

Heat kernel estimates for anomalous heavy-tailed random walks

Mathav Murugan^{a,1} and Laurent Saloff-Coste^{b,2}

^a*Department of Mathematics, University of British Columbia and Pacific Institute for the Mathematical Sciences, Vancouver, BC V6T 1Z2, Canada. E-mail: mathav@math.ubc.ca*

^b*Department of Mathematics, Cornell University, Ithaca, NY 14853, USA. E-mail: lsc@math.cornell.edu*

Received 27 May 2016; revised 4 September 2017; accepted 27 February 2018

Abstract. Sub-Gaussian estimates for the natural random walk is typical of many regular fractal graphs. Subordination shows that there exist heavy tailed jump processes whose jump indices are greater than or equal to two. However, the existing machinery used to prove heat kernel bounds for such heavy tailed random walks fail in this case. In this work we extend Davies' perturbation method to obtain transition probability bounds for these anomalous heavy tailed random walks. We prove global upper and lower bounds on the transition probability density that are sharp up to constants. An important feature of our work is that the methods we develop are robust to small perturbations of the symmetric jump kernel.

Résumé. Pour de nombreux graphes réguliers de type fractal, la marche aléatoire simple satisfait des estimations de type sous-Gaussiennes. La technique de la subordination montre alors qu'il existe des processus de saut à queue lourde dont l'indice des sauts est supérieur ou égale à 2. Pour de tels processus, les techniques usuelles pour les estimations loin de la diagonale ne fonctionnent pas. Nous étendons la célèbre méthode de Davies dans le cas de ces processus à sauts « anormaux. » Nous obtenons des bornes supérieures et inférieures précises sur le noyau de transition par des méthodes qui sont stables sous de petites perturbations des sauts.

MSC: 60J10; 60J75

Keywords: Heavy-tailed random walks; Jump process; Fractals; Heat kernel

1. Introduction

The motivation behind this work is to obtain transition probability estimates for a class of random walks with heavy tailed jumps on fractal-like graphs. Many fractals satisfy the following sub-Gaussian estimates on the transition probability density p_t for the natural diffusion: there exists $c_1, c_2, C_3, C_4 > 0$ such that

$$\begin{aligned} \frac{c_1}{t^{d_f/d_w}} \exp\left(-\left(\frac{d(x, y)^{d_w}}{c_2 t}\right)^{1/(d_w-1)}\right) \\ \leq p_t(x, y) \leq \frac{C_3}{t^{d_f/d_w}} \exp\left(-\left(\frac{d(x, y)^{d_w}}{C_4 t}\right)^{1/(d_w-1)}\right) \end{aligned} \quad (1)$$

for all points x, y in the underlying space and for all times $t > 0$ [3,5,11,28,31]. The parameters $d_f, d_w > 0$ are intrinsic and depend on the geometry of the underlying space. Sub-Gaussian estimates similar to (1) were obtained in

¹Partially supported by NSERC (Canada) and the Pacific Institute for the Mathematical Sciences.

²Partially supported by NSF grant DMS 1404435.

the discrete time setting for the nearest neighbor walks on fractal like graphs [4,6,9,26,27,30]. A precise formulation of these sub-Gaussian estimates for graphs is provided in $USG(d_f, d_w)$ and $LSG(d_f, d_w)$ in Section 3.1.

If \mathcal{L} denotes the generator of a diffusion whose heat kernel satisfies the sub-Gaussian estimate (1), then for any $\beta \in (0, d_w)$, the operator $-(-\mathcal{L})^{\beta/d_w}$, is the generator of a non-local Dirichlet form, and its heat kernel admits the estimate

$$p_t(x, y) \asymp \frac{1}{t^{d_f/\beta}} \left(1 + \frac{d(x, y)}{t^{1/\beta}} \right)^{-(d_f+\beta)} \tag{2}$$

for all times t and for all points x, y in the underlying space (see, for example, [32,38], [22, Lemma 5.4]). Here \asymp means that the ratio between both sides of the expression is bounded above and below by finite positive constants that do not depend on x, y, t . Estimates similar to (2) can be established using subordination in the discrete time context as well [36, Section 5]. Furthermore, it was shown in [25] that (1) and (2) exhaust all possible two-sided estimates of heat kernels of self-similar type. We call the parameter β in (2) the *jump index*.

Obtaining estimates of the form (2) for the case $\beta \in (0, 2)$ has received much attention recently, both in the context of continuous time jump processes (see, for example, [8,10,14,15]) and discrete time heavy tailed random walks (see [12,35]). For the case $\beta \in (0, 2)$, Davies’ method is used to obtain estimates of the form (2) which are stable under bounded perturbations of the corresponding Dirichlet form. However it is known that the use of Davies’ method to obtain the upper bound in (2) is no longer adequate for the case $\beta \geq 2$ (see [24, Section 1], [36, Remark 1.4(d)]).

Recently characterizations of bounds of the form (2) for the case $\beta \geq 2$ have been given in [24, Theorem 2.1 and Theorem 2.3]. However these characterizations rely on probabilistic estimates on the exit times (called survival estimates and tail estimates in [24]). It is not known if these survival and tail estimates are stable under bounded perturbation of the jump kernel of the jump process in the continuous time case or the conductance of the heavy tailed random walk in the discrete time case. Therefore there is a need to develop techniques to obtain (2) that are robust under bounded perturbation of the Dirichlet form for the case $\beta \geq 2$ – see Example 1 for a concrete example that compares our result with previous works.

The main goal of the work is to show that the estimate (2) is stable under bounded perturbation of conductances for a heavy tailed random walk in the case $\beta \geq 2$, under some natural hypothesis. This is carried out by modifying the Davies’ perturbation method using a cutoff Sobolev inequality. We extend the techniques developed in [37] to a non-local setting. Although our work concerns discrete time heavy tailed random walks, we expect that the techniques we develop extends to continuous time jump processes as well. For instance, in the proof of upper bound (see Proposition 3.1), we obtain corresponding continuous time bounds – cf. (71). Indeed, the modified Davies’ method developed in this work can be implemented for continuous time processes in the metric measure space setting without any new essential difficulties [29].

To state the results precisely, we recall some standard notions concerning graphs and Markov chains. Let $G = (X, E)$ be an infinite, simple, connected, locally finite graph. The elements of the set X are called vertices. Some of the vertices are connected by an edge, in which case we say that they are neighbors. Let $d(x, y)$ be the graph distance between points $x, y \in X$, that is the minimal number of edges in any edge path connecting x and y . Consider a measure μ on X . We sometimes abuse notation and consider μ as a function on X by setting $\mu(x)$ to be $\mu(\{x\})$. Denote the metric balls and their measures as follows

$$B(x, r) := \{y \in X : d(x, y) \leq r\} \quad \text{and} \quad V(x, r) := \mu(B(x, r))$$

for all $x \in X$ and $r \geq 0$. For convenience, we refer to the graph (X, E) endowed with the graph distance d and a measure μ as a metric measure space (X, d, μ) which we call a *vertex weighted graph*. We consider a vertex weighted graph (X, d, μ) that satisfies the polynomial volume growth assumption: there exists $d_f > 0, C_V > 0$ such that

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

for all $r \geq 1$ and for all $x \in X$. We use the notation $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ and $\mathbb{N}^* = \{1, 2, 3, \dots\}$.

Definition 1.1 (Symmetric Markov operator). We say that $K : \ell^\infty(X) \rightarrow \ell^\infty(X)$ is a μ -symmetric sub-Markov operator on a vertex weighted graph (X, d, μ) if there exists a kernel $k : X \times X \rightarrow [0, \infty)$ that satisfies

$$k(x, y) = k(y, x) \quad \text{for all } x, y \in X,$$

$$\sum_{y \in X} k(x, y)\mu(y) \leq 1 \quad \text{for all } x \in X,$$

$$Kf(x) = \sum_{y \in X} k(x, y)f(y)\mu(y) \quad \text{for all } x \in X, f \in \ell^\infty(X).$$

If the inequality above is replaced by an equality we say that K is a μ -symmetric Markov operator on a vertex weighted graph.

Given a set $U \subset X$ and a μ -symmetric Markov operator K on X with kernel k , we define the μ -symmetric sub-Markov operator K_U as the operator with kernel $k_U(x, y) = k(x, y)\mathbf{1}_U(x)\mathbf{1}_U(y)$.

There is a natural random walk $(Y_n)_{n \in \mathbb{N}}$ on X associated with a μ -symmetric Markov operator K with kernel k . The Markov chain $(Y_n)_{n \in \mathbb{N}}$ is defined by the following one-step transition probability

$$\mathbb{P}_x(Y_1 = y) := \mathbb{P}(Y_1 = y \mid Y_0 = x) = k(x, y)\mu(y).$$

The sub-Markov operator K_U then corresponds the Markov chain (Y_n) killed upon exiting U .

We say that a μ -symmetric Markov operator K on a vertex weighted graph (X, d, μ) satisfies $\mathbf{J}(\beta)$, if there exists a constant $C > 0$ such that the corresponding kernel k satisfies

$$C^{-1} \frac{1}{(1 + d(x, y))^{d_f + \beta}} \leq k(x, y) = k(y, x) \leq C \frac{1}{(1 + d(x, y))^{d_f + \beta}} \tag{J(\beta)}$$

for all $x, y \in X$, where d_f is the volume growth exponent given in $\mathbf{V}(d_f)$. If K satisfies $\mathbf{J}(\beta)$ for some $\beta > 0$, then we say that β is the *jump index* of the random walk driven by K .

Let $k_n(x, y)$ denote the kernel of the iterated power K^n with respect to the measure μ . We are interested in obtaining estimates on $k_n(x, y)$ for all values of $n \in \mathbb{N}^*$ and for all $x, y \in X$. We say that a μ -symmetric Markov operator K satisfies $\mathbf{HKP}(d_f, \beta)$, if there is a constant $C > 0$ such that the iterated kernel k_n satisfies the estimate

$$C^{-1} \min\left(\frac{1}{n^{d_f/\beta}}, \frac{n}{(d(x, y))^{d_f + \beta}}\right) \leq k_n(x, y) \leq C \min\left(\frac{1}{n^{d_f/\beta}}, \frac{n}{(d(x, y))^{d_f + \beta}}\right) \tag{HKP}(d_f, \beta)$$

for all $x, y \in X$ and for all $n \in \mathbb{N}^*$, where d_f is the volume group exponent given by $\mathbf{V}(d_f)$. We say that K satisfies $\mathbf{UHKP}(d_f, \beta)$, if the iterated kernel k_n satisfies the upper bound in $\mathbf{HKP}(d_f, \beta)$. Similarly, we say that K satisfies $\mathbf{LHKP}(d_f, \beta)$, if the iterated kernel k_n satisfies the lower bound in $\mathbf{HKP}(d_f, \beta)$. Note that $\mathbf{HKP}(d_f, \beta)$ is same as the bounds described in (2), since

$$\frac{1}{n^{d_f/\beta}} \left(1 + \frac{d(x, y)}{n^{1/\beta}}\right)^{-(d_f + \beta)} \leq \min\left(\frac{1}{n^{d_f/\beta}}, \frac{n}{(d(x, y))^{d_f + \beta}}\right),$$

$$\min\left(\frac{1}{n^{d_f/\beta}}, \frac{n}{(d(x, y))^{d_f + \beta}}\right) \leq 2^{d_f + \beta} \frac{1}{n^{d_f/\beta}} \left(1 + \frac{d(x, y)}{n^{1/\beta}}\right)^{-(d_f + \beta)}.$$

The goal of this work is to develop methods to obtain the bound $\mathbf{HKP}(d_f, \beta)$ that are robust to small perturbations of Dirichlet form in the sense given by $\mathbf{J}(\beta)$. Similar to the anomalous diffusion setting [37], we rely on a cutoff Sobolev inequality to implement Davies' method. To introduce cutoff Sobolev inequality, we first need to define cutoff functions and energy measure.

Definition 1.2 (Cutoff function). Let $U \subset V$ be finite sets in X . We say that $\phi : X \rightarrow \mathbb{R}$ is a cutoff function for $U \subset V$ if $\phi \equiv 1$ on U and $\phi \equiv 0$ on V^c .

Definition 1.3 (Energy measure). For a μ -symmetric Markov operator K with kernel $k(x, y)$ with respect to μ , and for functions $f, g \in L^\infty(X, \mu)$, we define the energy measure corresponding to K as the function

$$\Gamma(f, g)(x) = \frac{1}{2} \sum_{y \in X} (f(x) - f(y))(g(x) - g(y))k(x, y)\mu(y)\mu(x)$$

for all $x \in X$. The function $\Gamma(f, g)$ can be considered as a (signed) measure where the measure of the singleton $\{x\}$ is $\Gamma(f, g)(x)$.

We introduce a cutoff Sobolev inequality that plays an important role in obtaining $\text{HKP}(d_f, \beta)$. Cutoff Sobolev inequalities were first introduced by Barlow and Bass in [7] for graphs and then extended by Barlow, Bass and Kumagai [8] to metric measure spaces. Recently Andres and Barlow simplified the cutoff Sobolev inequalities in [1]. We obtain transition probability estimates using the following cutoff Sobolev inequality. The motivation behind studying cutoff Sobolev inequalities in [1,7,8] is that they provide a method to obtain sub-Gaussian estimates that is robust to bounded perturbation of the Dirichlet form. We are motivated by similar reasons to formulate a version of cutoff Sobolev inequality for heavy tailed random walks. Our definition below is inspired by the cutoff Sobolev annulus inequality in [1] and a self-improving property of cutoff Sobolev inequality [1, Lemma 5.1].

Definition 1.4 (Cutoff Sobolev inequality). Let K be a μ -symmetric Markov operator on a vertex weighted graph (X, d, μ) and let Γ denote the corresponding energy measure. Let $\beta \geq 2$. We say that K satisfies the cutoff Sobolev inequality $\text{CSJ}(\beta)$ if it satisfies the following property: for any $x \in X$ and for any $R, r > 0$, there exists a cutoff function ϕ for $B(x, R) \subset B(x, R+r)$ such that

$$\phi \equiv 1 \quad \text{in } B(x, R+r/2), \tag{3}$$

$$\phi \equiv 0 \quad \text{in } B(x, R+9r/10)^c, \tag{4}$$

and for any function $f \in L^2(X, \mu)$

$$\sum_{y \in U} f^2(y) \Gamma_U(\phi, \phi)(y) \leq C_1 \sum_{y \in U} \Gamma_U(f, f)(y) + \frac{C_2}{r^\beta} \sum_{y \in U} f^2(y) \mu(y), \tag{5}$$

where $U = B(x, R+r) \setminus B(x, R)$ and Γ_U is the energy measure corresponding to the sub-Markov operator K_U .

The condition $\text{CSJ}(\beta)$ is useful because it is preserved under bounded perturbation of the Markov kernel. The appearance of the expression r^β in the second term of (5) signifies the role played by β in the space-time scaling relation of the walk.

The main result is the following characterization of $\text{UHKP}(d_f, \beta)$ using $\text{CSJ}(\beta)$.

Theorem 1.5. Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$ with volume growth exponent d_f . Let K be μ -symmetric Markov operators whose kernel with respect to μ satisfies $\mathbf{J}(\beta)$ for some $\beta \geq 2$. Then K satisfies $\text{UHKP}(d_f, \beta)$ if and only if K satisfies $\text{CSJ}(\beta)$.

As a corollary we show that $\text{HKP}(d_f, \beta)$ is stable under bounded perturbation of the Markov kernel.

Corollary 1.6 (Stability of $\text{HKP}(d_f, \beta)$). Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$ with volume growth exponent d_f . Let K_1 and K_2 be two μ -symmetric Markov operators whose kernels with respect to μ satisfy $\mathbf{J}(\beta)$ for some $\beta \geq 2$. Then K_1 satisfies $\text{HKP}(d_f, \beta)$ if and only if K_2 satisfies $\text{HKP}(d_f, \beta)$.

In Section 3.1, we show how Theorem 1.5 and Corollary 1.6 apply to a large family of graphs that satisfy sub-Gaussian estimates for the simple random walk.

2. Cutoff Sobolev inequalities for heavy-tailed random walks

In this section, we prove the implication $\text{UHKP}(d_f, \beta) \implies \text{CSJ}(\beta)$ in Theorem 1.5. Then we show a self-improving property of $\text{CSJ}(\beta)$.

Proposition 2.1. Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$. Let K be a μ -symmetric Markov operator satisfying $\text{UHKP}(d_f, \beta)$, for some $\beta \geq 2$. Then K satisfies $\text{CSJ}(\beta)$.

We follow the approach of Andres and Barlow in [1, Section 5] to prove the cutoff Sobolev inequality in Proposition 2.1. The first step is to obtain estimates on exit times. We define the exit time for the Markov chain $(Y_n)_{n \in \mathbb{N}}$ as

$$\tau_{B(x,r)}^Y = \min \{k \in \mathbb{N} : Y_k \notin B(x, r)\}.$$

The exit time $\tau_{B(x,r)}^Y$ satisfies the following survival estimate.

Lemma 2.2. *Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$ with volume growth exponent d_f . Let K be μ -symmetric Markov operator satisfying $\text{UHKP}(d_f, \beta)$, for some $\beta \geq 2$. Let $(Y_n)_{n \in \mathbb{N}}$ denote the Markov chain driven by the operator K . There exist constants $\epsilon, \delta \in (0, 1)$ such that for all $x \in X$ and for all $r \geq 1$,*

$$\mathbb{P}_x(\tau_{B(x,r)}^Y \leq \delta r^\beta) \leq \epsilon. \tag{6}$$

Proof. We follow the argument in [10, p. 15]. By $\text{UHKP}(d_f, \beta)$ and $\mathbf{V}(d_f)$, there exists $C_1, C_2 \geq 1$ such that

$$\mathbb{P}_x(d(Y_n, x) \geq r) = \sum_{y \notin B(x,r)} k_n(x, y)\mu(y) \leq C_1 \sum_{y \in B(x,r)^c} \frac{n\mu(y)}{d(x, y)^{d_f+\beta}} \leq C_2 \frac{n}{r^\beta} \tag{7}$$

for all $n \in \mathbb{N}^*$ and for all $x \in X$. By (7) and the strong Markov property of $\{Y_k\}$ at time $\tau = \tau_{B(x,r)}^Y$, there exists $C_3 \geq 1$ such that

$$\begin{aligned} \mathbb{P}_x(\tau \leq n) &\leq \mathbb{P}_x(\tau \leq n, d(Y_{2n}, x) \leq r/2) + \mathbb{P}_x(d(Y_{2n}, x) > r/2) \\ &\leq \mathbb{P}_x(\tau \leq n, d(Y_{2n}, Y_\tau) \geq r/2) + 2^{1+\beta} C_2 n / r^\beta \\ &= \mathbb{E}^x(\mathbf{1}_{\{\tau \leq n\}} \mathbb{P}_{Y_\tau}(d(Y_{2n-\tau}, Y_0) \geq r/2)) + 2^{1+\beta} C_2 n / r^\beta \\ &\leq \sup_{y \in B(x,r)^c} \sup_{s \leq n} \mathbb{P}_y(d(Y_{2n-s}, y) \geq r/2) + 2^{1+\beta} C_2 n / r^\beta \\ &\leq C_3 n / r^\beta \end{aligned}$$

for all $x \in X, k \in \mathbb{N}^*$ and for all $r > 0$. This immediately implies the desired bound (6). □

For $D \subset X$ and a μ -symmetric Markov operator K with kernel k , recall from Definition 1.1 that K_D denotes the sub-Markov operator corresponding to the walk killed upon exiting D . As before, we define the exit time of D as

$$\tau_D = \tau_D^Y = \min \{n \in \mathbb{N} : Y_n \notin D\},$$

where $(Y_n)_{n \in \mathbb{N}}$ is the Markov chain corresponding to the operator K . For $D \subset X, \lambda > 0$, we define corresponding the ‘resolvent operator’ as

$$G_\lambda^D f(x) = \left(I - \frac{K_D}{1+\lambda}\right)^{-1} f(x) = \sum_{i=0}^{\infty} (1+\lambda)^{-i} K_D^i f(x) = \mathbb{E}_x \sum_{i=0}^{\tau_D} (1+\lambda)^{-i} f(Y_i). \tag{8}$$

Lemma 2.3. *Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$ with volume growth exponent d_f . Let K be a Markov operator satisfying $\text{UHKP}(d_f, \beta)$. Let $x_0 \in X, r > 10, R > 0$ and define the annuli $D_0 = B(x_0, R + 9r/10) \setminus B(x_0, R + r/10), D_1 = B(x_0, R + 4r/5) \setminus B(x_0, R + r/5), D_2 = B(x_0, R + 3r/5) \setminus B(x_0, R + 2r/5)$. Let $\lambda = r^{-\beta}$ and set*

$$h = G_\lambda^{D_0} \mathbf{1}_{D_1}, \tag{9}$$

where $G_\lambda^{D_0}$ is as defined in (8). Then h is supported in D_0 and satisfies

$$h(x) \leq 2r^\beta \quad \text{for all } x \in X, \quad (10)$$

$$h(x) \geq c_1 r^\beta \quad \text{for all } x \in D_2. \quad (11)$$

Proof. Since K_{D_0} is a contraction in L^∞ , we have

$$h(x) = \sum_{i=0}^{\infty} (1+\lambda)^{-i} K_{D_0}^i \mathbf{1}_{D_1}(x) \leq \sum_{i=0}^{\infty} (1+\lambda)^{-i} = (1+\lambda)\lambda^{-1} \leq 2r^\beta$$

for all $r \geq 1$, for all $x_0, x \in X$.

Let $(Y_n)_{n \in \mathbb{N}}$ denote the Markov chain driven by K . Let $\epsilon, \delta \in (0, 1)$ be given by Lemma 2.2. Let $r_0 = r/5$, $x \in D_2$ and $B_1 = B(x, r_0) \subset D_1$. Then there exists $c_1 > 0$ such that for all $x_0 \in X$, for all $r > 10$, for all $x \in D_2$, we have

$$\begin{aligned} h(x) &= \sum_{i=0}^{\infty} (1+\lambda)^{-i} K_{D_0}^i \mathbf{1}_{D_1}(x) \geq \sum_{i=0}^{\infty} (1+\lambda)^{-i} K_{B_1}^i \mathbf{1}_{B_1}(x) \\ &\geq (1+r^{-\beta})^{-\delta r^\beta} \mathbb{E}_x \left[\sum_{i=0}^{\lfloor \delta r^\beta \rfloor \wedge \tau_{B_1}^Y} \mathbf{1}_{B_1}(Y_i) \right] \\ &\geq (1+r^{-\beta})^{-\delta r^\beta} \delta r^\beta \mathbb{P}_x(\tau_{B_1}^Y > \delta r^\beta) \geq (1+r^{-\beta})^{-\delta r^\beta} \delta r^\beta (1-\epsilon) \geq c_1 r^\beta. \quad \square \end{aligned}$$

The Dirichlet form corresponding to a μ -symmetric sub-Markov operator P is defined as

$$\mathcal{E}_P(f, g) := \langle f, (I - P)g \rangle$$

for all $f, g \in L^2(X, \mu)$, where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L^2(X, \mu)$.

Lemma 2.4 (Folklore). *Let P be a μ -symmetric sub-Markov operator with kernel p and let \mathcal{E}_P and Γ denote the corresponding Dirichlet form and energy measure respectively.*

(a) *We have*

$$\mathcal{E}_P(f, g) = \sum_{x \in X} \Gamma(f, g)(x) + \sum_{x \in X} f(x)g(x)(1 - P\mathbf{1}(x))\mu(x) \quad (12)$$

for all $f, g \in L^2(X, \mu)$. In particular, if P is a Markov operator we have

$$\mathcal{E}_P(f, g) = \langle f, (I - P)g \rangle = \sum_{x \in X} \Gamma(f, g)(x). \quad (13)$$

(b) *The energy measure Γ satisfies the integrated version of Leibnitz rule*

$$\sum_{x \in X} \Gamma(fg, h)(x) = \sum_{x \in X} [f(x)\Gamma(g, h)(x) + g(x)\Gamma(f, h)(x)] \quad (14)$$

for all bounded functions f, g, h , as long as the above sums are absolutely convergent.

Proof. Note that

$$\begin{aligned} \mathcal{E}_P(f, g) &= \langle f, (I - P)g \rangle = \sum_{x \in X} f(x)g(x)\mu(x) - \sum_{x \in X} f(x)Pg(x)\mu(x) \\ &= \sum_{x \in X} f(x)g(x)\mu(x)(1 - P\mathbf{1}(x) + P\mathbf{1}(x)) - \sum_{x \in X} f(x)Pg(x)\mu(x). \quad (15) \end{aligned}$$

By symmetry $p(x, y) = p(y, x)$, we have

$$\begin{aligned} \sum_{x \in X} \Gamma(f, g)(x) &= \frac{1}{2} \sum_{x, y} (f(x) - f(y))(g(x) - g(y))p(x, y)\mu(x)\mu(y) \\ &= \frac{1}{2} \sum_{x \in X} f(x)g(x)P\mathbf{1}(x)\mu(x) + \frac{1}{2} \sum_{y \in X} f(y)g(y)P\mathbf{1}(y)\mu(y) \\ &\quad - \frac{1}{2} \sum_{x \in X} f(x)Pg(x)\mu(x) - \frac{1}{2} \sum_{y \in X} f(y)Pg(y)\mu(y) \\ &= \sum_{x \in X} f(x)g(x)P\mathbf{1}(x)\mu(x) - \sum_{x \in X} f(x)Pg(x)\mu(x). \end{aligned} \tag{16}$$

Combining (15) and (16), we obtain (12).

To prove (b), we follow [13, Theorem 3.7] and write

$$\begin{aligned} &(f(x)g(x) - f(y)g(y))(h(x) - h(y)) \\ &= \frac{1}{2}(g(x) + g(y))(f(x) - f(y))(h(x) - h(y)) \\ &\quad + \frac{1}{2}(f(x) + f(y))(g(x) - g(y))(h(x) - h(y)). \end{aligned}$$

Then an application of symmetry $p(x, y) = p(y, x)$ similar to (a) yields the desired result (14). □

The following technical lemma is a consequence of Leibnitz rule above.

Lemma 2.5. ([13, Lemma 3.5]) *Let (X, d, μ) be a vertex weighted graph and let P with a μ -symmetric sub-Markov operator with kernel p with respect to μ . Let Γ denote the corresponding energy measure. If $f, h \in L^2(X, \mu)$ and $g \in L^\infty(X, \mu)$, we have*

$$\sum_{x \in X} g(x)\Gamma(f, h)(x) = \frac{1}{2} \sum_{x \in X} [\Gamma(gh, f)(x) + \Gamma(gf, h)(x) - \Gamma(g, fh)(x)].$$

Proof. We use Leibnitz rule (Lemma 2.4(b)) to all terms in the right hand side to obtain the desired equality. The absolute convergence of the various sums are a consequence of Hölder inequality. □

Remark. Using the observation $\Gamma(f_1, f_2)(x) = \Gamma(f_1 + c_1, f_2 + c_2)(x)$, for any $c_1, c_2 \in \mathbb{R}$ and for all $f_1, f_2 \in L^\infty$, we could slightly generalize Lemma 2.5.

Proof of Proposition 2.1. If $r \leq 10$, we may assume $\phi = \mathbf{1}_{B(x, R+r/2)}$. Let $r > 10$ and let h, c_1 be as defined in Lemma 2.3. Set

$$g(y) = \frac{h(y)}{c_1 r^\beta} \quad \text{for all } y \in X, \tag{17}$$

$$\phi(y) = \begin{cases} 1 \wedge g(y) & \text{if } y \in B(x, R + r/2)^c, \\ 1 & \text{if } y \in B(x, R + r/2). \end{cases} \tag{18}$$

By Lemma 2.3, we have (3) and (4). It remains to verify (5). By Leibnitz rule (Lemma 2.4(b), (14)) and the fact that $\Gamma_U(g, g)$ is supported in U , we obtain

$$\sum_{y \in U} f^2(y)\Gamma_U(g, g)(y) = \sum_{y \in X} f^2(y)\Gamma_U(g, g)(y)$$

$$= \sum_{y \in X} \Gamma_U(f^2 g, g)(y) - \sum_{y \in X} g(y) \Gamma_U(f^2, g)(y). \quad (19)$$

As in the Lemma 2.3, we set $\lambda = r^{-\beta}$ and $D_1 = B(x_0, R + 4r/5) \setminus B(x_0, R + r/5)$. Since g is supported in $D_0 = B(x, R + 9r/10) \setminus B(x, R + r/10)$, by (12) of Lemma 2.4(a) and Lemma 2.3, we obtain

$$\begin{aligned} \sum_{y \in X} \Gamma_U(f^2 g, g) &\leq \mathcal{E}_{K_U}(f^2 g, g) \leq \langle f^2 g, (I - K_U)g \rangle_{L^2(U, \mu)} + \lambda \langle f^2 g, g \rangle_{L^2(U, \mu)} \\ &= \langle f^2 g, ((1 + \lambda)I - K_U)g \rangle_{L^2(U, \mu)} \\ &= \langle f^2 g, ((1 + \lambda)I - K_{D_0})g \rangle_{L^2(D_0, \mu)} \\ &= (1 + \lambda)(c_1 r^\beta)^{-1} \langle f^2 g, (I - (1 + \lambda)^{-1} K_{D_0}) G_\lambda^{D_0} \mathbf{1}_{D_1} \rangle_{L^2(D_0, \mu)} \\ &= (1 + \lambda)(c_1 r^\beta)^{-1} \langle f^2 g, \mathbf{1}_{D_1} \rangle_{L^2(D_0, \mu)} \\ &\leq 4c_1^{-2} r^{-\beta} \sum_{y \in D_1} f^2(y) \mu(y). \end{aligned} \quad (20)$$

We use Cauchy–Schwarz inequality, the μ -symmetry of K_U and $ab \leq a^2/4 + b^2$ to obtain

$$\begin{aligned} &\left| \sum_{y \in X} g(y) \Gamma_U(f^2, g)(y) \right| \\ &= \frac{1}{2} \left| \sum_{y, z \in U} g(y)(f^2(y) - f^2(z))(g(y) - g(z))k(y, z)\mu(y)\mu(z) \right| \\ &\leq \frac{1}{2} \left| \sum_{y, z \in U} g(y)f(y)(f(y) - f(z))(g(y) - g(z))k(y, z)\mu(y)\mu(z) \right| \\ &\quad + \frac{1}{2} \left| \sum_{y, z \in U} g(y)f(z)(f(y) - f(z))(g(y) - g(z))k(y, z)\mu(y)\mu(z) \right| \\ &\leq \frac{1}{4} \sum_{y \in U} f^2(y) \Gamma_U(g, g)(y) + \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y) \\ &\quad + \frac{1}{4} \sum_{z \in U} f^2(z) \Gamma_U(g, g)(z) + \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y) \\ &= \frac{1}{2} \sum_{y \in U} f^2(y) \Gamma_U(g, g)(y) + 2 \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y). \end{aligned} \quad (21)$$

Combining the above, we obtain

$$\begin{aligned} \sum_{y \in U} f^2(y) \Gamma(\phi, \phi)(y) &\leq \sum_{y \in U} f^2(y) \Gamma(g, g)(y) \\ &\leq 2 \sum_{y \in X} \Gamma_U(f^2 g, g)(y) + 4 \sum_{y \in U} g^2(y) \Gamma_U(f, f)(y) \\ &\leq 4c_1^{-2} \sum_{y \in U} \Gamma_U(f, f)(y) + 8c_1^{-2} r^{-\beta} \sum_{y \in U} f^2(y) \mu(y). \end{aligned}$$

For the first line above we use $|\phi(y) - \phi(z)| \leq |g(y) - g(z)|$, the second line follows from (19) and (21), and the last line follows from (20) and $g \leq c_1^{-1}$. \square

The “linear cutoff function” for $B(x, r) \subset B(x, 2r)$

$$\psi(y) = \max\left(\min\left(1, \frac{2r - d(x, y)}{r}\right), 0\right) \tag{22}$$

is commonly used to obtain off-diagonal estimates using Davies’ method (see, for example, [10,20]). We will first see how this linear cutoff functions compares to the ones obtained in Proposition 2.1 satisfying inequality (5).

Lemma 2.6. *Let (X, d, μ) be a vertex weighted graph satisfying $V(d_f)$. Let K be a μ -symmetric Markov operator satisfying $J(\beta)$, for some $\beta \geq 2$. Let ψ is the cutoff function in (22) for $B(x, R) \subset B(x, R + r)$ for some $x \in X$ and $r \geq 1$. Then there exists $C_1 > 0$ such that the corresponding energy measure $\Gamma(\psi, \psi)$ satisfies the inequality*

$$\Gamma(\psi, \psi)(y) \leq \begin{cases} \frac{C_1}{r^2} & \text{if } \beta > 2, \\ \frac{C_1 \log(1+r)}{r^2} & \text{if } \beta = 2 \end{cases} \tag{23}$$

for all $y \in X$, where C_1 does not depend on $x \in X, r \geq 1$ or $R > 0$.

Proof. Let k denote the μ -symmetric kernel of K . Note that ψ is $1/r$ -Lipschitz function and $|\psi(y) - \psi(z)| \leq 1$ for all $y, z \in X$. Therefore

$$\begin{aligned} \Gamma(\psi, \psi)(y) &= \sum_{z \in X} (\psi(y) - \psi(z))^2 k(y, z) \mu(z) \mu(y) \\ &\leq \sum_{z \in B(y, r)} r^{-2} d(y, z)^2 k(y, z) \mu(z) \mu(y) + \sum_{z \notin B(y, r)} k(y, z) \mu(z) \mu(y). \end{aligned} \tag{24}$$

To bound the second term above, by $J(\beta)$ and $V(d_f)$ there exists $C_2, C_3 > 0$ such that

$$\begin{aligned} \sum_{z \notin B(y, r)} k(y, z) \mu(z) \mu(y) &\leq \sum_{z \notin B(y, r)} \frac{C_2 \mu(z)}{(d(y, z))^{d_f + \beta}} \\ &= \sum_{i=1}^{\infty} \sum_{z: 2^{i-1}r < d(y, z) \leq 2^i r} \frac{C_2 \mu(z)}{(d(y, z))^{d_f + \beta}} \leq \sum_{i=1}^{\infty} \frac{C_2 V(y, 2^i r)}{(2^{i-1}r)^{d_f + \beta}} \\ &\leq C_3 r^{-\beta} \leq C_3 r^{-2} \end{aligned} \tag{25}$$

for all $y \in X$ and for all $r \geq 1$.

For the first term in (24), by $J(\beta)$ and $V(d_f)$ there exists $C_4, C_5 > 0$ such that

$$\begin{aligned} \sum_{z \in B(y, r)} d(y, z)^2 k(y, z) \mu(z) \mu(y) &\leq \sum_{i=0}^{\lceil \log_2 r \rceil} \sum_{z \in X: 2^i \leq d(y, z) < 2^{i+1}} d(y, z)^2 k(y, z) \mu(z) \mu(y) \\ &\leq C_4 \sum_{i=0}^{\lceil \log_2 r \rceil} 2^{i(2-\beta)} \\ &\leq \begin{cases} C_5 \log(1+r) & \text{if } \beta = 2 \\ C_5 & \text{if } \beta > 2 \end{cases} \end{aligned} \tag{26}$$

for all $y \in X$ and for all $r \geq 1$. Combining (24), (25) and (26), we obtain the desired bound (23). \square

For a μ -symmetric Markov operator K with kernel $k(\cdot, \cdot)$ and for bounded function f , we define the energy measure corresponding to a truncation at scale $L > 0$ as

$$\Gamma_{(L)}(f, f)(x) = \frac{1}{2} \sum_{y \in B(x, L)} (f(x) - f(y))^2 k(x, y) \mu(y) \mu(x). \quad (27)$$

Next, we show a self-improving property of the cutoff-Sobolev inequality that will play an important role in the next section.

Proposition 2.7. *Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$ with volume growth exponent d_f . Let K be μ -symmetric Markov operator whose kernel with respect to μ satisfies $\mathbf{J}(\beta)$ for some $\beta \geq 2$. If K satisfies $\mathbf{CSJ}(\beta)$, then K satisfies the following estimate: there exists $C_1, C_2 > 0$ such that for all $n \in \mathbb{N}^*$, for all $x \in X$, for all $r \geq 1$, there exists a cut-off function ϕ_n for $B(x, r) \subset B(x, 2r)$ such that*

$$\sum_{y \in X} f^2(y) \Gamma(\phi_n, \phi_n)(y) \leq \frac{C_1}{n^2} \sum_{y \in X} \Gamma(f, f)(y) + \frac{C_2 \mathcal{G}_\beta(n)}{r^\beta} \sum_{y \in X} f^2(y) \mu(y), \quad (28)$$

where the function \mathcal{G}_β is given by

$$\mathcal{G}_\beta(n) = \begin{cases} n^{\beta-2} & \text{if } \beta > 2, \\ \log(1+n) & \text{if } \beta = 2. \end{cases} \quad (29)$$

Further the function ϕ_n above satisfies

$$\|\phi_n - \psi\|_\infty \leq n^{-1}, \quad (30)$$

where ψ is the linear cutoff function given by

$$\psi(y) = \max\left(\min\left(1, \frac{2r - d(x, y)}{r}\right), 0\right).$$

Proof. Let Γ denote the energy measure of K . By Lemma 2.6, without loss of generality we may assume $r > 10n$. We divide the annulus $U = B(x, 2r) \setminus B(x, r)$ into n -annuli U_1, U_2, \dots, U_n of equal width, where

$$U_i := B(x, r + ir/n) \setminus B(x, r + (i-1)r/n), \quad i = 1, 2, \dots, n.$$

By Proposition 2.1, there exists $C_3, C_4 > 0$ and cutoff functions $\phi_{(i)}$ for $B(x, r + (i-1)r/n) \subset B(x, r + ir/n)$ satisfying $0 \leq \phi_{(i)} \leq 1$,

$$\phi_{(i)} \equiv 1 \quad \text{in } B(x, r + (i-1)r/n + r/(2n)), \quad (31)$$

$$\phi_{(i)} \equiv 0 \quad \text{in } B(x, r + (i-1)r/n + 9r/(10n))^{\mathbb{C}}, \quad (32)$$

$$\sum_{y \in U_i} f^2(y) \Gamma_{U_i}(\phi_{(i)}, \phi_{(i)})(y) \leq C_3 \sum_{y \in U_i} \Gamma_{U_i}(f, f)(y) + \frac{C_4}{(r/n)^\beta} \sum_{U_i} f^2(y) \mu(y) \quad (33)$$

for $i = 1, 2, \dots, n$. We define $\phi_n = n^{-1} \sum_{i=1}^n \phi_{(i)}$, $s = r/(10n)$ and $\Gamma_{(s)}$ truncated energy measure at scale s as given by (27). Note that ϕ_n satisfies (30) because $\phi_n(y) \in [1 - i/n, 1 - (i-1)/n]$ for all $y \in U_i$ and for all $i = 1, 2, \dots, n$. By (31) and (32), we have

$$\Gamma_{(s)}(\phi_n, \phi_n)(y) = n^{-2} \sum_{i=1}^n \Gamma_{(s)}(\phi_{(i)}, \phi_{(i)})(y) \leq n^{-2} \sum_{i=1}^n \Gamma_{U_i}(\phi_{(i)}, \phi_{(i)})(y) \quad (34)$$

for all $y \in X$. Combining (33) and (34), we obtain

$$\sum_{y \in X} f^2(y) \Gamma_{(s)}(\phi_n, \phi_n)(y) \leq C_3 n^{-2} \sum_{y \in X} \Gamma(f, f)(y) + \frac{C_4 n^{\beta-2}}{r^\beta} \sum_{y \in X} f^2(y) \mu(y). \quad (35)$$

To bound $\Gamma(\phi_n, \phi_n)(y) - \Gamma_{(s)}(\phi_n, \phi_n)(y)$, we write

$$\Gamma(\phi_n, \phi_n) - \Gamma_{(s)}(\phi_n, \phi_n) = \Gamma(\phi_n, \phi_n) - \Gamma_{(r)}(\phi_n, \phi_n) + \Gamma_{(r)}(\phi_n, \phi_n) - \Gamma_{(s)}(\phi_n, \phi_n). \quad (36)$$

Since $0 \leq \phi_n \leq 1$ by (25) there exists $C_5 > 0$ such that

$$\Gamma(\phi_n, \phi_n)(y) - \Gamma_{(r)}(\phi_n, \phi_n)(y) \leq \sum_{y \in B(x, r)^c} k(x, y) \mu(x) \mu(y) \leq C_5 r^{-\beta} \quad (37)$$

for all $r > 0$ and $n \in \mathbb{N}^*$. Note that for all $y, z \in X$ such that $d(y, z) \geq s = r/(10n)$, by (30) we have $|\phi_n(y) - \phi_n(z)| \leq r^{-1}d(y, z) + 2/n \leq 21r^{-1}d(y, z)$. Therefore by a similar calculation as (26), there exists $C_6 > 0$ such that, for all $y \in X$, for all $n \in \mathbb{N}^*$ and for all $r > 10n$, we have

$$\begin{aligned} & \Gamma_{(r)}(\phi_n, \phi_n)(y) - \Gamma_{(s)}(\phi_n, \phi_n)(y) \\ &= \frac{1}{2} \sum_{z \in B(y, r) \setminus B(y, s)} (\phi_n(y) - \phi_n(z))^2 k(y, z) \mu(z) \mu(y) \\ &\leq \frac{21^2}{2} \sum_{z \in B(y, r) \setminus B(y, s)} r^{-2} d(y, z)^2 k(y, z) \mu(z) \mu(y) \\ &\leq \frac{21^2}{2} \sum_{i=0}^{\lfloor \log_2(10n) \rfloor} \sum_{z \in B(y, 2^{i+1}s) \setminus B(y, 2^i s)} r^{-2} d(y, z)^2 k(y, z) \mu(z) \mu(y) \\ &\leq C_6 \frac{\mathcal{G}_\beta(n)}{r^\beta}. \end{aligned} \quad (38)$$

Combining (35), (36), (37) and (38), we obtain

$$\sum_{y \in X} f^2(y) \Gamma(\phi_n, \phi_n)(y) \leq \frac{C_3}{n^2} \mathcal{E}_K(f, f) + (C_4 + C_6 + C_7) \frac{\mathcal{G}_\beta(n)}{r^\beta} \sum_{y \in X} f^2(y) \mu(y) \quad (39)$$

for all $x \in X$, for all $n \in \mathbb{N}^*$ and for all $r > 10n$. Combining (36) with Lemma 2.6 yields the desired result. \square

3. Davies' method

In this section, we carry out the Davies perturbation method to obtain heat kernel upper bounds $\text{UHKP}(d_f, \beta)$ for heavy tailed walks satisfying $\mathbf{J}(\beta)$ and the cutoff Sobolev inequality $\text{CSJ}(\beta)$, for some $\beta \geq 2$. For the case the $\beta \in (0, 2)$, in [10] the heat kernel upper bounds for the corresponding continuous time process was obtained. The corresponding discrete time bounds were obtained in [35].

The idea behind the approach of [10] is to use Meyer's construction [33] to split the jump kernel into small and large sums and then apply Davies' method for the smaller jumps (see [10, Section 3]). However as mentioned in the introduction (see [24, Section 1]), the above approach is no longer adequate to obtain $\text{UHKP}(d_f, \beta)$ for the case $\beta \geq 2$. The goal of this section is to modify Davies' perturbation method to obtain upper bounds for heavy tailed jump processes satisfying $\mathbf{J}(\beta)$ for the case $\beta \geq 2$. The following Proposition is the converse of Proposition 2.7.

Proposition 3.1. *Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$ with volume growth exponent d_f . Let K be μ -symmetric Markov operator whose kernel with respect to μ satisfies $\mathbf{J}(\beta)$ for some $\beta \geq 2$. If K satisfies the cutoff Sobolev inequality $\text{CSJ}(\beta)$, then K satisfies the transition probability upper bounds $\text{UHKP}(d_f, \beta)$.*

We introduce two definitions below.

Definition 3.2. We say that a μ -symmetric sub-Markov operator T on a graph (X, d) is L -local if its corresponding kernel t satisfies $t(x, y) = 0$ whenever $x, y \in X$ satisfies $d(x, y) > L$.

Definition 3.3. Let ψ be a function on a graph (X, d) and let $L > 0$. We define the oscillation of ψ at scale L as

$$\text{osc}(\psi, L) := \sup_{x, y \in X: d(x, y) \leq L} |\psi(x) - \psi(y)|.$$

The following Lemma is an analogue of [13, Theorem 3.9]. The computations are similar to the ones in [13] but we will use a different strategy to control the energy measure at various places.

Lemma 3.4. Let T be a μ -symmetric, L -local, sub-Markov operator on a vertex weighted graph (X, d, μ) and let Γ denote the corresponding energy measure. Then for any function $\psi \in L^\infty(X, \mu)$ with bounded support, for all $p \geq 1$ and for all $f \in L^2(X, \mu)$ with $f \geq 0$, we have

$$\begin{aligned} & \sum_{x \in X} \Gamma(e^{-\psi} f, e^\psi f^{2p-1})(x) \\ & \geq \frac{1}{2p} \sum_{x \in X} \Gamma(f^p, f^p)(x) - 9pe^{2\text{osc}(\psi, L)} \sum_{x \in X} f^{2p}(x) \Gamma(\psi, \psi)(x). \end{aligned} \quad (40)$$

Proof. Let t denote the kernel of T with respect to μ . Then by Lemma 2.5, we have

$$\begin{aligned} \sum_{x \in X} \Gamma(e^{-\psi} f, e^\psi f^{2p-1})(x) &= \sum_{x \in X} (\Gamma(e^{-\psi} f^{2p}, e^\psi)(x) + \Gamma(f, f^{2p-1})(x)) \\ &\quad - 2 \sum_{x \in X} e^{-\psi(x)} f(x) \Gamma(e^\psi, f^{2p-1})(x). \end{aligned} \quad (41)$$

A diligent reader will observe that one cannot directly apply Lemma 2.5 because $e^\psi \notin L^2(X, \mu)$. However, since ψ has bounded support, $e^\psi - 1 \in L^2(X, \mu)$ and we can apply Lemma 2.5 with f, g, h replaced by $e^\psi - 1, e^{-\psi} f, f^{2p-1}$ respectively. Then we use the remark following Lemma 2.5 to obtain (41).

We use Leibniz rule (Lemma 2.4(b)) for the first term in the right in (41) to obtain

$$\begin{aligned} \sum_{x \in X} \Gamma(e^{-\psi} f, e^\psi f^{2p-1})(x) &= \sum_{x \in X} f^{2p-1}(x) \Gamma(e^{-\psi} f, e^\psi)(x) \\ &\quad + \sum_{x \in X} (\Gamma(f, f^{2p-1})(x) - e^{-\psi(x)} f(x) \Gamma(e^\psi, f^{2p-1})(x)). \end{aligned} \quad (42)$$

Note that

$$\begin{aligned} & \sum_{x \in X} (f^{2p-1}(x) \Gamma(e^{-\psi} f, e^\psi)(x) - e^{-\psi(x)} f(x) \Gamma(e^\psi, f^{2p-1})(x)) \\ &= \frac{1}{2} \sum_{x, y \in X} \{ (e^{-\psi(x)} f(x) f^{2p-1}(y) - e^{-\psi(y)} f(y) f^{2p-1}(x)) \\ &\quad \times (e^{\psi(x)} - e^{\psi(y)}) t(x, y) \mu(x) \mu(y) \}. \end{aligned} \quad (43)$$

We define $a(x, y) := t(x, y) \mu(x) \mu(y)$ for all $x, y \in X$. Next, we rewrite the right side of (43) as

$$\frac{1}{2} \sum_{x, y \in X} (f^{2p}(y) - f^{2p}(x)) e^{-\psi(x)} (e^{\psi(x)} - e^{\psi(y)}) a(x, y)$$

$$\begin{aligned}
 & + \frac{1}{2} \sum_{x,y \in X} f^{2p}(x)(e^{-\psi(x)} - e^{-\psi(y)})(e^{\psi(x)} - e^{\psi(y)})a(x, y) \\
 & + \sum_{x,y \in X} f^{2p-1}(y)(f(x) - f(y))e^{\psi(y)}(e^{-\psi(y)} - e^{-\psi(x)})a(x, y).
 \end{aligned} \tag{44}$$

For the first term in (44), we use Cauchy–Schwartz inequality to obtain

$$\begin{aligned}
 & \frac{1}{2} \sum_{x,y \in X} (f^{2p}(y) - f^{2p}(x))e^{-\psi(x)}(e^{\psi(x)} - e^{\psi(y)})a(x, y) \\
 & = \frac{1}{2} \sum_{x,y \in X} f^p(y)(f^p(y) - f^p(x))e^{\psi(y)}(e^{-\psi(y)} - e^{-\psi(x)})a(x, y) \\
 & + \frac{1}{2} \sum_{x,y \in X} f^p(x)(f^p(y) - f^p(x))e^{-\psi(x)}(e^{\psi(x)} - e^{\psi(y)})a(x, y) \\
 & \geq - \left(\sum_{x \in X} \Gamma(f^p, f^p)(x) \right)^{1/2} \left[\left(\sum_{y \in X} f^{2p}(y)e^{2\psi(y)}\Gamma(e^{-\psi}, e^{-\psi})(y) \right)^{1/2} \right. \\
 & \quad \left. + \left(\sum_{x \in X} f^{2p}(x)e^{-2\psi(x)}\Gamma(e^{\psi}, e^{\psi})(x) \right)^{1/2} \right].
 \end{aligned} \tag{45}$$

For the second term in (44), we use Cauchy–Schwartz inequality to obtain

$$\begin{aligned}
 & \frac{1}{2} \sum_{x,y \in X} f^{2p}(x)(e^{-\psi(x)} - e^{-\psi(y)})(e^{\psi(x)} - e^{\psi(y)})a(x, y) \\
 & = \sum_{x \in X} f^{2p}(x)\Gamma(e^{-\psi}, e^{\psi})(x) \\
 & \geq - \left(\sum_{x \in X} f^{2p}(x)e^{2\psi(x)}\Gamma(e^{-\psi}, e^{-\psi})(x) \right)^{1/2} \left(\sum_{x \in X} f^{2p}(x)e^{-2\psi(x)}\Gamma(e^{\psi}, e^{\psi})(x) \right)^{1/2}.
 \end{aligned} \tag{46}$$

For the last term in (44), we use Cauchy–Schwartz inequality to obtain

$$\begin{aligned}
 & \sum_{x,y \in X} f^{2p-1}(y)(f(x) - f(y))e^{\psi(y)}(e^{-\psi(y)} - e^{-\psi(x)})a(x, y) \\
 & \geq -2 \left(\sum_{x \in X} f^{2p-2}(x)\Gamma(f, f)(x) \right)^{1/2} \left(\sum_{x \in X} f^{2p}(x)e^{2\psi(x)}\Gamma(e^{-\psi}, e^{-\psi})(x) \right)^{1/2}.
 \end{aligned} \tag{47}$$

We need two more elementary inequalities (Cf. (3.16) and (3.17) in [13]). For any non-negative function $f \in L^2(X, \mu)$ and for all $p \geq 1$, we have

$$\sum_{x \in X} \Gamma(f^{2p-1}, f)(x) \geq \sum_{x \in X} f^{2p-2}(x)\Gamma(f, f)(x) \geq \frac{1}{2p-1} \sum_{x \in X} \Gamma(f^{2p-1}, f)(x) \tag{48}$$

and

$$\sum_{x \in X} \Gamma(f^p, f^p)(x) \geq \sum_{x \in X} \Gamma(f^{2p-1}, f)(x) \geq \frac{2p-1}{p^2} \sum_{x \in X} \Gamma(f^p, f^p)(x). \tag{49}$$

Since T is L -local, we use the inequality $(e^a - 1)^2 \leq e^{2|a|}a^2$ to obtain

$$\max(e^{2\psi(x)}\Gamma(e^{-\psi}, e^{-\psi})(x), e^{-2\psi(x)}\Gamma(e^\psi, e^\psi)(x)) \leq e^{2\text{osc}(\psi, L)}\Gamma(\psi, \psi)(x) \tag{50}$$

for all $x \in X$. Combining equations (42) through (50), we have

$$\begin{aligned} & \sum_{x \in X} \Gamma(e^{-\psi} f, e^\psi f^{2p-1})(x) \\ & \geq \frac{2p-1}{p^2} \sum_{x \in X} \Gamma(f^p, f^p)(x) - e^{2\text{osc}(\psi, L)} \sum_{x \in X} f^{2p}(x)\Gamma(\psi, \psi)(x) \\ & \quad - 4e^{\text{osc}(\psi, L)} \left(\sum_{x \in X} \Gamma(f^p, f^p)(x) \right)^{1/2} \left(\sum_{x \in X} f^{2p}(x)\Gamma(\psi, \psi)(x) \right)^{1/2}. \end{aligned}$$

We use the inequality $4ab \leq a^2/(2p) + 8pb^2$ and $p \geq 1$ to obtain

$$\begin{aligned} & \sum_{x \in X} \Gamma(e^{-\psi} f, e^\psi f^{2p-1})(x) \\ & \geq \frac{1}{2p} \sum_{x \in X} \Gamma(f^p, f^p)(x) - 9pe^{2\text{osc}(\psi, L)} \sum_{x \in X} f^{2p}(x)\Gamma(\psi, \psi)(x) \end{aligned}$$

for all bounded functions ψ and for all non-negative functions $f \in L^2(X, \mu)$. □

Remark. The elementary inequalities (48) and (49) are sometimes called Stroock–Varopolous inequalities. An unified approach to such inequalities is provided in [34, Lemma 2.4].

The next step is to bound the second term in the right side of (40) using cutoff Sobolev inequalities developed in Section 2.

Lemma 3.5. *Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$ with volume growth exponent d_f . Let K be a μ -symmetric Markov operator whose kernel $k = k_1$ satisfies $\mathbf{J}(\beta)$ for some $\beta \geq 2$. Further assume that K satisfies the cutoff Sobolev inequality $\text{CSJ}(\beta)$. Let Γ denote the energy measure corresponding to K and let $\Gamma_{(L)}$ correspond to the truncated version of Γ for some $L > 0$. Define*

$$\vartheta := \frac{1}{8(d_f + \beta)}.$$

Then there exists $\lambda_0 \geq 1, C_0 \geq 1$ such that the following property holds: For all $\lambda \geq \lambda_0$, for all $p \geq 1$, for all $x \in X$, for all $r \geq 1$, there exists a cut-off function $\phi = \phi_{p,\lambda}$ for $B(x, r) \subset B(x, 2r)$ such that for all non-negative functions $f \in L^2(X, \mu)$ we have

$$\sum_{y \in X} \Gamma_{(\vartheta r)}(e^{-\lambda\phi} f, e^{\lambda\phi} f^{2p-1}) \geq \frac{1}{4p} \sum_{y \in X} \Gamma(f^p, f^p)(y) - \frac{C_0 p^\beta e^{8\beta\vartheta\lambda}}{r^\beta} \sum_{y \in X} f^{2p}(y)\mu(y), \tag{51}$$

where the cutoff function ϕ above satisfies

$$\|\phi - \psi\|_\infty \leq \frac{1}{\lambda p}, \tag{52}$$

where ψ is given by (22).

Proof. By Lemma 3.4, for any cutoff function ϕ for $B(x, r) \subset B(x, 2r)$, for all $p \geq 1$, for all $\lambda > 0$, and for all non-negative $f \in L^2(X, \mu)$ we have

$$\begin{aligned} & \sum_{y \in X} \Gamma_{(\vartheta r)}(e^{-\lambda \phi} f, e^{\lambda \phi} f^{2p-1}) \\ & \geq \frac{1}{2p} \sum_{y \in X} \Gamma_{(\vartheta r)}(f^p, f^p)(y) - 9p\lambda^2 e^{2\lambda \operatorname{osc}(\phi, \vartheta r)} \sum_{y \in X} f^{2p}(y) \Gamma(\phi, \phi)(y). \end{aligned} \tag{53}$$

Using $J(\beta)$ and a similar computation as [35, Equation (33)], there exists $C_2 > 0$ such that

$$\sum_{y \in X} (\Gamma(f^p, f^p)(y) - \Gamma_{(\vartheta r)}(f^p, f^p)(y)) \leq C_2 r^{-\beta} \sum_{y \in X} f^{2p}(y) \mu(y) \tag{54}$$

for all $r \geq 1$, for all $p \geq 1$ and for all $f \in L^2(X, \mu)$. By CSJ(β), there exists $C_3, C_4 > 1$ such that for any $n \in \mathbb{N}^*$ there exists a cutoff function $\phi = \phi_n$ for $B(x, r) \subset B(x, 2r)$ such that for all $f \in L^2(X, \mu)$ with $f \geq 0$, we have

$$\sum_{y \in X} f^2(y) \Gamma(\phi, \phi)(y) \leq \frac{C_3}{n^2} \sum_{y \in X} \Gamma(f, f)(y) + \frac{C_4 \mathcal{G}_\beta(n)}{r^\beta} \sum_{y \in X} f^2(y) \mu(y), \tag{55}$$

where the function $\phi = \phi_n$ above satisfies

$$\|\phi - \psi\|_\infty \leq n^{-1}. \tag{56}$$

We make the choice

$$n = \lceil 6p\lambda \exp(3\lambda\vartheta) \sqrt{C_3} \rceil \tag{57}$$

and $\lambda_0 \geq 1$ such that $n \geq \vartheta^{-1}$ for all $\lambda \geq \lambda_0$ and for all $p \geq 1$. We will verify ϕ satisfies the desired properties (51) and (52). Using (57) and (56), we immediately have (52). By (56), triangle inequality, $n \geq \vartheta^{-1}$ and by the fact that ψ is r^{-1} -Lipschitz, we have

$$\operatorname{osc}(\phi, \vartheta r) \leq \operatorname{osc}(\psi, \vartheta r) + 2n^{-1} \leq \vartheta + 2\vartheta = 3\vartheta. \tag{58}$$

By (57), (55), (58), we have

$$\begin{aligned} & 9p\lambda^2 e^{2\lambda \operatorname{osc}(\phi, \vartheta r)} \sum_{y \in X} f^{2p}(y) \Gamma(\phi, \phi)(y) \\ & \leq \frac{1}{4p} \sum_{y \in X} \Gamma(f^p, f^p)(y) + \frac{9C_4 p \lambda^2 e^{6\vartheta\lambda} \mathcal{G}_\beta(n)}{r^\beta} \sum_{y \in X} f^{2p}(y) \mu(y), \end{aligned} \tag{59}$$

for all non-negative $f \in L^2(X, \mu)$. We use $n \leq 12p\lambda e^{3\lambda\vartheta} \sqrt{C_3}$, $\lambda \leq \vartheta^{-1} e^{\vartheta\lambda}$ and $\mathcal{G}_\beta(n) \leq n^{\beta-1}$ for all $n \in \mathbb{N}^*$ to obtain $C_5 > 0$ (that depends on ϑ, β) such that

$$9C_4 p \lambda^2 e^{6\vartheta\lambda} \mathcal{G}_\beta(n) \leq 9C_4 p \lambda^2 e^{6\vartheta\lambda} (12p\lambda e^{3\lambda\vartheta} \sqrt{C_3})^{\beta-1} \leq C_5 p^\beta e^{4(\beta+1)\lambda\vartheta} \leq C_5 p^\beta e^{8\beta\lambda\vartheta} \tag{60}$$

for all $p, n \geq 1$ and for all $\lambda > \lambda_0$. Combining (53), (54), (59) and (60), we obtain (51). □

We need the following Nash inequality to obtain off-diagonal estimates using Davies' method.

Proposition 3.6 (Nash inequality). *Let (X, d, μ) be a vertex weighted graph satisfying satisfies $V(d_f)$ with volume growth exponent d_f . Let K be a μ -symmetric Markov operator whose kernel k satisfies $J(\beta)$ for some $\beta > 0$. Let $\mathcal{E}(\cdot, \cdot)$ denote the corresponding Dirichlet form. Then there exists $C_N > 0$ such that*

$$\|f\|_2^{2(1+\beta/d_f)} \leq C_N \mathcal{E}(f, f) \|f\|_1^{2\beta/d_f} \tag{61}$$

for all $f \in L^1(X, \mu)$, where $\|\cdot\|_p$ denotes the $L^p(X, \mu)$ norm.

Proof. The proof of the Nash inequality (61) is essentially contained in [2, p. 1064]. We repeat the proof for convenience.

For $r > 0$ and $f \in L^1(X, \mu)$, we denote by $f_r : X \rightarrow \mathbb{R}$ the μ -weighted average

$$f_r(x) = \frac{1}{V(x, r)} \sum_{y \in B(x, r)} f(y) \mu(y).$$

We first bound $\|f_r\|_1$. There exists $C_V > 0$ such that

$$\begin{aligned} \|f_r\|_1 &= \sum_{x \in X} |f_r(x)| \mu(x) \\ &\leq \sum_{x \in X} \sum_{y \in B(x, r)} |f(y)| V(x, r)^{-1} \mu(y) \mu(x) = \sum_{y \in X} |f(y)| \mu(y) \sum_{x \in B(y, r)} V(x, r)^{-1} \mu(x) \\ &\leq C_V^2 \|f\|_1 \end{aligned} \tag{62}$$

for all $f \in L^1(X, \mu)$ and for all $r > 0$. The second line above follows from triangle inequality and Fubini's theorem and the last line follows from $V(d_f)$. By $V(d_f)$, there exists $C_1 > 0$ such that

$$\|f_r\|_\infty \leq \|f\|_1 \sup_{x \in X} V(x, r)^{-1} \leq C_1 r^{-d_f} \|f\|_1 \tag{63}$$

$f \in L^1(X, \mu)$ and for all $r > 0$. Combining (62) and (63), there exists $C_2 > 0$ such that

$$\|f_r\|_2 \leq \|f_r\|_1^{1/2} \|f_r\|_\infty^{1/2} \leq C_2 r^{-d_f/2} \|f\|_1 \tag{64}$$

for all $f \in L^1(X, \mu)$ and for all $r > 0$.

There exists $C_3 > 0$ such that for all $f \in L^1(X, \mu)$ and for all $r > 0$, we have

$$\begin{aligned} \|f - f_r\|_2^2 &= \sum_{x \in X} |f(x) - f_r(x)|^2 \mu(x) \\ &\leq \sum_{x \in X} \frac{1}{V(x, r)} \sum_{y \in B(x, r)} (f(x) - f(y))^2 \mu(y) \mu(x) \\ &\leq C_1 \sum_{x, y \in X} (f(x) - f(y))^2 \mathbf{1}_{\{d(x, y) \leq r\}} r^{-d_f} \mu(x) \mu(y) \\ &\leq 2^{d_f + \beta} C_1 r^\beta \sum_{x, y \in X} (f(x) - f(y))^2 \frac{1}{(1 + d(x, y))^{d_f + \beta}} \mu(x) \mu(y) \\ &\leq \frac{C_3 r^\beta}{2} \sum_{x, y \in X} (f(x) - f(y))^2 k(x, y) \mu(x) \mu(y) = C_3 r^\beta \mathcal{E}(f, f). \end{aligned} \tag{65}$$

The second line above follows from Jensen's inequality, the third line follows from an application of $V(d_f)$ similar to (63) and the last line follows from $J(\beta)$ and Lemma 2.4(a).

By triangle inequality, (64) and (65), we have

$$\|f\|_2 \leq \|f_r\|_2 + \|f - f_r\|_2 \leq C_2 r^{-d_f/2} \|f\|_1 + C_3^{1/2} r^{\beta/2} \mathcal{E}(f, f)^{1/2} \tag{66}$$

for all $f \in L^1(X, \mu)$ and all $r > 0$. The Nash inequality (61) follows from (66) and the choice

$$r = \left(\frac{\|f\|_1^2}{\mathcal{E}(f, f)} \right)^{1/(d_f + \beta)}. \quad \square$$

The following lemma is analogous to [13, Lemma 3.21] but the statement and its proof is slightly modified to suit our context.

Lemma 3.7. *Let $w : [0, \infty) \rightarrow (0, \infty)$ be a non-decreasing function and suppose that $u \in C^1([0, \infty); (0, \infty))$ satisfies*

$$u'(t) \leq -\frac{\epsilon}{p} \left(\frac{t^{(p-2)/\theta p}}{w(t)} \right)^{\theta p} u^{1+\theta p}(t) + \delta p^\beta u(t) \quad (67)$$

for some positive ϵ, θ and $\delta, \beta \in [2, \infty)$ and $p = 2^k$ for some $k \in \mathbb{N}^*$. Then u satisfies

$$u(t) \leq \left(\frac{2p^{\beta+1}}{\epsilon\theta} \right)^{1/\theta p} t^{(1-p)/\theta p} w(t) e^{\delta t/p}. \quad (68)$$

Proof. Set $v(t) = e^{-\delta p^\beta t} u(t)$. By (67), we have

$$v'(t) = e^{-\delta p^\beta t} (u'(t) - \delta p^\beta u(t)) \leq -\frac{\epsilon t^{p-2}}{p w(t)^{\theta p}} e^{\theta \delta p^{\beta+1} t} v(t)^{1+\theta p}.$$

Hence

$$\frac{d}{dt} (v(t))^{-\theta p} \geq \epsilon \theta t^{p-2} w(t)^{-\theta p} e^{\theta \delta p^{\beta+1} t}$$

and so, since w is non-decreasing

$$e^{\delta \theta p^{\beta+1} t} u(t)^{-\theta p} \geq \epsilon \theta w(t)^{-\theta p} \int_0^t s^{(p-2)} e^{\theta \delta p^{\beta+1} s} ds. \quad (69)$$

Note that

$$\begin{aligned} \int_0^t s^{(p-2)} e^{\theta \delta p^{\beta+1} s} ds &\geq (t/\delta \theta p^{\beta+1})^{p-1} \int_{\delta \theta p^{\beta+1} (1-1/p^{\beta+1})}^{\delta \theta p^{\beta+1}} y^{(p-2)} e^{ty} dy \\ &\geq \frac{t^{p-1}}{p-1} \exp(\delta \theta p^{\beta+1} t - \delta \theta t) [1 - (1 - p^{-\beta-1})^{p-1}] \\ &\geq \frac{t^{p-1}}{2p^{\beta+1}} \exp(\delta \theta p^{\beta+1} t - \delta \theta t). \end{aligned} \quad (70)$$

In the last line above, we used the bound $1 - (1 - p^{-\beta-1})^{p-1} \geq \frac{1}{2} p^{-\beta-1} (p-1)$ for all $p, \beta \geq 2$. Combining (69) and (70) yields (68). \square

We now have all the necessary ingredients to prove Proposition 3.1.

Proof of Proposition 3.1. Let $H_t := \exp(t(K - I))$ denote the corresponding continuous time semigroup. Note that H_t is μ -symmetric and let h_t be the kernel of H_t with respect to μ . As explained in [35, Remark 3], by [21, Theorem 3.6] it suffices to show the following bound on h_t : there exists $C_1 > 0$ such that

$$h_t(x, y) \leq C_1 \min \left(\frac{1}{t^{d_f/\beta}}, \frac{t}{(d(x, y))^{d_f + \beta}} \right) \quad (71)$$

for all $x, y \in X$ and for all $t \geq 1$. By [19, Proposition II.1] the Nash inequality (61) in Proposition 3.6 implies that there exists $C_2 > 0$ such that

$$h_t(x, y) \leq \frac{C_2}{t^{d_f/\beta}} \quad (72)$$

for all $x, y \in X$ and for all $t \geq 1$.

Let $t \geq 1$ and let $x, y \in X$ such that $d(x, y) \geq 2$. Let $\vartheta = 1/(8(d_f + \beta))$ and $\lambda_0, C_0 \geq 1$ be constants given by Lemma 3.5. We define parameters L and r that depend only on x and y as

$$r = d(x, y)/2, \quad L = \vartheta r. \quad (73)$$

Let $\lambda \geq \lambda_0$ be arbitrary. Let Γ denote the energy measure corresponding to K and let $\Gamma_{(L)}$ denote its truncation. Define the truncated Dirichlet form

$$\mathcal{E}_L(f, f) = \sum_{y \in X} \Gamma_{(L)}(f, f)(y)$$

for all $f \in L^2(X, \mu)$. Let $H_t^{(L)}$ denote the continuous time kernel with respect to μ for the corresponding jump process and let $H_t^{(L)}$ denote the corresponding Markov semigroup. Let $p_k = 2^k$ for $k \in \mathbb{N}$ and let $\psi_k = \lambda \phi_{p_k, \lambda}$ where $\phi_{p_k, \lambda}$ is the cutoff function for $B(x, r) \subset B(x, 2r)$ given by Lemma 3.5. Define the ‘perturbed semigroup’

$$H_t^{L, \psi_k} f(x) = e^{\psi_k(x)} (H_t^{(L)}(e^{-\psi_k} f))(x). \quad (74)$$

We now pick $f \in L^2(X, \mu)$ and $f \geq 0$ with $\|f\|_2 = 1$ and define

$$f_{t,k} := H_t^{L, \psi_k} f \quad (75)$$

for all $k \in \mathbb{N}$, and $P_t^{\psi_k}$ denotes the perturbed semigroup defined in (74). We remark that the constants below do not depend on the choice of $x, y \in X$ with $d(x, y) \geq 2$ or the choice of $f \in L^2(X, \mu)$, $t \geq 1$ or $\lambda \geq \lambda_0$.

Using Lemma 3.5, we have

$$\begin{aligned} \frac{d}{dt} \|f_{t,0}\|_2^2 &= -2\mathcal{E}_L(e^{\psi_0} f_{t,0}, e^{-\psi_0} f_{t,0}) \\ &\leq 2C_0 \frac{e^{8\beta\vartheta\lambda}}{r^\beta} \|f_{t,0}\|_2^2 \end{aligned} \quad (76)$$

and

$$\begin{aligned} \frac{d}{dt} \|f_{t,k}\|_{2p_k}^{2p_k} &= -2p_k \mathcal{E}_L(e^{\psi_k} f_{t,k}^{2p_k-1}, e^{-\psi_k} f_{t,k}) \\ &\leq -\frac{1}{2} \mathcal{E}(f_{t,k}^{p_k}, f_{t,k}^{p_k}) + 2C_0 \frac{e^{8\beta\vartheta\lambda} p_k^{\beta+1}}{r^\beta} \|f_{t,k}\|_{2p_k}^{2p_k} \end{aligned} \quad (77)$$

for all $k \in \mathbb{N}^*$. By solving (76), we obtain

$$\|f_{t,0}\|_{p_1} = \|f_{t,0}\|_2 \leq \exp(C_0 e^{8\beta\vartheta\lambda} t / r^\beta) \|f\|_2 = \exp(C_0 e^{8\beta\vartheta\lambda} t / r^\beta). \quad (78)$$

Using (77) and Nash inequality (61), we obtain

$$\frac{d}{dt} \|f_{t,k}\|_{2p_k} \leq -\frac{1}{4C_N p_k} \|f_{t,k}\|_{2p_k}^{1+2\beta p_k/d_f} \|f_{t,k}\|_{p_k}^{-2\beta p_k/d_f} + C_0 p_k^\beta \frac{e^{8\beta\vartheta\lambda}}{r^\beta} \|f_{t,k}\|_{2p_k} \quad (79)$$

for all $k \in \mathbb{N}^*$. By (52), we have $\|\psi_k - \psi_{k-1}\|_\infty \leq 3/p_k$. This along the fact that $H_t^{(L)}$ is a contraction on L^∞ yields

$$\exp(-6/p_k) f_{t,k-1} \leq f_{t,k} \leq \exp(6/p_k) f_{t,k-1} \quad (80)$$

for all $k \in \mathbb{N}^*$. Combining (79) and (80), we obtain

$$\frac{d}{dt} \|f_{t,k}\|_{2p_k} \leq -\frac{1}{C_A p_k} \|f_{t,k}\|_{2p_k}^{1+2\beta p_k/d_f} \|f_{t,k-1}\|_{p_k}^{-2\beta p_k/d_f} + C_0 p_k^\beta \frac{e^{8\beta\vartheta\lambda}}{r^\beta} \|f_{t,k}\|_{2p_k} \tag{81}$$

for all $k \in \mathbb{N}^*$, where $C_A = 4C_N \exp(12\beta/d_f)$.

Let $u_k(t) = \|f_{t,k-1}\|_{p_k}$ and let

$$u_k(t) = \sup\{s^{d_f(p_k-2)/(2\beta p_k)} u_k(s) : s \in (0, t]\}.$$

By (78), $w_1(t) \leq \exp(C_0 e^{8\beta\vartheta\lambda} t/r^\beta)$. Further by (81), u_{k+1} satisfies (67) with $\epsilon = 1/C_A$, $\theta = 2\beta/d_f$, $\delta = C_0(e^{8\beta\vartheta\lambda}/r^\beta)$, $w = w_k$, $p = p_k$. Hence by Lemma 3.7,

$$u_{k+1}(t) \leq (2^{(\beta+1)k+1}/\epsilon\theta)^{1/(\theta p_k)} t^{(1-p_k)/\theta p_k} e^{\delta t/p_k} w_k(t).$$

Therefore

$$w_{k+1}(t)/w_k(t) \leq (2^{(\beta+1)k+1}/\epsilon\theta)^{1/(\theta 2^k)} e^{\delta t/2^k}$$

for $k \in \mathbb{N}^*$. Hence, we obtain

$$\lim_{k \rightarrow \infty} w_k(t) \leq C_3 e^{\delta t} w_1(t) \leq C_2 \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta),$$

where $C_3 = C_3(\beta, \epsilon, \theta) > 0$. Hence

$$\lim_{k \rightarrow \infty} u_k(t) = \|H_t^{L, \psi_\infty} f\|_\infty \leq \frac{C_2}{t^{d_f/2\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta),$$

where $\psi_\infty = \lim_{k \rightarrow \infty} \psi_k$. Since the above bound holds for all non-negative $f \in L^2(X, \mu)$, we obtain

$$\|H_t^{L, \psi_\infty}\|_{2 \rightarrow \infty} \leq \frac{C_3}{t^{d_f/2\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta).$$

The estimate is unchanged if we replace ψ_k 's by $-\psi_k$. Since $H_t^{L, -\psi}$ is the adjoint of $H_t^{L, \psi}$, by duality we have that for all $t > 0$

$$\|H_t^{L, \psi_\infty}\|_{1 \rightarrow 2} \leq \frac{C_3}{t^{d_f/2\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta).$$

Combining the above, we have

$$\|H_t^{L, \psi_\infty}\|_{1 \rightarrow \infty} \leq \frac{C_3 2^{d_f/\beta}}{t^{d_f/\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta)$$

for all $\lambda \geq \lambda_0$, for all $x, y \in M$ with $d(x, y) \geq 2$, where L is as defined in (73). Therefore

$$h_t^{(L)}(x, y) \leq \frac{C_2 2^{d_f/\beta}}{t^{d_f/\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta + \psi_\infty(y) - \psi_\infty(x)) \tag{82}$$

for all $x, y \in X$ satisfying $d(x, y) \geq 2$, for all $t \geq 1$ and for all $\lambda \geq \lambda_0$.

Now we choose λ as

$$\lambda = \frac{d_f + \beta}{\beta} \log(r^\beta/t). \tag{83}$$

By (72), it suffices to show (71) for the case $d(x, y) \geq 2$ and $\lambda \geq \lambda_0$. By the choices in (73), we have $\psi_\infty(y) - \psi_\infty(x) = -\lambda$ and hence by (82), we obtain

$$h_t^{(L)}(x, y) \leq \frac{C_2 2^{d_f/\beta}}{t^{d_f/\beta}} \exp(2C_0 e^{8\beta\vartheta\lambda} t/r^\beta - \lambda) = \frac{C_2 2^{d_f/\beta}}{t^{d_f/\beta}} \exp(2C_0) \frac{t^{1+(d_f/\beta)}}{r^{d_f+\beta}}. \tag{84}$$

By (84), (73) along with Meyer decomposition bound in [10, Lemma 3.1], there exists $C_4, C_5 > 0$ such that for all $x, y \in X$ and for all $t \geq 1$ satisfying $d(x, y)^\beta \geq C_5 t$, we have

$$h_t(x, y) \leq \frac{C_4 t}{(d(x, y))^{d_f+\beta}}.$$

The above equation along with (72) yields (71) which in turn implies $\text{UHKP}(d_f, \beta)$. □

We will now prove the main results stated in the Introduction.

Proof of Theorem 1.5. Proposition 3.1 and Proposition 2.7 are the two desired implications. □

Proof of Corollary 1.6. Let K_1 satisfy $\text{HKP}(d_f, \beta)$. Then by Theorem 1.5, we have that K_1 satisfies the cutoff Sobolev inequality $\text{CSJ}(\beta)$. Since two μ -symmetric Markov operators satisfying $\mathbf{J}(\beta)$ have comparable energy measures, K_2 satisfies $\text{CSJ}(\beta)$. Therefore by Theorem 1.5, K_2 satisfies $\text{UHKP}(d_f, \beta)$. The upper bounds $\text{UHKP}(d_f, \beta)$, (71), and the lower bound of the kernel in $\mathbf{J}(\beta)$ are sufficient to show the matching lower bounds $\text{LHKP}(d_f, \beta)$ using an iteration argument due to Bass and Levin [12]. The argument in [35, Sections 4 and 5] can be directly adapted to this setting. Therefore K_2 satisfies $\text{HKP}(d_f, \beta)$. □

3.1. Applications

The following transition probability estimate is the main application of our stability result and provides several examples.

Theorem 3.8. *Let (X, d, μ) be a vertex weighted graph satisfying $\mathbf{V}(d_f)$ with volume growth exponent d_f . Let $(S_n)_{n \in \mathbb{N}}$ denote the simple random on X with transition probability $P_n(x, y) = \mathbb{P}(X_n = y \mid X_0 = x)$. Suppose that the transition probability P_n satisfies the following sub-Gaussian estimates $\text{USG}(d_f, d_w)$ and $\text{LSG}(d_f, d_w)$ with walk dimension d_w : there exists constants $c, C > 0$ such that, for all $x, y \in X$*

$$P_n(x, y) \leq \frac{C}{n^{d_f/d_w}} \exp\left[-\left(\frac{d(x, y)^{d_w}}{Cn}\right)^{\frac{1}{d_w-1}}\right], \quad \forall n \geq 1; \tag{USG}(d_f, d_w)$$

and

$$(P_n + P_{n+1})(x, y) \geq \frac{c}{n^{d_f/d_w}} \exp\left[-\left(\frac{d(x, y)^{d_w}}{cn}\right)^{\frac{1}{d_w-1}}\right], \quad \forall n \geq 1 \vee d(x, y). \tag{LSG}(d_f, d_w)$$

Let K be a μ -symmetric Markov operator whose kernel $k = k_1$ satisfies $\mathbf{J}(\beta)$ for some $\beta \in [2, d_w)$. Then the corresponding iterated kernel k_n satisfies $\text{HKP}(d_f, \beta)$.

Remark.

- (a) For the case $\beta \in (0, 2)$, $\text{HKP}(d_f, \beta)$ follows from $\mathbf{V}(d_f)$ and $\mathbf{J}(\beta)$ (see [35, Theorem 1.1]). In this case there is no assumption required on the heat kernel of (X, d, μ) .
- (b) For the case $\beta \geq d_w$, $\text{HKP}(d_f, \beta)$ does not hold. We refer the reader to [36, Theorem 1.2] for on-diagonal estimates. We do not know off-diagonal estimates in this case.
- (c) We obtained matching two sided estimates on $\sup_{x \in X} k_{2n}(x, x)$ in [36] for all $\beta > 0$. Due to Theorem 3.8, we have analogous pointwise on-diagonal lower bounds on $k_n(x, x)$ for all $x \in X$ for the case $\beta \in [2, d_w)$.

- (d) The hypothesis of sub-Gaussian estimate for the simple random walk can be generalized to any random walk satisfying certain uniform ellipticity hypothesis – see [36, Theorem 1.2] for such a set up.
- (e) The application of Davies' method in literature (see for example [10, Section 3.2]) is inadequate to obtain $\text{UHKP}(d_f, \beta)$ for the case $\beta \geq 2$. If $\beta > 2$, the existing methods yields the off-diagonal upper bound corresponding to the off-diagonal estimate in $\text{UHKP}(d_f, 2)$ and therefore not optimal by a factor of $(1 + d(x, y))^{\beta-2}$. Even in the case $\beta = 2$, the existing method gives an off-diagonal upper bound that is not optimal by the logarithmic factor $\log(2 + d(x, y))$.

Proof of Theorem 3.8. By a known subordination argument (see [36, Theorem 5.1]), there exists a μ -symmetric Markov operator satisfying $\mathbf{J}(\beta)$ and $\text{HKP}(d_f, \beta)$. The desired result then follows from Corollary 1.6. \square

In the remainder of this section, we will elaborate on a rich family of examples using Theorem 3.8.

Example 1. For any $d_f \in [1, \infty)$ and for any $d_w \in [2, d_f + 1]$, Barlow constructs graphs that satisfy sub-Gaussian estimates $\text{USG}(d_f, d_w)$ and $\text{LSG}(d_f, d_w)$ – see [4, Theorem 2] and [27, Theorem 3.1]. Moreover, these are the complete range of d_f and d_w for which sub-Gaussian estimates $\text{USG}(d_f, d_w)$ and $\text{LSG}(d_f, d_w)$ could possibly hold for graphs.

Let us fix one such graph (X, d) satisfying $\text{USG}(d_f, d_w)$ and $\text{LSG}(d_f, d_w)$ for some $d_f \in [1, \infty)$, $d_w \in (2, d_f + 1]$. Consider any function $J : X \times X \rightarrow (0, \infty)$ satisfying

$$\frac{C^{-1}}{(1 + d(x, y))^{d_f + \beta}} \leq J(x, y) = J(y, x) \leq \frac{C}{(1 + d(x, y))^{d_f + \beta}} \quad \forall x, y \in X. \quad (85)$$

Any function J satisfying (85) defines a measure $\mu(\{x\}) = \sum_{y \in X} J(x, y)$ and a μ -symmetric Markov operator K with kernel $k(x, y) = J(x, y)(\mu(x)\mu(y))^{-1}$. If $\beta \in [2, d_w)$, we obtain $\text{HKP}(d_f, \beta)$ for the Markov operator K .

To compare with some earlier works in the continuous time case, let us point out that we obtain continuous time heat kernel estimates corresponding to $\text{UHKP}(d_f, \beta)$ in (71). Consider a continuous time jump process with symmetric measure ν satisfying $\mathbf{V}(d_f)$ and the Dirichlet form

$$\mathcal{E}(f, f) = \int_X \int_X (f(x) - f(y))^2 J(x, y) \nu(dx) \nu(dy) \quad (86)$$

on $L^2(X, \nu)$, where J satisfies (85). The condition (85) on the jump kernel J of the Dirichlet form should be interpreted as the continuous time analog of $\mathbf{J}(\beta)$. By comparison of energy measures and the measures μ and ν , we obtain Nash and cut-off Sobolev inequalities for the Dirichlet form \mathcal{E} with symmetric measure ν . By the same argument as the proof of (71), we obtain the continuous time analogue of $\text{UHKP}(d_f, \beta)$ for the heat kernel of the continuous time process associated with the above Dirichlet form \mathcal{E} on $L^2(X, \nu)$.

Now, let us consider a graph (X, d) satisfying $\text{USG}(d_f, d_w)$ and $\text{LSG}(d_f, d_w)$ with $d_f = 100$, $d_w = 101$. As mentioned above, the heat kernel corresponding to $(\mathcal{E}, L^2(X, \nu))$ satisfies the continuous time analogue of $\text{UHKP}(d_f, \beta)$ for all $\beta \in [2, 101)$. We note that the results of [24] imply the continuous time analogue of $\text{UHKP}(d_f, \beta)$ for the case $\beta \in (100, 101)$ – see [24, Corollary 6.14]. However, in the case $\beta \in [2, 100)$ the exit time and survival time estimates appearing in [24, Theorems 2.1 and 2.3] are *a priori* difficult to verify.

Acknowledgements

We thank Martin Barlow for inspiring conversations and for his helpful suggestions on an earlier draft. We thank Tom Hutchcroft for proofreading part of the manuscript.

As the authors were working on this project, another group, Zhen-Qing Chen, Takashi Kumagai, and Jian Wang, were pursuing similar goals based on somewhat different ideas [16]. Both groups obtain similar final results regarding the heat kernel estimates $\text{HKP}(d_f, \beta)$, independently. We believe that both works will prove useful in future progress. The work [16] concerns similar estimates in the continuous time setting and allows for more general volume growth and space-time scaling. Further their work involves a weaker version of cutoff Sobolev inequality. A crucial step in

their proof involves a mean value inequality proved using an iteration procedure by repeated application of Faber–Krahn and cutoff-Sobolev inequalities.

Added in revision: A more recent preprint of Grigor’yan, Hu, and Hu [23] also independently addresses similar questions using a different approach. We would also like to point out that [17,18] addresses the stability of Harnack inequalities and the relationship between Harnack inequalities and heat kernel estimates in the context of jump processes.

References

- [1] S. Andres and M. T. Barlow. Energy inequalities for cutoff-functions and some applications. *J. Reine Angew. Math.* **699** (2015) 183–215. [MR3305925](#)
- [2] D. Bakry, T. Coulhon, M. Ledoux, Saloff-Coste and L. Sobolev. Inequalities in disguise. *Indiana Univ. Math. J.* **44** (4) (1995) 1033–1074. [MR1386760](#)
- [3] M. T. Barlow. *Diffusions on Fractals, Lectures on Probability Theory and Statistics. Lecture Notes in Math.* **1690**, 1–121. Springer, Berlin, 1998. [MR1668115](#)
- [4] M. T. Barlow. Which values of the volume growth and escape time exponent are possible for a graph? *Rev. Mat. Iberoam.* **20** (1) (2004) 1–31. [MR2076770](#)
- [5] M. T. Barlow and R. F. Bass. Brownian motion and harmonic analysis on Sierpinski carpets. *Canad. J. Math.* **51** (4) (1999) 673–744. [MR1701339](#)
- [6] M. T. Barlow and R. F. Bass. *Random Walks on Graphical Sierpinski Carpets. Random Walks and Discrete Potential Theory. Sympos. Math.* **XXXIX**, 26–55. Cambridge Univ. Press, Cambridge, 1999. [MR1802425](#)
- [7] M. T. Barlow and R. F. Bass. Stability of parabolic Harnack inequalities. *Trans. Amer. Math. Soc.* **356** (4) (2004) 1501–1533. [MR2034316](#)
- [8] M. T. Barlow, R. F. Bass and T. Kumagai. Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps. *Math. Z.* **261** (2009) 297–320. [MR2457301](#)
- [9] M. T. Barlow, T. Coulhon and T. Kumagai. Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs. *Comm. Pure Appl. Math.* **58** (12) (2005) 1642–1677. [MR2177164](#)
- [10] M. T. Barlow, A. Grigor’yan and T. Kumagai. Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.* **626** (2009) 135–157. [MR2492992](#)
- [11] M. T. Barlow and E. A. Perkins. Brownian motion on the Sierpinski gasket. *Probab. Theory Related Fields* **79** (4) (1988) 543–623. [MR0966175](#)
- [12] R. F. Bass and D. A. Levin. Transition probabilities for symmetric jump processes. *Trans. Amer. Math. Soc.* **354** (7) (2002) 2933–2953. [MR1895210](#)
- [13] E. A. Carlen, S. Kusuoka and D. W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. Henri Poincaré Probab. Statist.* **23** (1987) 245–287. [MR0898496](#)
- [14] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for stable-like processes on d -sets. *Stoch. Model. Appl.* **108** (2003) 27–62. [MR2008600](#)
- [15] Z.-Q. Chen and T. Kumagai. Heat kernel estimates for jump processes of mixed types on metric measure spaces. *Stoch. Model. Appl.* **108** (2003) 27–62. [MR2357678](#)
- [16] Z.-Q. Chen, T. Kumagai and J. Wang. Stability of heat kernel estimates for symmetric jump processes on metric measure spaces. Preprint. Available at [arXiv:1604.04035](https://arxiv.org/abs/1604.04035).
- [17] Z.-Q. Chen, T. Kumagai and J. Wang. Stability of parabolic Harnack inequalities for symmetric non-local Dirichlet forms. Preprint. Available at [arXiv:1609.07594](https://arxiv.org/abs/1609.07594).
- [18] Z.-Q. Chen, T. Kumagai and J. Wang. Elliptic Harnack inequalities for symmetric non-local Dirichlet forms. Preprint. Available at [arXiv:1703.09385](https://arxiv.org/abs/1703.09385).
- [19] T. Coulhon. Ultracontractivity and Nash type inequalities. *J. Funct. Anal.* **141**, (1996) 510–539. [MR1418518](#)
- [20] E. B. Davies. Uniformly elliptic operators with measurable coefficients. *J. Funct. Anal.* **132** (1) (1995) 141–169. [MR1346221](#)
- [21] T. Delmotte. Parabolic Harnack inequality and estimates of Markov chains on graphs. *Revista Mat. Iberoamericana* **15** (1999). [MR1681641](#)
- [22] A. Grigor’yan. *Heat Kernels and Function Theory on Metric Measure Spaces. Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces. Contemp. Math.* **338**, 143–172. Amer. Math. Soc., Providence, RI, 2003. [MR2039954](#)
- [23] A. Grigor’yan, E. Hu and J. Hu. Two-sided estimates of heat kernels of jump type Dirichlet forms, 2016. Preprint. Available at <https://www.math.uni-bielefeld.de/~grigor/gcap.pdf>. [MR3267014](#)
- [24] A. Grigor’yan, J. Hu and K.-S. Lau. Estimates of heat kernels for non-local regular Dirichlet forms. *Trans. Amer. Math. Soc.* **366** (12) (2014) 6397–6441. [MR3267014](#)
- [25] A. Grigor’yan and T. Kumagai. On the dichotomy in the heat kernel two sided estimates, analysis on graphs and its applications. In *Proc. Sympos. Pure Math.* 199–210, **77**. Amer. Math. Soc., Providence, RI, 2008. [MR2459870](#).
- [26] A. Grigor’yan and A. Telcs. Sub-Gaussian estimates of heat kernels on infinite graphs. *Duke Math. J.* **109** (3) (2001) 451–510. [MR1853353](#)
- [27] A. Grigor’yan and A. Telcs. Harnack inequalities and sub-Gaussian estimates for random walks. *Math. Ann.* **324** (2002) 551–556. [MR1938457](#)
- [28] B. M. Hambly and T. Kumagai. Transition density estimates for diffusion processes on post critically finite self-similar fractals. *Proc. Lond. Math. Soc.* (3) **78** (2) (1999) 431–458. [MR1665249](#)

- [29] J. Hu and X. Li The Davies method revisited for heat kernel upper bounds of regular Dirichlet forms on metric measure spaces. Preprint. Available at [arXiv:1605.05548](https://arxiv.org/abs/1605.05548).
- [30] O. D. Jones. Transition probabilities for the simple random walk on the Sierpinski graph. *Stochastic Process. Appl.* **61** (1) (1996) 45–69. [MR1378848](#)
- [31] J. Kigami. *Analysis on Fractals. Cambridge Tracts in Mathematics* **143**. Cambridge University Press, Cambridge, 2001. viii+226 pp. [MR1840042](#)
- [32] T. Kumagai. Some remarks for stable-like jump processes on fractals. In *Fractals in Graz 2001* 185–196. *Trends Math.* Birkhäuser, Basel, 2003. [MR2091704](#)
- [33] P.-A. Meyer. Renaissance, recollements, mélanges, ralentissement de processus de Markov. *Ann. Inst. Fourier (Grenoble)* **25** (3–4) (1975) 464–497. [MR0415784](#).
- [34] E. Mossel, K. Oleszkiewicz and A. Sen. On reverse hypercontractivity. *Geom. Funct. Anal.* **23** (3) (2013) 1062–1097. [MR3061780](#)
- [35] M. Murugan and L. Saloff-Coste. Transition probability estimates for long range random walks. *New York J. Math.* **21** (2015) 723–757. [MR3386544](#)
- [36] M. Murugan and L. Saloff-Coste. Anomalous threshold behavior of long range random walks. *Electron. J. Probab.* **20** (2015) Article ID 74. [MR3371433](#)
- [37] M. Murugan and L. Saloff-Coste. Davies’ method for anomalous diffusions. *Proc. Amer. Math. Soc.* **145** (4) (2017) 1793–1804. [MR3601569](#)
- [38] A. Stós. Symmetric α -stable processes on d-sets. *Bull. Pol. Acad. Sci. Math.* **48** (3) (2000) 237–245. [MR1779007](#)