists  $t_1 > 0$  such that if  $|x_0 - y_0| = R$  and  $R \geq 1/16$ , and  $t \leq t_1$ , then

$$p(t, x_0, y_0) \leq c_{24}te^{-c_{25}R} \quad \text{for } t \in (0, t_1]. \quad (3.31)$$

Combining this with (1.3), (3.23) follows.  $\square$

**Proof of Theorem 1.1(b):** Let  $X^n$  be the Hunt process associated with the regular Dirichlet form  $(\mathcal{E}^{(1/n)}, \mathcal{F}^{(1/n)})$ , whose transition semigroup will be denoted as  $\{P_t^n, t \geq 0\}$ . Note that  $X^n$  has jump kernel  $J_{1/n}$  as defined by (1.5). Since the symmetric  $\beta$ -stable process is conservative, it follows from Remark 3.4 on Meyer's construction that  $X^n$  is conservative. By Theorem 3.1 and (3.23),  $X^n$  admits a quasi-continuous transition density function  $p_n(t, x, y)$  with a properly exceptional set  $\mathcal{N}_n \subset \mathbb{R}^d$  of  $X^n$  such that

$$p_n(t, x, y) \leq c_1 t e^{-c_2|x-y|} \quad (3.32)$$

for  $t \in (0, 2]$  and  $x, y \in \mathbb{R}^d \setminus \mathcal{N}_n$  with  $|x - y| \geq 1$ . Here  $c_1$  and  $c_2$  are two positive constants independent of  $n \geq 1$ . Let  $p(t, x, y)$  be the quasi-continuous transition density function of  $X$  with properly exceptional set  $\mathcal{N}$ . Since the norm  $\mathcal{E}_1^{(1/n)}$  dominates the norm  $\mathcal{E}_1$ , each  $\mathcal{N}_n$  is  $\mathcal{E}$ -exceptional. Thus without loss of generality, we may assume that  $\cup_{n=1}^\infty \mathcal{N}_n \subset \mathcal{N}$ . Let  $B_r := B(0, r)$ . Then the conservativeness of  $X^n$  and (3.32) implies that for every  $\varepsilon \in (0, 1)$ , there is  $r_\varepsilon \geq 1$  so that for every  $n \geq 1$  and every  $x \in B_1 \setminus \mathcal{N}$ ,

$$P_1^n \mathbb{1}_{B_{r_\varepsilon}}(x) = 1 - \int_{B_{r_\varepsilon}^c} p_n(1, x, y) dy \geq 1 - \varepsilon. \quad (3.33)$$

On the other hand, by Theorem 1.5,  $X^n$  converges to  $X$  in the sense of Mosco. Thus we have for every  $r > 0$ ,

$$\lim_{n \rightarrow \infty} \|P_1^n \mathbb{1}_{B_r} - P_1 \mathbb{1}_{B_r}\|_{L^2(B_1)} \leq \lim_{n \rightarrow \infty} \|P_1^n \mathbb{1}_{B_r} - P_1 \mathbb{1}_{B_r}\|_{L^2(\mathbb{R}^d)} = 0$$

It follows then from (3.33), for every  $\varepsilon \in (0, 1)$ ,

$$\limsup_{n \rightarrow \infty} \|1 - P_1 \mathbb{1}\|_{L^2(B_1)} \leq \limsup_{n \rightarrow \infty} \|1 - P_1 \mathbb{1}_{B_r}\|_{L^2(B_1)} \leq \varepsilon |B_1|,$$

where  $|B_1|$  denotes the Lebesgue measure of the unit ball  $B_1$ . Since  $\varepsilon \in (0, 1)$  can be arbitrarily small, we have  $P_1 \mathbb{1}(x) = 1$  for a.e.  $x \in B_1$ . By the quasi-continuity of  $P_1 \mathbb{1}$ , we have  $P_1 \mathbb{1}(x) = 1$  for every  $x \in B_1 \setminus \mathcal{N}$  and hence for every  $x \in \mathbb{R}^d \setminus \mathcal{N}$ . This proves the conservativeness of  $X$ .  $\square$

**Proposition 3.10** *There exist constants  $c_1, c_2$  and  $t_0 > 0$  such that*

(a)  $\mathbb{P}^x\left(\sup_{s \leq t} |X_s - x| > r\right) < c_1 t e^{-c_2 r}$  for every  $r \geq 1/8, t \leq t_0$ .

(b) *There exists  $t_1 > 0$  such that*

$$\mathbb{P}^x\left(\sup_{s \leq t_0} |X_s - x| > 1/4\right) < 1/4. \quad (3.34)$$

**Proof.** Since by Theorem 1.1(b)  $X$  is conservative, by integrating (3.23) over  $B(x, R)^c$ , we obtain, for  $x \notin \mathcal{N}$ ,

$$\mathbb{P}^x(X_t \notin B(x, R)) \leq c_{26} t e^{-c_{27} R} \quad \text{if } R \geq 1/16 \text{ and } t \in (0, t_1],$$

and (b) follows by Lemma 3.8.  $\square$

## 4 Lower bounds on the heat kernel

Throughout this section we will assume that the jump kernel  $J$  satisfies (A1)–(A3) and also (A4)( $\xi$ ) for some  $\xi \in (0, 1)$ , unless otherwise specified.

In view of Remark 3.4 and the fact the symmetric  $\alpha$ -stable process is a Lévy process that starts from every point in  $\mathbb{R}^d$ , we can take the properly exceptional set  $\mathcal{N}$  to be the empty set throughout this section.

Let  $B$  be a ball of radius  $R$  centered at 0 for  $R \in [1, 4]$ . Let  $p^B(t, x, y)$  be the transition density function of the subprocess  $X^B$  of  $X$  killed upon exiting the ball  $B$ .

**Lemma 4.1** *There exist positive constants  $c_1, c_2$  independent of  $\xi$  such that*

$$p^B(t, x, y) \leq c_1 t^{-d/\alpha} e^{c_2 t} \quad \text{and} \quad \left| \frac{\partial p^B(t, x, y)}{\partial t} \right| \leq c_1 t^{-1-\frac{d}{\alpha}} e^{c_2 t} \quad (4.1)$$

for every  $x, y \in B$  and  $t > 0$ .

**Proof.** The first inequality follows immediately from Theorem 1.2.

As

$$\int_{B \times B} p^B(t, x, y)^2 dx dy = \int_B p^B(2t, x, x) dx < \infty,$$

the symmetric semigroup  $P_t^B$  of  $X^B$  is a Hilbert-Schmidt operator on  $L^2(B, dx)$  and so it is compact and has a discrete spectrum  $\{e^{-\lambda_i t}, i \geq 1\}$ , with repetitions according to multiplicity. Let  $\{\psi_i, i \geq 1\}$  be the corresponding eigenfunctions normalized to have unit  $L^2$ -norm on  $B$  and to be orthogonal to each other. Then it is well known (see, e.g., [Bas97], Section VII.6) that

$$p^B(t, x, y) = \sum_{i=1}^{\infty} e^{-\lambda_i t} \psi_i(x) \psi_i(y).$$

Hence

$$\frac{\partial p^B(t, x, y)}{\partial t} = - \sum \lambda_i e^{-\lambda_i t} \psi_i(x) \psi_i(y), \quad (4.2)$$

where the convergence is in the  $L^2$ -sense. Note that for any given  $\delta \in (0, 1/2)$ , by Cauchy-Schwarz and the Chapman-Kolmogorov equations,

$$\begin{aligned} \psi_i(x) &= e^{\lambda_i \delta} \int_B p^B(\delta, x, y) \psi_i(y) dy \\ &\leq e^{\lambda_i \delta} \left( \int_B p^B(\delta, x, y)^2 dy \right)^{1/2} \left( \int_B (\psi_i(y))^2 dy \right)^{1/2} \\ &\leq e^{\lambda_i \delta} p(2\delta, x, x)^{1/2} \leq c_3 \delta^{-d/\alpha} e^{(c_4 + \lambda_i) \delta} \end{aligned}$$

for every  $x \in B$  by (1.3). It follows that the series on the right hand side of (4.2) converges uniformly on  $(3\delta, \delta^{-1}) \times B \times B$ . The function  $x \rightarrow x e^{-xt/2}$  is bounded over nonnegative  $x$  by  $c_5/t$ , so for  $x, y \in B$ , the right hand side of (4.2) is bounded in absolute value by

$$\begin{aligned} \frac{c_5}{t} \sum_i e^{-\lambda_i t/2} |\psi_i(x)| |\psi_i(y)| &\leq \frac{c_5}{t} \left( \sum_i e^{-\lambda_i t/2} \psi_i(x)^2 \right)^{1/2} \left( \sum_i e^{-\lambda_i t/2} \psi_i(y)^2 \right)^{1/2} \\ &= \frac{c_5}{t} p^B(t/2, x, x)^{1/2} p^B(t/2, y, y)^{1/2}. \end{aligned}$$

The upper bound for  $p^B(t/2, x, y)$  in (4.1) now yields the desired bound on  $|\partial p^B(t, x, y)/\partial t|$ .

□

**Lemma 4.2** *Let  $Y$  be a pure jump symmetric process on  $\mathbb{R}^d$  with jump kernel  $J_Y(x, y)$  satisfying*

$$c_1 |x - y|^{-d-\beta} \leq J_Y(x, y) \leq c_2 |x - y|^{-d-\beta}, \quad x, y \in \mathbb{R}^d. \quad (4.3)$$

*Let  $D = B(0, \frac{1}{2})$  and  $p_Y^D(t, x, y)$  be the transition density of the subprocess of  $Y$  killed upon exiting  $D$ . Then there exist constants  $c_3, c_4$  (depending on  $c_1, c_2$ ) such that*

$$p_Y^D(t, x, y) \geq c_3 e^{-c_4 t} \quad \text{for } t \geq \frac{3}{4} \text{ and } x, y \in B(0, \frac{1}{4}).$$

**Proof.** It is known from [CK03] that such a process  $Y$  is a Feller process that can start from every point in  $\mathbb{R}^d$  and has a Hölder continuous transition density function  $p(t, x, y)$ . This implies in particular that the transition density function  $p_Y^D(t, x, y)$  for the subprocess of  $Y$  killed upon leaving  $D$  exists and is continuous on  $(0, \infty) \times D \times D$ . For each fixed  $x \in D$ ,  $(t, y) \mapsto p_Y^D(t, x, y)$  is a caloric function of  $Y$  in  $(0, \infty) \times D$ .

Let  $B' = B(0, \frac{1}{4})$ . Then  $\mathbb{P}^0(\tau_{B'}^Y \geq \frac{1}{2}) \geq c_5 > 0$ . By [CK03] a parabolic Harnack inequality holds for  $Y$ . Therefore for every  $x \in D$ ,  $\frac{1}{4} \leq t_1 \leq \frac{1}{2}$  and  $\frac{3}{4} \leq t_2 \leq 1$ ,

$$p_Y^D(t_1, x, y) \leq c_6 p_Y^D(t_2, x, y') \quad \text{for } y, y' \in B'.$$

So if  $x_1, y_1 \in B'$ ,

$$\begin{aligned} c_5 &\leq \int_{B'} p_Y^{B'}(t_1, x_1, y) dy \\ &\leq \int_{B'} p_Y^D(t_1, x_1, y) dy \\ &\leq c_6 \int_{B'} p_Y^D(t_2, x_1, y_1) dy = c_6 |B'| p_Y^D(t_2, x_1, y_1). \end{aligned}$$

This proves the result for  $t \in [\frac{3}{4}, 1]$ . An easy iteration argument now gives the result for  $t > 1$ .  $\square$

**Proposition 4.3** *Let  $B = B(0, R)$  with  $R \in [1, 4]$  and  $0 < t_0 < t_1 < \infty$ . There is a constant  $c_1 = c_1(\xi, t_0, t_1) > 0$  such that*

$$p^B(t, x, y) \geq c_1 (R - |x|)^\beta (R - |y|)^\beta \quad \text{for every } t \in [t_0, t_1] \text{ and } x, y \in B.$$

**Proof.** Recall we are working under (A4)( $\xi$ ); the constants  $c_i$  in this proof will depend on  $\xi$ . We can assume without loss of generality that  $\xi \leq \frac{1}{8}$ . By considering the process  $X'_t = X_{2t/t_0}$ , which satisfies (A1)–(A3) and (A4)( $\xi$ ) (but with different constants  $\kappa_i$ ), we can assume that  $t_0 \geq 2$ .

Let  $V$  be a Lévy process with jump kernel

$$J_0(x, y) = \begin{cases} \kappa_2 |y - x|^{-d-\beta} & \text{if } |y - x| \leq \xi, \\ \kappa_1 |y - x|^{-d-\alpha} & \text{if } \xi < |y - x| \leq 1, \\ 0 & \text{if } |y - x| > 1. \end{cases}$$

We can construct  $X$  from  $V$  by using the construction of Remark 3.4. Since  $\mathcal{J}(x) := \int_{\mathbb{R}^n} (J(x, y) - J_0(x, y)) dy$  is bounded, by Lemma 3.6 there is a positive constant  $c_2 = c_2(t_1) > 0$  such that

$$\mathbb{P}^x (X_t \in A \text{ and } \tau_B^X > t) \geq c_2 \mathbb{P}^x (V_t \in A \text{ and } \tau_B^V > t)$$

for every  $t \in (0, t_1]$  and  $A \subset B$ . It thus suffices to get a lower bound on  $v^B(t, x, y)$ , the transition density of the subprocess of  $V$  killed upon exiting  $B$ .

By the Chapman-Kolmogorov equations,

$$v^B(t, x, y) \geq \int_{B(0, 1/4)} v^B(t/2, x, z) v^B(t/2, z, y) dz.$$

Therefore it is enough to prove that

$$v^B(t, y, z) \geq c_3(R - |y|)^\beta \quad \text{for } y \in B, z \in B(0, 1/4), t \in [1, t_1]. \quad (4.4)$$

Let  $y \in B$ ,  $\varepsilon = R - |y|$ , and  $\eta = (\varepsilon \wedge \xi)/2$ . Let  $y_0 \in B(0, R - \frac{1}{2})$  with  $|y_0 - y| \leq \frac{3}{4}$ . Our first estimate is to prove that there exist  $c_4, c_5 > 0$  such that

$$\mathbb{P}^y(V_{t\eta^\beta} \in B(y_0, 1/4) \text{ and } \tau_B^V > t\eta^\beta) \geq c_5 t\eta^\beta \quad \text{for } t \leq c_4. \quad (4.5)$$

To prove this, define

$$V_t^\eta = V_t - \sum_{0 < s \leq t} \Delta V_s \mathbb{1}_{(|\Delta V_s| > \eta)}.$$

Note that  $V^\eta$  is a Lévy process with jump kernel  $\kappa_2|x - w|^{-d-\beta} \mathbb{1}_{(|x-w| \leq \eta)}$  and that the processes  $V^\eta$  and  $V - V^\eta$  are independent. We write  $J_1(x, w)$  for the jump kernel of  $V - V^\eta$ .

Each coordinate of  $V^\eta$  is a square integrable martingale. Applying Doob's maximal inequality to each coordinate of  $V^\eta$  separately yields

$$\begin{aligned} \mathbb{P}^y\left(\sup_{s \leq t\eta^\beta} |V_s^\eta - V_0^\eta| \geq \eta\right) &\leq 4d(\eta/d)^{-2} \mathbb{E}\left[|V_{t\eta^\beta}^\eta - V_0^\eta|^2\right] \\ &= 4d^3 t\eta^\beta \eta^{-2} \int_{|h| \leq \eta} \kappa_2 |h|^{2-d-\beta} dh \\ &\leq c_6 t. \end{aligned} \quad (4.6)$$

Let  $c_4 = \min\left\{\frac{1}{4c_6}, \frac{1}{24}\right\}$ .

There are constants  $c_8 > c_7 > 0$  such that the total Lévy measure of  $V - V^\eta$  is bounded by

$$c_7 \eta^{-\beta} \leq \int_{\mathbb{R}^d} J_1(0, x) dx \leq c_8 \eta^{-\beta}.$$

Since  $|y_0 - y| \leq \frac{3}{4}$ ,

$$\int_{B(y_0 - y, 1/8)} J_1(0, w) dw \geq c_9.$$

For  $t > 0$  let  $F(t)$  be the complement of the event on the left side of (4.6). Take  $c_4 > 0$  small so that for any  $t \leq c_4$ ,  $\mathbb{P}(F(t)) \geq 3/4$ . Let  $G(t)$  be the event that  $V - V^\eta$  makes exactly one jump in the time interval  $[0, t\eta^\beta]$ , and that the size falls within the ball  $B(y_0 - y, 1/8)$ . Then

$$\mathbb{P}^y(G(t)) \geq (c_7 \eta^{-\beta})(t\eta^\beta) e^{-(c_8 \eta^{-\beta})(t\eta^\beta)} \frac{c_9}{c_8 \eta^{-\beta}} \geq c_{10} t\eta^\beta.$$



Note that  $\eta \leq \xi/2 \leq 1/16$ . If both  $F(t)$  and  $G(t)$  occur, since  $V^\eta = V_0 = y$ , then  $V_{t\eta^\beta} \in B(y_0, 1/4)$  and  $\tau_B^V > t\eta^\beta$ . As  $V^\eta$  and  $V - V^\eta$  are independent, we obtain (4.5). In particular, we have for every  $y \in B$ , with  $s_1 := c_4 2^{-\beta} ((R - |y|) \wedge \xi)^\beta$ ,

$$\begin{aligned} \mathbb{P}^y (V_{s_1} \in B(0, R - \tfrac{1}{4}) \text{ and } \tau_B^V > s_1) &\geq \mathbb{P}^y (V_{s_1} \in B(y_0, 1/4) \text{ and } \tau_B^V > s_1) \\ &\geq c_5 s_1. \end{aligned}$$

Similarly, for every  $z \in B(0, R - \frac{1}{4})$ , let  $z_0 \in B(0, R - \frac{3}{4})$  with  $|z - z_0| < 3/4$ . Note that in this case  $R - |z| \geq 1/4$  and  $\xi \leq 1/8$  and so  $\eta := \frac{(R - |z|) \wedge \xi}{2} = \frac{\xi}{2}$ . The same argument as above shows that, with  $s_2 := c_4 2^{-\beta} \xi^\beta$ ,

$$\begin{aligned} \mathbb{P}^z (V_{s_2} \in B(0, R - \tfrac{1}{2}) \text{ and } \tau_B^V > s_2) &\geq \mathbb{P}^z (V_{s_2} \in B(z_0, 1/4) \text{ and } \tau_B^V > s_2) \\ &\geq c_5 s_2. \end{aligned}$$

Applying the Markov property of  $V$  at time  $s_1$ , we have for  $y \in B$ ,

$$\begin{aligned} &\mathbb{P}^y (V_{s_1+s_2} \in B(0, R - \tfrac{1}{2}) \text{ and } \tau_B^V > s_1 + s_2) \\ &\geq \mathbb{P}^y (\mathbb{P}^{V_{s_1}} (V_{s_2} \in B(0, R - \tfrac{1}{2}) \text{ and } \tau_B^V > s_2); V_{s_1} \in B(0, R - \tfrac{1}{4}) \text{ and } \tau_B^V > s_1) \\ &\geq c_5 s_2 \mathbb{P}^y (V_{s_1} \in B(0, R - \tfrac{1}{4}) \text{ and } \tau_B^V > s_1) \\ &\geq c_5^2 s_2 s_1. \end{aligned}$$

Repeating this at most a number  $[4R]$  of times, for any  $y \in B$ , there exists  $s_0 \leq \frac{1}{4}$  such that

$$\mathbb{P}^y (V_{s_0} \in B(0, 1/4) \text{ and } \tau_B^V > s_0) \geq c_{11} (R - |y|)^\beta. \quad (4.7)$$

Now let

$$J_2(x, y) = J_0(x, y) + \kappa_2 |x - y|^{-d-\beta} \mathbb{1}_{(|x-y|>1)},$$

and  $Y$  be the associated Lévy process; note that  $J_2$  satisfies (4.3). We can use Remark 3.4 to construct  $Y$  from  $V$ . Let  $T$  be the time of the first added jump, so that  $V_s = Y_s$ ,  $0 \leq s < T$  and  $|\Delta Y_T| > 1$ . Let  $B' = B(0, \frac{1}{2})$ . If  $x \in B'$  and  $A \subset B'$  then  $T \geq \tau_{B'}^Y \geq \tau_{B'}^V$ . Therefore

$$\mathbb{P}^x (V_t \in A \text{ and } \tau_{B'}^V > t) \geq \mathbb{P}^x (Y_t \in A \text{ and } \tau_{B'}^Y > t),$$

from which it follows that

$$v^B(t, x, y) \geq v^{B'}(t, x, y) \geq p_Y^{B'}(t, x, y),$$

for  $x, y \in B'$ . So, using Lemma 4.2

$$v^B(t, x, y) \geq c_{12} \quad \text{for } x, y \in B(0, \tfrac{1}{4}) \text{ and } \tfrac{3}{4} \leq t \leq t_1. \quad (4.8)$$

Hence for  $t \in [t_0, t_1]$ ,  $y \in B$  and  $z \in B(0, \frac{1}{4})$ , with  $s_0$  the time corresponding to  $y$  in (4.7), we have from (4.8)

$$\begin{aligned} v^B(t, y, z) &\geq \int_{B(0, 1/4)} v^B(s_0, y, w) v^B(t - s_0, w, z) dw \\ &\geq c_{12} \int_{B(0, 1/4)} v^B(s_0, y, w) dw \\ &= c_{12} \mathbb{P}^y (V_{s_0} \in B(0, 1/4) \text{ and } \tau_B^V > \frac{1}{4}) \\ &\geq c_{12} c_{11} (R - |y|)^\beta, \end{aligned}$$

which proves (4.4).  $\square$

Define

$$\varphi(x) = \left( (R^2 - |x|^2)^+ \right)^{12/(2-\beta)}. \quad (4.9)$$

The following corollary follows immediately from Proposition 4.3.

**Corollary 4.4** *For every  $y \in B$ ,  $\delta \in (0, 1)$  and  $\gamma \in (\frac{2-\beta}{6}, 1]$ , there is a constant  $C = C(y, \beta, \delta, \xi) > 0$  such that*

$$\varphi(x)^\gamma / p^B(t, x, y) \leq C, \quad \text{for every } t \in [\delta, 4] \text{ and } x \in B.$$

Recall the definition of  $\mathcal{F}^B$  from Section 3.

**Lemma 4.5** *For every  $t > 0$  and  $y \in B$ ,  $p^B(t, x, y)$  as a function of  $x \in B$  is in  $\mathcal{F}^B$ .*

**Proof.** Fix  $y \in B$ . By Lemma 4.1, if  $t > 0$ ,  $f(x) := p^B(t/2, x, y) \in L^2(B, dx)$  and so (see [FOT94, Lemma 1.3.3]),  $p^B(t, x, y) = P_{t/2}^B f(x)$  as a function of  $x \in B$  is in  $\mathcal{F}^B$ .  $\square$

**Lemma 4.6** *For each  $t > 0$  and  $y_0 \in B$ , the function  $\varphi(\cdot)/p^B(t, \cdot, y_0)$  is in  $\mathcal{F}^B$ .*

**Proof.** Without loss of generality, we may assume that  $t < 4$ . By Corollary 4.4,  $\varphi(\cdot)^{1/2}/p^B(t, \cdot, y_0)$  is bounded on  $B$ . By extending the function  $x \mapsto \varphi(\cdot)/p^B(t, \cdot, y_0)$  to be zero on  $B^c$ , we see that  $x \mapsto u(x) := \varphi(\cdot)/p^B(t, \cdot, y_0)$  vanishes continuously on  $B^c$ . Recall the killing function  $\kappa_B$  from (3.15). It is clear that  $\int_B u(x)^2 \kappa_B(dx) < \infty$ . So by (3.13)–(3.14) and Lemma 2.5, it suffices to show

$$\int_{B \times B} \left( \frac{\varphi(x)}{p^B(t, x, y_0)} - \frac{\varphi(y)}{p^B(t, y, y_0)} \right)^2 |x - y|^{-d-\beta} dx dy < \infty. \quad (4.10)$$

We look at

$$\int \int_{x, y \in B, \varphi(y) \geq \varphi(x)} \left( \frac{\varphi(y)}{p^B(t, y, y_0)} - \frac{\varphi(x)}{p^B(t, x, y_0)} \right)^2 |x - y|^{-d-\beta} dx dy;$$

if we show this is finite, we will have finiteness of the integral over  $\varphi(x) > \varphi(y)$  by the same argument, and combining the two estimates shows (4.10). We need to bound

$$\begin{aligned} & \int \int_{x,y \in B, \varphi(y) \geq \varphi(x)} \left( \frac{\varphi(y) - \varphi(x)}{p^B(t, y, y_0)} \right)^2 |x - y|^{-d-\beta} dx dy \\ & + \int \int_{x,y \in B, \varphi(y) \geq \varphi(x)} \left( \varphi(x) \frac{p^B(t, y, y_0) - p^B(t, x, y_0)}{p^B(t, x, y_0)p^B(t, y, y_0)} \right)^2 |x - y|^{-d-\beta} dx dy \\ & = I_1 + I_2. \end{aligned}$$

On the set  $\{(x, y) \in B \times B : \varphi(y) > \varphi(x)\}$ , since  $\varphi(x) \leq \varphi(x)^{1/2}\varphi(y)^{1/2}$  and  $\varphi(x)^{1/2}/p^B(t, x, y_0)$  is bounded in  $x$  by Corollary 4.4, the finiteness of  $I_2$  follows from Lemma 4.5. To handle  $I_1$ , let  $\gamma = (2 - \beta)/3$ . Note that  $|\varphi(y) - \varphi(x)|^{2\gamma} \leq |\varphi(y)|^{2\gamma}$ , and that  $\varphi(x)^{2\gamma}/p^B(t, x, y_0)^2$  is bounded in  $x$  by Lemma 4.3. Since  $\varphi$  has a bounded derivative,  $|\varphi(y) - \varphi(x)|^{2-2\gamma} \leq c_1|y - x|^{2-2\gamma}$ . As  $|y - x|^{2-2\gamma-d-\beta}$  is integrable over  $B \times B$  since  $2\gamma < 2 - \beta$ , the finiteness of  $I_1$  follows.  $\square$

**Lemma 4.7** Fix  $y_0 \in B$  and let  $G(t) = \int_B \varphi(x) \log p^B(t, x, y_0) dx$ . Then for every  $t > 0$ ,

$$G'(t) = -\mathcal{E} \left( p^B(t, \cdot, y_0), \frac{\varphi(\cdot)}{p^B(t, \cdot, y_0)} \right).$$

**Proof.** Write  $(f, g)$  for  $\int_B f(x)g(x) dx$ . Using Lemmas 4.5 and 4.6, we know by [FOT94, Lemma 1.3.4] that

$$\begin{aligned} -\mathcal{E} \left( p^B(t, \cdot, y_0), \frac{\varphi(\cdot)}{p^B(t, \cdot, y_0)} \right) &= \lim_{h \downarrow 0} \frac{1}{h} \left( P_h^B p^B(t, \cdot, y_0) - p^B(t, \cdot, y_0), \frac{\varphi(\cdot)}{p^B(t, \cdot, y_0)} \right) \\ &= \lim_{h \downarrow 0} \frac{1}{h} \left( p^B(t + h, \cdot, y_0), -p^B(t, \cdot, y_0), \frac{\varphi(\cdot)}{p^B(t, \cdot, y_0)} \right) \\ &= \lim_{h \downarrow 0} \frac{1}{h} \int_B \varphi(x) \left( \frac{p^B(t + h, x, y_0)}{p^B(t, x, y_0)} - 1 \right) dx. \end{aligned}$$

Also

$$G'(t) = \lim_{h \rightarrow 0} \frac{1}{h} \int_B (\log p^B(t + h, x, y_0) - \log p^B(t, x, y_0)) \varphi(x) dx.$$

Let

$$F(h) = \left[ \log p^B(t + h, x, y_0) - \log p^B(t, x, y_0) - \left( \frac{p^B(t + h, x, y_0)}{p^B(t, x, y_0)} - 1 \right) \right] \varphi(x).$$

Then

$$\begin{aligned} F'(h) &= \frac{\partial}{\partial t} p^B(t + h, x, y_0) \left( \frac{1}{p^B(t + h, x, y_0)} - \frac{1}{p^B(t, x, y_0)} \right) \varphi(x) \\ &= \frac{\partial}{\partial t} p^B(t + h, x, y_0) (p^B(t + h, x, y_0) - p^B(t, x, y_0)) \frac{\varphi(x)}{p^B(t + h, x, y_0)p^B(t, x, y_0)}. \end{aligned}$$

By the mean value theorem,  $F(h)/h = F'(h^*)$  for some  $h^* = h^*(x, y_0, h) \in (0, h)$ . Hence by Lemma 4.1 and Corollary 4.4,  $F(h)/h$  tends to 0 uniformly in  $x \in B$  as  $h \rightarrow 0$ . The lemma now follows from the dominated convergence theorem.  $\square$

We need a weighted Poincaré inequality, which we derive along the lines of the appendix to [SCS91].

**Proposition 4.8** *Let  $R \in [1, 4]$ ,  $y_0 \in \mathbb{R}^d$ ,  $B = B(y_0, R)$ ,*

$$\varphi_R(x) = c_1 (R^2 - |x - y_0|^2)^{12/(2-\beta)} \mathbb{1}_B(x) \quad (4.11)$$

*with normalizing constant  $c_1 > 0$  chosen so that  $\int_B \varphi_R(x) dx = 1$ , and set*

$$\bar{f} = \int_B f(x) \varphi_R(x) dx.$$

*There exists a constant  $c_2$  depending on  $R$  but not  $f$  or  $y_0$  such that*

$$\int_B |f(x) - \bar{f}|^2 \varphi_R(x) dx \leq c_2 \int_B \int_B (f(x) - f(y))^2 \varphi_R(x) \wedge \varphi_R(y) J(x, y) dx dy.$$

**Proof.** If  $B$  is any ball, let

$$\mathcal{E}_B(f, f) = \int_B \int_B (f(y) - f(x))^2 J(x, y) dx dy.$$

Set  $f_B = |B|^{-1} \int_B f(x) dx$ . If  $B$  is any ball of radius  $r \leq 1$ , using (A3) we have

$$\begin{aligned} \int_B |f(x) - f_B|^2 dx &= \int_B (f(x)^2 - (f_B)^2) dx \\ &= \frac{1}{2} |B|^{-1} \int_B \int_B (f(x) - f(y))^2 dx dy \\ &\leq \frac{1}{2} |B|^{-1} (\kappa_1 r^{-d-\alpha})^{-1} \int_B \int_B (f(x) - f(y))^2 J(x, y) dx dy \\ &= c_1 r^\alpha \mathcal{E}_B(f, f). \end{aligned}$$

We now follow the proof in the appendix of [SCS91] closely, with the principal changes being to use  $\mathcal{E}_B(f, f)$  in place of  $\int_B |\nabla f(x)|^2 dx$ ,  $\int_B \int_B (f(x) - f(y))^2 (\varphi(x) \wedge \varphi(y)) J(x, y) dx dy$  in place of  $\int_B |\nabla f(x)|^2 \varphi(x) dx$ , and  $r^\alpha$  in place of  $r^2$ .  $\square$

**Proposition 4.9** *Let  $J$  satisfy the conditions (A1)–(A3). Let  $\xi \in (0, 1)$ ,  $J_\xi$  be defined by (1.5), and  $X^{(\xi)}$  be the Hunt process associated with  $(\mathcal{E}^{(\xi)}, \mathcal{F}^{(\xi)})$ . Let  $y_0 \in \mathbb{R}^d$  and  $\delta \in (0, 1/2)$ . Let  $R \in [1, 4]$ ,  $B = B(y_0, R)$ , and  $p_\xi^B(t, x, y)$  be the transition density of  $X^{(\xi)}$  killed on exiting  $B$ . Then there exists a positive constant  $C$  that depends on  $\alpha, \beta, \kappa_1, \kappa_2$ , and  $\delta$ , but not on  $\xi$  or  $y_0$  such that for all  $t \in [\delta, 2]$*

$$p_\xi^B(t, x, y) \geq C \quad (4.12)$$

*for every  $(x, y) \in B(y_0, 3R/4) \times B(y_0, 3R/4)$ .*

**Proof.** By a change of coordinate systems, without loss of generality we may assume that  $y_0 = 0$  and so  $B = B(0, R)$ . Fix an arbitrary  $x_0 \in B(0, 3R/4)$  and write  $u(t, x) = p_\xi^B(t, x_0, x)$ . Let  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}_+$  be equal to  $\varphi_R$  as defined in (4.11). Note that both  $u(t, \cdot)$  and  $\varphi$  are positive on  $B$ . Set for  $t \in (0, \infty)$

$$\begin{aligned} r(t, x) &= u(t, x)/\varphi(x)^{1/2}, \\ H(t) &= \int_B \varphi(y) \log u(t, y) dy, \\ G(t) &= \int_B \varphi(y) \log r(t, y) dy = \int_B \varphi(y) \log u(t, y) dy - c_1 = H(t) - c_1. \end{aligned}$$

Then by Lemma 4.7

$$G'(t) = -\mathcal{E}^{(\xi)}\left(u(t, \cdot), \frac{\varphi}{u(t, \cdot)}\right). \quad (4.13)$$

The reason we work with  $\mathcal{E}^{(\xi)}$  rather than  $\mathcal{E}$  is so that we can use Lemma 4.7 to obtain (4.13). The remainder of the argument does not use the condition (A4)( $\xi$ ), and in particular the constants can be taken to be independent of  $\xi$ .

By (3.14),

$$\begin{aligned} G'(t) &= - \int_B \int_B \frac{[u(t, y) - u(t, x)]}{u(t, x)u(t, y)} [u(t, x)\varphi(y) - \varphi(x)u(t, y)] J^{(\xi)}(x, y) dy dx \\ &\quad - \int_B \varphi(x) \kappa_B(x) dx. \end{aligned}$$

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

$$G'(t) \geq c_2 \int_B \int_B [\log u(t, y) - \log u(t, x)]^2 (\varphi(x) \wedge \varphi(y)) J^{(\xi)}(x, y) dx dy - c_3 \quad (4.14)$$

for positive constants  $c_2, c_3$ .

Setting  $a = u(t, y)/u(t, x)$  and  $b = \varphi(y)/\varphi(x)$ , we see that

$$\begin{aligned} &\frac{[u(t, y) - u(t, x)]}{u(t, x)u(t, y)} [u(t, x)\varphi(y) - \varphi(x)u(t, y)] \\ &= \varphi(x) \left( b - \frac{b}{a} - a + 1 \right) \\ &= \varphi(x) \left[ \left( (1 - b^{1/2})^2 - b^{1/2} \left( \frac{a}{b^{1/2}} + \frac{b^{1/2}}{a} - 2 \right) \right) \right]. \end{aligned} \quad (4.15)$$

Using the inequality

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

with  $A = a/\sqrt{b}$ , the right hand side of (4.15) is bounded above by

$$(\varphi(x)^{1/2} - \varphi(y)^{1/2})^2 - (\varphi(x) \wedge \varphi(y)) (\log r(t, y) - \log r(t, x))^2.$$

Substituting in the formula for  $G'(t)$  and using Proposition 4.8 and (3.15),

$$\begin{aligned} H'(t) = G'(t) &\geq -c_4 + \int_B \int_B (\log r(t, y) - \log r(t, x))^2 (\varphi(x) \wedge \varphi(y)) J^{(\xi)}(x, y) dx dy \\ &\geq -c_4 + c_5 \int_B (\log r(t, y) - G(t))^2 \varphi(y) dy \\ &\geq -c_6 + c_7 \int_B (\log u(t, y) - H(t))^2 \varphi(y) dy. \end{aligned}$$

In the first inequality we used the fact

$$\int_B \int_B (\varphi(x)^{1/2} - \varphi(y)^{1/2})^2 J^{(\xi)}(x, y) dx dy \leq c_8,$$

which follows from (A3). Recall the constant  $t_0$  from Proposition 3.10. We may assume that  $\delta \leq t_0$ . By Proposition 3.10, for every  $t \leq t_0$ ,

$$\begin{aligned} \int_{B(x_0, 1/4)} u(t, x) dx &\geq \mathbb{P}^{x_0} \left( \sup_{s \in [0, t_0]} |X_s - X_0| < 1/4 \right) \\ &\geq 1 - \mathbb{P}^{x_0} \left( \sup_{s \in [0, t_0]} |X_s - x_0| \geq 1/4 \right) \geq \frac{3}{4}. \end{aligned}$$

Choose  $K$  such that  $|B(x_0, 1/4)|e^{-K} = \frac{1}{4}$  and define

$$D_t = \{x \in B(x_0, 1/4) : u(t, x) \geq e^{-K}\}.$$

By Theorem 1.2, if  $t \leq t_0$

$$\begin{aligned} \frac{3}{4} &\leq \int_{B(x_0, 1/4)} u(t, x) dx = \int_{D_t} u(t, x) dx + \int_{B(x_0, 1/4) \setminus D_t} u(t, x) dx \\ &\leq c_9 |D_t| t^{-d/\alpha} + |B(x_0, 1/4)| e^{-K}. \end{aligned}$$

Therefore

$$|D_t| \geq \frac{t^{d/\alpha}}{2c_9} \geq c_{10} > 0$$

if  $t \in [\delta/4, t_0]$ . Note that the positive constant  $c_{10}$  can be chosen to be independent of the  $\xi$  in condition (A4)( $\xi$ ).

Jensen's inequality tells us that if  $t \leq t_0$

$$H(t) = \int_B (\log u(t, x)) \varphi(x) dx \leq \log \int_B u(t, x) \varphi(x) dx \leq \log \|\varphi\|_\infty := \bar{H}.$$

On  $D_t$ ,  $\log u(t, x) \geq -K$  so there are only four possible cases:

- (a) If  $\log u(t, x) > 0$  and  $H(t) \leq 0$ , then  $(\log u(t, x) - H(t))^2 \geq H(t)^2$ .

(b) If  $\log u(t, x) > 0$  and  $0 < H(t) \leq \bar{H}$ , then

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

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

(d) If  $-K \leq \log u(t, x) \leq 0$  and  $|H(t)| < 2K$ , then

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

Thus we conclude

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

Since  $\varphi$  is bounded below by  $c_{11} > 0$  on  $B(x_0, 1/4)$ , then

$$\begin{aligned} c_7 \int_B (\log u(t, x) - H(t))^2 \varphi(x) dx - c_6 &\geq c_7 \int_{D_t} (\log u(t, x) - H(t))^2 \varphi(x) dx - c_6 \\ &\geq c_{12} |D_t| \left( \frac{1}{4} H(t)^2 - (\bar{H} \vee K)^2 \right) - c_6. \end{aligned}$$

We therefore have

$$H'(t) \geq FH(t)^2 - E, \quad t \in [\delta/4, t_0]$$

for some positive constants  $E$  and  $F$  that are independent of  $\xi$ .

Now we do some calculus. Let  $t_2 \in [\delta/2, t_0 \wedge 2]$ . Let  $Q := \max(16E, (16E/F)^{1/2})$ . Suppose  $H(t_2) \leq -Q$ . Since  $H'(t) \geq -E$  and  $t_2 - t < t_0 \wedge 2 \leq 2$ ,

$$H(t_2) - H(t) \geq -2E \quad \text{for } t \in [\delta/4, t_2], \quad (4.16)$$

This implies  $H(t) \leq -Q/2$  for  $t \in [\delta/4, t_2]$ . Since  $FQ^2/4 \geq 4E$ ,  $E < \frac{F}{2}H(t)^2$  and hence

$$H'(t) \geq \frac{F}{2}H(t)^2.$$

Integrating  $H'/H^2 \geq F/2$  over  $[\frac{\delta}{4}, t_2]$  yields

$$\frac{1}{H(t_2)} - \frac{1}{H(\delta/4)} \leq -\frac{F}{2}(t_2 - \delta/4) \leq -\frac{F\delta}{8}.$$

Since  $H(\delta/4) \leq -Q/2 < 0$ , we have  $1/H(t_2) \leq -F\delta/16$ , that is,

$$H(t_2) \geq -\frac{16}{F\delta}.$$

This proves that either  $H(t_2) \geq -Q$  or  $H(t_2) \geq -16/(F\delta)$ . Thus in either case,

$$H(t_2) \geq -U \quad \text{for every } t_2 \in [\delta/2, t_0 \wedge 2],$$

where  $U = U(\delta) := \max\{Q, 16/(F\delta)\} > 0$ , and so

$$G(t_2) \geq -U + c_1 \quad \text{for every } t_2 \in [\delta/2, t_0 \wedge 2].$$

Now for every  $x_0, x_1 \in B(0, 3R/4)$ , applying the above first with  $x_0$  and then with  $x_0$  replaced by  $x_1$ , we have for  $t \in [\delta, (2t_0) \wedge 4]$ ,

$$\begin{aligned} \log p_\xi^B(t, x_0, x_1) &= \log \int p_\xi^B(t/2, x_0, z) p_\xi^B(t/2, x_1, z) dz \\ &\geq \log \int_B p_\xi^B(t/2, x_0, z) p_\xi^B(t/2, x_1, z) \varphi(z) dz - \log \|\varphi\|_\infty \\ &\geq \int_B \log \left( p_\xi^B(t/2, x_0, z) p_\xi^B(t/2, x_1, z) \right) \varphi(z) dz - \log \|\varphi\|_\infty \\ &= \int_B \log p_\xi^B(t/2, x_0, z) \varphi(z) dz + \int_B \log p_\xi^B(t/2, x_1, z) \varphi(z) dz \\ &\quad - \log \|\varphi\|_\infty \\ &\geq -2(U + c_{12}), \end{aligned}$$

that is,  $p_\xi^B(t, x_0, x_1) \geq e^{-2(U+c_{12})}$  for every  $t \in [\delta, (2t_0) \wedge 4]$ . A repeated use of the semigroup property (but at most  $1/t_0$  more times) then shows  $p_\xi^B(t, x_0, x_1) \geq c_{13}(\delta)$  for every  $t \in [\delta, 2]$ .  $\square$

**Theorem 4.10** *Let  $J$  satisfy the conditions (A1)–(A3). Let  $\xi \in (0, 1)$ ,  $J_\xi$  be defined by (1.5), and  $X^{(\xi)}$  be the Hunt process associated with  $(\mathcal{E}^{(\xi)}, \mathcal{F}^{(\xi)})$ . Let  $y_0 \in \mathbb{R}^d$  and  $\delta \in (0, 1/2)$ . Let  $R > 0, T > 1/2$ ,  $B = B(y_0, R)$ , and  $p_\xi^B(t, x, y)$  be the transition density of  $X^{(\xi)}$  killed on exiting  $B$ . Then there exists a positive constant  $C = C(R)$  that depends on  $\alpha, \beta, \kappa_1, \kappa_2, R$  and  $\delta$ , but not on  $\xi$  or  $y_0$  such that for all  $t \in [\delta, T]$*

$$p_\xi^B(t, x, y) \geq C \tag{4.17}$$

for every  $(x, y) \in (B(y_0, 3R/4) \times (B(y_0, 3R/4))$ .

**Proof.** By a change of coordinate system, we assume without loss of generality that  $y_0 = 0$ . Suppose first that  $T \geq 2$  and  $R \in [1, 4]$ . Let  $r = 3R/4, n = [T]$ , and  $\delta = 1/n$ . So if  $t \in [1, T]$ , then  $t/n \in [\delta, 2]$ . By the semigroup property, if  $x, y \in B(0, 3R/4) \setminus \mathcal{N}$ , then

$$\begin{aligned} p_\xi^B(t, x, y) &\geq \int_{B(0, r)} \cdots \int_{B(0, r)} p_\xi^B(t/n, x, z_1) p_\xi^B(t/n, z_1, z_2) \cdots \\ &\quad p_\xi^B(t/n, z_{n-2}, z_{n-1}) p_\xi^B(t/n, z_{n-1}, y) dz_1 \cdots dz_{n-1} \\ &\geq c_1 |B(0, r)|^{n-1} \geq c_2 \end{aligned} \tag{4.18}$$

by Proposition 4.9. We therefore have the conclusion of the theorem for all  $T \geq 2$ .

Next suppose  $R > 4$ . Suppose  $x, y \in B(0, 3R/4) \setminus \mathcal{N}$ ,  $t \in [\delta, T]$ ,  $n = 2[|x - y|] + 1$ , and let  $z_0 = x, z_1, \dots, z_{n-1}, z_n = y$  be equally spaced points on the line segment joining  $x$  and



$y$ . Here  $\llbracket x - y \rrbracket$  is the integer part of  $|x - y|$ . Then  $|z_{i+1} - z_i| \leq 1/2 \leq R/8$ . Set  $r = 1$ . Using (4.18) and Proposition 4.9 with  $\delta$  replaced by  $\delta/n$ , we again obtain our conclusion.

Finally, suppose  $R < 1$ . Fix  $\delta$  and  $T$ . Consider the process  $Z = R^{-1}X^{(\xi)}$  with corresponding jump kernel  $J_Z$ . By a change of variables, we see that it suffices to obtain a lower bound on  $p_Z^B(t, x, y)$  for  $x, y \in B(0, 3/4)$ , and  $p_Z^B$  is the transition density for  $Z$  killed on exiting  $B(0, 1)$ . The jump kernel corresponding to  $Z$  is  $J_Z(x, y) = R^{-d}J(Rx, Ry)$ . Let  $J_{Z^{(0)}}(x, y) = J_Z(x, y)\mathbb{1}_{(|x-y|<1)}$ . It is easy to see that  $J_{Z^{(0)}}$  satisfies (A1)–(A3) and (A4)( $\xi$ ) (but with different constants  $\kappa_1, \kappa_2, \xi$ ). Let  $Z^{(0)}$  be the process corresponding to  $J_{Z^{(0)}}$  and construct  $Z$  from  $Z^{(0)}$  using Remark 3.4. Then if  $A \subset B(0, 3/4)$ ,  $\tau_B = \inf\{t : Z_t \notin B\}$ , and  $\tau_B^0 = \inf\{t : Z_t^{(0)} \notin B\}$ , by independence, and using Lemma 3.6,

$$\begin{aligned} \int_A p_Z^B(t, x, y) dy &= \mathbb{P}^x(Z_t \in A, \tau_B > t) \\ &\geq e^{-(\sup \mathcal{J})T} \int_A p_{Z^{(0)}}^B(t, x, y) dy \geq c_3|A|, \end{aligned}$$

which proves the theorem in this case as well.  $\square$

**Proof of Theorem 1.3:** Let  $\delta_k$  be a sequence of positive numbers decreasing to 0. Define  $J_k(x, y)$  by (2.4) and define  $(\mathcal{E}^k, \mathcal{F}^k)$  as in the proof of Theorem 1.5. Clearly  $J_k$  satisfies the conditions (A1)–(A3) as well as (A4)( $\delta_k$ ) with the same  $\kappa_1$  and  $\kappa_2$  as  $J$ .

Let  $p^{k,B}(t, x, y)$  and  $p^B(t, x, y)$  denote the transition density functions of  $X^{k,B}$  and  $X^B$  respectively. It follows from Theorem 4.10 and Proposition 2.2 that for any given  $\delta \in (0, 1)$ , there is a constant  $c = c(\delta) > 0$  such that for any bounded non-negative functions  $f$  and  $g$  on  $B$  and  $t \in [\delta, 2]$ ,

$$\begin{aligned} \int_{B(y_0, 3R/4) \times B(y_0, 3R/4)} p^B(t, x, y) f(x) g(y) dx dy \\ = \lim_{k \rightarrow \infty} \int_{B(y_0, 3R/4) \times B(y_0, 3R/4)} p^{k,B}(t, x, y) f(x) g(y) dx dy \\ \geq c \int_{B(y_0, 3R/4) \times B(y_0, 3R/4)} f(x) g(y) dx dy. \end{aligned}$$

This implies that  $p(t, x, y) \geq c$  for almost every  $x, y \in B(y_0, 3R/4)$ . On the other hand, it follows from the proof of Theorem 1.2 that there is a properly exceptional set  $\mathcal{N}$  so that  $p(t, x, y)$  is well-defined on  $(B \setminus \mathcal{N}) \times (B \setminus \mathcal{N})$  and that for each fixed  $y \in B \setminus \mathcal{N}$ ,  $x \mapsto p^B(t, x, y)$  is  $X^B$ -quasi-continuous (and hence  $X$ -quasi-continuous). It follows that  $p(t, x, y) \geq c$  for every  $t \in [\delta, 2]$  and every  $x, y \in B(y_0, 3R/4) \setminus \mathcal{N}$ .  $\square$

## 5 Parabolic Harnack inequality

In this section we prove Theorem 1.7 in the case  $R \geq 1$ . The argument uses balayage; see [BG68], Chapter VI, for details.

**Proof of Theorem 1.7:** Without loss of generality we may assume the following: by a change of coordinate system, we may assume  $x_0 = 0$ ; by the Markov property we may assume  $t_0 = 0$ ; by looking at the process  $X'_t = X_{t/T}$ , we see that the jump kernel corresponding to  $X'$  satisfies (A1)–(A3) (but with different  $\kappa_1, \kappa_2$ ), so we may assume  $T = 1$ . With these assumptions  $Q = (0, 5) \times B(0, 4R)$ . Recall the notation for hitting and exit times given in (1.6). Let  $E = (\frac{1}{2}, \frac{9}{2}) \times B(0, 3R)$ ,  $D = (\frac{3}{4}, \frac{17}{4}) \times B(0, 2R)$ , and write  $A = \overline{B(0, 3R)} \setminus B(0, 2R)$  and  $B = B(0, 4R)$ . By the martingale property,

$$P_{t-s}^B u(s, x) \leq u(t, x), \quad \text{for } s < t \text{ with } (s, x), (t, x) \in Q. \quad (5.1)$$

This says that the function  $u$  is excessive with respect to the space-time subprocess  $(V^Q, X^Q)$  of  $(V, X)$  killed upon exiting  $Q$ , where  $V_t = V_0 - t$ . We can define  $u_E$ , the réduite of  $u$  with respect to  $E$ , by

$$u_E(s, x) = \mathbb{E}^{(s, x)}[u(V_{T_E}, X_{T_E}); T_E < \tau_Q].$$

The function  $u_E$  is again excessive with respect to the killed process  $(V^Q, X^Q)$ , is 0 on  $Q^c$ , and agrees with  $u$  on  $E$ ; see [BG68]. Note that the process  $(t, X_t)$  is in duality with the process  $(V_t, X_t)$  in the sense of [BG68, Chapter VI]. By the Riesz decomposition theorem (cf. [BG68, Theorem VI.2.11]),  $u_E$  is the potential of a measure  $\nu_E$  supported on  $\overline{E}$ . This means that if  $(t, x) \in Q$  then

$$u_E(t, x) = \int_E p^B(t - r, x, z) \nu_E(dr, dz). \quad (5.2)$$

Here we have  $p^B(s, x, y) = 0$  if  $s < 0$ .

Since the jumps of the process  $X$  are bounded by 1,  $u_E$  is caloric on  $(1/2, 9/2) \times B(0, 2R)$ . It follows that the support of  $\nu_E$  is contained in  $\overline{E} \setminus ((1/2, 9/2) \times B(0, 2R))$ . For  $t \in (1/2, 5)$ , let

$$\begin{aligned} F_1(t) &= [1/2, t] \times (A \setminus \mathcal{N}), \\ F_2(t) &= \{\frac{1}{2}\} \times (\overline{B(0, 2R)} \setminus \mathcal{N}), \\ F(t) &= F_1(t) \cup F_2(t). \end{aligned}$$

Thus if  $(t, x) \in D$  then we can write (5.2) as

$$u(t, x) = u_E(t, x) = \int_{F(t)} p^B(t - r, x, z) \nu_E(dr, dz). \quad (5.3)$$

Since  $\nu_E$  is an equilibrium measure (i.e., a capacitary measure), it does not charge polar sets; in particular, it does not charge  $[0, 5] \times \mathcal{N}$ .

Consider (5.3) when  $(t_1, x_1) \in Q_- = (1, 2) \times (B(0, R) \setminus \mathcal{N})$ . If  $(r, z) \in F_1(t)$  then  $|x_1 - z| \geq R$ , and thus by Proposition 3.9(a),

$$p^B(t_1 - r, x_1, z) \leq p(t_1 - r, x_1, z) \leq c_1. \quad (5.4)$$

If  $(r, z) \in F_2(t)$ , then  $t_1 - r \geq \frac{1}{2}$  and by Theorem 1.2 again (5.4) holds.

Now let  $(t_2, x_2) \in Q_+ = (3, 4) \times (B(0, R) \setminus \mathcal{N})$ . If  $(r, z) \in F(t_1)$  then  $t_2 - r \geq t_2 - t_1 \geq 1$ , and  $|x_2 - z| \leq 4$ , so by Theorem 1.3

$$p^B(t_2 - r, x_2, z) \geq c_2. \quad (5.5)$$

Hence

$$\begin{aligned} u(t_2, x_2) &= \int_{F(t_2)} p^B(t_2 - r, x_2, z) \nu_E(dr, dz) \\ &\geq \int_{F(t_1)} p^B(t_2 - r, x_2, z) \nu_E(dr, dz) \\ &\geq c_2 \nu_E(F(t_1)) \\ &\geq (c_2/c_1) \int_{F(t_1)} p^B(t_1 - r, x, z) \nu_E(dr, dz) \\ &= (c_2/c_1) u(t_1, x_1), \end{aligned}$$

giving the parabolic Harnack inequality with constant  $C = c_1/c_2$ .  $\square$

## 6 Harmonic functions need not be continuous

One of the applications of scale invariant Harnack inequalities is that they imply regularity, e.g., Hölder continuity of harmonic and caloric functions, and resolvents. This can be used in order to remove properly exceptional sets. It is therefore interesting to see that such regularity can fail, even when a Harnack inequality holds. We say a function  $h$  is harmonic in a domain  $D \subset \mathbb{R}^d$  if  $h(X_{t \wedge \tau_{D_1}})$  is a right continuous martingale for every subdomain  $D_1$  with  $\bar{D}_1 \subset D$ , where  $\tau_{D_1} = \inf\{t > 0 : X_t \notin D_1\}$ .

In this section we construct a class of symmetric jump processes satisfying our hypotheses where there exist bounded harmonic functions that are not continuous. An interesting side result related to this example is that the martingale problem for variable order jump processes is not always well posed. See Remark 6.8(d) and the results in [HKa05] for other examples which are similar but lead to continuous harmonic functions and Feller semigroups.

Using the integral conditions given in Theorems 11.2 and 11.5 of [Fr74] we obtain:

**Lemma 6.1** *Let  $X_t$  be a one-dimensional stable process of index  $\alpha \in (0, 1)$  and  $\varepsilon > 0$ . Then*

$$\liminf_{t \rightarrow 0} \frac{|X_t|}{t^{(1/\alpha)+\varepsilon}} = \infty \quad \text{and} \quad \limsup_{t \rightarrow 0} \frac{|X_t|}{t^{(1/\alpha)-\varepsilon}} = 0, \quad \mathbb{P}^0\text{-a.s.}$$

Before constructing the main counterexample on  $\mathbb{R}^2$ , we need to look at an auxiliary process  $Y$ . Let  $0 < a < b < 2$  and set for  $z \in \mathbb{R}^2$ ,  $z_1 \neq z_2$  :

$$m(z_1, z_2) = \begin{cases} \min(|z_1|^{-a-2}, |z_2|^{-b-2}) & \text{if } |z_1| \vee |z_2| \leq 1, \\ 0 & \text{if } |z_1| \vee |z_2| > 1. \end{cases} \quad (6.1)$$

Assume  $|z_1| \vee |z_2| \leq 1$ . Note that

$$\frac{1}{|z_1|^{a+2} + |z_2|^{b+2}} \leq \min\left(\frac{1}{|z_1|^{a+2}}, \frac{1}{|z_2|^{b+2}}\right) \leq \frac{2}{|z_1|^{a+2} + |z_2|^{b+2}}.$$

This implies

$$c_0|z|^{-a-2} \leq \min(|z_1|^{-a-2}, |z_2|^{-b-2}) \leq c_1|z|^{-b-2}, \quad (6.2)$$

where  $c_0, c_1$  are independent of  $z$ . Now for  $x, y \in \mathbb{R}^2$ ,  $x \neq y$  let

$$J_0(x, y) = \begin{cases} m(|x_1 - y_1|, |x_2 - y_2|) & \text{if } x - y \in [-1, 1]^2, \\ 0 & \text{if } x - y \notin [-1, 1]^2. \end{cases}$$

Furthermore, set

$$\alpha = \frac{(a+1)(b+1) - 1}{b+2}, \quad \beta = \frac{(a+1)(b+1) - 1}{a+2}. \quad (6.3)$$

The following facts are now obvious.

**Lemma 6.2** (a)  $J_0(x, y)$  is symmetric in  $(x, y)$  and

$$c_1|x - y|^{-a-2} \leq J_0(x, y) \leq c_2|x - y|^{-b-2} \quad \text{for } |x_1 - y_1| \vee |x_2 - y_2| \leq 1. \quad (6.4)$$

(b)  $\alpha < \beta$ .

(c)  $a < (2 - b)/b$  if and only if  $\beta < 1$ .

Now choose  $a, b$  with  $0 < a < b < 2$  and  $a < (2 - b)/b$ , so that  $0 < \alpha < \beta < 1$ . Let  $Y_t = (Y_t^1, Y_t^2) \in \mathbb{R}^2$  be the pure jump symmetric Lévy process with jump kernel  $J_0$ . The following lemma explains the behavior of the marginals  $Y_t^1, Y_t^2$ .

**Lemma 6.3** We have

$$n_1(z_1) := \int_{-1}^1 m(z_1, z_2) dz_2 = \frac{2}{b+1} + \frac{2b}{b+1}|z_1|^{-\alpha-1} \quad \text{for } z_1 \in [-1, 1],$$

$$n_2(z_2) := \int_{-1}^1 m(z_1, z_2) dz_1 = \frac{2}{a+1} + \frac{2a}{a+1}|z_2|^{-\beta-1} \quad \text{for } z_2 \in [-1, 1].$$

**Proof.**

$$\begin{aligned} \frac{1}{2} \int_{-1}^1 m(z_1, z_2) dz_2 &= \int_0^{|z_1|^{\frac{a+2}{b+2}}} |z_1|^{-a-2} dz_2 + \int_{|z_1|^{\frac{a+2}{b+2}}}^1 |z_2|^{-b-2} dz_2 \\ &= |z_1|^{-\frac{(a+2)(b+1)}{b+2}} + \frac{1}{b+1} - \frac{1}{b+1} |z_1|^{-\frac{(a+2)(b+1)}{b+2}}. \end{aligned}$$

Note that  $\frac{(a+2)(b+1)}{b+2} = \alpha + 1$ . The first assertion of the lemma follows. The second one is proved analogously.  $\square$

The coordinate processes  $Y^i$  are one-dimensional Lévy processes with jump measure  $n_i$ ; note however that  $Y^1$  and  $Y^2$  are not independent.

Let

$$V(\lambda) = \{(x_1, x_2) : |x_1| < \lambda|x_2|\}, \quad \tau_{V(\lambda)} = \inf\{t > 0 : Y_t \notin V(\lambda)\}.$$

and write  $V = V(1)$ .

**Lemma 6.4** *Let  $\lambda > 0$ ; then  $\mathbb{P}^0(\tau_{V(\lambda)} > 0) = 1$ .*

**Proof.** By Lemma 6.3,  $n_1$  differs from the jump measure of a stable process of index  $\alpha$  by a finite measure. Therefore  $Y^1$  has the same local behavior at time 0 as a stable process of index  $\alpha$ ; similarly  $Y^2$  has the same local behavior at time 0 as a stable process of index  $\beta$ . (Note that points are polar for these two processes.) Choose  $\varepsilon$  such that  $\varepsilon + \frac{1}{\beta} < \frac{1}{\alpha} - \varepsilon$ . By Lemma 6.1 for all sufficiently small  $t > 0$ ,

$$0 < |Y_t^1| \leq t^{(1/\alpha)-\varepsilon} \leq \frac{1}{2}\lambda t^{(1/\beta)+\varepsilon} \leq \frac{1}{2}\lambda |Y_t^2|.$$

$\square$

**Remark 6.5** By Lemma 6.4 we can find  $t_0 > 0$  such that  $\mathbb{P}^0(\tau_{V(1/3)} \leq t_0) < 1/20$ . Let  $D = D(r) = (-r, r)^2$ , and choose  $r$  small enough so that  $\mathbb{P}^0(Y_{t_0} \in D) < 1/20$ . Let

$$\tau_D = \inf\{t > 0 : Y_t \notin D\} \quad \text{and} \quad F(\lambda) = \{Y_{\tau_D} \in V(\lambda)\}.$$

Note that if  $Y_{\tau_D} \notin V(\lambda)$  and  $\tau_{V(\lambda)} > t_0$  then we have  $\tau_D \geq \tau_{V(\lambda)} > t_0$ . So,

$$\mathbb{P}^0(F(\frac{1}{3})^c) \leq \mathbb{P}^0(\tau_{V(\frac{1}{3})} \leq t_0) + \mathbb{P}^0(Y_{t_0} \in D) \leq 1/10.$$

Note also that the events  $F(\lambda)$  are increasing in  $\lambda$ .

**Lemma 6.6** *There exist sequences  $x_n \rightarrow 0$  and  $\delta_n \rightarrow 0$  such that  $\bigcup_{n=1}^{\infty} B(x_n, \delta_n) \subset V(\frac{1}{2})$  and*

$$\mathbb{P}^y(F(\frac{1}{2})) \geq 7/10 \quad \text{for } y \in \bigcup_{n=1}^{\infty} B(x_n, \delta_n).$$

**Proof.** Let  $h_Y(x) = \mathbb{P}^x(F(\frac{1}{2}))$ . Then  $h_Y(0) > 9/10$  and  $h_Y(Y_{t \wedge \tau_D})$  is a right-continuous martingale. Using the right-continuity of  $h_Y(Y)$  and  $Y$ , and the fact that  $Y_t \in V(1/4)$  for all small times  $t$ , we deduce that there exist  $x_n \in V(1/4)$  with  $x_n \rightarrow 0$  such that  $h_Y(x_n) \geq 8/10$  for all  $n$ .

Note that, since the coordinate processes  $Y^i$  are symmetric stable processes, each point  $x \in \partial D$  is regular for  $D^c$ . Therefore, if the event  $F(\frac{1}{3})$  occurs for the process  $Y(\omega)$  (with

$Y_0 = x_n$ ) then  $F(\frac{1}{2})$  occurs, a.s., for the process  $u + Y$  for all sufficiently small  $u$ . It follows that, for each  $n$ ,

$$\liminf_{y \rightarrow x_n} \mathbb{P}^y(F(\frac{1}{2})) \geq 8/10.$$

We now take  $\delta_n > 0$  small enough so that  $B(x_n, \delta_n) \in V(\frac{1}{2})$  and  $\mathbb{P}^y(F(\frac{1}{2})) \geq 7/10$  on  $B(x_n, \delta_n)$ .  $\square$

We now define a second jump kernel  $J_1$  follows. Write  $B = B(0, 1)$ . If  $x, y \in V$  we set

$$J_1(x, y) = m(|x_1 - y_1|, |x_2 - y_2|) \mathbb{1}_{(x-y \in B)}.$$

If  $x, y \in V^c$  we set

$$J_1(x, y) = m(|x_2 - y_2|, |x_1 - y_1|) \mathbb{1}_{(x-y \in B)}.$$

If  $x \in V$  and  $y \in V^c$  or vice versa and  $|x - y| \leq 1$ , we define

$$J_1(x, y) = (|x_1 - y_1|^{-2-a} \wedge |x_2 - y_2|^{-2-a}) \mathbb{1}_{(x-y \in B)}.$$

It is easy to see that  $J_1$  satisfies (A1)–(A3).

**Proof of Theorem 1.9:** Let  $X = \{X_t, t \geq 0\}$  be the symmetric jump process associated with the Dirichlet form given by (1.1)–(1.2) with  $J_1$  as above in place of  $J$ . Note that if  $x \in V$ , then  $J_1(x, y) = J_0(x, y)$  for  $y \in V$ , while  $J_1(x, y) \leq J_0(x, y)$  for  $y \in V^c$ . Thus  $X$  makes as many jumps within  $V$  as  $Y$  does, but makes fewer jumps from  $V$  to  $V^c$ . This can be made more precise by using Remark 3.4 to construct  $Y$  from  $X$ . Although when  $X$  enters  $V^c$ , the positive continuous additive functional  $C = \{C_t, t \geq 0\}$  defined in Remark 3.4 becomes infinite immediately, we will only be looking at time intervals  $[0, \tau_V(X))$ , so this will not be an issue for us. In particular, all that we need is that  $X_s = Y_s$  for  $0 \leq s < \tau_V(Y)$ .

Let  $\sigma = \inf\{t > 0 : X_t \notin D\}$ , and set

$$h(x) = \mathbb{P}^x(X_\sigma \in V(\frac{1}{2})).$$

Then as  $X = Y$  on  $[0, \tau_V(Y))$ , by a similar argument as that for Lemma 6.6, there exists  $E := \cup_{n \geq 1} B(x_n, \delta_n) \subset V(\frac{1}{2})$ , where  $x_n \rightarrow 0$  and  $\delta_n \rightarrow 0$ , such that

$$h(x) = \mathbb{P}^x(X_\sigma \in V(\frac{1}{2})) \geq 7/10 \quad \text{for } x \in E.$$

Let  $\Theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be defined by

$$\Theta((x_1, x_2)) = (x_2, x_1). \tag{6.5}$$

Then the reflection symmetry of the law of  $X$  gives

$$\mathbb{P}^x(X_\sigma \in \Theta(V(\frac{1}{2}))) \geq 7/10, \quad x \in \Theta(E).$$

Hence we have

$$h(x) \leq 3/10, \quad x \in \Theta(E).$$

Since  $E$  and  $\Theta(E)$  are open sets and 0 is an accumulation point of both sets, we deduce that  $h$  is not continuous at 0.  $\square$

**Corollary 6.7** *There exist a bounded continuous function  $H$  and  $t_0 > 0$  such that  $x \mapsto \mathbb{E}^x [H(X_{t_0})]$  is not continuous at 0.*

**Proof.** Recall that  $Y$  is the Lévy process with jump kernel  $J_0$ . By Remark 6.5 there exist  $t_0 > 0$  and  $r$  such that

$$\mathbb{P}^0 (Y_{t_0} \in V(1/3) \setminus D(r)) \geq \frac{9}{10}.$$

Let  $H$  be a continuous function bounded by  $-1$  and  $1$  such that  $H$  is 1 on  $V(\frac{1}{3}) \setminus D(r)$ ,  $H \geq 0$  on  $V(1)$ , and  $H(x_2, x_1) = -H(x_1, x_2)$ . Since  $H$  is bounded below by  $-1$ , then  $\mathbb{E}^0 [H(Y_{t_0})] \geq \frac{9}{10} - \frac{1}{10} = \frac{8}{10}$ . Since  $Y$  is a Lévy process, it has the Feller property, and so  $x \mapsto \mathbb{E}^x [H(Y_{t_0})]$  is continuous. Therefore  $\lim_{x \rightarrow 0} \mathbb{E}^x [H(Y_{t_0})] \geq \frac{8}{10}$ .

As in the proof of Theorem 1.9, there exist sequences  $x_n \rightarrow 0$ ,  $\delta_n \rightarrow 0$  such that  $E := \cup_n B(x_n, \delta_n) \subset V(1)$  and  $\mathbb{E}^x [H(X_{t_0})] \geq \frac{7}{10}$  for  $x \in E$ . By the antisymmetry of  $H$ ,  $\mathbb{E}^x [H(X_{t_0})] \leq -\frac{7}{10}$  for  $x \in \Theta(E)$ , where  $\Theta$  is defined by (6.5). We conclude  $x \mapsto \mathbb{E}^x [H(X_{t_0})]$  is not continuous at 0.  $\square$

**Remark 6.8** (i) If we wish, we can replace the double cone  $V(\lambda)$  by the half-space  $H_- = \{x_1 \leq 0\}$ . Use polar coordinates in  $\mathbb{R}^2$  and let  $\theta_0 \in (0, \pi/2)$  satisfy  $\tan \theta_0 = 1/\lambda$ . Let  $\psi : [0, 2\pi] \rightarrow [0, 2\pi]$  be strictly increasing and piecewise linear with  $\psi(0) = 0$ ,  $\psi(\theta_0) = \pi/2$ ,  $\psi(2\pi - \theta_0) = 3\pi/2$  and  $\psi(2\pi) = 2\pi$ . Let  $\Psi(r, \theta) = (r, \psi(\theta))$ ; note that  $\Psi(V(\lambda)) \subset H_-$ .

Let  $Z = \Psi(X)$ . Then  $Z$  is a symmetric jump process associated with a Dirichlet form  $\mathcal{E}'$  and jump measure  $n'$  which satisfies (A1)–(A3). By the construction of  $Z$  and using Remark 6.5 we see that there exist points  $x_n \rightarrow 0$  with  $x_n \in H_-$ , and  $t_0 > 0$  such that, for all  $n$ ,

$$\mathbb{P}^{x_n} (\tau_{H_-}^Z \geq t_0) \geq 9/10. \quad (6.6)$$

If we time change  $Z$  so that the associated Dirichlet form is on  $L^2(\mathbb{R}^2, dx)$ , the new process  $\tilde{Z}$  still satisfies (6.6) with a different value of  $t_0$ .

In particular we see that for the processes considered in this paper, if  $H$  is a half space, points on  $\partial H$  need not be regular for  $H$ .

(ii) In this example we needed  $\alpha < \beta < 1$  because points are not polar for the symmetric stable process with  $\alpha > 1$ . If we look at similar constructions in higher dimensions such as  $\mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2$ , then it seems likely that we could construct a similar example for any  $0 < \alpha < \beta < 2$ .

(iii) Note that (6.1) allows for the choice of  $b = a + \varepsilon$  for any value of  $\varepsilon > 0$ . Define

$$\tilde{m}(z_1, z_2) = \begin{cases} \min \left\{ |z_1|^{-a-2}, \log\left(\frac{3}{|z_2|}\right) |z_2|^{-a-2} \right\} & \text{if } |z_1| \vee |z_2| \leq 1, \\ 0 & \text{if } |z_1| \vee |z_2| > 1. \end{cases} \quad (6.7)$$

Then  $\tilde{m}$  is very similar to  $m$  in (6.1). Now, proceed as in the example above but this time construct  $J_1$  with the help of  $\tilde{m}$  instead of  $m$ . Then it is shown as a corollary in [HKa05] that the Dirichlet form corresponds to a Feller semigroup and that harmonic functions are continuous satisfying certain *a priori* estimates. Therefore when trying to construct discontinuous harmonic functions, one cannot modify our class of examples much.

**Proposition 6.9** *With  $J$  as above, the martingale problem for the operator*

$$\mathcal{L}f(x) = \int_{\mathbb{R}^2} (f(x+h) - f(x))J_1(x, x+h) dh$$

*acting on  $C_c^2$  functions is not well-posed.*

**Proof.** The function  $\mathcal{L}f$  is bounded by the  $C^2$ -norm of  $f$  when  $f \in C_c^2(\mathbb{R}^2)$  by (A1)–(A3). Section 3 of [Bas88] (trivially modified to handle the case of dimensions larger than one) shows that  $\{\mathbb{P}^x : x \in \mathbb{R}^2\}$  is tight.

We claim that any subsequential weak limit point of  $\mathbb{P}^x$  as  $x \rightarrow 0$  is a solution to the martingale problem for  $\mathcal{L}$  started at 0. If  $\mathbb{P}$  is such a limit point, it is easy to see that  $\mathbb{P}(X_0 = 0) = 1$ . If  $f \in C_c^2$  and  $r_1 \leq \dots \leq r_n \leq s \leq t$  and  $g_1, \dots, g_n$  are continuous functions with compact support, then

$$\begin{aligned} \mathbb{E}^{x_n} \left[ \left\{ f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_u) du \right\} Y \right] \\ = \mathbb{E}^{x_n} \left[ \left\{ f(X_s) - f(X_0) - \int_0^s \mathcal{L}f(X_u) du \right\} Y \right], \end{aligned} \quad (6.8)$$

where  $Y = \prod_{i=1}^n g_i(X_{r_i})$ . Since  $X$  has no jumps larger than 1, we see that if  $f \in C_c^2(\mathbb{R}^2)$ , then  $\mathcal{L}f$  will be zero at points that are a distance more than one from the support of  $f$ . Therefore  $\mathcal{L}f$  also has compact support. Let  $\varepsilon > 0$  and let  $I$  be a continuous function that equals  $\mathcal{L}f$  except on a set  $A$  of Lebesgue measure at most  $\varepsilon$ . Since  $f(X_t)Y$ ,  $f(Y_s)Y$ ,  $f(X_0)Y$ ,  $Y \int_0^t I(X_u) du$ , and  $Y \int_0^s I(X_u) du$  are continuous functionals of the path, their expectation under  $\mathbb{P}^{x_n}$  converges to the corresponding expectation under  $\mathbb{P}$ . We have the estimate

$$\mathbb{E}^{X_n} \left[ \int_0^t |\mathcal{L}f(X_u) - I(X_u)| du \right] \leq c_1 t_1 + c_1 \mathbb{E}^{x_n} \left[ \int_{t_1}^t \mathbb{1}_A(X_u) du \right] \leq c_1 t_1 + c_2 t_1^{-d/\alpha} |A|$$

using Theorem 1.2. If we set  $t_1 = |A|^{\alpha/(\alpha+d)}$ , we obtain the upper bound  $c_3 |A|^{\alpha/(\alpha+d)}$ . A limit argument yields the same bound for the expectation with respect to  $\mathbb{P}$ . Therefore

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left| \mathbb{E} \left[ \left\{ f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_u) du \right\} Y \right] \right. \\ \left. - \mathbb{E} \left[ \left\{ f(X_s) - f(X_0) - \int_0^s \mathcal{L}f(X_u) du \right\} Y \right] \right| \\ \leq c_4 \varepsilon^{\alpha/(\alpha+d)}. \end{aligned}$$



Since  $\varepsilon$  is arbitrary, we have (6.8) with  $\mathbb{E}^{x_n}$  replaced by  $\mathbb{E}$ . This shows that  $\mathbb{P}$  is a solution to the martingale problem started at 0.

Suppose now that we had uniqueness for the martingale problem for  $\mathcal{L}$  started at 0. We conclude  $\mathbb{P}^x \rightarrow \mathbb{P}^0$  as  $x \rightarrow 0$ . In particular, if  $H$  is a bounded continuous function,  $\mathbb{E}^x [H(X_t)] \rightarrow \mathbb{E}^0 [H(X_t)]$  as  $x \rightarrow 0$  for all  $t > 0$ . But this contradicts Corollary 6.7.  $\square$

## References

- [BB92] M. T. Barlow and R. F. Bass. Transition densities for Brownian motion on the Sierpiński carpet. *Probab. Theory Rel. Fields*, **91** (1992), 307–330.
- [BGK] M.T. Barlow, A. Grigor’yan and T. Kumagai. Heat kernel upper bounds for jump processes and the first exit time. preprint 2006.
- [Bas88] R. F. Bass. Uniqueness in law for pure jump processes. *Probab. Theory Rel. Fields*, **79** (1988), 271–287.
- [Bas95] R. F. Bass. *Probabilistic Techniques in Analysis*. Probability and its Applications (New York). Springer-Verlag, New York, 1995.
- [Bas97] R. F. Bass. *Diffusions and Elliptic Operators*. Probability and its Applications (New York). Springer-Verlag, New York, 1997.
- [Bas02] R. F. Bass. On Aronson’s upper bounds for heat kernels. *Bull. London Math. Soc.*, **34** (2002), 415–419.
- [BK05a] R. Bass and M. Kassmann. Harnack inequalities for non-local operators of variable order. *Trans. A.M.S*, **357** (2005), 837–850.
- [BK05b] R. Bass and M. Kassmann. Hölder continuity of harmonic functions with respect to operators of variable order. *Comm. Part. Diff. Eq.*, **30** (2005), 1249–1259.
- [BL02a] R. F. Bass and D. A. Levin. Harnack inequalities for jump processes. *Potential Anal.*, **17** (2002), 375–388.
- [BL02b] R. F. Bass and D. A. Levin. Transition probabilities for symmetric jump processes. *Trans. Amer. Math. Soc.*, **354** (2002), 2933–2953.
- [BG68] R. M. Blumenthal and R. K. Gettoor. *Markov Processes and Potential Theory*. Academic Press, New York, 1968.
- [CKS87] E. A. Carlen, S. Kusuoka, and D. W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. H. Poincaré Probab. Statist.*, **23(2, suppl.)** (1987), 245–287.
- [Ch92] Z.-Q. Chen. On reflected Dirichlet spaces. *Probab. Theor. Rel. Fields*, **94**, 135–162, 1999.

- [CK03] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for stable-like processes on  $d$ -sets. *Stoch. Proc. Applic.*, **108** (2003), 27–62.
- [CK06] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for jump processes of mixed type on metric measure spaces. To appear in *Probab. Theor. Rel. Fields*.
- [CS98] Z.-Q. Chen and R. Song. Estimates on Green functions and Poisson kernels of symmetric stable processes. *Math. Ann.*, **312** (1998), 465–501.
- [CW86] S. Chanillo and R. L. Wheeden. Harnack’s inequality and mean-value inequalities for solutions of degenerate elliptic equations. *Comm. Part. Diff. Eq.*, **11** (1986), 1111–1134.
- [Dav87] E. B. Davies. Explicit constants for Gaussian upper bounds on heat kernels. *Amer. J. Math.*, **109** (1987), 319–333.
- [DG57] E. DeGiorgi. Sulla differenziabilità e l’analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)*, **3** (1957), 25–43.
- [Fab93] E. B. Fabes. Gaussian upper bounds on fundamental solutions of parabolic equations; the method of Nash. In *Dirichlet forms (Varenna, 1992)*, volume 1563 of *Lecture Notes in Math.*, pages 1–20. Springer, Berlin, 1993.
- [FS86] E. B. Fabes and D. W. Stroock. A new proof of Moser’s parabolic Harnack inequality using the old ideas of Nash. *Arch. Rat. Mech. Anal.*, **96** (1986), 327–338.
- [Fr74] B. Fristedt. Sample functions of stochastic processes with stationary, independent increments. In: *Advances in probability and related topics, Vol. 3*, pp. 241–396. Dekker, New York, 1974.
- [FOT94] M. Fukushima, Y. Ōshima, and M. Takeda. *Dirichlet Forms and Symmetric Markov Processes*, volume 19 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1994.
- [GW90] C. E. Gutiérrez and R. L. Wheeden. Mean value and Harnack inequalities for degenerate parabolic equations. *Colloq. Math.*, **60/61** (1990), 157–194.
- [HKu05] J. Hu and T. Kumagai. Nash-type inequalities and heat kernels for non-local Dirichlet forms. *Kyushu J. Math* **60** (2006) 245–265.
- [HKa05] R. Hussein and M. Kassmann. Jump processes,  $\mathcal{L}$ -harmonic functions and continuity estimates. *preprint*; see <http://www.iam.uni-bonn.de/~kassmann/publications.html>.
- [INW66] N. Ikeda, N. Nagasawa, and S. Watanabe. A construction of Markov processes by piecing out. *Proc. Japan Acad.*, **42** (1966), 370–375.

- [Kol00] V. Kolokoltsov. Symmetric stable laws and stable-like jump-diffusions. *Proc. London Math Soc.*, **80** (2000), 725–768.
- [KS80] N. V. Krylov and M. V. Safonov. A property of the solutions of parabolic equations with measurable coefficients. *Izv. Akad. Nauk SSSR Ser. Mat.*, **44** (1980), 161–175 and 239.
- [Lev72] T. Leviatan. Perturbations of Markov processes. *J. Funct. Anal.*, **10** (1972), 309–325.
- [Mey75] P.-A. Meyer. Renaissance, recollements, mélanges, ralentissement de processus de Markov. *Ann. Inst. Fourier*, **25** (1975), 464–497.
- [Mo94] U. Mosco. Composite media and asymptotic Dirichlet forms. *J. Funct. Anal.*, **123** (1994), 368–421.
- [Mos61] J. Moser. On Harnack’s theorem for elliptic differential equations. *Comm. Pure Appl. Math.*, **14** (1961), 577–591.
- [Mos64] J. Moser. A Harnack inequality for parabolic differential equations. *Comm. Pure Appl. Math.*, **17** (1964), 101–134.
- [Mos71] J. Moser. On a pointwise estimate for parabolic differential equations. *Comm. Pure Appl. Math.*, **24** (1971), 727–740.
- [Nas58] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, **80** (1958), 931–954.
- [SCS91] L. Saloff-Coste and D. W. Stroock. Opérateurs uniformément sous-elliptiques sure les groupes de Lie. *J. Funct. Anal.*, **98** (1991), 97–121.
- [SU05] R.L. Schilling and T. Uemura, On the Feller property of Dirichlet forms generated by pseudo differential operators. *Tohoku Math. J.*, to appear.
- [SS06] R. Schneider and C. Schwab. Wavelet solutions of variable order pseudodifferential equations. preprint.
- [Si74] M.L. Silverstein, *Symmetric Markov Processes*. Springer-Verlag, Berlin-New York, 1974.
- [SV04] R. Song and Z. Vondracek. Harnack inequality for some classes of Markov processes. *Math. Z.*, **246** (2004), 177–202.
- [Str88] D. W. Stroock. Diffusion semigroups corresponding to uniformly elliptic divergence form operators. In *Séminaire de Probabilités, XXII*, volume 1321 of *Lecture Notes in Math.*, pages 316–347. Springer, Berlin, 1988.
- [TT07] M. Takeda and K. Tsuchida, Differentiability of spectral functions for symmetric  $\alpha$ -stable process. To appear in *Trans. Amer. Math. Soc.*

**Martin T. Barlow**  
Department of Mathematics  
University of British Columbia  
Vancouver, BC, V6T 1Z2, Canada  
barlow@math.ubc.ca

**Richard. F. Bass**  
Department of Mathematics  
University of Connecticut  
Storrs, CT 06269-3009, USA  
bass@math.uconn.edu

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

**Moritz Kassmann**  
Institut für Angewandte Mathematik  
Universität Bonn  
Beringstrasse 6  
D-53115 Bonn, Germany  
kassmann@iam.uni-bonn.de