DAVIES’ METHOD FOR ANOMALOUS DIFFUSIONS

MATHAV MURUGAN AND LAURENT SALOFF-COSTE†

Abstract. Davies’ method of perturbed semigroups is a classical technique to obtain off-diagonal upper bounds on the heat kernel. However, Davies’ method does not apply to anomalous diffusions due to the singularity of energy measures. In this note, we overcome the difficulty by modifying the Davies’ perturbation method to obtain sub-Gaussian upper bounds on the heat kernel. Our computations closely follow the seminal work of Carlen, Kusuoka and Stroock [8]. However, a cutoff Sobolev inequality due to Andres and Barlow [1] is used to bound the energy measure.

1. Introduction

Davies’ method of perturbed semigroups is a well-known method to obtain off-diagonal upper bounds on the heat kernel. It was introduced by E. B. Davies to obtain the explicit constants in the exponential term for Gaussian upper bounds [9] using the logarithmic Sobolev inequality. Davies’ method was extended by Carlen, Kusuoka and Stroock to a non-local setting [8, Section 3] using Nash inequality. Moreover, Davies extended this technique to higher order elliptic operators on $\mathbb{R}^n$ [10, Section 6 and 7]. More recently Barlow, Grigor’yan and Kumagai applied Davies’ method as presented in [8] to obtain off-diagonal upper bounds for the heat kernel of heavy tailed jump processes [6, Section 3].

Despite these triumphs, Davies’ perturbation method has not yet been made to work in the following contexts:

(a) Anomalous diffusions (See [3, Section 4.2]).
(b) Jump processes with jump index greater than or equal to 2 (See [16, Remark 1(d)] and [11, Section 1]).

The goal of this work is to extend Davies’ method to anomalous diffusions in order to obtain sub-Gaussian upper bounds. In the anomalous diffusion setting, we use cutoff functions satisfying a cutoff Sobolev inequality to perturb the corresponding heat semigroup. We use a recent work of Andres and Barlow [1] to construct these cutoff functions. We extend the techniques developed here in a sequel to a non-local setting for the jump processes mentioned in (b) above [17]. In [17], we consider the analogue of symmetric stable processes on fractals while in this work we are motivated by Brownian motion on fractals.

Date: March 10, 2016.

2010 Mathematics Subject Classification. 60J60, 60J35.

‡Both the authors were partially supported by NSF grant DMS 1404435.
Before we proceed, we briefly outline Davies’ method as presented in [8] and point out the main difficulty in extending it to the anomalous diffusion setting. Consider a metric measure space \((M, d, \mu)\) and a Markov semigroup \((P_t)_{t \geq 0}\) symmetric with respect to \(\mu\). Instead of considering the original Markov semigroup \((P_t)_{t \geq 0}\), we consider the perturbed semigroup
\[
(P_t^\psi f)(x) = e^{\psi(x)} (P_t (e^{-\psi} f))(x)
\]
where \(\psi\) is a ‘sufficiently nice function’. Given an ultracontractive estimate
\[
\|P_t\|_{1 \to \infty} \leq m(t)
\]
for the diffusion semigroup, Davies’ method yields an ultracontractive estimate for the perturbed semigroup
\[
\|P_t^\psi\|_{1 \to \infty} \leq m_\psi(t).
\]
If \(p_t(x,y)\) is the kernel of \(P_t\), then the kernel of \(P_t^\psi\) is
\[
p_t^\psi(x,y) = e^{-\psi(x)} p_t(x,y) e^{\psi(y)}.
\]
Therefore by (3), we obtain the off-diagonal estimate
\[
p_t(x,y) \leq m_\psi(t) \exp (\psi(y) - \psi(x)).
\]
By varying \(\psi\) over a class of ‘nice functions’ to minimize the right hand side of (4), Davies obtained off-diagonal upper bounds. In Davies’ method as presented in [9, 8], it is crucial that the function \(\psi\) satisfies
\[
e^{-2\psi} \Gamma(e^\psi, e^\psi) \ll \mu \quad \text{and} \quad e^{2\psi} \Gamma(e^{-\psi}, e^{-\psi}) \ll \mu,
\]
where \(\Gamma(\cdot, \cdot)\) denotes the corresponding energy measure. In fact the expression of \(m_\psi\) in (3) depends on the uniform bound on the Radon-Nikodym derivatives of the energy measure given by (See [8, Theorem 3.25])
\[
\Gamma(\psi) := \left\| \frac{d e^{-2\psi} \Gamma(e^\psi, e^\psi)}{d\mu} \right\|_\infty \left\| \frac{d e^{2\psi} \Gamma(e^{-\psi}, e^{-\psi})}{d\mu} \right\|_\infty.
\]

The main difficulty in extending Davies’ method to anomalous diffusions is that, for many ‘typical fractals’ that satisfy a sub-Gaussian estimate, the energy measure \(\Gamma(\cdot, \cdot)\) is singular with respect to the underlying symmetric measure \(\mu\) [12, 14, 7]. This difficulty is well-known to experts (for instance, [4, p. 1507] or [13, p. 86]). In this context, the condition \(e^{-2\psi} \Gamma(e^\psi, e^\psi) \ll \mu\) implies that \(\psi\) is necessarily a constant, in which case the off-diagonal estimate of (4) is not an improvement over the diagonal estimate (3).

Let \((M, d, \mu)\) be a locally compact metric measure space where \(\mu\) is a positive Radon measure on \(M\) with \(\text{supp}(\mu) = M\). We denote by \(\langle \cdot, \cdot \rangle\) the inner product on \(L^2(M, \mu)\). We consider a regular strongly local Dirichlet form \((\mathcal{E}, \mathcal{F})\) with generator \(-\mathcal{L}\), where \(\mathcal{L}\) is a positive definite, self-adjoint operator. That is
\[
\mathcal{E}(f,g) = -\langle \mathcal{L}f, g \rangle \quad \text{for all} \quad f \in \mathcal{D}(\mathcal{L}), f \in \mathcal{F}.
\]
Let \((P_t)_{t \geq 0}\) denote the associated semigroup and let \(p_t(\cdot, \cdot)\) be the (regularised) kernel of \(P_t\) with respect to \(\mu\) [1, eq. (1.10)]. We denote by \(B(x,r) := \{y \in M : d(x,y) < r\}\) the ball centered at \(x\) with radius \(r\) and by
\[
V(x,r) := \mu(B(x,r))
\]
DAVIES’ METHOD FOR ANOMALOUS DIFFUSIONS

the corresponding volume. We assume that the metric measure space is Ahlfors-regular: meaning that there exist $C_1 > 0$ and $d_f > 0$ such that

$$C_1^{-1} r^{d_f} \leq V(x, r) \leq C_1 r^{d_f}$$

for all $x \in M$ and for all $r \geq 0$. The quantity $d_f > 0$ is called the volume growth exponent or fractal dimension. We are interested in obtaining sub-Gaussian upper bounds of the form

$$p_t(x, y) \leq C_1 \frac{t^{d_f/d_w}}{t^{d_w/d_f}} \exp \left( -C_2 \left( \frac{d(x, y)^{d_w}}{t} \right)^{1/(d_w-1)} \right)$$

USG($d_f, d_w$)

where $d_w \geq 2$ is the escape time exponent or walk dimension. Such sub-Gaussian estimates are typical of many fractals. We assume the on-diagonal bound corresponding to the sub-Gaussian estimate of USG($d_f, d_w$). That is, we assume that there exists $C_1 > 0$ such that

$$p_t(x, x) \leq C_1 \frac{t^{d_f/d_w}}{t^{d_w/d_f}}$$

(5)

for all $x \in M$ and for all $t > 0$. The on-diagonal estimate of (5) is equivalent to the following Nash inequality ([8, Theorem 2.1]): there exists $C_N > 0$ such that

$$\|f\|_2^{2(1+d_w/d_f)} \leq C_N E(f, f) \|f\|_1^{2d_w/d_f}$$

N($d_f, d_w$)

for all $f \in F \cap L^1(M, \mu)$. The Nash inequality N($d_f, d_w$) may be replaced by an equivalent Sobolev inequality, a logarithmic Sobolev inequality or a Faber-Krahn inequality (See [2]). However, we will follow the approach of [8] and use the Nash inequality. Such a Nash inequality can be obtained from geometric assumptions like a Poincaré inequality and a volume growth assumption like $V(d_f)$.

Since $E$ is regular, it follows that $E(f, g)$ can be written in terms of a signed measure $\Gamma(f, g)$ as

$$E(f, g) = \int_M \Gamma(f, g),$$

where the energy measure $\Gamma$ is defined as follows. For any essentially bounded $f \in F$, $\Gamma(f, f)$ is the unique Borel measure on $M$ (called the energy measure) on $M$ satisfying

$$\int_M g d\Gamma(f, f) = E(f, fg) - \frac{1}{2} E(f^2, g)$$

for all essentially bounded $g \in F$; $\Gamma(f, g)$ is then defined by polarization. We shall use the following properties of the energy measure.

(i) **Locality:** For all functions $f, g \in F$ and all measurable sets $G \subset M$ on which $f$ is constant

$$1_G d\Gamma(f, g) = 0$$

(ii) **Leibniz and chain rules:** For $f, g \in F$ essentially bounded and $\phi \in C^1(\mathbb{R})$,

$$d\Gamma(fg, h) = fd\Gamma(g, h) + gd\Gamma(f, h)$$

(6)

$$f\Gamma(\phi(f), g) = \phi'(f) d\Gamma(f, g).$$

(7)
We wish to obtain an off-diagonal estimate using Davies’ perturbation method. The main difference from the previous implementations of the method is that, in addition to an on-diagonal upper bound (or equivalently Nash inequality), we also require a cutoff Sobolev inequality. Spaces satisfying the sub-Gaussian upper bound given in USG$(d_f,d_w)$ necessarily satisfy the cutoff Sobolev annulus inequality CSA$(d_w)$, a condition introduced by Andres and Barlow [1]. The condition CSA simplifies the cut-off Sobolev inequalities CS which were originally introduced by Barlow and Bass [4] for weighted graphs. The significance of the cut-off Sobolev inequalities CS and CSA is that they are stable under bounded perturbations of the Dirichlet form (Cf. [1, Corollary 5.2]). Moreover, the condition CS is stable under quasi-isometries of the underlying space [5, Theorem 2.21(b)]. Therefore cutoff Sobolev inequalities provide a robust method to obtain heat kernel estimates with anomalous time-space scaling. We now define the cutoff Sobolev inequality CSA$(d_w)$.

**Definition 1.1.** Let $U \subset V$ be open sets in $M$ with $U \subset \overline{U} \subset V$. We say that a continuous function $\phi$ is a cutoff function for $U \subset V$ if $\phi \equiv 1$ on $U$ and $\phi \equiv 0$ on $V^c$.

**Definition 1.2.** ([1, Definition 1.10]) We say CSA$(d_w)$ holds if there exists $C_1,C_2>0$ such that for every $x \in M$, $R>0$, $r>0$, there exists a cutoff function $\phi$ for $B(x,R) \subset B(x,R+r)$ such that if $f \in F$, then
\[
\int_{U} f^2 d\Gamma(\phi,\phi) \leq C_1 \int_{U} \phi^2 d\Gamma(f,f) + \frac{C_2}{r^{d_w}} \int_{U} f^2 d\mu, \quad \text{CSA}(d_w)
\]

where $U = B(x,R+r) \setminus \overline{B(x,r)}$.

It is clear that the condition CSA$(d_w)$ is preserved by bounded perturbations of the Dirichlet form. The above definition is slightly different to the one introduced in [1, Definition 1.10]. However both definitions are equivalent due to a ‘self-improving’ property of CSA$(d_w)$ [1, Lemma 5.1].

Our main result is that the Nash inequality N$(d_f,d_w)$ and the cutoff Sobolev inequality CSA$(d_w)$ imply the desired sub-Gaussian estimate USG$(d_f,d_w)$. More precisely,

**Theorem 1.3.** Let $(M,d,\mu)$ be a locally compact metric measure space that satisfies $V(d_f)$ with volume growth exponent $d_f$. Let $(\mathcal{E},\mathcal{F})$ be a strongly local, regular, Dirichlet form whose energy measure $\Gamma$ satisfies the cutoff Sobolev inequality CSA$(d_w)$ for some $d_w > 2$. Then the Nash inequality N$(d_f,d_w)$ implies the sub-Gaussian upper bound USG$(d_f,d_w)$.

**Remark.** The above properties given by $V(d_f)$ and USG$(d_f,d_w)$ are a special case of the more general assumptions of volume doubling and heat kernel upper bounds with a general time-space scaling of [1]. In fact, Theorem 1.3 is subsumed by [1, Theorem 1.12]. A recent work of Lierl provides an alternate proof of the sub-Gaussian estimates in [1] using Moser’s iteration method and extends the results to certain time-dependent, non-symmetric local bilinear forms [15]. Like earlier work
by Andres and Barlow and the present work, Lierl’s arguments involve improved
closest control on some cutoff functions.

Our methods give an alternate proof to [1, Theorem 1.12] in a restricted setting.
Moreover we show in [17] that this techniques can be adapted to the non-local
setting to provide new result and resolve the conjecture posed in [16, Remark
1(d)].

2. Off diagonal estimates using Davies’ method

Spaces satisfying have a rich class of cutoff functions with low energy. We start
by studying energy estimates of these cutoff functions.

2.1. Self-improving property of CSA. The cutoff Sobolev inequality CSA(d_w)
has a self-improving property which states that the constants C_1, C_2 in CSA(d_w)
are flexible. For example, we can decrease the value of C_1 in CSA(d_w) by increasing
C_2 appropriately. This is quantified in Lemma 2.1. Lemma 2.1 is essentially
contained in [1]; we simplify the proof and obtain a slightly stronger result.

Lemma 2.1. Let (M, d, µ) satisfy V(d_f). Let (E, F) denote a strongly local, reg-
ular, Dirichlet form with energy measure Γ that satisfies CSA(d_w). There ex-
ists C > 0 such that for each ρ ∈ (0, 1], there exists a cutoff function φ_ρ for
B(x, R) ⊂ B(x, R + r) that satisfies

\[ \int_U f^2 d Γ(φ_ρ, φ_ρ) \leq 4 C_1 ρ^2 \int_U d Γ(f, f) + \frac{C_2 ρ^{2-d_w}}{ρ^{d_w}} \int_U f^2 d µ \]  

(8)

for all f ∈ F, where C_1, C_2 are the constants in CSA(d_w). Further the cutoff
function φ_ρ above satisfies

\[ \left| φ_ρ(y) - \left( \frac{R + r - d(x, y)}{r} \right) \right| \leq 2ρ \]  

(9)

for all y ∈ B(x, R + r) \ B(x, R).

Remark. Lemma 2.1 is essentially contained in work of Andres and Barlow [1,
Lemma 5.1]. More recently, following [1, Lemma 5.1], Lierl obtained a cutoff
Sobolev inequality [15, Lemma 2.3] that is similar to Lemma 2.1. However, the
estimate (9) is new and it shows that the cutoff functions converge in L^∞ norm as ρ → 0 to the “linear cutoff function”. The constructions in [1, Lemma 5.1] and
[15, Lemma 2.3] converge in L^∞ norm as ρ → 0 to a somewhat more complicated
cutoff function that depends on d_w. The proof below was suggested to us by
Martin Barlow.

Proof of Lemma 2.1. Let x ∈ M, r > 0, R > 0, ρ > 0. Define n := [ρ^{-1}] ∈
[ρ^{-1/2}, ρ^{-1}]. We divide the annulus U = B(x, R + r) \ B(x, R) into n-annuli
U_1, U_2, ..., U_n of equal width, where

\[ U_i := B(x, R + ir/n) \setminus B(x, R + (i - 1)r/n), \quad i = 1, 2, ..., n. \]
By CSA($d_w$), there exists a cutoff function $\phi_i$ for $B(x, R + (i - 1)r/n) \subset B(x, R + ir/n)$ satisfying
\[
\int_{U_i} f^2 \, d\Gamma(\phi_i, \phi_i) \leq C_1 \int_{U_i} d\Gamma(f, f) + \frac{C_2}{(r/n)^d_w} \int_{U_i} f^2 \, d\mu \quad (10)
\]
for $i = 1, 2, \ldots, n$. We define $\phi = n^{-1} \sum_{i=1}^n \phi_i$. By locality, we have
\[
d\Gamma(\phi, \phi) = \frac{1}{n^2} \sum_{i=1}^n d\Gamma(\phi_i, \phi_i). \quad (11)
\]
Therefore by (11), (10) and $\rho^{-1}/2 \leq n = |\rho^{-1}| \leq \rho^{-1}$, we obtain
\[
\int_{U} f^2 \, d\Gamma(\phi, \phi) = n^{-2} \sum_{i=1}^n \int_{U_i} f^2 \, d\Gamma(\phi, \phi) \leq n^{-2} \sum_{i=1}^n \left( C_1 \int_{U_i} d\Gamma(f, f) + \frac{C_2}{(r/n)^d_w} \int_{U_i} f^2 \, d\mu \right) \leq C_1 n^{-2} \int_{U} d\Gamma(f, f) + \frac{C_2 n^{d_w-2}}{r^{d_w}} \int_{U_i} f^2 \, d\mu \leq 4C_1 \rho^2 \int_{U} d\Gamma(f, f) + \frac{C_2 \rho^{2-d_w}}{r^{d_w}} \int_{U_i} f^2 \, d\mu.
\]
This completes the proof of (8).

Observe that by (9), the cutoff function $\phi_\rho$ for $B(x, r) \subset B(x, R + r)$ satisfies
\[
\lim_{\rho \downarrow 0} \phi_\rho(y) = 1 \wedge \left( \frac{(R + r - d(x, y))_+}{r} \right),
\]
where the above limit is with respect to $L^\infty(M, \mu)$ norm. Besides the shorter proof, an important motivation behind the above construction of the cutoff functions is that this approach extends to jump processes (as opposed to the previous constructions in [1, 15]). For an extension of the Lemma 2.1 to a non-local setting (i.e. jump process), we refer the reader to [17, Proposition 2.1].

2.2. Estimates on perturbed forms. The key to carry out Davies’ method is the following elementary inequality.

**Lemma 2.2.** Let $(\mathcal{E}, \mathcal{F})$ be a strongly local, regular, Dirichlet form. Then
\[
\mathcal{E}(e^{\psi} f^{2p-1}, e^{-\psi} f) \geq \frac{1}{p} \mathcal{E}(f^p, f^p) - p \int_M f^2 \, d\Gamma(\psi, \psi) \quad (12)
\]
for all $f \in \mathcal{F}$, $\psi \in \mathcal{F}$ and $p \in [1, \infty)$. 

Note that if $y \in U_i$, then $1 - i/n \leq \phi_i(y) \leq 1 - (i - 1)/n$ and $R + (i - 1)r/n \leq d(x, y) < R + ir/n$, for each $1 \leq i \leq n$. This along with $n^{-1} \leq 2\rho$ implies (9). □
Proof. Using Leibniz rule (6) and chain rule (7), we obtain
\[
\Gamma(e^\psi f^{2p-1}, e^{-\psi} f) - \frac{1}{p} \Gamma(f^p, f^p) = \int_M f^{2p-1} \Gamma(f, f) - \frac{1}{p} \Gamma(f^p, f^p) \Gamma(\psi, \psi)
\]
\[
= (p-1) \left( f^{2(p-1)} \Gamma(f, f) + f^{2p} \Gamma(\psi, \psi) - 2 f^{2(p-1)} \Gamma(f, \psi) \right). \tag{13}
\]
By \cite[Theorem 3.7]{GU} and Cauchy-Schwarz inequality, we have
\[
\int_M f^{2p-1} \Gamma(f, f) \leq \left( \int_M f^{2(p-1)} d\Gamma(f, f) \cdot \int_M f^{2p} d\Gamma(\psi, \psi) \right)^{1/2}.
\]
Therefore
\[
2 \int_M f^{2p-1} d\Gamma(f, f) \leq \int_M f^{2(p-1)} d\Gamma(f, f) + \int_M f^{2p} d\Gamma(\psi, \psi). \tag{14}
\]
By integrating (13) and using (14), we obtain (12). \hfill \Box

Davies used the bound
\[
\int_M f^{2p} d\Gamma(\psi, \psi) \leq \left\| \frac{d\Gamma(\psi, \psi)}{d\mu} \right\|_\infty \left\| f \right\|_{2p}^2
\]
to control a term in (12). However for anomalous diffusions, the energy measure is singular to \( \mu \). We will instead use \( \text{CSA}(d_w) \) to bound \( \int_M f^{2p} d\Gamma(\psi, \psi) \) by choosing \( \psi \) to be a multiple of the cutoff function satisfying \( \text{CSA}(d_w) \). The following estimate is analogous to \cite[Theorem 3.9]{GU} but unlike in \cite{GU}, the cutoff functions depend on both \( p \) and \( \lambda \). This raises new difficulties is the implementation of Davies’ method.

**Proposition 2.3.** Let \((M, d, \mu)\) be a metric measure space. Let \((\mathcal{E}, \mathcal{F})\) be a strongly local, regular, Dirichlet form on \( M \) satisfying \( \text{CSA}(d_w) \). There exists \( C > 0 \) such that, for all \( \lambda > 1 \), for all \( r > 0 \), for all \( x \in M \) and for all \( p \in [1, \infty) \), there exists a cutoff function \( \phi = \phi_{p, \lambda} \) on \( B(x, r) \subset B(x, 2r) \) such that
\[
\mathcal{E}(e^{\psi f} f^{2p-1}, e^{-\psi} f) \geq \frac{1}{2p} \mathcal{E}(f^p, f^p) - C \frac{\lambda^d w^{d-1}}{r^{d_w}} \left\| f \right\|_{2p}^2.
\]
for all \( f \in \mathcal{F} \). There exists \( C' > 0 \) such that the cutoff functions \( \phi_{p, \lambda} \) above satisfy
\[
\left\| \exp \left( \lambda (\phi_{p, \lambda} - \phi_{2p, \lambda}) \right) \right\|_\infty \vee \left\| \exp \left( -\lambda (\phi_{p, \lambda} - \phi_{2p, \lambda}) \right) \right\|_\infty \leq \exp(C' / p)
\]
for all \( \lambda \geq 1 \) and for all \( p \geq 1 \).

**Proof.** This Theorem follows from Lemma 2.2 and Lemma 2.1. Let \( x \in M \) and \( r > 0 \) be arbitrary. Using (12), we obtain
\[
\mathcal{E}(e^{\psi f} f^{2p-1}, e^{-\psi} f) \geq \frac{1}{p} \left( \mathcal{E}(f^p, f^p) - (p\lambda)^2 \int_M f^{2p} d\Gamma(\phi, \phi) \right) \tag{17}
\]
By Lemma 2.1 and fixing \( p^2 = (p\lambda)^{-2}/(8C_1) \) in (8), we obtain a cutoff function \( \phi = \phi_{p, \lambda} \) for \( B(x, r) \subset B(x, 2r) \) and \( C > 0 \) such that
\[
(p\lambda)^2 \int_M f^{2p} d\Gamma(\phi, \phi) \leq \frac{1}{2} \mathcal{E}(f^p, f^p) + C \frac{(\lambda p)^{d_w}}{r^{d_w}} \int_M f^{2p} d\mu \tag{18}
\]
By (17) and (18), we obtain (15).

By (9) and the above calculations, there exists $C' > 0$ such that the cutoff functions $\phi_{p,\lambda}$ satisfy

$$\| \phi_{p,\lambda} - \phi_{2p,\lambda} \|_\infty \leq \frac{C'}{p\lambda}$$

for all $p \geq 1$, for all $\lambda \geq 1$, for all $x \in M$ and for all $r > 0$. This immediately implies (16).

□

Remark. Estimates similar to (15), were introduced by Davies in [10, equation (3)] to obtain off-diagonal estimates for higher order (order greater than 2) elliptic operators. Roughly speaking, the generator $L$ for anomalous diffusion with walk dimension $d_w$ behaves like an ‘elliptic operator of order $d_w$’. However the theory presented in [10] is complete only when the ‘order’ $d_w$ is bigger than the volume growth exponent $d_f$, i.e. in the strongly recurrent case. This is because the method in [10] relies on a Gagliardo-Nirenberg inequality which is true only in the strongly recurrent setting. We believe that one can adapt the methods of [10] to obtain an easier proof for the strongly recurrent case. However, we will not impose any such restrictions and our proof will closely follow the one in [8].

2.3. Proof of Theorem 1.3: Let $\lambda \geq 1$ and $x \in M$ and $r > 0$. Let $p_k = 2^k$ and let $\psi_k = \lambda \phi_{p_k,\lambda}$, where $\phi_{p_k,\lambda}$ is a cutoff function on $B(x, r) \subset B(x, 2r)$ given by Proposition 2.3. We write

$$f_{t, k} := P_{t}^{\psi_k} f$$

for all $k \in \mathbb{N}$, where $f \in \mathcal{F}$ and $P_{t}^{\psi_k}$ denotes the perturbed semigroup as in (1).

Using (15), there exists $C_0 > 0$ such that

$$\frac{d}{dt} \| f_{t, 0} \|_2^2 = -2\mathcal{E} \left( e^{\psi} f_{t, 0}, e^{-\psi} f_{t, 0} \right)$$

$$\leq 2C_0 \frac{\lambda d_w}{r d_w} \| f_{t, 0} \|_2^2$$

(20)

and

$$\frac{d}{dt} \| f_{t, k} \|_{2p_k}^2 = -2p_k \mathcal{E} \left( e^{\psi_k} f_{t, k}^{2p_k-1}, e^{-\psi_k} f_{t, k} \right)$$

$$\leq -\mathcal{E} \left( f_{t, k}^{p_k}, f_{t, k}^{p_k} \right) + 2C_0 \left( \frac{\lambda d_w}{r} \right)^{d_w} \| f_{t, k} \|_{2p_k}^2$$

(21)

for all $k \in \mathbb{N}^*$. By (20), we obtain

$$\| f_{t, 0} \|_{p_1} = \| f_{t, 0} \|_2 \leq \exp \left( C_0 \lambda^{d_w} \frac{t}{r} \frac{d_w}{d_f} \right) \| f \|_2.$$

(22)

Using (21) and Nash inequality $N(d_f, d_w)$, we obtain

$$\frac{d}{dt} \| f_{t, k} \|_{2p_k} \leq -\frac{1}{2C_N p_k} \| f_{t, k} \|_{2p_k}^{1+2d_w p_k/df} \| f_{t, k} \|_{p_k}^{-2d_w p_k/df} + C_0 p_k^{d_w-1} \left( \frac{\lambda}{r} \right)^{d_w} \| f_{t, k} \|_{2p_k}$$

(23)
for all $k \in \mathbb{N}^*$. By (16) and the fact that $P_t$ is a contraction on $L^\infty$, we have
\[ \exp(-2C_1/p_k)f_{t,k+1} \leq f_{t,k} \leq \exp(2C_1/p_k)f_{t,k+1} \] (24)
for all $k \in \mathbb{N}_{\geq 0}$. Combining (23) and (24), we obtain
\[ \frac{d}{dt}\|f_{t,k}\|_{2p_k} \leq -\frac{1}{CAp_k}\|f_{t,k}\|_{1+2dwp_k/d_f}\|f_{t,k-1}\|_{-2dwp_k/d_f} + C_0p_d^{d_w-1}\left(\frac{\lambda}{\tau}\right)^{d_w}\|f_{t,k}\|_{2p_k} \] (25)
for all $k \in \mathbb{N}^*$, where $C_A = 2C_N \exp(8d_wC_1/d_f)$.

To obtain off-diagonal estimates using the differential inequalities (25) we use the following lemma. The following lemma is analogous to [8, Lemma 3.21] but the statement and its proof is slightly modified to suit our anomalous diffusion context with walk dimension $d_w$.

**Lemma 2.4.** Let $w : [0, \infty) \to (0, \infty)$ be a non-decreasing function and suppose that $u \in C^1([0, \infty); (0, \infty))$ satisfies
\[ u'(t) \leq -\frac{\epsilon}{p}\left(\frac{t^{(p-2)/\theta p}}{w(t)}\right)^{\theta p}u^{1+\theta p}(t) + \delta p^{d_w-1}u(t) \] (26)
for some positive $\epsilon, \theta$ and $\delta$, $d_w \in [2, \infty)$ and $p = 2^k$ for some $k \in \mathbb{N}^*$. Then $u$ satisfies
\[ u(t) \leq \left(\frac{2p_d^{d_w}}{\epsilon \theta}\right)^{1/\theta p}t^{(1-p)/\theta p}w(t)\exp(\delta t/p) \] (27)

**Proof.** Set $v(t) = e^{-\delta p^{d_w-1}t}u(t)$. By (26), we have
\[ v'(t) = e^{-\delta p^{d_w-1}t}\left(u(t) - \delta p^{d_w-1}u(t)\right) \leq -\frac{\epsilon t^{p-2}}{pw(t)^{\theta p}}e^{\theta p^{d_w}t}v(t)^{1+\theta p}. \]
Hence
\[ \frac{d}{dt}(v(t))^{-\theta p} \geq \epsilon \theta t^{p-2}w(t)^{-\theta p}e^{\delta p^{d_w}t} \]
and so, since $w$ is non-decreasing
\[ e^{\delta p^{d_w}t}u(t)^{-\theta p} \geq \epsilon \theta w(t)^{-\theta p}\int_0^t s^{(p-2)}e^{\delta p^{d_w}s}ds. \] (28)

Note that
\[ \int_0^t s^{(p-2)}e^{\delta p^{d_w}s}ds \geq \frac{t^{p-1}}{p-1} \int_{\delta p^{d_w}(1-1/p^{d_w})}^{\delta p^{d_w}} s^{(p-2)}e^{ty}dy \]
\[ \geq \frac{t^{p-1}}{p-1} \exp\left(\delta p^{d_w}t - \delta t\right) \left[1 - (1 - p^{-d_w})^{p-1}\right] \]
\[ \geq \frac{t^{p-1}}{2p^{d_w}} \exp\left(\delta p^{d_w}t - \delta t\right) \] (29)

In the last line above, we used the bound $(1 - p^{-d_w})^{p-1} \geq 1 - p^{-d_w}(p-1)$ for all $p, d_w \geq 2$. Combining (28) and (29) yields (27).
We now pick \( f \in L^2(M, \mu) \) and \( f \geq 0 \) with \( \|f\|_2 = 1 \). Let \( u_k(t) = \|f_{t,k-1}\|_{p_k} \) and let
\[
w_k(t) = \sup\{s^{d_f(p_k-2)/(2d_0p_k)}u_k(s) : s \in (0, t]\}.
\]
By (22), \( w_1(t) \leq \exp(2C_0\lambda^{d_w}t/r^{d_w}) \). Further by (25), \( u_{k+1} \) satisfies (26) with \( \epsilon = 1/C_A, \theta = 2d_w/d_f, \delta = C_0(\lambda/r)^{d_w}, \) \( w = w_k \) and \( p = p_k \). Hence by (27),
\[
u_{k+1}(t) \leq (2d_wk+1/\epsilon\theta)^{1/(\theta p_k)}(1-p_k)^{p_k}k\theta \varepsilon_t k w_k(t).
\]
Therefore
\[
w_{k+1}(t)/w_k(t) \leq (2d_k\theta + 1/\epsilon\theta)^{1/(\theta p_k)}\varepsilon_t k^{2k}
\]
for \( k \in \mathbb{N}^* \). Hence, we obtain
\[
\lim_{k \to \infty} w_k(t) \leq C_2 e^{\delta t} w_1(t) \leq C_2 \exp(2C_0\lambda^{d_w}t/r^{d_w})
\]
where \( C_2 = C_2(d_w, \epsilon, \theta) \). Since \( P_t \) is a contraction on all \( L^p(M, \mu) \) for \( 1 \leq p \leq \infty \), we obtain
\[
\lim_{k \to \infty} u_k(t) = \left\| P_t^{\psi_\infty} f \right\|_{\infty} \leq \frac{C_2}{t^{d_f/2d_w}} \exp(3C_0\lambda^{d_w}t/r^{d_w})
\]
where \( \psi_\infty = \lim_{k \to \infty} \psi_k \). Since the above bound holds for all \( f \in L^2(M, \mu) \) with \( \|f\|_2 = 1 \), we have
\[
\left\| P_t^{\psi_\infty} \right\|_{2 \to \infty} \leq \frac{C_2}{t^{d_f/2d_w}} \exp(2C_0\lambda^{d_w}t/r^{d_w})
\]
The estimate is unchanged if we replace \( \psi_k \)'s by \( -\psi_k \). Since \( P_t^{-\psi} \) is the adjoint of \( P_t^{\psi} \), by duality we have that
\[
\left\| P_t^{\psi_\infty} \right\|_{1 \to 2} \leq \frac{C_2}{t^{d_f/2d_w}} \exp(2C_0\lambda^{d_w}t/r^{d_w})
\]
Combining the above, we have
\[
\left\| P_t^{\psi_\infty} \right\|_{1 \to \infty} \leq \frac{C_2 2^{d_f/d_w}}{t^{d_f/d_w}} \exp(2C_0\lambda^{d_w}t/r^{d_w})
\]
Therefore
\[
p_t(x, y) \leq \frac{C_2 2^{d_f/d_w}}{t^{d_f/d_w}} \exp(2C_0\lambda^{d_w}t/r^{d_w} + \psi_\infty(y) - \psi_\infty(x)).
\]
for all \( x, y \in M \) and for all \( r, t > 0 \) and \( \lambda \geq 1 \). If we choose \( r = d(x, y)/2 \), we have \( \psi_\infty(y) - \psi_\infty(x) = -\lambda \). This yields
\[
p_t(x, y) \leq \frac{C_3}{t^{d_f/d_w}} \exp(C_4\lambda^{d_w}t/d(x, y)^{d_w} - \lambda)
\]
where \( C_3, C_4 > 1 \). Assume \( \lambda = C_4^{-1/(d_w-1)}(d(x, y)^{d_w}/t)^{1/(d_w-1)} \geq 1 \) in the above equation, we obtain
\[
p_t(x, y) \leq \frac{C_3}{t^{d_f/d_w}} \exp\left(-\left(-\frac{d(x, y)^{d_w}}{C_3 t}\right)^{1/(d_w-1)}\right)
\]
for all \( x, y \in M \) and for all \( t > 0 \) such that \( d(x, y)^{d_w} \geq C_4 t \).
If \(d(x, y)^{d_\omega} < C_4 t\), the on-diagonal estimate (5) suffices to obtain the desired sub-Gaussian upper bound.

Acknowledgement. We thank Martin Barlow for his enlightening remarks on the cutoff Sobolev inequalities and for suggesting the proof of Lemma 2.1 that replaced a more complicated proof in an earlier draft. We thank Evan Randles for clarifying various aspects of Davies’ work [10] on higher order operators. We thank Tom Hutchcroft for proofreading part of the manuscript.

References
