Heat kernel estimates for symmetric diffusion processes, stability of parabolic Harnack inequalities and applications

Takashi Kumagai *
RIMS, Kyoto University
http://www.kurims.kyoto-u.ac.jp/~kumagai

April 29, 2011

Abstract

These are lecture notes for the summer school of probability held on 17-20 August 2005 at Kyushu University, Japan.

Contents

*Research partially supported by Ministry of Education, Japan, Grant-in-Aid for Scientific Research (A)(1) 17204009 (Head Investigator: H. Sugita) and (A)(1) 14204010 (Head Investigator: S. Taniguchi).

1 Introduction

******

2 Classical metheods

2.1 History in brief

Before explaining the results for sub-diffusive cases, let us very briefly overview the history for diffusive cases. See [30, 72] etc. for details.

For any divergence operator $\mathcal{L} = \sum_{i,j=1}^n \frac{\partial}{\partial x_i}(a_{ij}(x)\frac{\partial}{\partial x_j})$ on $\mathbb{R}^n$ satisfying a uniform elliptic condition, Aronson ([2]) proved (2.1) with $\mu(B(x,t^{1/2})) \asymp t^{d/2}$. Later in the 20th century, there were various outstanding results in the field of global analysis on manifolds. Let Δ be the Laplace-Beltrami operator on a complete Riemannian manifold X with the Riemannian metric d and with the Riemannian measure μ. Li-Yau ([65]) proved the remarkable fact that if X has non-negative Ricci curvature, then the heat kernel $p_t(x,y)$ satisfies

$$\frac{c_1}{\mu(B(x,t^{1/2}))}\exp(-\frac{d(x,y)^2}{c_1 t}) \le p_t(x,y) \le \frac{c_2}{\mu(B(x,t^{1/2}))}\exp(-\frac{d(x,y)^2}{c_2 t}). \tag{2.1}$$

A few years later, Grigor'yan ([39]) and Saloff-Coste ([73]) elegantly refined the result and proved, in conjunction with the results by Fabes-Stroock ([34]) and Kusuoka-Stroock ([63]), that (2.1) is equivalent to a volume doubling condition (VD) plus Poincaré inequalities (PI(2)) –see Appendix for definition. The results were then extended to the framework of Dirichlet forms in [75, 76, 20], to the framework of graphs in [32]. Detailed heat kernel estimates are strongly related to the control of harmonic functions, i.e. elliptic and parabolic Harnack inequalities (EHI), (PHI(2)) on X. The origin of ideas and techniques used in this field go back to Nash ([70]), Moser ([68, 69]) and there are many other significant works in this area. Summarizing, the following equivalence holds.

$$(2.1) \Leftrightarrow (VD) + (PI(2)) \Leftrightarrow (PHI(2)). \tag{2.2}$$

An important corollary of this fact is, since (VD) and (PI(2)) are stable under certain perturbations of the operator, that (2.1) and (PHI(2)) are also stable under these perturbations.

2.2 The Nash inequality

Let X be a locally compact separable metric space and let $(\mathcal{E}, \mathcal{F})$ be a Dirichlet form on $L^2(X, \mu)$. Let $-\Delta$, $\{P_t\}$ be the corresponding non-negative self-adjoint operator and the semigroup respectively.

The next theorem was proved by Carlen-Kusuoka-Stroock ([25]), where the original idea of the proof of 1) $\Rightarrow$ 2) was due to Nash [70].

Theorem 2.1 (The Nash inequality, [25])
The following are equivalent for any $\delta > 0$.
1) There exist $c_1, \theta > 0$ such that for all $f \in \mathcal{F} \cap L^1$,

$$\|f\|_2^{2+4/\theta} \le c_1(\mathcal{E}(f,f) + \delta\|f\|_2^2)\|f\|_1^{4/\theta}, \tag{$Nash$}$$

where $\|f\|_p := (\int_X |f|^p d\mu)^{1/p}$.
2) For all $t > 0$, $P_t(L^1) \subset L^\infty$ and it is a bounded operator. Moreover, there exist $c_2, \theta > 0$ such that

$$\|P_t\|_{1\to\infty} \le c_2 e^{\delta t} t^{-\theta/2}, \qquad \forall t > 0.$$

Here $\|P_t\|_{1\to\infty}$ is an operator norm of $P_t : L^1 \to L^\infty$.

In order to prove the theorem, we prepare a lemma. For the lemma, $\mathcal{E}$ should merely be a symmetric closed form on a Hilbert space $\mathcal{H}$. Set $\mathcal{E}_1(\cdot,\cdot) = \mathcal{E}(\cdot,\cdot) + (\cdot,\cdot)$, where $(\cdot,\cdot)$ is the inner product of $\mathcal{H}$. (Then $(\mathcal{E}_1, \mathcal{F})$ is a Hilbert space.) Throughout this subsection, we refer to [Kig].

Lemma 2.2
a) For all $f \in Dom(-\Delta)$, $\mathcal{E}(P_t f, P_t f)$ is monotonically decreasing on $t > 0$ and $\lim_{t\downarrow 0} \mathcal{E}(P_t f, P_t f) = \mathcal{E}(f,f)$.
b) $\{P_t\}$ is a strongly continuous semigroup on $(\mathcal{E}_1, \mathcal{F})$.
c) Assume that $\{P_t\}$ is a Markovian semigroup on $L^2(X,\mu)$. Then $\|P_t f\|_1 \le \|f\|_1$ for all $f \in L^2 \cap L^1$.

Proof. a) Note that Δ is the generator of $\{P_t\}$, so that $P_t f \in Dom(-\Delta)$. Note also that for $f, g \in Dom(-\Delta)$,

$$\begin{aligned} \mathcal{E}(P_t f, g) &= -(P_t f, \Delta g) = -(\Delta P_t f, g) = -\lim_{h\downarrow 0}(\frac{P_h - I}{h} P_t f, g) = -\lim_{h\downarrow 0}(P_t \frac{P_h - I}{h} f, g) \\ &= -(P_t \Delta f, g) = -(\Delta f, P_t g) = \mathcal{E}(f, P_t g). \end{aligned} \tag{2.3}$$

Now let $u(t) = \mathcal{E}(P_{t/2} f, P_{t/2} f)$. Then, using (2.3), $u(t) = \mathcal{E}(f, P_t f) = -(\Delta f, P_t f)$, so that $u'(t) = -(\Delta f, \Delta P_t f) = -(\Delta f, P_t \Delta f) = -(P_{t/2}\Delta f, P_{t/2}\Delta f) \le 0$. Thus, $u(t)$ is monotonically decreasing. Since $\{P_t\}$ is strongly continuous, $u(t) = -(\Delta f, P_t f) \to -(\Delta f, f) = \mathcal{E}(f,f)$ as $t \downarrow 0$.

b) The semigroup property is clear, so we first prove the contraction property. Note that $Dom(-\Delta)$ is dense in $\mathcal{F}$ w.r.t. $\mathcal{E}_1$. For any $f \in \mathcal{F}$, take $\{f_n\} \subset Dom(-\Delta)$ so that $f_n \to f$ in $\mathcal{E}_1$. By a), $\mathcal{E}(P_t f_n, P_t f_n) \le \mathcal{E}(f_n, f_n)$ and $\{P_t f_n\}_n$ is an $\mathcal{E}$-Cauchy sequence. Since $\{P_t f_n\}_n$ is an $\mathcal{H}$-Cauchy sequence as well, and $P_t f_n \to P_t f$ in $\mathcal{H}$, it follows that $P_t f_n \to P_t f$ in $\mathcal{E}_1$. Hence $\mathcal{E}(P_t f, P_t f) \le \mathcal{E}(f, f)$. Strong continuity of $\{P_t\}$ can be proved using a) and the approximation by a sequence in $Dom(-\Delta)$.
c) First, we show that if $0 \le f \in L^2$, then $0 \le P_t f$. Indeed, if we let $f_n = f \cdot 1_{f^{-1}([0,n])}$, then $f_n \to f$ in L^2. Since $0 \le f_n \le n$, the Markov property of $\{P_t\}$ implies that $0 \le P_t f_n \le n$. Taking $n \to \infty$, we obtain $0 \le P_t f$. Using this, we have $P_t|f| \ge | P_t f|$, since $-|f| \le f \le |f|$. Using this fact and the Markov property, we have for all $f \in L^2 \cap L^1$ and all Borel set $A \subset X$,

$$(|P_t f|, 1_A)_2 \le (P_t|f|, 1_A)_2 = (|f|, P_t 1_A)_2 \le \|f\|_1,$$

where $(f,g)_2 := \int_X f(x)g(x)d\mu(x)$ for $f, g \in L^2$. Hence we see that $P_t f \in L^1$ and $\|P_t f\|_1 \le \|f\|_1$. $\square$

Proof of Theorem 2.1: 1) $\Rightarrow$ 2) : Let $f \in L^2 \cap L^1$ with $\|f\|_1 = 1$ and $u(t) := (P_t f, P_t f)_2$. Then,

$$\begin{aligned} \frac{u(t+h) - u(t)}{h} &= \frac{1}{h}(P_{t+h}f + P_t f, P_{t+h}f - P_t f)_2 = (P_{t+h}f + P_t f, \frac{(P_h - I)P_t f}{h})_2 \\ &\xrightarrow{h \downarrow 0} 2(P_t f, \Delta P_t f)_2 = -2\mathcal{E}(P_t f, P_t f). \end{aligned}$$

Hence $u'(t) = -2\mathcal{E}(P_t f . P_t f)$. Now by 1),

$$2u(t)^{1+2/\theta} \le c_1(-u'(t) + 2\delta u(t))\|P_t f\|_1^{4/\theta} \le c_1(-u'(t) + 2\delta u(t)),$$

because $\|P_t f\|_1 \le \|f\|_1 = 1$ (by Lemma 2.2 c)). Thus,

$$2(e^{-2\delta t}u(t))^{1+2/\theta} \le 2e^{-2\delta t}u(t)^{1+2/\theta} \le -c_1(e^{-2\delta t}u(t))'.$$

Set $v(t) = (e^{-2\delta t}u(t))^{-2/\theta}$, then we obtain $v'(t) \ge 4/(c_1\theta)$. Since $\lim_{t\downarrow 0} v(t) = u(0)^{-2/\theta} > 0$, it follows that $v(t) \ge 4t/(c_1\theta)$. This means $u(t) \le c_2 e^{2\delta t} t^{-\theta/2}$ where $c_2 = (c_1\theta/4)^{\theta/2}$. Hence

$$\|P_t f\|_2 \le c_3 e^{\delta t} t^{-\theta/4}\|f\|_1, \qquad \forall f \in L^2 \cap L^1,$$

which implies $\|P_t\|_{1\to 2} \le c_3 e^{\delta t} t^{-\theta/4}$. Since $P_t = P_{t/2} \circ P_{t/2}$ and $\|P_{t/2}\|_{1\to 2} = \|P_{t/2}\|_{2\to\infty}$, we obtain 2).
2) $\Rightarrow$ 1) : Let $f \in \mathcal{F} \cap L^1$. Then, for $0 < \epsilon < t$,

$$\begin{aligned} (e^{-\delta t}P_t f, f)_2 &= (e^{-\delta\epsilon}P_\epsilon f, f)_2 + \int_\epsilon^t (\frac{\partial}{\partial s}(e^{-\delta s}P_s f), f)_2 ds \\ &= (e^{-\delta\epsilon}P_\epsilon f, f)_2 - \int_\epsilon^t e^{-\delta s}((\delta I - \Delta)P_s f, f)_2 ds. \end{aligned}$$

Using Lemma 2.2 b),

$$\begin{aligned} e^{-\delta s}((\delta I - \Delta)P_s f, f)_2 &= \delta e^{-\delta s}(P_{s/2}f, P_{s/2}f)_2 - e^{-\delta s}(P_{s/2}\Delta P_{s/2}f, f)_2 \\ &= \delta e^{-\delta s}(P_{s/2}f, P_{s/2}f)_2 + e^{-\delta s}\mathcal{E}(P_{s/2}f, P_{s/2}f)_2 \le \delta\|f\|_2^2 + \mathcal{E}(f,f). \end{aligned}$$

On the other hand,

$$(P_t f, f)_2 \le \|P_t\|_{1\to\infty}\|f\|_1^2 \le c_4 e^{\delta t} t^{-\theta/2}\|f\|_1^2,$$

where we used 2) in the second inequality. Combining these, we have

$$c_4\|f\|_1^2 t^{-\theta/2} \ge (e^{-\delta\epsilon}P_\epsilon f, f)_2 - (t-\epsilon)(\delta\|f\|_2^2 + \mathcal{E}(f,f)).$$

Letting $\epsilon \downarrow 0$, we obtain

$$c_4\|f\|_1^2 t^{-\theta/2} + t(\delta\|f\|_2^2 + \mathcal{E}(f,f)) \geq \|f\|_2^2 \qquad \forall t > 0.$$

Now taking $t = \{c_4\|f\|_1^2/(\delta\|f\|_2^2 + \mathcal{E}(f,f))\}^{2/(2+\theta)}$, we obtain 1). $\square$

Corollary 2.3 *Suppose the Nash inequality (Theorem 2.1) holds. Let φ be an eigenfunction of $-\Delta$ with eigenvalue $\lambda \geq 1$. Then*

$$\|\varphi\|_\infty \leq c_3\lambda^{\theta/4}\|\varphi\|_2,$$

where $c_3 > 0$ is a constant independent of φ and λ.

PROOF. Since $-\Delta\varphi = \lambda\varphi$, $P_t\varphi = e^{-tH}\varphi = e^{-\lambda t}\varphi$. By Theorem 2.1, $\|P_t\|_{2\to\infty} = \|P_t\|_{1\to\infty}^{1/2} \leq c_1 t^{-\theta/4}$ for $t \leq 1$. Thus

$$e^{-\lambda t}\|\varphi\|_\infty = \|P_t\varphi\|_\infty \leq c_1 t^{-\theta/4}\|\varphi\|_2.$$

Taking $t = \lambda^{-1}$ and $c_3 = c_1 e$, we obtain the result. $\square$

Remark. Generalizations of Theorem 2.1 are given in [28, 77] etc. In subsection 8.1, we give a localized version of such generalizations.

2.3 The Davies method

In [31] (see also [30]), E.B. Davies gave a general method to obtain the Gaussian off-diagonal estimate from the Nash inequality. This method also gives the explicit constant in the exponential term of the estimates.

Let $\hat{\mathcal{F}} := \{h + c : h \in \mathcal{F}_b, c \in \mathbb{R}\}$ and $\hat{\mathcal{F}}_\infty := \{\psi \in \hat{\mathcal{F}} : e^{-2\psi}\Gamma(e^\psi, e^\psi) \ll \mu, e^{2\psi}\Gamma(e^{-\psi}, e^{-\psi}) \ll \mu\}$. The following version of is due to Carlen-Kusuoka-Stroock ([25]).

Theorem 2.4 ([25] Theorem 3.25) *Assume* (Nash). *Then, there is a constant $c > 0$ such that for each $\rho \in (0,1]$,*

$$p_t(x,y) \leq c\,(\rho t)^{-\theta/2} e^{-E((1+\rho)t,x,y)+\delta\rho t} \qquad \text{for } t > 0 \text{ and } x, y \in X, \tag{2.4}$$

where

$$E(t,x,y) := \sup\{|\psi(x) - \psi(y)| - t\Lambda(\psi)^2 : \Lambda(\psi) < \infty\}$$

with

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

The key inequality for the proof is

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

which holds for all $f \in \hat{\mathcal{F}}$ and all $p \in [1,\infty)$ (see Theorem 3.9 [25]). Indeed, let $f_t(x) := e^{\psi(x)}[P_t(e^{-\psi}f)](x)$ and apply this inequality and (Nash) to

$$\frac{\partial}{\partial t}\|f_t\|_{2p}^{2p} = -2p\mathcal{E}(e^\psi f_t^{2p-1}, e^{-\psi} f_t).$$

One then obtains a differential inequality. Handling the inequality in a suitable way (Lemma 3.21 in [25]), (2.4) can be obtained.

Now consider the divergence operator $\mathcal{L} = \sum_{i,j=1}^n \frac{\partial}{\partial x_i}(a_{ij}(x)\frac{\partial}{\partial x_j})$ on $\mathbb{R}^n$ satisfying a uniform elliptic condition; $\sigma^{-1}I \le a(\cdot) \le \sigma I$ for some $\sigma \ge 1$. In this case, (Nash) holds with $\theta = n$, $\delta = 0$ and

$$\Lambda(\psi)^2 = \sup_x (\nabla\psi(x), a(x)\nabla\psi(x)).$$

Let $\rho = 1$. Taking $\psi(x) = \theta \cdot x$ for some $\theta \in \mathbb{R}^n$ in (2.4), we get

$$p_t(x,y) \le c_1 t^{-n/2} \exp(\theta \cdot (x-y) + 2\|\theta\|^2 \sigma t).$$

Taking $\theta = (y-x)/(4\sigma t)$, we obtain

$$p_t(x,y) \le c_1 t^{-n/2} \exp(-\frac{|y-x|^2}{8\sigma t}),$$

and the Gaussian upper bound is obtained.

In fact, we can get much sharper estimate. Let

$$d_{\mathcal{E}}(x,y) := \sup\{\psi(x) - \psi(y) : \psi \in \hat{\mathcal{F}}_\infty \cap C(X), \Lambda(\psi) \le 1\}.$$

This is a metric and sometimes called an *intrinsic metric*. By a simple computation, we see

$$E((1+\rho)t, x, y) = \frac{d_{\mathcal{E}}(x,y)^2}{4(1+\rho)t}.$$

So, we conclude

$$p_t(x,y) \le c_1(\rho t)^{-n/2} \exp(-\frac{d_{\mathcal{E}}(x,y)^2}{4(1+\rho)t}).$$

Remark. For the case discussed from Section 3 (when $\beta > 2$), this method does not work. Indeed, it is known that for diffusions on 'typical' fractals, the energy measure is singular to the Hausdorff measure ([47, 61]) so $d_{\mathcal{E}}(x,y) \equiv 0$.

2.4 Moser's arguments

In [69], J. Moser proved elliptic Harnack inequalities ((EHI) – see subsection 3.2 for definition) for harmonic functions of some class of differential operators (uniform elliptic divergence forms). There the famous Moser's iteration arguments were used. He then extended the methods and proved the parabolic Harnack inequalities in [68]. Later, the arguments were simplified in [67]. In this subsection, we will overview his arguments.

For simplicity we give the argument for the Laplace-Beltrami operator on a Riemannian manifold X satisfying (VD), (PI(β)) (see subsection 3.2 for definition) and with regular volume growth

$$c_1 r^\alpha \le \mu(B(x,r)) \le c_2 r^\alpha, \qquad x \in X,\ r \ge 1.$$

Let μ be the Riemannian measure on X, and write

$$\fint_B f = \mu(B)^{-1} \int_B f d\mu.$$

From (PI(β)) one obtains (see [72], [73] Section 5.2) the Sobolev inequality

$$\Big(\fint_B |f|^{2\kappa}\Big)^{1/\kappa} \le c_1 R^\beta \fint_B |\nabla f|^2, \tag{2.5}$$

for $f \in C_0^\infty(B)$, where B has radius $R \geq 1$ and $\kappa = \bar{\alpha}/(\bar{\alpha}-2)$ where $\bar{\alpha} = 3 \vee \alpha$.

Since we are now treating the Laplace-Beltrami operator, $d\Gamma(f,f) = |\nabla f|^2 d\mu$ for $f \in \mathcal{F}$. Let $u > 0$ be harmonic on B (note that u is continuous in B in this case). Let $v = u^p$ for $p > 0$, $1/2 < a_2 < a_1 < 1$, $B_i := B(x_0, a_i R)$ and $\varphi \in C_0^\infty(B_1)$ be a cut-off function for $B_2 \subset B_1$. By "converse to the Poincaré inequality" (see Lemma 4.6 below),

$$\int_{B_1} |\varphi \nabla v|^2 \leq c_2 \|\nabla \varphi\|_\infty^2 \int_{B_1} v^2. \tag{2.6}$$

Using (2.5) with $f = v$ and (2.6),

$$(\fint_{B_2} u^{2\kappa p})^{1/\kappa} \leq c_3 R^\beta \fint_{B_2} |\nabla v|^2 \leq c_3 R^\beta \fint_{B_1} \varphi^2 |\nabla v|^2 \leq c_4 R^\beta \|\nabla \varphi\|_\infty^2 \int_{B_1} v^2.$$

Taking the "classical" cut-off function $\varphi(x) = \frac{d(x,B^c)}{R(a_1-a_2)}$, we have $\|\nabla\varphi\|_\infty^2 \leq \frac{c_5}{(a_1-a_2)^2 R^2}$. Thus

$$(\fint_{B_2} u^{2\kappa p})^{1/\kappa} \leq c_6 R^{\beta-2}(a_1-a_2)^{-2} \fint_{B_1} u^{2p}. \tag{2.7}$$

Now, let $a_k = (1+2^{-k})/2, p_k = p\kappa^k$ and $B_k = B(x_0, a_k R)$. (Then $a_k - a_{k+1} = 2^{-k-2}$.) Set $I_k = (\fint_{B_{k+1}} u^{2p_k})^{1/(2p_k)}$. Then, by (2.7) we have

$$I_{k+1} \leq (c_7 R^{\beta-2} 2^{2k})^{1/(2p_k)} I_k.$$

By iteration (this part is the first part of Moser's argument), we have

$$I_k \leq \Pi_{l=0}^{k-1} (c_7 R^{\beta-2} 2^{2l})^{1/(2p_l)} I_0 \leq c_8 R^{c'(\beta-2)} I_0.$$

Here the last inequality is due to the fact $\sum_l \kappa^{-l} < \infty$ and $\sum_l l\kappa^{-l} < \infty$, because $\kappa > 1$. Take $k \to \infty$. Since $p_k \to \infty$ and u is continuous, we have

$$\sup_{y \in B(x_0,R/2)} u(y) \leq c_8 R^{c'(\beta-2)} (\fint_B u^{2p})^{1/(2p)} =: c_8 R^{c'(\beta-2)} \Phi(2p, B).$$

Thus, when $\beta = 2$, by the second part of Moser's argument (which gives the comparison between $\Phi(2p,B)$ and $\Phi(-2p,B)$) gives

$$\sup_{B(x_0,R/2)} u \leq c_1 \Phi(2p,B) \leq c_2 \Phi(-2p,B) \leq c_3 \inf_{B(x_0,R/2)} u$$

and (EHI) is proved.

Remark. If $\beta > 2$, one still obtains an L^∞ bound on u in $B(x,R/2)$, but the constant now depends on R, so that the final constant in the (EHI) will also depend on R! Similar problems would arise if one tried other approaches, such as that in [34]. As we see, the problem arises in the first ('easy') part of Moser's argument. Instead of the linear cut-off functions, one needs cut-off functions such that the term $R^{\beta-2}$ in the right hand side of (2.7) disappears.

3 Framework and main theorem

3.1 Framework

We will consider two classes of spaces, namely metric measure Dirichlet spaces and weighted graphs.

Metric measure Dirichlet spaces Let (X,d) be a connected locally compact complete separable metric space. We assume that the metric d is geodesic: for each $x,y \in X$ there exists a (not necessarily unique) geodesic path $\gamma(x,y)$ such that for each $z \in \gamma(x,y)$, we have $d(x,z)+d(z,y) = d(x,y)$. Let μ be a Borel measure on X such that $0 < \mu(B) < \infty$ for every ball B in X. We write $B(x,r) = \{y : d(x,y) < r\}$, and $V(x,r) = \mu(B(x,r))$. Note that under the assumptions above, the closure of $B(x,r)$ is compact for all $x \in X$ and $0 < r < \infty$. For simplicity in what follows, we will also assume that X has infinite diameter, but similar results (with obvious modifications to the statements and the proofs) hold when the diameter of X is finite. We will call such a space a *metric measure space*, or a MM space.

Now let $(\mathcal{E},\mathcal{F})$ be a regular, strong local Dirichlet form on $L^2(X,\mu)$: see [35] for details. We denote by Δ the corresponding (non-positive) self-adjoint operator; that is, we say h is in the domain of Δ and $\Delta h = f$ if $h \in \mathcal{F}$ and $\mathcal{E}(h,g) = -\int fg\,d\mu$ for every $g \in \mathcal{F}$. Let $\{P_t\}$ be the corresponding semigroup. $(\mathcal{E},\mathcal{F})$ is called *conservative* (or *stochastically complete*) if $P_t 1 = 1$ for all $t > 0$. Throughout the paper, we assume that $(\mathcal{E},\mathcal{F})$ is conservative. Since $\mathcal{E}$ is regular, $\mathcal{E}(f,g)$ can be written in terms of a signed measure $\Gamma(f,g)$. To be more precise, for $f \in \mathcal{F}_b$ (the collection $\mathcal{F}_b$ is the set of functions in $\mathcal{F}$ that are essentially bounded) $\Gamma(f,f)$ is the unique smooth Borel measure (called the energy measure) on X satisfying

$$\int_X \tilde{g}\,d\Gamma(f,f) = 2\mathcal{E}(f,fg) - \mathcal{E}(f^2,g), \qquad g \in \mathcal{F}_b,$$

where $\tilde{g}$ is the quasi-continuous modification of $g \in \mathcal{F}$. (Recall that $u : X \to \mathbb{R}$ is called quasi-continuous if for any $\varepsilon > 0$, there exists an open set $G \subset X$ such that $\mathrm{Cap}(G) < \varepsilon$ and $u|_{X\setminus G}$ is continuous. It is known that each $u \in \mathcal{F}$ admits a quasi-continuous modification $\tilde{u}$ – see [35], Theorem 2.1.3.) Throughout the paper, we will abuse notation and take the quasi-continuous modification of $g \in \mathcal{F}_b$ without writing $\tilde{g}$. $\Gamma(f,g)$ is defined by

$$\Gamma(f,g) = \frac{1}{2}(\Gamma(f+g,f+g) - \Gamma(f,f) - \Gamma(g,g)), \qquad f,g \in \mathcal{F}.$$

$\Gamma(f,g)$ is also local, linear in f and g, and satisfies the Leibniz and chain rules – see [35], p. 115-116. That is, if $f_1,\ldots,f_m,g$, and $\varphi(f_1,\ldots,f_m)$ are in $\mathcal{F}_b$, and φ_i denotes the partial derivative of φ in the i^{th} direction, we have:

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

$$d\Gamma(\varphi(f_1,\ldots,f_m),g) = \sum_{i=1}^m \varphi_i(f_1,\ldots,f_m)\,d\Gamma(f_i,g).$$

We call $(X,d,\mu,\mathcal{E})$ a metric measure Dirichlet space, or a MMD space.

Let $Y = (Y_t, t \ge 0, \mathbb{P}^x, x \in X\}$ be the Hunt process associated with the Dirichlet form $\mathcal{E}$ on $L^2(X,\mu)$ – see [35], Theorem 7.2.1. Since $\mathcal{E}$ is strongly local, by [35], Theorem 7.2.2 Y is a diffusion.

Examples. 1. If X is a Riemannian manifold, we can take d to be the Riemannian metric and μ the Riemannian measure. The Dirichlet form $\mathcal{E}$ is defined by taking its core $\mathcal{C}$ to be the C^∞ functions on X with compact support, and defining

$$\mathcal{E}(f,f)=\int_X |\nabla f|^2 d\mu, \quad f\in\mathcal{C}.$$

The domain $\mathcal{F}$ of $\mathcal{E}$ is then the completion of $\mathcal{C}$ with respect to the norm $||f||_2+\mathcal{E}(f,f)^{1/2}$, and $d\Gamma(f,g)=\nabla f\cdot\nabla g\, d\mu$.

2. Cable system of a graph. Given a weighted graph (G,E,ν) (see Definition 2.13 below) we can define the *cable system* G_C by replacing each edge of G by a copy of $(0,1)$, joined together in the obvious way at the vertices. For further details see [9] etc. Let μ be the measure on G_C given by taking $d\mu(t)=\nu_{xy}\,dt$ for t in the cable connecting x and y, where ν_{xy} is the conductance of the edge connecting x and y; see [9]. One takes as the core $\mathcal{C}$ the functions in $C(G_C)$ which have compact support and are C^1 on each cable, and sets

$$\mathcal{E}(f,f)=\int_{G_C}|f'(t)|^2 d\mu(t).$$

One use of this construction is that the restriction to G of a harmonic function h on G_C yields a harmonic function on G.

3. Let D be a domain in $\mathbb{R}^d$ with a smooth boundary. Then let $\mathcal{C}=C_0^2(\overline{D})$, μ be Lebesgue measure, and

$$\mathcal{E}(f,f)=\tfrac{1}{2}\int_D |\nabla f|^2 d\mu.$$

The associated Markov process Y is Brownian motion on D with normal reflection on ∂D. For the extension of this construction to piecewise smooth domains such as the pre-Sierpinski carpet, see [10].

4. For fractal sets it is not as easy to describe $\mathcal{E}$. However, let $F\subset\mathbb{R}^d$ be a connected set with diameter 1, and suppose that there exists a geodesic metric d on F. Let μ be the Hausdorff α-measure on F (with respect to d) and suppose that

$$c_1 r^\alpha \le \mu(B(x,r))\le c_2 r^\alpha, \quad x\in F,\ r>0.$$

Let

$$\begin{aligned} N_{\sigma,\infty}(f) &= \sup_{0<r\le 1} r^{-\alpha-2\sigma}\int_F\int_F 1_{B(y,r)}(x)|f(x)-f(y)|^2 d\mu(x)d\mu(y),\\ \Lambda^\sigma_{2,\infty}(F) &= \{u\in L^2(F,\mu): N_{\sigma,\infty}(u)<\infty\}. \end{aligned}$$

There exist many fractals satisfying the above with a Dirichlet form $\mathcal{E}$ on $L^2(F,\mu)$ for which the domain $\mathcal{F}$ of $\mathcal{E}$ is given by $\Lambda^{\beta/2}_{2,\infty}$, and $c_1 N_{\sigma,\infty}(f)\le\mathcal{E}(f,f)\le c_2 N_{\sigma,\infty}(f)$; see [37, 60] etc.

In the particular case of the (compact) Sierpinski gasket $F=F_{SG}$, let F_n be the set of vertices of triangles of side 2^{-n}; regard F_n as a graph with $x\sim y$ if and only if x and y are in some triangle of side 2^{-n}. Then for $f\in\Lambda^{\beta/2}_{2,\infty}$ with $\beta=\log 5/\log 2$, one has

$$\mathcal{E}(f,f)=c\lim_{n\to\infty}(5/3)^n\sum_{x\sim y}(f(x)-f(y))^2.$$

Weighted graphs Let (G,E) be an infinite locally finite connected graph. We write $x\sim y$ if $(x,y)\in E$, i.e., there is an edge connecting x and y. Define edge weights (conductances) $\mu_{xy}=\mu_{yx}\ge 0$, $x,y\in G$, and assume that μ is adapted to the graph structure by requiring that $\mu_{xy}>0$ if and only if $x\sim y$. Let $\mu_x=\sum_y \mu_{xy}$, and define a measure μ on G by $\mu(A)=\sum_{x\in A}\mu_x$. We call (G,μ) a *weighted graph*.

We write $d(x,y)$ for the graph distance, and define the balls

$$B_G(x,r)=\{y: d(x,y)<r\}.$$

Given $A \subset G$ write $\partial A = \{y \in A^c : d(x,y) = 1 \text{ for some } x \in A\}$ for the exterior boundary of A, and let $\overline{A} = A \cup \partial A$.

A weighted graph (G, μ) has *controlled weights* if there exists $p_0 > 0$ such that for all $x, y \in G$

$$\frac{\mu_{xy}}{\mu_x} \geq p_0, \qquad x \sim y.$$

This was called the p_0-condition in [41].

The Laplacian is defined on (G, μ) by

$$\Delta f(x) = \frac{1}{\mu_x} \sum_y \mu_{xy}(f(y) - f(x)).$$

We also define a Dirichlet form $(\mathcal{E}, \mathcal{F})$ by taking $\mathcal{F} = L^2(G, \mu)$, and

$$\mathcal{E}(f,g) = \tfrac{1}{2} \sum_x \sum_y (f(x) - f(y))(g(x) - g(y))\mu_{xy}, \quad f, g \in \mathcal{F}.$$

If $f \in \mathcal{F}$ we define the measure $\Gamma_G(f,f)$ on G by setting

$$\Gamma_G(f,f)(x) = \sum_{y \sim x} (f(x) - f(y))^2 \mu_{xy}.$$

Let $Y = \{Y_t\}_{t \geq 0}$ be the continuous time random walk on G associated with $\mathcal{E}$ and the measure μ. When the natural weights are given on G, Y is called the simple random walk on G. Y waits at x for an exponential mean 1 random time and then moves to a neighbour y of x with probability proportional to μ_{xy}. We define the transition density (heat kernel density) of Y with respect to μ by

$$q_t(x,y) = \mathbb{P}^x(Y_t = y)/\mu_y. \tag{3.1}$$

3.2 Inequalities

In this subsection, we will define various inequalities for later use. Here we state under the framework of MMD spaces. Similar definition can be given for weighted graphs. For weighted graphs case, we will consider only global structures, so, for example $R \geq 1$, $t \geq 1$ in the following inequalities.

Let $\beta, \bar{\beta} \geq 2$ and

$$\Psi(s) = \Psi_{\bar{\beta},\beta}(s) = \begin{cases} s^{\bar{\beta}} & \text{if } s \leq 1 \\ s^{\beta} & \text{if } s > 1. \end{cases} \tag{3.2}$$

$\Psi(s)$ will give the space/time scaling on the space X. Generalization of this time scaling factor (for instance, simply assuming (8.1)) may be possible, but we do not pursue it here.

(I) X satisfies volume doubling (VD) if there exists a constant c_1 such that

$$V(x, 2R) \leq c_1 V(x, R) \quad \text{ for all } x \in X,\ R \geq 0. \tag{VD}$$

(II) X satisfies the *Poincaré inequality* (PI(Ψ)) if there exists a constant c_2 such that for any ball $B = B(x,R) \subset X$ and $f \in \mathcal{F}$,

$$\int_B (f(x) - \overline{f}_B)^2 d\mu(x) \leq c_2 \Psi(R) \int_B d\Gamma(f,f). \tag{PI(Ψ)}$$

Here $\overline{f}_B = \mu(B)^{-1} \int_B f(x) d\mu(x)$.

(III) We say a function u is *harmonic* on a domain D if $u \in \mathcal{F}_{loc}$ and $\mathcal{E}(u,g) = 0$ for all $g \in \mathcal{F}$ with support in D. Here $u \in \mathcal{F}_{loc}$ if and only if for any relatively compact open set G, there exists a function $w \in \mathcal{F}$

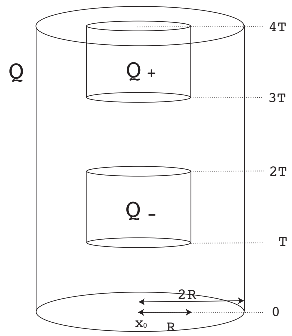

Figure 1: Parabolic Harnack inequality

such that $u = w$ μ-a.e. on G. See page 117 in [35] for the definition of $\mathcal{E}(u, g)$ for $u \in \mathcal{F}_{loc}$ when $(\mathcal{E}, \mathcal{F})$ is a regular, strong local Dirichlet form. Functions in $\mathcal{F}$ are only defined up to quasi-everywhere equivalence; we use a quasi-continuous modification of u. X satisfies the *elliptic Harnack inequality* (EHI) if there exists a constant c_3 such that, for any ball $B(x, R)$, whenever u is a non-negative harmonic function on $B(x, R)$ then there is a quasi-continuous modification $\tilde{u}$ of u that satisfies

$$\sup_{B(x,R/2)} \tilde{u} \leq c_3 \inf_{B(x,R/2)} \tilde{u}. \tag{EHI}$$

Note that by a standard argument (see subsection 9.3) (EHI) implies that $\tilde{u}$ is Hölder continuous.

(IV) Let $Q = Q(x_0, T, R) = (0, 4T) \times B(x_0, 2R) =: I \times B_{2R}$. Let $u(t, x) : Q \to \mathbb{R}$.

- We define $u_t = \frac{\partial u}{\partial t} \in L^2(dt \times \mu)$ as the derivative in the Schwartz' distribution sense. That is, we define u_t to be the function f in $L^2(dt \times \mu)$ so that for any function $g : Q \to \mathbb{R}$ such that $g(x, \cdot) \in C_K^\infty(0, 4T)$ for each $x \in B(x_0, 2R)$ and $g_t = \frac{\partial g}{\partial t} \in L^2(dt \times \mu)$, then

$$\int_Q (f(x,t)g(x,t) + u(x,t)g_t(x,t))\, dt\, d\mu(x) = 0.$$

- Let $H(I \to \mathcal{F}^*)$ be the space of functions $u \in L^2(I \to \mathcal{F}^*)$ with the distributional time derivative $u_t \in L^2(I \to \mathcal{F}^*)$ equipped with the norm

$$\Big(\int_I \|u(t,\cdot)\|^2_{\mathcal{F}^*} + \|u_t(t,\cdot)\|^2_{\mathcal{F}^*} dt \Big)^{1/2}.$$

  Here we identify $L^2(X, \mu)$ with its own dual and denote the dual of $\mathcal{F}$ by $\mathcal{F}^*$. So, $\mathcal{F} \subset L^2(X, \mu) \subset \mathcal{F}^*$ with continuous and dense embeddings.

  Let $\mathcal{F}(I \times X) = L^2(I \to \mathcal{F}) \cap H(I \to \mathcal{F}^*)$ be a Hilbert space with norm

$$\|u\|_{\mathcal{F}(I\times X)} = \Big(\int_I \|u(t,\cdot)\|^2_{\mathcal{F}} + \|u_t(t,\cdot)\|^2_{\mathcal{F}^*} dt \Big)^{1/2}.$$

- We define $\mathcal{F}_{loc}(Q)$ to be the set of $dt \otimes d\mu$-measurable functions on Q such that for every relatively compact open set $D \subset\subset B_{2R}$ and every open interval $I' \subset\subset I$, there exists a function $u' \in \mathcal{F}(I \times X)$ with $u = u'$ on $I' \times D$. We define

$$\mathcal{F}_c(Q) := \{u \in \mathcal{F}(I \times X) : u(t,\cdot) \text{ has compact support in } B_{2R} \text{ for a.e. } t \in I\}.$$

We say a function $u(t,x): Q \to \mathbb{R}$ is a solution of the heat equation in Q if $u \in \mathcal{F}_{loc}(Q)$ and

$$\int_J \Big[\int f(t,x) u_t(t,x) \mu(dx) + \mathcal{E}(f(t,\cdot), u(t,\cdot)) \Big] dt = 0, \quad \forall J \subset\subset I, \ \forall f \in \mathcal{F}_c(Q). \tag{3.3}$$

X satisfies the *parabolic Harnack inequality* (PHI(Ψ)), if there exists a constant c_4 such that the following holds. Let $x_0 \in X$, $R > 0$, $T = \Psi(R)$, and $u = u(t,x)$ be a non-negative solution of the heat equation in $Q(x_0, T, R)$. Write $Q_- = (T, 2T) \times B(x_0, R)$ and $Q_+ = (3T, 4T) \times B(x_0, R)$; then there exists $\tilde{u} = \tilde{u}(t,x)$ such that $\tilde{u}(t,\cdot)$ is a quasi-continuous modification of $u(t,\cdot)$ for each t and

$$\sup_{Q_-} \tilde{u} \le c_4 \inf_{Q_+} \tilde{u}. \tag{PHI(Ψ)}$$

Given this (PHI(Ψ)), a standard oscillation argument implies that $\tilde{u}$ is jointly continuous.

Remark. In the case of general MMD spaces we can only define harmonic functions up to quasi-everywhere equivalence. This is why we needed to be careful in our definitions of (EHI) and (PHI(Ψ)).

(V) Let A, B be disjoint subsets of X. We define the effective resistance $R(A,B)$ by

$$R(A,B)^{-1} = \inf\Big\{ \int_X d\Gamma(f,f) : f = 0 \text{ on } A \text{ and } f = 1 \text{ on B}, \ f \in \mathcal{F} \Big\}. \tag{3.4}$$

X satisfies the condition (RES(Ψ)) if there exist constants c_1, c_2 such that for any $x_0 \in X$, $R \ge 0$,

$$c_1 \frac{\Psi(R)}{V(x_0,R)} \le R(B(x_0,R), B(x_0,2R)^c) \le c_2 \frac{\Psi(R)}{V(x_0,R)}. \tag{RES(Ψ)}$$

(VI) X satisfies (CS(Ψ)) if there exist $\theta \in (0,1]$ and constants c_1, c_2 such that the following holds. For every $x_0 \in X$, $R > 0$ there exists a cut-off function $\varphi (= \varphi_{x_0,R})$ with the properties:
(a) $\varphi(x) \ge 1$ for $x \in B(x_0, R/2)$.
(b) $\varphi(x) = 0$ for $x \in B(x_0,R)^c$.
(c) $|\varphi(x) - \varphi(y)| \le c_1 (d(x,y)/R)^\theta$ for all x,y.
(d) For any ball $B(x,s)$ with $0 < s \le R$ and $f \in \mathcal{F}$,

$$\int_{B(x,s)} f^2 d\Gamma(\varphi,\varphi) \le c_2 (s/R)^{2\theta} \Big(\int_{B(x,2s)} d\Gamma(f,f) + \Psi(s)^{-1} \int_{B(x,2s)} f^2 d\mu \Big). \tag{3.5}$$

Remarks. 1. We call (3.5) a weighted Sobolev inequality. It is clear that to prove (3.5) it is enough to consider nonnegative f.
2. Suppose (CS(Ψ)) holds for X, but with (a) above replaced by

$$\varphi(x) \ge 1 \text{ for } x \in B(x_0, \delta R),$$

for some $\delta < \frac{1}{2}$. Then an easy covering argument (using (VD)) gives (CS(Ψ)) with $\delta = \frac{1}{2}$.
3. Let $\lambda > 1$. Suppose that (CS(Ψ)) holds, except that instead of (3.5) we have

$$\int_{B(x,s)} f^2 d\Gamma(\varphi,\varphi) \le c_2 (s/R)^{2\theta} \Big(\int_{B(x,\lambda s)} d\Gamma(f,f) + \Psi(s)^{-1} \int_{B(x,\lambda s)} f^2 d\mu \Big).$$

Then once again it is easy to obtain (CS(Ψ)) with $\lambda = 2$ by a covering argument.
4. Any operation on the cut-off function φ which reduces $d\Gamma(\varphi,\varphi)$ while keeping properties (a), (b) and (c) of (VI) will generate a new cut-off function which still satisfies (3.5). We can therefore assume that any cut-off function φ satisfies the following: (a) $0 \le \varphi \le 1$. (b) For each $t \in (0,1)$ the set $\{x : \varphi(x) > t\}$ is connected and contains $B(x_0, R/2)$. (c) Each connected component A of $\{x : \varphi(x) < t\}$ intersects $B(x_0,R)^c$.

5. Note that if (CS(Ψ)) holds for $\Psi = \Psi_{\bar\beta,\beta}$, then (CS($\Psi_{\bar\beta',\beta'}$)) holds if $\beta' \ge \beta$ and $\bar\beta' \le \bar\beta$.
(VII) For $(t,r) \in (0,\infty)\times[0,\infty)$, let

$$\Lambda_1 = \{(t,r) : t \le 1 \vee r\}, \qquad \Lambda_2 = \{(t,r) : t \ge 1 \vee r\}, \qquad \text{and} \qquad g_\beta(r,t) = \exp\Big(-\Big(\frac{r^\beta}{t}\Big)^{1/(\beta-1)}\Big).$$

We say X satisfies (HK(Ψ)) if the heat kernel $p_t(x,y)$ on X exists and satisfies

$$\frac{c_1 g_{\bar\beta}(c_2 d(x,y),t)}{V(x,t^{1/\bar\beta})} \le p_t(x,y) \le \frac{c_3 g_{\bar\beta}(c_4 d(x,y),t)}{V(x,t^{1/\bar\beta})}, \tag{3.6}$$

for $x,y\in X$ and $t\in(0,\infty)$ with $(t,d(x,y))\in\Lambda_1$, and

$$\frac{c_1 g_{\beta}(c_2 d(x,y),t)}{V(x,t^{1/\beta})} \le p_t(x,y) \le \frac{c_3 g_{\beta}(c_4 d(x,y),t)}{V(x,t^{1/\beta})}, \tag{3.7}$$

for $x,y\in X$ and $t\in(0,\infty)$ with $(t,d(x,y))\in\Lambda_2$.

Let $h(r) = \Psi(r)/r$. It is easy to see that (HK(Ψ)) is equivalent to the following:

$$\frac{c_1}{V(x,\Psi^{-1}(t))}\exp\Big(-\frac{c_2 d(x,y)}{h^{-1}(t/d(x,y))}\Big) \le p_t(x,y) \le \frac{c_3}{V(x,\Psi^{-1}(t))}\exp\Big(-\frac{c_4 d(x,y)}{h^{-1}(t/d(x,y))}\Big), \tag{3.8}$$

for all $x,y\in X$ and $t\in(0,\infty)$ where we let $d(x,y)/h^{-1}(t/d(x,y)) = 0$ if $d(x,y)=0$. We sometimes refer the first inequality of (3.8) as $(LHK(\Psi))$ and the second inequality of (3.8) as $(UHK(\Psi))$.

Remark. To understand why the crossover takes the form it does, it is useful to consider the contribution to $p_t(x,y)$ from various types of paths in X. Let $r = d(x,y)$. First, if $0 < t \le 1$ and $r < 1$ then the behaviour is essentially local.

If $r \ge t$ then we are in the 'large deviations' regime: the main contribution to $p_t(x,y)$ is from those paths of the Markov process Y which are within a distance $O(t/r)$ of a geodesic from x to y. So, once the length of the geodesic is given, only the local structure of X plays a role. Note that in this case the term in the exponential is smaller than e^{-ct}, so that the volume term $V(x,t^{1/\bar\beta})^{-1}$ could be absorbed into the exponential with a suitable modification of the constants c_2 and c_4.

Finally, if $t > 1$ and $r < t$, then the paths which contribute to $p_t(x,y)$ fill out a much larger part of X: those which lie in $B(x,t^{1/\beta})$ if $r < t^{1/\beta}$, and those which are within a distance $O(t/r^{\beta-1})$ of a geodesic from x to y in the case when $t^{1/\beta} \le r \le t$.

(VIII) We say X satisfies $(\mathrm{VD})_{\mathrm{loc}}$ if (VD) holds for $x\in X$, $0 < R \le 1$. Similarly we define $(\mathrm{PI}(\bar\beta))_{\mathrm{loc}}$, $(\mathrm{EHI})_{\mathrm{loc}}$, $(\mathrm{CS}(\bar\beta))_{\mathrm{loc}}$ and $(\mathrm{PHI}(\bar\beta))_{\mathrm{loc}}$ by requiring the conditions only for $0 < R \le 1$. For $(\mathrm{HK}(\bar\beta))_{\mathrm{loc}}$ we require the bounds only for $t\in(0,1)$ – so only (3.6) is involved. The value 1 here is for simplicity: each of the local conditions implies an analogous local condition for $0 < R \le R_0$ for any (fixed) $R_0 > 1$ – see Section 2 of [46].

Finally, we introduce two local notions which do not include any scaling order.
(IX) (a) We call φ a *cut-off function* for $A_1 \subset A_2$ if $\varphi = 1$ on A_1 and is zero on A_2^c.
(b) We say X satisfies $(\mathrm{PI})_{\mathrm{loc}}$ if for each $c_1 > 0$, there exists $c_2 > 0$ such that

$$\int_B (f(x) - \overline{f}_B)^2 d\mu(x) \le c_2 \int_B d\Gamma(f,f)$$

for any ball $B = B(x,c_1) \subset X$ and $f\in\mathcal{F}$.
(c) We say X satisfies $(\mathrm{CC})_{\mathrm{loc}}$ if for every $x_0\in X$, there exists a cut-off function $\varphi(=\varphi_{x_0})$ for $B(x_0,1/2)\subset B(x_0,1)$ such that

$$\int_{B(x_0,1)} d\Gamma(\varphi,\varphi) \le c_3 V(x_0,1),$$

where $c_3 > 0$ is independent of x_0 and φ.

Remark. (CC) stands for 'controlled cut-off' functions. Clearly $(\mathrm{PI}(\bar\beta))_{\mathrm{loc}}$ for any $\bar\beta \geq 2$ implies $(\mathrm{PI})_{\mathrm{loc}}$ and $(\mathrm{CS}(\bar\beta))_{\mathrm{loc}}$ for any $\bar\beta > 0$ implies $(\mathrm{CC})_{\mathrm{loc}}$.
(X) X satisfies the condition $(\mathrm{E}(\Psi))$ if for any $x_0 \in X$, $R \geq 0$,

$$c_1\Psi(R) \leq \mathbb{E}^{x_0}[\tau_{B(x_0,R)}] \leq c_2\Psi(R), \tag{E(Ψ)}$$

where $\tau_A = \inf\{t \geq 0 : Y_t \notin A\}$, Y_t is the strong Markov process associated to the Dirichlet form $(\mathcal{E}, \mathcal{F})$, and $\mathbb{E}^{x_0}$ denotes the expectation starting from the point x_0. The first inequality in $(\mathrm{E}(\Psi))$ is referred as $(\mathrm{E}(\Psi)_{\geq})$ and the second is referred as $(\mathrm{E}(\Psi)_{\leq})$.

Remark. The conditions (VD), (EHI) and $(\mathrm{PHI}(\Psi))$ for graphs are defined in exactly the same way as for manifolds; see [9]. The definitions of $(\mathrm{PI}(\Psi))$ and $(\mathrm{RES}(\Psi))$ are also the same. For the bound $(\mathrm{HK}(\Psi))$ we only require (3.7). The condition $(\mathrm{CS}(\Psi))$ is also the same; the weighted Sobolev inequality (3.5) takes the form

$$\sum_{x \in B_G(x_1,s)} f(x)^2 \Gamma_G(\varphi,\varphi)(x) \leq c_2 (\frac{s}{R})^{2\theta} \Big(\sum_{x \in B_G(x_1,2s)} \Gamma_G(f,f)(x) + \Psi(s)^{-1} \sum_{x \in B_G(x_1,2s)} \nu_x f(x)^2 \Big).$$

It is easy to check that $(\mathrm{PI})_{\mathrm{loc}}$ and $(\mathrm{CC})_{\mathrm{loc}}$ hold for any weighted graph with controlled weights. In fact, $(\mathrm{PI}(\bar\beta))_{\mathrm{loc}}$ and $(\mathrm{CS}(\bar\beta))_{\mathrm{loc}}$ hold for any choice of $\bar\beta \geq 2$ on such graphs, since it is irrelevant to treat $R < 1$ for graphs.

We summarize the conditions we have introduced:

(VD)	Volume doubling
$(\mathrm{PI}(\Psi))$	Poincaré inequality
(EHI)	Elliptic Harnack inequality
$(\mathrm{PHI}(\Psi))$	Parabolic Harnack inequality
$(\mathrm{RES}(\Psi))$	Resistance exponent
$(\mathrm{CS}(\Psi))$	Cut-off Sobolev inequality
(CC)	Controlled cut-off functions
$(\mathrm{HK}(\Psi))$	Heat kernel estimates
$(\mathrm{E}(\Psi))$	Walk dimension

When $\bar\beta = \beta$, we would write $(...(\beta))$ instead of $(...(\Psi))$, for instance $(\mathrm{PI}(\beta))$ instead of $(\mathrm{PI}(\Psi))$.

3.3 Main Theorems

Our main theorem in this section is the following.

Theorem 3.1 *Suppose that X is either an infinite connected weighted graph with controlled weights, or a MMD space. The following are equivalent:*
(a) X satisfies $(PHI(\Psi))$.
(b) X satisfies $(HK(\Psi))$.
(c) X satisfies (VD), $(PI(\Psi))$ and $(CS(\Psi))$.
(d) X satisfies (VD), (EHI) and $(RES(\Psi))$.
(e) X satisfies (VD), (EHI) and $(E(\Psi))$.

Stability We now discuss the stability of $(\mathrm{PHI}(\Psi))$. We will actually discuss two kinds of stability.

Definition 3.2 *A property P is stable under bounded perturbation if whenever P holds for $(\mathcal{E}^{(1)}, \mathcal{F})$, then it holds for $(\mathcal{E}^{(2)}, \mathcal{F})$, provided*

$$c_1\mathcal{E}^{(1)}(f,f) \leq \mathcal{E}^{(2)}(f,f) \leq c_2\mathcal{E}^{(1)}(f,f), \quad \textit{for all } f \in \mathcal{F}. \tag{3.9}$$

The following result is due to Le Jan ([64], Proposition 1.5.5(b)). A simple proof is given in [66] p. 389.

Lemma 3.3 *Let X be a MMD space. Suppose $(\mathcal{E}^{(i)}, \mathcal{F}), i = 1, 2$ are strong local regular Dirichlet forms that satisfy (3.9). Then the energy measures $\Gamma^{(i)}$ satisfy*

$$c_1 d\Gamma^{(1)}(f,f) \le d\Gamma^{(2)}(f,f) \le c_2 d\Gamma^{(1)}(f,f), \qquad \textit{for all } f \in \mathcal{F}.$$

It is immediate from Lemma 3.3 that the conditions $PI(\Psi)$ and $CS(\Psi)$ are stable under bounded perturbations. So we deduce:

Theorem 3.4 *Let X be a MMD space. Then (PHI(Ψ)) and (HK(Ψ)) are stable under bounded perturbations.*

The second kind of stability is stability under rough isometries.

Definition 3.5 *For each $i = 1, 2$, let (X_i, d_i, μ_i) be either a metric measure space or a weighted graph. A map $\varphi : X_1 \to X_2$ is a rough isometry if there exist constants $c_1 > 0$ and $c_2, c_3 > 1$ such that*

$$X_2 = \bigcup_{x \in X_1} B_{d_2}(\varphi(x), c_1),$$

$$c_2^{-1}(d_1(x,y) - c_1) \le d_2(\varphi(x), \varphi(y)) \le c_2(d_1(x,y) + c_1),$$

and

$$c_3^{-1}\mu_1(B_{d_1}(x, c_1)) \le \mu_2(B_{d_2}(\varphi(x), c_1)) \le c_3\mu_1(B_{d_1}(x, c_1)).$$

If there exists a rough isometry between two spaces they are said to be roughly isometric. (One can check this is an equivalence relation.)

This concept was introduced by Kanai in [53, 52]. A rough isometry between X_1 and X_2 means that the global structure of the two spaces is the same. However, to have stability of Harnack inequalities, we also require some control over the local structure. In the case of graphs it is enough to have controlled weights, but for metric measure spaces more regularity is needed. (In [53, 52] this local control was obtained by geometrical assumptions on the manifolds).

The following theorem concerns the stability of (PHI(Ψ)) under rough isometries.

Theorem 3.6 *Let X_i be either a MMD space satisfying $(VD)_{\rm loc}$ and $(PI)_{\rm loc}$ or a graph with controlled weights, and suppose there exists a rough isometry $\varphi : X_1 \to X_2$. Let $\Psi_i(s) = s^{\bar\beta_i} 1_{\{s \le 1\}} + s^{\beta} 1_{\{s \ge 1\}}$.*
(a) Suppose that X_2 satisfies $(PI(\bar\beta_2))_{\rm loc}$. If X_1 satisfies (VD), $(CC)_{\rm loc}$ and $(PI(\Psi_1))$ then X_2 satisfies (VD) and $(PI(\Psi_2))$.
(b) Suppose that X_2 satisfies $(CS(\bar\beta_2))_{\rm loc}$. If X_1 satisfies (VD) and $(CS(\Psi_1))$ then X_2 satisfies (VD) and $(CS(\Psi_2))$.

The proof of thie theorem is given in [14] ([44] for the case of weighted graphs).

By this theorem together with Theorem 3.1, we see that (PHI(Ψ)) is stable under rough isometries, given suitable local regularity of the two spaces.

Examples 1) It is known that the simple random walk on the S.G. graph (the left of Figure 2) satisfies $(HK(\log 5/\log 2))$ for $t \ge 1$. The graph on the right of Figure 1 is an image of the S.G. graph by a rough isometry. So the simple random walk on the graph also satisfies $(HK(\log 5/\log 2))$, and thus satisfies $(PHI(\log 5/\log 2))$ for $R \ge 1$.
2) Figure 3 is a 2-dimensional Riemannian manifold whose global structure is like that of the S.G.. This can be constructed from the left of Figure 1 by changing each bond to the cylinder and putting projections and dents locally. The diffusion corresponding to the Dirichlet form moves on the surface of the cylinders. Using Theorem 3.6, one can show that any divergence operator $\mathcal{L} = \sum_{i,j=1}^2 \frac{\partial}{\partial x_i}(a_{ij}(x)\frac{\partial}{\partial x_j})$ on the manifold which satisfies the uniform elliptic condition enjoys $(HK(2))$ for $t \le 1 \vee d(x,y)$ and $(HK(\log 5/\log 2))$ for $t \ge 1 \vee d(x,y)$.

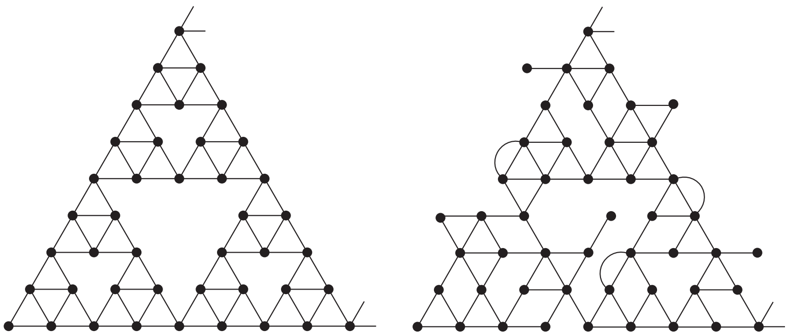

Figure 2: S.G. graph and its image by a rough isometry

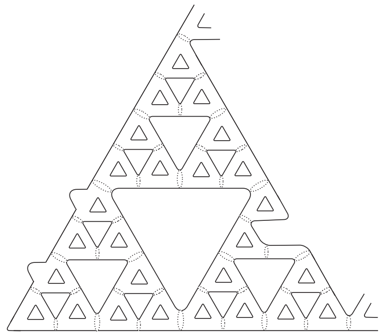

Figure 3: Fractal-like manifold

4 Proof of Theorem 3.1

In this section, we will give the proof of the key part of the theorem. The proof of $(b) \Leftrightarrow (a)$ and $(d) \Rightarrow (e)$ will be given in Appendix 2 (Section 9). Recall that $h(r) = \Psi(r)/r$. We give some inequalities.

$$p_t(x,y) \le \frac{C_1}{V(x,\Psi^{-1}(t))}, \qquad \forall x,y \in X, t>0. \qquad (DUHK(\Psi))$$

$$P^x(\tau_{B(x,r)} \le t) \le C_2 \exp\Big(-\frac{C_3 r}{h^{-1}(t/r)}\Big), \qquad \forall x \in X, r,t>0. \qquad (ELD(\Psi))$$

$$p_t(x,x) \ge \frac{C_4}{V(x,\Psi^{-1}(t))}, \qquad \forall x \in X, t>0. \qquad (DLHK(\Psi))$$

$$p_t(x,y) \ge \frac{C_5}{V(x,\Psi^{-1}(t))}, \qquad \forall x,y \in X, t>0 \text{ with } \Psi(d(x,y)) \le C_6 t. \qquad (NLHK(\Psi))$$

4.1 Proof of $(e) \Rightarrow (b)$

This is one of the most important part. Note that the existence of the heat kernel (especially the continuous one) is highly non-trivial in this general setting. With extra work, we can prove the existence, but here we will assume it to avoid the proof (which is already quite involved) more complicated.

For the proof, we first prove the following.

Proposition 4.1

$$(VD) + (DUHK(\Psi)) + (EHI) + (E(\Psi)) \Rightarrow (HK(\Psi)).$$

This proposition will be proved through several steps.

STEP 1: PROOF OF $(E(\Psi)) \Rightarrow (ELD(\Psi))$. We first give the following key lemma due to Barlow-Bass.

Lemma 4.2 *Let $\{\xi_i\}$ be non-negative random variables. Suppose there exist $0 < p < 1$ and $a > 0$ such that*

$$P(\xi_i \le t | \sigma(\xi_1, \cdots, \xi_{i-1})) \le p + at, \qquad \forall t > 0.$$

Then,

$$\log P(\sum_{i=1}^n \xi_i \le t) \le 2\Big(\frac{ant}{p}\Big)^{1/2} - n \log \frac{1}{p}.$$

Proof. We follow [7]. Let η be a random variable with distribution $P(\eta \le t) = (p + at) \wedge 1$. Then,

$$E(e^{-\lambda \xi_i} | \sigma(\xi_1, \cdots, \xi_{i-1})) \le Ee^{-\lambda\eta} = p + \int_0^{(1-p)/a} e^{-\lambda t} a dt \le p + a\lambda^{-1}.$$

So,

$$\begin{aligned} P(\sum_{i=1}^n \xi_i \le t) &= P(e^{-\lambda \sum_{i=1}^n \xi_i} \ge e^{-\lambda t}) \le e^{\lambda t} E e^{-\lambda \sum_{i=1}^n \xi_i} \\ &\le e^{\lambda t}(p + a\lambda^{-1})^n \le p^n \exp(\lambda t + \frac{an}{\lambda p}). \end{aligned}$$

The result follows on setting $\lambda = (an/(pt))^{1/2}$. □

Proof of $(E(\Psi)) \Rightarrow (ELD(\Psi))$. We first prove that there exists $0 < c_1 < 1$ and $c_2 > 0$ such that

$$P^x(\tau_{B(x,r)} \le s) \le 1 - c_1 + c_2 s/\Psi(r) \qquad \text{for all } \; x \in X, \; s \ge 0. \tag{4.1}$$

Indeed, by the Markov property, for each $x \in X$ we have

$$E^x \tau_{B(x,r)} \le s + E^x[1_{\{\tau_{B(x,r)} > s\}} E^{Y_s} \tau_{B(x,r)}] \le s + E^x[1_{\{\tau_{B(x,r)} > s\}} E^{Y_s} \tau_{B(X_s, 2r)}]. \tag{4.2}$$

Applying $(E(\Psi))$ and using the doubling property of h, which is due to the definition of Ψ, we have

$$c_3 \Psi(r) \le s + c_4 \Psi(2r) P^x(\tau_{B(x,r)} > s) = s + c_5 \Psi(r)(1 - P^x(\tau_{B(x,r)} \le s)). \tag{4.3}$$

Rearranging gives (4.1).

Next, let $l \ge 1$, $b = r/l$, and define stopping times $\sigma_i, \; i \ge 0$ by

$$\sigma_0 = 0, \quad \sigma_{i+1} = \inf\{t \ge \sigma_i : d(Y_{\sigma_i}, Y_t) \ge b\}.$$

Let $\xi_i = \sigma_i - \sigma_{i-1}, \; i \ge 1$. Let $\mathcal{F}_t$ be the filtration generated by $\{Y_s : s \le t\}$ and let $\mathcal{G}_m = \mathcal{F}_{\sigma_m}$. We have by (4.1)

$$P^x(\xi_{i+1} \le t | \mathcal{G}_i) = P^{Y_{\sigma_i}}(\tau_{B(Y_{\sigma_i}, b)} \le t) \le p + c_2 t/\Psi(b),$$

where $0 < p < 1$. As $d(Y_{\sigma_i}, Y_{\sigma_{i+1}}) = b$, we have $d(Y_0, Y_{\sigma_l}) \le r$, so that $\sigma_l = \sum_{i=1}^l \xi_i \le \tau_{B(Y_0, r)}$. So, by Lemma 4.2,

$$\log P^x(\tau_{B(x,r)} \le t) \le 2p^{-1/2}(\frac{c_2 l t}{\Psi(r/l)})^{1/2} - l \log(1/p) = c_6(\frac{lt}{\Psi(r/l)})^{1/2} - c_7 l.$$

Now take $l_0 \in \mathbb{N}$ the largest integer l that satisfies

$$c_7 l/2 > c_6(\frac{lt}{\Psi(r/l)})^{1/2}. \tag{4.4}$$

This is equivalent to $r/l > h^{-1}(c_8 t/r)$ where $c_8 = 4c_6^2/c_7^2$. Note that if $r \le h^{-1}(c_8 t/r)$, then $(ELD(\Psi))$ clearly holds by taking $c_1 > 0$ large, so we may assume that (4.4) holds for small $l \in \mathbb{N}$. Then

$$l_0 < \frac{r}{h^{-1}(c_8 t/r)} \le l_0 + 1, \quad \text{and} \quad \log P^x(\tau_{B(x,r)} \le t) \le -c_7 l_0/2.$$

We thus obtain $(ELD(\Psi))$. $\square$

STEP 2: PROOF OF $(VD)+(DUHK(\Psi))+(ELD(\Psi))\Rightarrow(UHK(\Psi))$. Fix $x\neq y$ and t and let $r:=d(x,y)$, $\epsilon<r/6$. For $a\in X$, set $B_\epsilon(a)=\{b\in X: d(a,b)<\epsilon\}$. Let $\bar\mu_x=\mu|_{B_\epsilon(x)}$, $A_1=\{z\in X: d(z,x)\le d(z,y)\}$ and $A_2=X-A_1$. Then

$$\begin{aligned} P^{\bar\mu_x}(Y_t\in B_\epsilon(y)) &= P^{\bar\mu_x}(Y_t\in B_\epsilon(y), Y_{\frac{t}{2}}\in A_1)\\ &\quad +P^{\bar\mu_x}(Y_t\in B_\epsilon(y), Y_{\frac{t}{2}}\in A_2)\equiv I_1+I_2. \end{aligned}$$

Now, letting $\tau:=\tau_{B(x,r/2)}$, we have

$$\begin{aligned} I_2 &\le P^{\bar\mu_x}(Y_t\in B_\epsilon(y),\tau<\frac{t}{2})=E^{\bar\mu_x}(1_{\tau<t/2}\int_{B_\epsilon(y)}p_{t-\tau}(Y_\tau,w)d\mu(w))\\ &\le P^{\bar\mu_x}(\tau<t/2)\sup_{z\in B(x,r/2)\cup B_\epsilon(y)}p_{t/2}(z,z)\mu(B_\epsilon(y)). \end{aligned}$$

This is OK!! For $z\in B_\epsilon(x)$, by $(ELD(\Psi))$,

$$P^z(\tau_{B(z,r/3)}<\frac{t}{2})\le c_1\exp\Big(-\frac{c_2r}{h^{-1}(t/r)}\Big).$$

Thus,

$$I_2\le c_1\Big(\sup_{z\in B(x,r/2)\cup B_\epsilon(y)}p_{t/2}(z,z)\Big)\mu(B_\epsilon(x))\mu(B_\epsilon(y))\exp\Big(-\frac{c_2r}{h^{-1}(t/r)}\Big).$$

For I_1, by the symmetry of $p_t(x,y)$,

$$P^{\bar\mu_x}(Y_t\in B_\epsilon(y), Y_{\frac{t}{2}}\in A_1)=P^{\bar\mu_y}(Y_t\in B_\epsilon(x), Y_{\frac{t}{2}}\in A_1)$$

which is bounded in exactly the same way as I_2,where x and y are changed. Adding the bounds for I_1 and I_2,

$$P^{\bar\mu_x}(Y_t\in B_\epsilon(y))\le c_1\Big(\sup_{z\in B(x,r/2)\cup B(y,r/2)}p_{t/2}(z,z)\Big)\mu(B_\epsilon(x))\mu(B_\epsilon(y))\exp\Big(-\frac{c_2r}{h^{-1}(t/r)}\Big).$$

By $(DUHK(\Psi))$ and (9.1),

$$\sup_{z\in B(x,r/2)\cup B(y,r/2)}p_{t/2}(z,z)\le\frac{c_3}{V(x,\Psi^{-1}(t))}\Big(\frac{r+\Psi^{-1}(t)}{\Psi^{-1}(t)}\Big)^\alpha.$$

If $\Psi(r)\le t$, this is bounded by $c_4V(x,\Psi^{-1}(t))^{-1}$. If $\Psi(r)>t$, then, for each $\epsilon>0$, there exists $c_\epsilon>0$ such that

$$\Big(\frac{r+\Psi^{-1}(t)}{\Psi^{-1}(t)}\Big)^\alpha\exp\Big(-\frac{\epsilon r}{h^{-1}(t/r)}\Big)\le c_\epsilon.$$

This is due to the following fact; $M=r/\Psi^{-1}(t)$ is equivalent to $h(r/M)=tM/r$, so that $M<r/h^{-1}(t/r)$. In any case, we obtain

$$P^{\bar\mu_x}(Y_t\in B_\epsilon(y))\le\frac{c_5}{V(x,\Psi^{-1}(t))}\mu(B_\epsilon(x))\mu(B_\epsilon(y))\exp\Big(-\frac{c_6r}{h^{-1}(t/r)}\Big).$$

Dividing both sides by $\mu(B_\epsilon(x))$, $\mu(B_\epsilon(y))$ and using the continuity of $p_t(x,y)$ gives $(UHK(\Psi))$. $\square$

STEP 3: PROOF OF $(VD)+(ELD(\Psi))\Rightarrow(DLHK(\Psi))$. Using $(ELD(\Psi))$ we have that

$$P^x(Y_t\notin B(x,r))\le P(\tau_{B(x,r)}\le t)\le c_1\exp\Big(-\frac{c_2r}{h^{-1}(t/r)}\Big).$$

Hence by choosing r such that $c_3\Psi(r) < t < c_4\Psi(r)$ for some $c_3, c_4 > 0$, we have

$$P^x(Y_t \notin B(x,r)) \le c_5 < 1.$$

Thus $P^x(Y_t \in B(x,r)) \ge 1 - c_5 > 0$. By Cauchy-Schwarz,

$$(1-c_5)^2 \le P^x(Y_t \in B(x,r))^2 = (\int_{B(x,r)} p_t(x,z)d\mu(z))^2 \le V(x,r)p_{2t}(x,x).$$

Now, using the lower bound of our choice of t and (VD), we obtain the result. □

Remark. By the same argument, we can obtain th following slightly stronger conclusion; Assume (VD) and $(ELD(\Psi))$. Then, there exist $c_1, c_2 > 0$ such that

$$p_t^{B(x,R)}(x,x) \ge \frac{c_1}{V(x,\Psi^{-1}(t))}, \qquad \forall x \in X, R > 0, t \in (0, c_2\Psi(R)]. \tag{4.5}$$

<u>STEP 4: PROOF OF (VD) + $(DUHK(\Psi))$ + (EHI) + (E(Ψ)) $\Rightarrow$ $(NLHK(\Psi))$.</u> We follow the arguments in [40, 42]. Fix $x \in X$, $t > 0$ and set $R := \Psi^{-1}(t/\varepsilon)$ where $\varepsilon > 0$ will be chosen later. We can assume $\varepsilon < c_2$ where c_2 is given in (4.5). Hence, by (4.5)

$$p_t^B(x,x) \ge \frac{c_1}{V(x,\Psi^{-1}(t))}, \tag{4.6}$$

where $B := B(x,R)$. Set $f(y) = \partial_t p_t^B(x,y)$. Applying Proposition 9.9 to p_t^B, we have, for $y \in B$,

$$|f(y)| \le \frac{2}{t}\sqrt{p_{t/2}^B(x,x)p_{t/2}^B(y,y)} \le \frac{2}{t}\sqrt{p_{t/2}(x,x)p_{t/2}(y,y)}.$$

By $(DUHK(\Psi))$, we have

$$p_{t/2}(x,x) \le \frac{c_1}{V(x,\Psi^{-1}(t))},$$

and

$$\begin{aligned} p_{t/2}(y,y) &\le \frac{c_1}{V(y,\Psi^{-1}(t))} \le \frac{c_1}{V(x,\Psi^{-1}(t))}\frac{V(x,\Psi^{-1}(t))}{V(y,\Psi^{-1}(t))} \\ &\le \frac{c_1}{V(x,\Psi^{-1}(t))}\Big(1 + \frac{d(x,y)}{\Psi^{-1}(t)}\Big)^{\alpha} \le \frac{c_1(1+\varepsilon^{-\alpha'})^{\alpha}}{V(x,\Psi^{-1}(t))}, \qquad \forall y \in B, \end{aligned}$$

for some $\alpha, \alpha' > 0$ where we used (9.1) and the definition of R and Ψ. Hence, by (VD), we have

$$|f(y)| \le \frac{c_2(1+\varepsilon^{-\alpha'})^{\alpha/2}}{tV(x,\Psi^{-1}(t))}, \qquad \forall y \in B. \tag{4.7}$$

Define $u(y) = p_t^B(x,y)$. Note that $\partial_t u = \Delta_B u$ and the Green operator G^B is a bounded operator in $L^2(B)$ and $G^B = (-\Delta_B)^{-1}$. Thus, $u = -G^B(\partial_t u) = -G^B f$. Let $\gamma > \alpha\alpha'/2$ and apply Proposition 9.6 with $\varepsilon^{\gamma+1}$ instead of ε. Then, there exists $\delta > 0$ such that for any $0 < r < R$,

$$\mathrm{Osc}_{B(x,\delta r)} u \le 2(\bar{E}(x,r) + \varepsilon^{\gamma+1}\bar{E}(x,R)\|f\|_\infty.$$

By (E(Ψ)), we have $\bar{E}(x,r) \le c_3\Psi(r)$ and $\bar{E}(x,R) \le c_3\Psi(R)$. Estimating $\|f\|_\infty$ by (4.7), we obtain

$$\mathrm{Osc}_{B(x,\delta r)} u \le \frac{\Psi(r) + \varepsilon^{\gamma+1}\Psi(R)}{t} \cdot \frac{c_4(1+\varepsilon^{-\alpha'})^{\alpha/2}}{V(x,\Psi^{-1}(t))}.$$

By definition of R, we have

$$\frac{\varepsilon^{\gamma+1}\Psi(R)}{t} = \varepsilon^{\gamma}.$$

Choose r by the equation $\Psi(r) = \varepsilon^{\gamma+1}\Psi(R)$, which implies, by definition of Ψ, $r \geq \delta' R$ for some $\delta' > 0$. Hence, we obtain

$$\mathrm{Osc}_{y\in B(x,\delta\delta' R)} p_t^B(x,y) \leq \mathrm{Osc}_{B(x,\delta r)} u \leq \frac{2c_4\varepsilon^{\gamma}(1+\varepsilon^{-\alpha'})^{\alpha/2}}{V(x,\Psi^{-1}(t))}. \tag{4.8}$$

By the choice of $\gamma > 0$, $\varepsilon^{\gamma}(1+\varepsilon^{-\alpha'})^{\alpha/2} \to 0$ as $\varepsilon \to 0$. So, choosing ε small enough and combining (4.8) with (4.6), we conclude that

$$p_t(x,y) \geq p_t^B(x,y) \geq \frac{c_1/2}{V(x,\Psi^{-1}(t))}, \qquad \forall y \in B(x,\delta\delta' R),$$

which proves $(NLHK(\Psi))$. □

STEP 5: PROOF OF (VD) + $(NLHK(\Psi)) \Rightarrow (LHK(\Psi))$. First, since $h(0) = 0$, $\lim_{t\to\infty} h(t) = \infty$ and h is increasing, for all $t > 0$ and $x \neq y \in X$, there exists $\varepsilon_0 := \varepsilon(t, d(x,y)) > 0$ such that

$$c_1 t \leq h(\varepsilon_0) d(x,y) \leq c_2 t. \tag{4.9}$$

Since there is nothing to prove when $\Psi(d(x,y)) \leq C_6 t$ due to $(NLHK(\Psi))$, we will consider the case $\Psi(d(x,y)) > C_6 t$, which means $\varepsilon_0 < c_3 d(x,y)$ for some $c_3 > 0$. From now on, we take $\varepsilon := \varepsilon(c_* t, d(x,y))$ where $c_* \in (0,1)$ will be chosen later. Since $\varepsilon \leq \varepsilon_0$, we still have $\varepsilon < c_3 d(x,y)$.

For $c_4 \geq 2c_3 \vee 1$, take $N \in \mathbb{N}$ such that

$$\frac{c_3 d(x,y)}{\varepsilon} \leq N \leq \frac{c_4 d(x,y)}{\varepsilon}, \tag{4.10}$$

and let $\{x_i\}_{i=0}^N$ be such that $x_0 = x, x_N = y$ and $d(x_i, x_{i+1}) \leq \varepsilon$ for $i = 0, 1, \cdots, N-1$. Such a sequence exists by the choice of N and by the fact that d is a geodesic metric. We then have

$$\begin{aligned} p_t(x,y) &= \int_X \cdots \int_X p_{t/N}(x,z_1) p_{t/N}(z_1,z_2) \cdots p_{t/N}(z_{N-1},y) d\mu(z_1)\cdots d\mu(z_{N-1}) \\ &\geq \int_{B(x_1,\varepsilon)} \cdots \int_{B(x_{N-1},\varepsilon))} p_{t/N}(x,z_1) p_{t/N}(z_1,z_2) \cdots p_{t/N}(z_{N-1},y) d\mu(z_1)\cdots d\mu(z_{N-1}). \end{aligned} \tag{4.11}$$

Clearly $d(z_i, z_{i+1}) \leq 3\varepsilon$. Now, by (4.9) applied to ε and by (4.10), we have

$$\Psi^{-1}\Big(\frac{c_1 c_3 c_* t}{N}\Big) \leq \varepsilon \leq \Psi^{-1}\Big(\frac{c_2 c_4 c_* t}{N}\Big).$$

By definition of Ψ, taking c_* small, we have $\Psi^{-1}(c_2 c_4 c_* t/N) \leq (C_6/3)\Psi^{-1}(t/N)$, so we conclude

$$\Psi^{-1}\Big(\frac{c_5 t}{N}\Big) \leq \varepsilon \leq \frac{C_6}{3}\Psi^{-1}\Big(\frac{t}{N}\Big). \tag{4.12}$$

Hence, by $(NLHK(\Psi))$, (VD) and (4.12), we have

$$p_{t/N}(z_i, z_{i+1}) \geq \frac{c_6}{V(z_i, \Psi^{-1}(t/N))} \geq \frac{c_7}{V(x_i, \Psi^{-1}(t/N))} \geq \frac{c_8}{V(x_i, \varepsilon)}.$$

Therefore, it follows form (4.11)

$$\begin{aligned} p_t(x,y) &\geq \frac{c_8}{V(x,\Psi^{-1}(t/N))} \prod_{i=1}^{N-1} \frac{c_8}{V(x_i,\varepsilon)} \cdot V(x_i,\varepsilon) \geq \frac{c_8^N}{V(x,\Psi^{-1}(t/N))} \\ &\geq \frac{\exp(-c_9 N)}{V(x,\Psi^{-1}(t))} \geq \frac{\exp(-c_{10} d(x,y)/\varepsilon)}{V(x,\Psi^{-1}(t))}. \end{aligned}$$

On the other hand, by (4.9) applied to ε, we have $h^{-1}(t/d(x,y)) \le c_{11}\varepsilon$, so that

$$\frac{d(x,y)}{\varepsilon} \le c_{11}\frac{d(x,y)}{h^{-1}(t/d(x,y))}.$$

We thus obtain $(LHK(\Psi))$. □

Combining Step 1 –5, the proof of Proposition 4.1 is completed.

Proposition 4.3

$$(VD)+(EHI)+(E(\Psi)) \Rightarrow (DUHK(\Psi)).$$

The proof is given for the case of weighted graphs in [41] and for the case of MMD spaces in [40]. Since the proof is long, here we will additionally assume (PI(Ψ)) and prove the result. (PI(Ψ)) implies $(FK(\Psi))$ – see subsection 8.1 for the definition, so we shall prove the following.

Proof of (VD) + $(FK(\Psi))$ + $(E(\Psi)) \Rightarrow (DUHK(\Psi))$. Fix $x_0 \in X$ and let $0 < r < \rho' < \rho < R$. If we denote $B_s := B(x_0, s)$, then, as in [36] (12.6), we have

$$\sup_{x,y\in B_r} p_t^{B_R}(x,y) \le \sup_{x,y\in B_{\rho'}} p_t^{B_{\rho'}}(x,y) + 2\sup_{x\in B_r}\varphi^{B_{\rho'}}(x_0,t/2)\sup_{t/2\le s\le t}\sup_{x,y\in B_\rho} p_s^{B_R}(x,y),$$

where we denote $\varphi^B(x,t) := \mathbb{P}^x(\tau_B \le t)$. Using the fact $(FK(\Psi)) \Rightarrow (UC(\Psi))$ in Theorem 8.1,

$$\sup_{x,y\in B_{\rho'}} p_t^{B_{\rho'}}(x,y) \le \sup_{x,y\in B_\rho} p_t^{B_\rho}(x,y) \le \frac{c_1}{V(x_0,\Psi^{-1}(t))} \qquad \forall t \le \Psi(\rho).$$

By (E(Ψ)) and $(E(\Psi)) \Rightarrow (ELD(\Psi))$ (Step 1 above), for $x \in B_r$,

$$\varphi^{B_{\rho'}}(x,t/2) \le \varphi^{B(x,\rho'-r)}(x,t/2) \le \frac{1}{4K}, \qquad \forall \rho'-r \ge M\Psi^{-1}(\frac{t}{2M}),$$

if M is large. This is the case if

$$\rho - r \ge M\Psi^{-1}(t) \tag{4.13}$$

and ρ' is sufficiently close to ρ. Noting that the function $s \mapsto \sup_{x,y\in B_r} p_s^{B_R}(x,y)$ is non-increasing, we obtain

$$\sup_{x,y\in B_r} p_t^{B_R}(x,y) \le \frac{c_1}{V(x_0,\Psi^{-1}(t))} + \frac{1}{2K}\sup_{t/2\le s\le t}\sup_{x,y\in B_\rho} p_s^{B_R}(x,y) \le \frac{c_1}{V(x_0,\Psi^{-1}(t))} + \frac{1}{2K}\sup_{x,y\in B_\rho} p_{t/2}^{B_R}(x,y), \tag{4.14}$$

for all $t \le \Psi(\rho)$.

Now, for a fixed $t > 0$, set $t_n := t/2^n$, $n \ge 0$ and

$$r_n := M\sum_{i=0}^{n-1}\Psi^{-1}(t_i), \qquad n \ge 1.$$

It follows by this and the definition of Ψ that

$$r_n \le 2M\int_0^{2t}\Psi^{-1}(s)\frac{ds}{s} =: I(t) < \infty.$$

Assume that $R \ge I(t)$ so that all the balls $B_n := B(x_0, r_n)$ are in B_R. Using the fact $r_{n+1} - r_n = M\Psi^{-1}(t_n)$, which matches (4.13) and the fact $t_n \le \Psi(r_{n+1})$, we obtain from (4.14)

$$\sup_{x,y\in B_n} p_{t_n}^{B_R}(x,y) \le \frac{c_1}{V(x_0,\Psi^{-1}(t_n))} + \frac{1}{2K}\sup_{x,y\in B_{n+1}} p_{t_{n+1}}^{B_R}(x,y). \tag{4.15}$$

By (VD), we have

$$\frac{c_1}{V(x_0,\Psi^{-1}(t_n))} \le \frac{c_1 K}{V(x_0,\Psi^{-1}(t_{n-1}))} \le \cdots \le K^n \frac{c_1}{V(x_0,\Psi^{-1}(t_0))} = K^n \frac{c_1}{V(x_0,\Psi^{-1}(t))}.$$

Thus, we have

$$\sup_{x,y\in B_n} p_{t_n}^{B_R}(x,y) \le K^n \frac{c_1}{V(x_0,\Psi^{-1}(t))} + \frac{1}{2K} \sup_{x,y\in B_{n+1}} p_{t_{n+1}}^{B_R}(x,y).$$

By iteration, we obtain

$$\sup_{x,y\in B_0} p_{t_n}^{B_R}(x,y) \le \frac{c_1}{V(x_0,\Psi^{-1}(t))}\sum_{i=0}^{n-1}(1/2)^i + (\frac{1}{2K})^n \sup_{x,y\in B_n} p_{t_n}^{B_R}(x,y). \tag{4.16}$$

Applying $(FK(\Psi))$, Theorem 8.1 and using (4.15),

$$\sup_{x,y\in B_n} p_{t_n}^{B_R}(x,y) \le \sup_{x,y\in B_R} p_{t_n}^{B_R}(x,y) \le \frac{c_1}{V(x_0,\Psi^{-1}(t_n))} \le \frac{c_1 K^n}{V(x_0,\Psi^{-1}(t))},$$

since $t_n \le \Psi(R)$. Hence, $\lim_{n\to\infty}(2K)^{-n}\sup_{x,y\in B_n} p_{t_n}^{B_R}(x,y) = 0$, and taking $n\to\infty$ in (4.16), we conclude

$$\sup_{x,y\in B_0} p_t^{B_R}(x,y) \le \frac{2c_1}{V(x_0,\Psi^{-1}(t))}.$$

Finally, taking $R\to\infty$ and noticing $p_t^{B_R} \to p_t$, we obtain the desired estimate. □

4.2 Proof of $(c) \Rightarrow (d)$

Lemma 4.4

$$(VD) + (PI(\Psi)) + (CS(\Psi)) \Rightarrow (RES(\Psi)).$$

Proof. We first prove the following. If X satisfy (VD) and (PI(Ψ)), then the following holds.

$$R(B(x_0,R),B(x_0,2R)^c) \le c_1 \frac{\Psi(R)}{V(x_0,R)}, \quad \forall x_0\in X, R\ge 0. \tag{4.17}$$

Let f be the function which attains the minimum on the right hand side of (3.4) when $A = B(x_0,R)$ and $B = B(x_0,2R)^c$. Let $\overline{f} = \int_{B(x_0,3R)} f d\mu/V(x_0,3R)$. Choose y_0 so that $d(x_0,y_0) = 5R/2$. Then by (9.1) we have $V(y_0,R/2) \ge c_2 V(x_0,R)$. Depending on whether $\overline{f} \ge 1/2$ or $\overline{f} < 1/2$, $|f-\overline{f}| \ge 1/2$ on either $B(x_0,R)$ or $B(y_0,R/2)$, and then using (PI(Ψ)) we have

$$\begin{aligned} V(x_0,R) &\le c_3\int_{B(x_0,3R)}(f-\overline{f})^2 d\mu \le c_4\Psi(R)\int_{B(x_0,3R)} d\Gamma(f,f) \\ &= c_4\Psi(R)R(B(x_0,R),B(x_0,2R)^c)^{-1}. \end{aligned}$$

So (4.17) is proved.

We next prove the following. If X satisfy (VD) and (CS(Ψ)), then the following holds.

$$R(B(x_0,R),B(x_0,2R)^c) \ge c_5 \frac{\Psi(R)}{V(x_0,R)}, \quad \forall x_0\in X, R\ge 0. \tag{4.18}$$

Let φ be a cut-off function for $B(x_0,R)$ given by (CS(Ψ)). Then taking $f\equiv 1$, $I = B(x_0,R)$ and $I^* = B(x_0,2R)$ in (3.5) we obtain

$$R(B(x_0,R/2),B(x_0,R)^c)^{-1} \le \int_I d\Gamma(\varphi,\varphi) \le c_6\Psi(R)^{-1}\int_{I^*} d\mu \le c_7\frac{V(x_0,R)}{\Psi(R)},$$

where (VD) was used in the last inequality. So (4.18) is proved. □

By Lemma 4.4, the rest is to show (VD) + (PI(Ψ)) + (CS(Ψ)) $\Rightarrow$ (EHI). This is the highlight of this section. Recall the Moser's argument in subsection 2.4. The crucial loss for the case $\beta \neq 2$ is in using the bound (2.6); one needs a cutoff function φ such that the final term in (2.7) can be controlled by a term of order $R^{-\beta}$. We shall now see how the (CS(Ψ)) enables one to do this. (Clearly, $(CS(\Psi))$ guarantees the existence of 'nice' cut-off functions $\varphi = \varphi_{x,R}$ that satisfies $\mathcal{E}(\varphi,\varphi) \le c_1 \Psi(R)^{-1} V(x,R)$ for each $x \in X$ and $R > 0$.)

For $x \in X$, $R > 0$ let $\varphi = \varphi_{x,R}$ be the cut-off function in (CS(Ψ)). We define the measure $\gamma = \gamma_{x,R}$ by

$$d\gamma = d\mu + \Psi(R) d\Gamma(\varphi,\varphi).$$

We remark that we do not know if the measure γ satisfies volume doubling. The first step in the argument is to use (CS(Ψ)) to obtain a weighted Sobolev inequality. For any set $J \subset X$ set

$$J^s = \{y : d(y,J) \le s\}.$$

Proposition 4.5 *Let $s \le R$ and $J \subset B(x_0,R)$ be a finite union of balls of radius s. There exist $\kappa > 1$ and $c_1 > 0$ such that*

$$\Big(\mu(J)^{-1}\int_J |f|^{2\kappa} d\gamma\Big)^{1/\kappa} \le c_1\Big(\Psi(R)\mu(J)^{-1}\int_{J^s} d\Gamma(f,f) + (s/R)^{-2\theta}\mu(J)^{-1}\int_J f^2 d\gamma\Big).$$

The strategy of the proof is to show weighted Poincaré inequalities first and then prove the weighted Nash inequality, which deduce the desired inequality. See subsection 9.8 for details.

The next result is the generalization of Lemma 4 of [69] to the case of a MMD space.

Lemma 4.6 *Let D be a domain in X, let u be positive and harmonic in D, $v = u^k$, where $k \in \mathbb{R}$, $k \neq \frac{1}{2}$, and let η be supported in D. Suppose $\int_D d\Gamma(\eta,\eta) < \infty$, then*

$$\int_D \eta^2 d\Gamma(v,v) \le \Big(\frac{2k}{2k-1}\Big)^2 \int_D v^2 d\Gamma(\eta,\eta).$$

Proof. Let $g \in \mathcal{F}$ be supported by D. Then if $u' = Gh$ where $h = 0$ on D we have

$$\int_D d\Gamma(gu',u') = \int_X d\Gamma(gu',u') = \int_X gu'h\,d\mu = 0.$$

Hence, approximating u by functions of the form u' we deduce that

$$\int_D d\Gamma(gu,u) = 0.$$

Using this, and taking $g = \eta^2 k^2 u^{2k-2}$, we conclude that

$$\int_D \eta^2\, d\Gamma(v,v) = \int_D g\, d\Gamma(u,u) = -\int_D u\, d\Gamma(g,u). \tag{4.19}$$

Using the Leibniz and chain rules, the right hand side is equal to

$$-2k\int_D \eta v\, d\Gamma(\eta,v) - (2k-2)\int_D \eta^2 d\Gamma(v,v).$$

Thus,

$$\begin{aligned}\int_D \eta^2 d\Gamma(v,v) &= -\frac{2k}{2k-1}\int_D \eta v\, d\Gamma(v,\eta)\\ &\le \frac{2|k|}{|2k-1|}\Big(\int_D \eta^2 d\Gamma(v,v)\Big)^{1/2}\Big(\int_D v^2 d\Gamma(\eta,\eta)\Big)^{1/2},\end{aligned}$$

where we used Cauchy-Schwarz. Dividing and squaring, we obtain the result. $\square$

Let u be harmonic and nonnegative in $B(x_0, 4R)$. By looking at $u+\varepsilon$ and letting $\varepsilon \downarrow 0$ we may without loss of generality suppose u is strictly positive. Note that, as for a general MMD space we do not initially have any *a priori* continuity for u, we do not obtain a pointwise bound in (4.20).

Proposition 4.7 *Let v be either u or u^{-1}. There exists c_1 such that if $B(x, 2r) \subset B(x_0, 4R)$ and $0 < q < 2$, then*

$$\text{ess sup}_{B(x,r/2)} v^{2q} \le c_1 V(x, 2r)^{-1} \int_{B(x,2r)} \Big(\Psi(r) d\Gamma(v^q, v^q) + v^{2q} d\mu \Big). \tag{4.20}$$

Proof. Let φ_0 be a (regularized) cut-off function given by (CS(Ψ)) for $B(x, r)$. Let $h_n = 1 - 2^{-n}$, $0 \le n \le \infty$, so that $0 = h_0 < h_\infty = 1$. For $k \ge 0$ set

$$\varphi_k(x) = (\varphi_0(x) - h_k)^+, \qquad d\gamma_0 = d\mu + \Psi(r) d\Gamma(\varphi_0, \varphi_0).$$

Set $A_k = \{x : \varphi_0(x) > h_k\}$, and note that $B(x, r/2) \subset A_{n_0} \subset A_0 \subset B(x, r)$ for every n_0. We therefore have, writing V for $V(x, r)$,

$$c_2 V \le \mu(A_k) \le V, \quad k \ge 0.$$

The Hölder condition on φ_0 given by (CS(Ψ)) implies that if $x \in A_{k+1}$ and $y \in A_k^c$, then $d(x, y) \ge c_3 r 2^{-k/\theta}$. Set $s_k = \frac{1}{2} c_3 r 2^{-k/\theta}$, and note that $\varphi_k > c_4 2^{-k}$ on $A_{k+1}^{s_k}$. Let $\{B_i\}$ be a cover of A_{k+1} by balls of radius $s_k/2$, and let $J_{k+1} = \cup_i B_i$. Write $J'_{k+1} = J_{k+1}^{s_k/2}, A'_{k+1} = A_{k+1}^{s_k}$ and note that $A_{k+1} \subset J_{k+1} \subset J'_{k+1} \subset A'_{k+1}$.

From Proposition 4.5 with $f = v^p$ and s replaced by $s_k/2$,

$$\begin{aligned} \Big(V^{-1} \int_{A_{k+1}} f^{2\kappa} d\gamma_0 \Big)^{1/\kappa} &\le \Big(V^{-1} \int_{J_{k+1}} f^{2\kappa} d\gamma_0 \Big)^{1/\kappa} \\ &\le c_5 V^{-1} \Big[\Psi(r) \int_{J'_{k+1}} d\Gamma(f, f) + (r/s_k)^{2\theta} \int_{J'_{k+1}} f^2 d\gamma_0 \Big] \\ &\le c_6 V^{-1} \Big[\Psi(r) \int_{A'_{k+1}} d\Gamma(f, f) + 2^{2k} \int_{A_k} f^2 d\gamma_0 \Big]. \end{aligned} \tag{4.21}$$

By Lemma 4.6, we have the 'converse to the Poincaré inequality' for $f = v^p$, which controls the first term in (4.21).

$$\begin{aligned} \Psi(r) \int_{A'_{k+1}} d\Gamma(f, f) &\le \Psi(r)(c_7 2^{-k})^{-2} \int_{A'_{k+1}} \varphi_k^2 d\Gamma(f, f) \le c_8 2^{2k} \Psi(r) \int_{A_k} \varphi_k^2 d\Gamma(f, f) \\ &\le c_9 2^{2k} \Psi(r) \Big(\frac{2p}{2p-1} \Big)^2 \int_{A_k} f^2 d\Gamma(\varphi_k, \varphi_k) \le c_{10} 2^{2k} \Big(\frac{2p}{2p-1} \Big)^2 \int_{A_k} f^2 d\gamma_0. \end{aligned}$$

We therefore deduce that

$$\Big(V^{-1} \int_{A_{k+1}} f^{2\kappa} d\gamma_0 \Big)^{1/\kappa} \le c_{11} \Big(\frac{2p}{2p-1} \Big)^2 2^{2k} V^{-1} \int_{A_k} f^2 d\gamma_0. \tag{4.22}$$

We now make an argument similar to the first part of Moser's argument [69] mentioned in subsection 2.4. Choose $q' > 0$ such that $\inf_{m \in \mathbb{Z}} |q' \kappa^m - \frac{1}{2}| \ge c_{12} > 0$. Suppose first that $q_0 = q' \kappa^{-i}$ for some i. Let $p_n = 2 q_0 \kappa^n$ for $n \ge 0$, and write

$$\Psi_k = \Big[\mu(A_k)^{-1} \int_{A_k} v^{p_k} d\gamma_0 \Big]^{1/p_k}.$$

Note that $p_{k+1}/2\kappa = p_k/2$. Applying (4.22) to $f = v^{p_{k+1}/(2\kappa)} = v^{p_k/2}$ we have

$$\Psi_{k+1}^{p_{k+1}/\kappa} = \Big(\mu(A_{k+1})^{-1} \int_{A_{k+1}} v^{p_{k+1}} d\gamma_0 \Big)^{1/\kappa} \le c_{13} 2^{2k} \Big(\mu(A_k)^{-1} \int_{A_k} v^{p_k} d\gamma_0 \Big) = c_{13} 2^{2k} \Psi_k^{p_k},$$

or

$$\Psi_{k+1} \le \Big(c_{13} 2^{2k}\Big)^{1/p_k} \Psi_k.$$

Hence for every m

$$\log \Psi_m \le \log \Psi_0 + \sum_{k=1}^m p_k^{-1} \log(c_{13} 2^{2k}). \tag{4.23}$$

As the sum in (4.23) converges and ess $\sup_{B(x,r/2)} v \le \limsup_{m\to\infty} \Psi_m$, we have

$$\text{ess sup}_{B(x,r/2)} v \le c_{14} \Psi_0 \le c_{15} \Big(V^{-1} \int_{B(x,r)} v^{2q_0} d\gamma_0\Big)^{1/(2q_0)}.$$

Now let $q \in (0,2)$. We can take $q_0 = q'\kappa^{-i} < q$. Then by Hölder's inequality, and Proposition 9.20 (d),

$$\begin{aligned} V^{-1} \int_{B(x,r)} v^{2q_0} d\gamma_0 &\le \Big(V^{-1} \int_{B(x,r)} v^{2q} d\gamma_0\Big)^{q_0/q} \Big(V^{-1} \int_{B(x,r)} d\gamma_0\Big)^{1-q_0/q} \\ &\le c_{16} \Big(V^{-1} \int_{B(x,r)} v^{2q} d\gamma_0\Big)^{q_0/q}. \end{aligned}$$

Thus

$$\text{ess sup}_{B(x,r/2)} v^{2q} \le c_{17} V^{-1} \int_{B(x,r)} v^{2q} d\gamma_0.$$

By Proposition 9.20 (a) with $R = s = r$ and (VD) this implies

$$\text{ess sup}_{B(x,r/2)} v^{2q} \le c_{18} V(x,2r)^{-1} \int_{B(x,2r)} (\Psi(r) d\Gamma(v^q, v^q) + v^{2q} d\mu).$$

□

Recall that φ is a cut-off function for $B(x_0,R)$ given by (CS(Ψ)). We define

$$Q(t) = \{x : \varphi(x) > t\}, \quad 0 < t < 1,$$

and write $Q(1)$ for the interior of $\{x : \varphi(x) \ge 1\}$.

Corollary 4.8 *Let $1 > s > t > 0$. There exists $\zeta > 2$ such that if $0 < q < \frac{1}{3}$,*

$$\textit{ess sup}_{Q(s)} v^{2q} \le c_1 (s-t)^{-\zeta} V(x_0,R)^{-1} \int_{Q(t)} v^{2q} d\gamma. \tag{4.24}$$

Proof. By the maximum principle the essential supremum of v^{2q} in $\overline{Q(s)}$ is equal to an essential supremum around a point $x' \in \partial Q(s)$. Let $\eta = \frac{1}{4}(s-t)$, $s' = s - 2\eta$. By the Hölder continuity of φ the sets $Q(s)$ and $Q(s')^c$ are separated by a distance of at least $\xi = c_2 R (s-t)^{1/\theta}$, so that $B(x',\xi) \subset Q(s')$. By Proposition 4.7,

$$\text{ess sup}_{B(x',\xi/4)} v^{2q} \le c_3 \Psi(\xi) V(x',\xi)^{-1} \int_{B(x',\xi)} d\Gamma(v^q,v^q) + c_3 V(x',\xi)^{-1} \int_{B(x',\xi)} v^{2q} d\mu. \tag{4.25}$$

Note that by (9.1) we have

$$\frac{V(x_0,R)}{V(x',\xi)} \le c_4 \Big(\frac{d(x',x_0)+R}{\xi}\Big)^{\alpha} \le c_5 (s-t)^{-\alpha/\theta}. \tag{4.26}$$

Using (4.25),

$$\text{ess sup}_{Q(s)} v^{2q} \le c_6 \xi^{\Psi} V(x',\xi)^{-1} \int_{Q(s')} d\Gamma(v^q,v^q) + c_6 V(x',\xi)^{-1} \int_{Q(s')} v^{2q} d\mu.$$

Let

$$\varphi_{st} = (s \wedge \varphi - t)^+$$

and observe that $\int_B d\Gamma(\varphi_{st}, \varphi_{st}) \le \int_B d\Gamma(\varphi, \varphi)$ for any B. Since $\varphi_{st} \ge c_7(s-t)$ on $Q(s')$, using Lemma 4.6, we have the "converse to the Poincaré inequality" for v^q;

$$\begin{aligned}\int_{Q(s')} d\Gamma(v^q, v^q) &\le c_7(s-t)^{-2} \int_{Q(s')} \varphi_{st}^2 d\Gamma(v^q, v^q) \le c_7(s-t)^{-2} \int_{Q(t)} \varphi_{st}^2 d\Gamma(v^q, v^q)\\ &\le c_8(s-t)^{-2} \int_{Q(t)} v^{2q} d\Gamma(\varphi_{st}, \varphi_{st}) \le c_9(s-t)^{-2}\Psi(R)^{-1} \int_{Q(t)} v^{2q} d\gamma.\end{aligned}$$

Thus, noting $\Psi(\xi/R) = \Psi(c_2(s-t)^{1/\theta}) \le c_{10}$,

$$\begin{aligned}\text{ess sup}_{Q(s)} v^{2q} &\le c_{11}\Psi(\xi/R)(s-t)^{-2} V(x', \xi)^{-1} \int_{Q(t)} v^{2q} d\gamma + c_{11} V(x', \xi)^{-1} \int_{Q(t)} v^{2q} d\mu\\ &\le c_{12} V(x', \xi)^{-1}(s-t)^{-2} \int_{Q(t)} v^{2q} d\gamma\\ &\le c_{13} V(x_0, R)^{-1}(s-t)^{-2-\alpha/\theta} \int_{Q(t)} v^{2q} d\gamma,\end{aligned}$$

where we used (4.26) in the last inequality. So taking $\zeta_1 = 2 + \alpha/\theta$ we obtain (4.24). □

Now our goal is to deduce the elliptic Harnack inequality. The following corresponds to the second part of Moser's arguments.

Let $w = \log u$, and write $\overline{w} = V(x_0, R)^{-1} \int_{B(x_0,R)} w \, d\mu$.

Proposition 4.9 *(a) There exists c_1 such that*

$$\int_{B(x_0,2R)} d\Gamma(w, w) \le c_1 \frac{V(x_0, R)}{\Psi(R)}.$$

(b) Let $1 \ge s > t > 0$. Then

$$\int_{\{|w-\overline{w}|>A\}\cap Q(s)} d\gamma \le c_2 \frac{V(x_0, R)}{A^2}.$$

Proof. Again, this is essentially Moser's proof. Let $\varphi_1(x)$ be a cut-off function given by (CS(Ψ)) for the ball $B^* := B(x_0, 4R)$. So

$$\int_{B(x_0,2R)} d\Gamma(w, w) \le c \int_{B^*} \varphi_1^2 d\Gamma(w, w).$$

Applying (4.19) with $\eta = \varphi_1$, $v = w$, $g = \varphi_1^2/u^2$ and $D = B^*$, we have

$$\int_{B^*} \varphi_1^2 d\Gamma(w, w) = -\int_{B^*} u d\Gamma(\varphi_1^2/u^2, u).$$

Using the Leibniz and chain rules, the right hand side is equal to

$$-2\int_{B^*} \varphi_1 d\Gamma(\varphi_1, w) + 2\int_{B^*} \varphi_1^2 d\Gamma(w, w).$$

Thus,

$$\int_{B^*} \varphi_1^2 d\Gamma(w, w) = 2\int_{B^*} \varphi_1 d\Gamma(\varphi_1, w) \le 2\Big(\int_{B^*} d\Gamma(\varphi_1, \varphi_1)\Big)^{1/2}\Big(\int_{B^*} \varphi_1^2 d\Gamma(w, w)\Big)^{1/2},$$

where we used Cauchy-Schwarz. Dividing and squaring,

$$\int_{B^*} \varphi_1^2 d\Gamma(w,w) \le 4 \int_{B^*} d\Gamma(\varphi_1,\varphi_1).$$

Finally, using (CS(Ψ)) in B^* with $f \in \mathcal{F}$ such that $f|_{B(x_0,8R)} \equiv 1$ (since $(\mathcal{E},\mathcal{F})$ is a regular Dirichlet form, such an f exists) and (VD) we deduce that

$$\int_{B^*} d\Gamma(\varphi_1,\varphi_1) \le c\Psi(R)^{-1}V(x_0,R).$$

(b) By Chebyshev's inequality, Proposition 9.20 (b) and (a)

$$\begin{aligned} A^2 \int_{\{|w-\overline{w}|>A\}\cap Q(s)} d\gamma &\le \int_{\{|w-\overline{w}|>A\}\cap Q(s)} |w-\overline{w}|^2 d\gamma \\ &\le \int_{Q(s)} |w-\overline{w}|^2 d\gamma \le \int_{B(x_0,R)} |w-\overline{w}|^2 d\gamma \\ &\le c_5\Psi(R) \int_{B(x_0,2R)} d\Gamma(w,w) \le c_6 V(x_0,R). \end{aligned}$$

□

In order to get the Harnack inequality the argument in [68] required a generalization of the John-Nirenberg inequality with a complicated proof. Bombieri [22] found a way to avoid such an argument for elliptic second order differential equations. Moser (Lemma 3 in [67]) carried the idea over to the parabolic case and Bombieri and Giusti (Theorem 4 in [23]) obtained the inequality in an abstract setting. (See also Lemma 2.2.6 in [72].) This argument can be applied to our setting (with suitable modifications) and we can show that Corollary 4.8 and Proposition 4.9 (b) give

$$\text{ess sup}_{B(x_0,R/2)} \log u \le c_1. \tag{4.27}$$

(For the sake of completeness, we will give the proof of (4.27) in subsection 9.9.) Let $v = u^{-1}$. The same argument implies ess $\sup_{B(x_0,R/2)} \log v \le c_1$, or ess $\inf_{B(x_0,R/2)} \log u \ge -c_1$. Combining we deduce

$$e^{-c_1} \le \text{ess inf}_{B(x_0,R/2)} u \le \text{ess sup}_{B(x_0,R/2)} u \le e^{c_1}.$$

We thus obtain the following.

Theorem 4.10 *There exists c_1 such that if u is nonnegative and harmonic in $B(x_0,4R)$, then*

$$\textit{ess sup}_{B(x_0,R/2)} u \le c_1 \textit{ess inf}_{B(x_0,R/2)} u.$$

Proof of $(c) \Rightarrow (d)$. As we mentioned in the beginning of this section, it is enough to show (VD) + (PI(Ψ)) + (CS(Ψ)) $\Rightarrow$ (EHI). But given Theorem 4.10, (EHI) can be proved as in subsection 9.3 □

4.3 Proof of $(b) \Rightarrow (c)$

In this subsection, we will use the equivalence $(a) \Leftrightarrow (b)$ which is proved in Appendix 2 (Section 9).

Assuming (b) or equivalently (a), (VD) and (PI(Ψ)) hold by standard arguments (which are partly discussed in subsection 9.7). So, we will prove (PHI(Ψ)) (equivalently (HK(Ψ))) $\Rightarrow$ (CS(Ψ)).

Let $D = B(x_0, R-\varepsilon)$ where $\varepsilon < R/10$, and $\lambda > 0$. Let Y be the process associated with the Dirichlet form $(\mathcal{E},\mathcal{F})$. Let G_λ^D be the resolvent associated with the process Y killed on exiting D; that is,

$$G_\lambda^D f(x) = E^x \int_0^{\tau_D} e^{-\lambda t} f(Y_t) dt,$$

for bounded measurable f, where $\tau_D = \inf\{t : Y_t \in X - D\}$. Let $p_t^D(\cdot,\cdot)$ be the heat kernel of Y killed on exiting D. Then the Green kernel of G_λ^D is given by

$$g_\lambda^D(x,y) = \int_0^\infty e^{-\lambda t} p_t^D(x,y)dt.$$

We use the Green kernel to build a cut-off function φ.

Lemma 4.11 *Let $x_0 \in X$. Then there exists $\delta > 0$ such that if $\lambda = c_0 \Psi(R)^{-1}$*

$$g_\lambda^D(x_0,y) \le C_1 \frac{\Psi(R)}{V(x_0,R)}, \qquad y \in B(x_0,\delta R)^c,$$

$$g_\lambda^D(x_0,y) \ge C_2 \frac{\Psi(R)}{V(x_0,R)}, \qquad y \in B(x_0,\delta R).$$

Proof. This follows easily from (HK(Ψ)) by integration. □

Lemma 4.12 *Let x_0 and R be as above, and let $x,y \in B(x_0,\delta R)^c$. Then there exists $\theta > 0$ such that*

$$|g_\lambda^D(x_0,x) - g_\lambda^D(x_0,y)| \le c_1 \Big(\frac{d(x,y)}{R}\Big)^\theta \sup_{B(x_0,\delta R)^c} g_\lambda^D(x_0,.). \tag{4.28}$$

Proof. The Hölder continuity of p_t^D follows from (PHI(Ψ)) by a standard argument. Integrating we obtain (4.28). □

Fix $x_0 \in X$ and let $B' = B(x_0,\delta R)$, $B = B(x_0,R)$, $D = B(x_0,R-\varepsilon)$ where $\varepsilon < R/10$. Let $\lambda = c_0\Psi(R)^{-1}$ and define

$$\varphi(x) = 1 \wedge (c\Psi(R)^{-1} G_\lambda^D 1_{B'}(x)),$$

where c is chosen so that $\varphi(x) = 1$ on $x \in B'$. Using Lemmas 4.11 and 4.12, it is easy to check that φ is a cut-off function for $B' \subset B$ that satisfies subsection 3.2 (VI) (a)–(c). To complete the proof of (CS(Ψ)), we need to establish (3.5).

Proposition 4.13 *Let $x_1 \in X$ and $f \in \mathcal{F}$. Let δ be defined by Lemma 4.11 and let $I = B(x_1,\delta s)$ with $0 < s \le R$ and $I^* = B(x_1,s)$. There exist $c_1,c_2 > 0$ such that for all $f \in \mathcal{F}$,*

$$\int_I f^2 d\Gamma(\varphi,\varphi) \le c_1 (s/R)^{2\theta} \Big(\int_{I^*} d\Gamma(f,f) + c_2\Psi(s)^{-1} \int_{I^*} f^2 d\mu \Big). \tag{4.29}$$

Proof. **Case 1.** We first consider the case where $s = R$ and $x_1 = x_0$. Let

$$\mathcal{F}_D = \{f \in \mathcal{F} : \widetilde{f} = 0 \text{ q.e. on } X - D\}.$$

Set

$$\mathcal{E}_\lambda(f,g) = \mathcal{E}(f,g) + \lambda \int fg\, d\mu.$$

Let $v = G_\lambda^D 1_{B'}$. Note that

$$v(x) \le \int_{B'} g^D(x,y) d\mu(y) \le E^x[\tau_D] \le c\Psi(R), \qquad x \in D, \tag{4.30}$$

by the fact (VD) + $(DUHK(\Psi)) \Rightarrow$ (E(Ψ)$_\le$) – see subsection 9.2. By [35] Theorem 4.4.1, $v \in \mathcal{F}_D$ and is quasi-continuous. Further, since Y is continuous, $v = 0$ on $\overline{D}^c$. Let $f \in \mathcal{F}$. Then

$$\int_B f^2 d\Gamma(v,v) \le \int_X f^2 d\Gamma(v,v) = \int_X d\Gamma(f^2 v, v) - \int_X 2fv d\Gamma(f,v).$$

Since $v \in \mathcal{F}_D$ we have $f^2 v \in \mathcal{F}_D$, so by [35] Theorem 4.4.1,

$$\int_X d\Gamma(f^2 v, v) = \mathcal{E}(f^2 v, G_\lambda^D 1_{B'}) \le \mathcal{E}_\lambda(f^2 v, G_\lambda^D 1_{B'}) = \int_X f^2 v 1_{B'} d\mu \le c\Psi(R) \int_{B'} f^2 d\mu,$$

where we used (4.30) in the last inequality. Using Cauchy-Schwarz and (4.30), we obtain

$$\begin{aligned} \Big|\int_X 2fv d\Gamma(f,v)\Big| &\le c\Big(\int_X v^2 d\Gamma(f,f)\Big)^{1/2}\Big(\int_X f^2 d\Gamma(v,v)\Big)^{1/2} \\ &\le c\Psi(R)\Big(\int_B d\Gamma(f,f)\Big)^{1/2}\Big(\int_X f^2 d\Gamma(v,v)\Big)^{1/2}. \end{aligned}$$

So, writing $H = \int_X f^2 d\Gamma(v,v)$, $J = \int_B d\Gamma(f,f)$, $K = \int_B f^2 d\mu$, we have

$$H \le c\Psi(R)K + c\Psi(R)J^{1/2}H^{1/2},$$

from which it follows that $H \le c\Psi(R)K + c\Psi(R)^2 J$. From this, (4.29) with $s = R$ follows easily.

Case 2. Define

$$Q(b) = Q(x_0, b) = \{y : g_\lambda^D(x_0, y) > b\}.$$

and let

$$h = C_2\Psi(R)/(2V(x_0,R)),$$

where C_2 is as in Lemma 4.11. Note that by Lemma 4.11 and the fact $g_\lambda^D(x_0,y) = 0$ for $y \notin D$,

$$B(x_0, \delta R) \subset Q(2h) \subset Q(h) \subset B(x_0, R).$$

In Case 2, we will consider the situation that either

$$I^* \subset Q(2h) \tag{4.31}$$

or

$$I^* \cap B(x_0, \delta R/2) = \emptyset \tag{4.32}$$

hold. Since $\varphi \equiv 1$ on $Q(2h)$, (4.29) is clear if (4.31) holds. Thus, we consider when (4.32) holds. Let $\psi_s(x) = 1 \wedge (c\Psi(s)^{-1} G_\lambda^{B(x_0, s-\epsilon)} 1_I(x))$ be a cut-off function for $I \subset I^*$ given by Case 1. Let $\varphi_0(x) = \Psi(R)^{-1} G_\lambda^D 1_{B''}(x)$ where $B'' = B(x_0, \delta R/2)$ and $\varphi_1(x) = \varphi_0(x) - \min_{y \in I^*} \varphi(y)$, then by Lemma 4.12,

$$\varphi_1(x) \le c(s/R)^\theta = L, \qquad x \in I^*.$$

Let

$$\begin{aligned} A &= \int_I f^2 d\Gamma(\varphi, \varphi), \\ D &= \int_{I^*} d\Gamma(f,f) + \Psi(s)^{-1} \int_{I^*} f^2, \\ F &= \int_{I^*} f^2 \psi_s^2 d\Gamma(\varphi_1, \varphi_1). \end{aligned}$$

Now as

$$d\Gamma(f^2\psi_s^2\varphi, \varphi) \le d\Gamma(f^2\psi_s^2\varphi_1, \varphi_0) = f^2\psi_s^2 d\Gamma(\varphi_1, \varphi_0) + \varphi_1 d\Gamma(f^2\psi_s^2, \varphi_0),$$

we have

$$A \le F = \int_{I^*} f^2\psi_s^2 d\Gamma(\varphi_1, \varphi_0) = \int_{I^*} d\Gamma(f^2\psi_s^2\varphi_1, \varphi_0) - \int_{I^*} \varphi_1 d\Gamma(f^2\psi_s^2, \varphi_0). \tag{4.33}$$

For the first term in (4.33)

$$\begin{aligned}\int_{I^*} d\Gamma(f^2\psi_s^2\varphi_1,\varphi_0) &= \int_X d\Gamma(f^2\psi_s^2\varphi_1,\varphi_0)\\ &= \mathcal{E}_\lambda(f^2\psi_s^2\varphi_1,\Psi(R)^{-1}G_\lambda^D 1_{B''}) - \lambda\int_X f^2\psi_s^2\varphi_1\varphi_0 d\mu\\ &\le \mathcal{E}_\lambda(f^2\psi_s^2\varphi_1,\Psi(R)^{-1}G_\lambda^D 1_{B''}) = \Psi(R)^{-1}\int_{B''} f^2\psi_s^2\varphi_1 d\mu = 0.\end{aligned}$$

Here we used the fact that $\varphi_1 \ge 0$ on I^* and that the support of ψ_s is in I^*, hence outside B'' (due to (4.32)).

The final term in (4.33) is handled, using the Leibniz and chain rules and Cauchy-Schwarz, as

$$\begin{aligned}&\Big|\int_{I^*}\varphi_1 d\Gamma(f^2\psi_s^2,\varphi_0)\Big| \le 2\Big|\int_{I^*}\varphi_1 f\psi_s^2 d\Gamma(f,\varphi_0)\Big| + 2\Big|\int_{I^*}\varphi_1 f^2\psi_s d\Gamma(\psi_s,\varphi_0)\Big|\\ \le\ & c\Big\{\Big(\int_{I^*}\psi_s^2 d\Gamma(f,f)\Big)^{1/2} + \Big(\int_{I^*} f^2 d\Gamma(\psi_s,\psi_s)\Big)^{1/2}\Big\}\Big(\int_{I^*}\varphi_1^2 f^2\psi_s^2 d\Gamma(\varphi_0,\varphi_0)\Big)^{1/2}\\ \le\ & cD^{1/2}LF^{1/2},\end{aligned}$$

where we used Case 1 in the final line. Thus we obtain $A \le F \le cDL^2$ so that (4.29) holds.

Case 3. We finally consider the general case. When either (4.31) or (4.32) holds, the result is already proved in Case 2. So assume that neither of them hold. Then I^* must intersect both $B(x_0,\delta R/2)$ and $B(x_0,\delta R)^c$, so $s \ge \delta R/4$. We use Lemma 9.2 to cover I with balls $B_i = B(x_i, c_1 R)$, where $c_1 \in (0,\delta/4)$ has been chosen small enough so that each $B_i^* := B(x_i, c_1 R/\delta)$ satisfies at least one of (4.31) or (4.32). We can then apply (4.29) with I replaced by each ball B_i: writing $s' = c_1 R$ we have

$$\int_{B_i} f^2 d\Gamma(\varphi,\varphi) \le c_2(s'/R)^{2\theta}\Big(\int_{B_i^*} d\Gamma(f,f) + \Psi(s')^{-1}\int_{B_i^*} f^2 d\mu\Big).$$

We then sum over i. Since no point of I^* is in more than L_0 (not depending on x_0 or R) of the B_i^*, and $s/c_1 \le s' \le s$, we obtain (4.29) for I. □

5 Strongly recurrent case

5.1 Framework and the main theorem

Let $(X,d,\mu,\mathcal{E})$ be the MMD space or the weighted graph. It is called a *resistance form* if $\mathcal{F}\subset C(X)$ and

$$\sup\Big\{\frac{|u(p)-u(q)|^2}{\mathcal{E}(u,u)} : u\in\mathcal{F}, \mathcal{E}(u,u)>0\Big\} < \infty, \qquad \forall p,q\in X. \tag{5.1}$$

Define $R(p,q) =$ (LHS of (5.1)) if $p\ne q$ and $R(p,p)=0$. One can prove that R is a metric and it is called a *resistance metric*. By (5.1), the following key inequality holds.

$$|f(x)-f(y)|^2 \le R(x,y)\mathcal{E}(f,f), \qquad \forall f\in\mathcal{F}. \tag{5.2}$$

The next lemma shows that $R(p,q)$ is the effective resistance between p and q.

Lemma 5.1

$$R(p,q) = \Big(\inf\{\mathcal{E}(f,f) : f(p)=1, f(q)=0, f\in\mathcal{F}\}\Big)^{-1}. \tag{5.3}$$

Proof. By linear transform $f(x) = au(x) + b$, we can take $f(x) = 1, f(y) = 0$ if u is not const. So,

$$\begin{aligned} R(x,y) &= \sup\left\{\frac{|u(x)-u(y)|^2}{\mathcal{E}(u,u)} : u \in \mathcal{F}, \mathcal{E}(u,u) > 0\right\} = \sup\left\{\frac{1}{\mathcal{E}(f,f)} : f \in \mathcal{F}, f(x) = 1, f(y) = 0\right\} \\ &= \Big(\inf\{\mathcal{E}(f,f) : f(x) = 1, f(y) = 0, f \in \mathcal{F}\}\Big)^{-1}, \end{aligned}$$

and the conclusion holds. □

Examples. Any weighted graphs are resistance forms. For the Dirichlet form on $\mathbb{R}^d$ that corresponds to Brownian motion, it is a resistance form only when $d = 1$. Dirichlet forms on the Sierpinski gasket, nested fractals are resistance forms. Dirichlet forms on the 2-dimensional Sierpinski carpet are resistance forms.

We now give several inequalities.

(I) We say X satisfies a volume growth condition $(VG(\Psi_-))$ if there exist $\alpha < \beta \vee \bar{\beta}$ and $C > 0$ such that the following holds,

$$V(x,r) \le C\left(\frac{r}{s}\right)^{\alpha} V(x,s) \qquad \forall x \in X,\ \forall r \ge s > 0. \tag{$VG(\Psi_-)$}$$

(II) We say X satisfies a resistance upper and lower bound of order Ψ $(RU(\Psi)), (RL(\Psi))$ if there exist $C_1, C_2 > 0$ such that for all $x, y \in X$,

$$R(x,y) \le C_1 \frac{\Psi(d(x,y))}{\mu(B(x,d(x,y)))}, \tag{$RU(\Psi)$}$$

$$C_2 \frac{\Psi(d(x,y))}{\mu(B(x,d(x,y)))} \le R(x,y). \tag{$RL(\Psi)$}$$

Theorem 5.2 *Let $(X, d, \mu, \mathcal{E})$ be a resistance form on a MMD space or a weighted graph. Assume $(VG(\Psi_-))$. Then,*

$$(HK(\Psi)) \Leftrightarrow (RU(\Psi)) + (RL(\Psi)) \Leftrightarrow (RL(\Psi)) + (PI(\Psi)). \tag{5.4}$$

When (5.4) holds, it is strongly recurrent in the following sense. There exists $p_1 > 0$ such that

$$P^x(\sigma_y < \tau_{B(x,2r)}) \ge p_1, \qquad \forall x \in X, r > 0,\ y \in B(x,r), \tag{5.5}$$

where $\sigma_A = \inf\{t \ge 0 : X_t \in A\}$ and $\tau_A = \inf\{t \ge 0 : X_t \notin A\}$.

When X is a tree, we have a simpler equivalence condition as follows.

Corollary 5.3 *Let (X, μ) be a weighted graph with $c_1 \le \mu_{xy} \le c_2$ for all $x \sim y$. Assume that X is a tree. Then,*

$$(VG(\beta_-)) + (HK(\beta)) \Leftrightarrow [V(x, d(x,y)) \asymp d(x,y)^{\beta-1} \ \ \forall x, y].$$

5.2 Proof of Theorem 5.2: $(RU(\Psi)) + (RL(\Psi)) \Rightarrow (HK(\Psi))$

The flowchart of the proof is similar to that of Proposition 4.1.

First, note that the following holds by $(VG(\Psi_-))$; there exists $c > 0$ such that

$$\frac{\Psi(s)}{V(x,s)} \le c\frac{\Psi(r)}{V(x,r)} \qquad \forall r > s > 0. \tag{5.6}$$

Indeed, by $(VG(\Psi_-))$, we have

$$\frac{V(x,r)}{V(x,s)} \le c\Big(\frac{r}{s}\Big)^{\alpha} < c\Big(\frac{r}{s}\Big)^{\beta\wedge\bar{\beta}} \le c\frac{\Psi(r)}{\Psi(s)}, \qquad \forall r > s > 0,$$

which implies (5.6).

We now give the proof of $(RU(\Psi)) + (RL(\Psi)) \Rightarrow (HK(\Psi))$ step by step.

STEP A: PROOF OF $(RU(\Psi)) \Rightarrow (DUHK(\Psi))$. Let $f_t(y) = p_t(x,y)$ and

$$\varphi(t) := ||f_t||_2^2 = p_{2t}(x,x) = f_{2t}(x). \tag{5.7}$$

Since $\int_{B(x,r)} f_t d\mu \le 1$ for $r > 0$, there exists $y = y(t,r) \in B(x,r)$ with $f_t(y) \le V(x,r)^{-1}$. Using (5.2),

$$\frac{1}{2} f_t(x)^2 \le f_t(y)^2 + |f_t(x) - f_t(y)|^2 \le \frac{1}{V(x,r)^2} + \mathcal{E}(f_t, f_t) R(x,y).$$

Since $R(x,y) < c_1 \Psi(r)/V(x,r)$, which is due to $(RU(\Psi))$, it follows that

$$\frac{c_1 \Psi(r)}{V(x,r)} \mathcal{E}(f_t, f_t) \ge \frac{1}{2} \varphi(t/2)^2 - \frac{1}{V(x,r)^2}.$$

Hence

$$\varphi'(t) = -2\mathcal{E}(f_t, f_t) \le \frac{2V(x,r)^{-1} - \varphi(t/2)^2 V(x,r)}{c_1 \Psi(r)}. \tag{5.8}$$

Noting that $-\varphi(t/2)^2 \le -\varphi(t)^2$, which is due to the fact $\varphi'(t) = -2\mathcal{E}(f_t, f_t) \le 0$, we integrate (5.8) over $[t, 2t]$. Then,

$$-\varphi(t) \le \varphi(2t) - \varphi(t) \le \frac{2t}{c_1 \Psi(r) V(x,r)} - \frac{t\varphi(t)^2 V(x,r)}{c_1 \Psi(r)}.$$

Rearranging this, we have

$$t\varphi(t)^2 V(x,r)^2 \le 2t + c_1 \Psi(r) V(x,r) \varphi(t) \le (4t) \vee (2c_1 \Psi(r) V(x,r) \varphi(t)).$$

Thus, we obtain $\varphi(t) \le (2/V(x,r)) \vee (2c_1 \Psi(r)/(tV(x,r)))$. Taking $r = \Psi^{-1}(t)$ and using the doubling properties of Ψ and V, we obtain $(DUHK(\Psi))$. $\square$

STEP B: PROOF OF $(VG(\Psi_-)) + (RU(\Psi)) + (RL(\Psi)) \Rightarrow (E(\Psi))$. In order to prove this, we first give a key lemma.

Lemma 5.4 *Assume $(VG(\Psi_-)), (RU(\Psi))$ and $(RL(\Psi))$. Then, the following holds.*

$$\frac{c_1 \Psi(r)}{V(x,r)} \le R(x, B(x,r)^c) \le \frac{c_2 \Psi(r)}{V(x,r)} \qquad \textit{for all } \; r > 0, \; x \in X. \tag{5.9}$$

PROOF. First, take $y, z \in B(x,r)$ with $d(y,z) = \lambda r$, $\lambda \le 1$. We have by (5.2) and $(RU(\Psi))$,

$$|f(y) - f(z)|^2 \le R(y,z)\mathcal{E}(f,f) \le \frac{c_2 \Psi(\lambda r) \mathcal{E}(f,f)}{V(x, \lambda r)}, \qquad \text{for all } \; f \in \mathcal{F}. \tag{5.10}$$

Let $z \in X$ be such that $c_* r \le d(x,z) \le r$ for some $c_* < 1$. If h_z is the harmonic function on $X \setminus \{x,z\}$ with $h_z(z) = 0$, $h_z(x) = 1$ then $\mathcal{E}(h_z, h_z) = R(x,z)^{-1}$. Applying (5.6), (5.10) and $(RL(\Psi))$, we have, if $d(y,z) = \lambda r$,

$$|h_z(y)|^2 = |h_z(y) - h_z(z)|^2 \le \frac{c_2 \Psi(\lambda r)}{V(x, \lambda r) R(x,z)} \le \frac{c_3 \Psi(\lambda r) V(x, c_* r)}{V(x, \lambda r) \Psi(c_* r)}.$$

So there exists a constant λ_1 such that $d(y,z) \le \lambda_1 r$ implies that $h_z(y) \le \frac{1}{2}$.

Now use (VD) to cover $B(x,r) \setminus B(x, c_* r)$ by balls $B(z_i, \lambda_1 r)$, $1 \le i \le M$, with $c_* r \le d(x, z_i) \le r$. Here, M depends only on the volume doubling constant. Let $g = \min h_{z_i}$, and $h = 2(g - \frac{1}{2})^+ \cdot 1_{B(x,r)}$. Then $h(x) = 1$, and $h = 0$ on $B(x, c_* r)^c$, so that

$$R(x, B(x,r)^c)^{-1} \le \mathcal{E}(h,h) \le 4 \sum_i \mathcal{E}(h_{z_i}, h_{z_i}) \le 4M(\min_i R(x, z_i))^{-1} \le \frac{c_4 V(x, c_* r)}{\Psi(c_* r)} \le \frac{c_5 V(x,r)}{\Psi(r)}.$$

We thus obtain the first inequality of (5.9). The second inequality of (5.9) is clear from $(RU(\Psi))$, because $R(x, B(x,r)^c) \le R(x,y)$ for all $y \in \partial B(x,r)$. $\square$

PROOF OF $(E(\Psi))$. Denote $B := B(x_0, r)$ and let $(\mathcal{E}_B, \mathcal{F}_B)$ be the part of the Dirichlet form in the sense of [35] section 4.4. By Theorem 4.4.3 of [35], it is a regular Dirichlet form on $L^2(B,\mu)$ with $\mathcal{F}_B \subset \{f \in \mathcal{F} : f(x) = 0 \text{ on } x \in B^c\}$. Let X_t^B be the corresponding Hunt process, which is a process with the killing condition outside B. Using (5.2) and $(RU(\Psi))$, we have

$$\sup_{x\in B} |f(x)|^2 \le \frac{c_1\Psi(r)}{V(x_0,r)}\mathcal{E}(f,f) \qquad \text{for all } \ f \in \mathcal{F}_B. \tag{5.11}$$

Thus, $(\mathcal{E}_B, \mathcal{F}_B)$ is a transient Dirichlet form so that the extended Dirichlet space $(\mathcal{E}_B, (\mathcal{F}_B)_e)$ is a Hilbert space (Theorem 1.5.3 in [35]). Using (5.11) and the Riesz representation theorem, there exists a Green kernel $g_B(\cdot,\cdot)$ with the reproducing property; $\mathcal{E}(g_B(x,\cdot), f) = f(x)$ for all $f \in \mathcal{F}_B$. Using the reproducing property and the irreducibility of the form, $g_B(x,y) = g_B(y,x)$ and $g_B(x,x) > 0$ for all $x, y \in B$. Set $p_x(y) := g_B(x,y)/g_B(x,x)$. Then p_x is an equilibrium potential for $R(x, B^c)$ and we have

$$R(x,B^c)^{-1} = \mathcal{E}(p_x,p_x) = g_B(x,x)^{-1}. \tag{5.12}$$

Since $p_x(y) \le 1$ for all $y \in X$,

$$g_B(x,y) \le g_B(x,x) \qquad \text{for all } \ x, y \in X. \tag{5.13}$$

On the other hand, by the definition of the resistance,

$$R(x,B^c) \le R(x,y) \qquad \text{for all } \ x,y \in X, y \in B^c,$$

so that $g_B(x,x) \le c_1\Psi(r)/V(x,r)$. Now, since

$$E^{x_0}[\tau_{B(x_0,r)}] = \int_B g_B(x_0,y)d\mu(y), \tag{5.14}$$

we have

$$E^{x_0}[\tau_{B(x_0,r)}] \le \frac{c_1\Psi(r)}{V(x_0,r)}V(x_0,r) \le c_1\Psi(r),$$

where we use (5.13). We thus obtain the second inequality of $(E(\Psi))$.

Next, by (5.2) and the reproducing property of g_B, we have for $y \in B$,

$$|g_B(x_0,x_0) - g_B(x_0,y)|^2 \le \mathcal{E}(g_B,g_B)R(x_0,y) = g_B(x_0,x_0)R(x_0,y).$$

Thus, by (5.12) we have

$$|1 - p_{x_0}(y)|^2 \le \frac{R(x_0,y)}{R(x_0,B^c)}.$$

Now using Lemma 5.4, we see that there exists $\delta > 0$ such that

$$p_{x_0}(y) = \frac{g_B(x_0,y)}{g_B(x_0,x_0)} \ge 1/2 \qquad \text{for all } \ y \in B(x_0,\delta r). \tag{5.15}$$

On the other hand, by (5.12) and Lemma 5.4, we have $g_B(x_0,x_0) = R(x_0,B^c) \ge c_2\Psi(r)/V(x_0,r)$. Combining this with (5.15), we have

$$g_B(x_0,y) \ge \frac{c_3\Psi(r)}{V(x_0,r)}, \qquad \text{for all } \ y \in B(x_0,\delta r).$$

Applying this with (5.14) and (VD), we have

$$\mathbb{E}^{x_0}[\tau_{B(x_0,r)}] = \int_B g_B(x_0,y)d\mu(y) \geq \frac{c_3\Psi(r)}{V(x_0,r)}V(x_0,\delta r) \geq c_4\Psi(r),$$

where $c_4 > 0$ depends on δ. We thus obtain the first inequality of $(E(\Psi))$. $\square$

Remark. (5.15) implies immediately (5.5). This implies (EHI) by Lemma 1.6 in [6]. Thus, $(RU(\Psi)) + (RL(\Psi)) \Rightarrow (HK(\Psi))$ is proved by Proposition 4.1 and Proposition 4.3 (Step A above was not needed). But we do not choose this way because several steps of the current proof are much simpler than those of Proposition 4.1 and Proposition 4.3, thanks to (5.2).

STEP C: PROOF OF $(VD) + (DUHK(\Psi)) + (E(\Psi)) \Rightarrow (UHK(\Psi))$. This step is the same as Step 1 and Step 2 in the proof of Proposition 4.1.

STEP D: PROOF OF $(VD) + (ELD(\Psi)) \Rightarrow (DLHK(\Psi))$. This step is the same as Step 3 in the proof of Proposition 4.1.

STEP E: PROOF OF $(VG(\Psi_-)) + (RU(\Psi)) + (DLHK(\Psi)) \Rightarrow (NLHK(\Psi))$.

First, note that $(RU(\Psi))$ implies $(DUHK(\Psi))$ as shown before. Note also that, since $p_t(x,x) = \|p_{t/2}(\cdot,x)\|_2^2$, we have $\partial_t p_t(x,x) = 2(\Delta p_{t/2}(\cdot,x), p_{t/2}(\cdot,x)) = -2\mathcal{E}(p_{t/2}(\cdot,x), p_{t/2}(\cdot,x))$. Thus, using (5.2) and Proposition 9.9, we have

$$|p_t(x,y) - p_t(x,y')|^2 \leq R(y,y')\mathcal{E}(p_t(\cdot,x),p_t(\cdot,x)) \leq \frac{\Psi(d(y,y'))}{V(y,d(y,y'))} \cdot \frac{c_1}{tV(x,\Psi^{-1}(t))}.$$

Using this and $(DLHK(\Psi))$,

$$\begin{aligned} p_t(x,y) &\geq p_t(x,x) - |p_t(x,x) - p_t(x,y)| \\ &\geq \frac{c_2}{V(x,\Psi^{-1}(t))} - \left\{\frac{\Psi(d(x,y))}{V(x,d(x,y))} \cdot \frac{c_1}{tV(x,\Psi^{-1}(t))}\right\}^{1/2} \\ &= \frac{c_2}{V(x,\Psi^{-1}(t))^{1/2}}\left(\frac{1}{V(x,\Psi^{-1}(t))^{1/2}} - c_3\left(\frac{\Psi(d(x,y))}{tV(x,d(x,y))}\right)^{1/2}\right). \end{aligned}$$

Now, taking c_4 large enough, we have $\frac{1}{2V(x,\Psi^{-1}(t))^{1/2}} \geq c_3\left(\frac{\Psi(d(x,y))}{tV(x,d(x,y))}\right)^{1/2}$ if $\Psi(d(x,y)) \leq c_4 t$ holds. Here we used (5.6). We thus obtain the result. $\square$

STEP F: PROOF OF $(NLHK(\Psi)) \Rightarrow (LHK(\Psi))$. This step is the same as Step 5 in the proof of Proposition 4.1.

Combining Step A–F, the proof of $(RU(\Psi)) + (RL(\Psi)) \Rightarrow (HK(\Psi))$ is completed.

5.3 Proof of Theorem 5.2: The rest

Since this will not be discussed in the summer school, we just give references. $(HK(\Psi)) \Rightarrow (RU(\Psi)) + (RL(\Psi))$ is proved in [15] Section 4. $(VG(\Psi_-))+(RU(\Psi))+(RL(\Psi)) \Rightarrow (PI(\Psi))$ and $(VG(\Psi_-))+(PI(\Psi)) \Rightarrow (RU(\Psi))$ are proved in [15] subsection 2.2. They are proved for the case of weighted graphs, but the translation to the current setting is easy.

6 Application: RW on critical branching processes

6.1 Background

We recall the bond percolation model on the lattice $\mathbb{Z}^d$: each bond is open with probability $p \in (0,1)$, independently of all the others. Let $\mathcal{C}(x)$ be the open cluster containing x; then if $\theta(p) = P_p(|\mathcal{C}(x)| = +\infty)$ it is well known (see [43]) that there exists $p_c = p_c(d)$ such that $\theta(p) = 0$ if $p < p_c$ and $\theta(p) > 0$ if $p > p_c$.

If $d = 2$ or $d \geq 19$ (or $d > 6$ for 'spread out' models) it is known (see [43, 51]) that $\theta(p_c) = 0$, and it is conjectured that this holds for all $d \geq 2$. At the critical probability $p = p_c$ it is believed that in any box of side n there exist with high probability open clusters of diameter of order n – see [24]. For large n the local properties of these large finite clusters can, in certain circumstances, be captured by regarding them as subsets of an infinite cluster $\widetilde{\mathcal{C}}$, called the 'incipient infinite cluster' (IIC).

This was constructed when $d = 2$ in [54], by taking the limit as $N \to \infty$ of the cluster $\mathcal{C}(0)$ conditioned to intersect the boundary of a box of side N with center at the origin. For large d a construction of the IIC in $\mathbb{Z}^d$ is given in [49], using the lace expansion. It is believed that the results there will hold for any $d > 6$. [49] also gives the existence and some properties of the IIC for all $d > 6$ for 'spread-out' models: these include the case when there is a bond between x and y with probability pL^{-d} whenever y is in a cube side L with center x, and the parameter L is large enough. Rather more is known about the IIC for oriented percolation on $\mathbb{Z}_+ \times \mathbb{Z}^d$ (see [50, 51]), but in this discussion, which mainly concerns what is conjectured rather than what is known, we specialize to the case of $\mathbb{Z}^d$. We write $\widetilde{\mathcal{C}}_d$ for the IIC in $\mathbb{Z}^d$. It is believed that the global properties of $\widetilde{\mathcal{C}}_d$ are the same for all $d > d_c$, both for nearest neighbour and spread-out models. In [49] it is proved for 'spread-out' models that $\widetilde{\mathcal{C}}_d$ has one end – that is that any two paths from 0 to infinity intersect infinitely often.

For large d, it is believed that the geometry of $\widetilde{\mathcal{C}}_d$ is also similar to that of the IIC when '$d = \infty$' – that is to the IIC on a regular tree; this is supported by the results in [50, 49]. For trees the construction of the IIC is much easier than for lattices, and there is a close connection between the IIC and a critical Bienaymé-Galton-Watson branching processes conditioned on non-extinction. In [55], Kesten gave the construction of the IIC $\mathcal{G}$ for critical branching processes. This is an infinite subtree, which contains only one path from the root to infinity. This tree is quite sparse, and has polynomial volume growth: in the case when the offspring distribution has finite variance, a ball $B(x, r)$ in $\mathcal{G}$ has roughly r^2 points. (This is when distance in $\mathcal{G}$ is measured using the natural graph distance).

Let $Y = (Y_t, t \geq 0)$ be the simple random walk on $\widetilde{\mathcal{C}}_d$, and $q_t(x, y)$ be its transition density. Define the *spectral dimension* of $\widetilde{\mathcal{C}}_d$ by

$$d_s(\widetilde{\mathcal{C}}_d) = -2 \lim_{t\to\infty} \frac{\log q_t(x,x)}{\log t},$$

(if this limit exists). Alexander and Orbach [1] conjectured that, for any $d \geq 2$, $d_s(\widetilde{\mathcal{C}}_d) = 4/3$. While it is now thought that this is unlikely to be true for small d, the results on the geometry of $\widetilde{\mathcal{C}}_d$ in [50, 49] are consistent with this holding for large d. (Or for any d above the critical dimension for spread-out models).

Random walks on supercritical clusters in $\mathbb{Z}^d$ are studied in [3] (transition density estimates) and [74] (invariance principle for the quenched case for $d \geq 4$; in the annealed case, invariance principle was proved in [33]). In these cases the large scale behaviour of the random walk approximates that of the random walk on $\mathbb{Z}^d$, and the unique infinite cluster has spectral dimension d.

In what follows, we will specialize to the case of critical percolation on a regular rooted tree with degree $n_0 + 1$. We keep n_0 fixed, but (in view of possible future applications) wish to obtain estimates which do not depend on n_0.

6.2 The model and main results

We will define the random graph $\mathcal{G}$ we will be working with. We could regard this either as critical percolation on the n_0-ary tree $\mathbb{B}$, conditioned on the cluster containing the root 0 being infinite, or as the (critical) Bienaymé-Galton-Watson process with $Bin(n_0, 1/n_0)$ offspring distribution, conditioned on non-extinction.

Let $\mathbb{B}$ be the n_0-ary tree, and let 0 be the root. A point x in the nth generation (or level) is written $x = (0, l_1, \cdots, l_n)$, where $l_i \in \{1, 2, \cdots, n_0\}$. Let $\mathbb{B}_n$ be the set of n_0^n points in the nth generation, and let $\mathbb{B}_{\leq n} = \cup_{i=0}^n \mathbb{B}_i$. If $x \in \mathbb{B}_k$ we write $|x| = k$. If $x = (0, l_1, \cdots, l_n) \in \mathbb{B}_n$, let $a(x, r) = (0, l_1, \cdots, l_{n-r})$ be the ancestor of x at level $|x| - r$.

We regard $\mathbb{B}$ as a graph (in fact a tree) with edge set $E(\mathbb{B}) = \{\{x, a(x, 1)\}, x \in \mathbb{B} - \{0\}\}$. Let η_e, $e \in E(\mathbb{B})$, be i.i.d. Bernoulli $1/n_0$ r.v. defined on a probability space $(\Omega, \mathcal{F}, P)$. If $\eta_e = 1$ we say the edge e

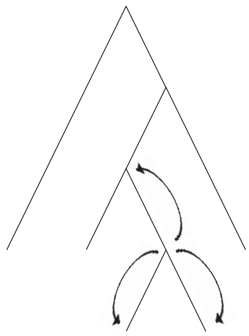

Figure 4: Random walk on the GW-tree

is *open*. Let

$$\mathcal{C}(0) = \{x \in \mathbb{B} : \text{ there exists an } \eta\text{–open path from 0 to } x\}$$

be the open cluster containing 0. It is clear that $Z_n = |\mathcal{C}(0) \cap \mathbb{B}_n|$ is a critical GW process with $Bin(n_0, 1/n_0)$ offspring distribution. Here and in the following, $|A|$ is a cardinality of the set A. As Z has extinction probability 1, the cluster $\mathcal{C}(0)$ is P–a.s. finite.

We have

Lemma 6.1 ([55], Lemma 1.14) *Let $A \subset \mathbb{B}_{\leq k}$. Then*

$$\lim_{n\to\infty} P(\mathcal{C}(0) \cap \mathbb{B}_{\leq k} = A | Z_n \neq 0) = |A \cap \mathbb{B}_k| P(\mathcal{C}(0) \cap \mathbb{B}_{\leq k} = A),$$

and writing $\mathbb{P}_0(A) = |A \cap \mathbb{B}_k| P(\mathcal{C}_{\leq k} = A)$, $\mathbb{P}_0$ has a unique extension to a probability measure $\mathbb{P}$ on the set of infinite connected subsets of $\mathbb{B}$ containing 0.

Let $\mathcal{G}'$ be a rooted labeled tree chosen with the distribution $\mathbb{P}$: we call this the *incipient infinite cluster* (IIC) on $\mathbb{B}$. For more information on $\mathcal{G}'$ see [48, 55] but we remark that $\mathbb{P}$–a.s. $\mathcal{G}'$ has exactly one infinite descending path from 0, which we call the *backbone*, and denote H.

It will be useful to give another construction of the IIC, obtained by modifying the cluster $\mathcal{C}(0)$ rather than its law. We can suppose the probability space $(\Omega, \mathcal{F}, P)$ carries i.i.d.r.v. ξ_i, $i \geq 1$ uniformly distributed on $\{1, 2, \cdots, n_0\}$, and independent of (η_e). For $n \geq 0$ let $\Xi_n = (0, \xi_1, \ldots, \xi_n)$, and let

$$\widetilde{\eta}_e = \begin{cases} 1 & \text{if } e = \{\Xi_n, \Xi_{n+1}\} \text{ for some } n \geq 0, \\ \eta_e & \text{otherwise.} \end{cases}$$

Then (see [48]) if

$$\mathcal{G} = \{x \in \mathbb{B} : \text{ there exists a } \widetilde{\eta}\text{–open path from 0 to } x\},$$

$\mathcal{G}$ has law $\mathbb{P}$. It is clear that the backbone of $\mathcal{G}$ is the set $H = \{\Xi_n, n \geq 0\}$.

For $x, y \in \mathbb{B}$ let

$$\mathbb{P}_x(\cdot) = \mathbb{P}(\cdot | x \in \mathcal{G}), \qquad \mathbb{P}_{xy}(\cdot) = \mathbb{P}(\cdot | x, y \in \mathcal{G}),$$

and let $\mathbb{E}_x$ and $\mathbb{E}_{xy}$ denote expectation with respect to $\mathbb{P}_x$ and $\mathbb{P}_{xy}$ respectively. Given a descending path $b = \{0, b_1, b_2, \ldots\}$, (which we call a *possible backbone*) let

$$\mathbb{P}_{x,b}(\cdot) = \mathbb{P}(\cdot | x \in \mathcal{G}, H = b),$$

and define $\mathbb{P}_{x,y,b}$ analogously.

For each x, $y \in \mathbb{B}$, let $\gamma(x, y)$ be the unique geodesic path connecting x and y. We say that z is a *middle point* of $\gamma(x, y)$ if $z \in \gamma(x, y)$ and $|d(x, z) - \frac{1}{2} d(x, y)| \leq \frac{1}{2}$. We remark that the construction of $\mathcal{G}$ makes it clear that $\mathbb{P}_{x,y,b}(\eta_e = 1) = 1$ if the edge e lies in any of the paths b, $\gamma(0,x)$ and $\gamma(0,y)$, and that under $\mathbb{P}_{x,y,b}$ the r.v. η_e, $e \notin b \cup \gamma(0,x) \cup \gamma(0,y)$ are i.i.d. with $\mathbb{P}_{x,y,b}(\eta_e = 1) = 1/n_0$.

For each fixed $\mathcal{G} = \mathcal{G}(\omega)$, we will consider the continuous time simple random walk $\{Y_t\}$ on $\mathcal{G}$ as in subsection 3.1 and define its heat kernel $q_t^\omega(x,y)$ as in (3.1).

Theorem 6.2 *(a) There exist c_0, c_1, c_2, $S(x)$ such that for each x,*

$$\mathbb{P}_x(S(x) \geq m) \leq c_0(\log m)^{-1},$$

and on $\{\omega : x \in \mathcal{G}(\omega)\}$

$$c_1 t^{-2/3}(\log\log t)^{-17} \leq q_t^\omega(x,x) \leq c_2 t^{-2/3}(\log\log t)^3 \text{ for all } t \geq S(x).$$

(b) $d_s(\mathcal{G}) = 4/3$ $\mathbb{P}$–a.s.

The cluster $\mathcal{G}$ contains large scale fluctuations, so that $q_t(x,x)$ does have oscillations of order $(\log\log t)^a$ as $t \to \infty$.

Proposition 6.3

$$\liminf_{t\to\infty}(\log\log t)^{1/6} t^{2/3} q_{2t}^\omega(0,0) \leq 2, \qquad P_\omega^0 - a.s.$$

Theorem 6.4 *(a) We have*

$$c_1 t^{1/3} \leq \mathbb{E}_x E_\omega^x d(x, Y_t) \leq \mathbb{E}_x E_\omega^x \sup_{0\leq s\leq t} d(x, Y_s) \leq c_2 t^{1/3}.$$

(b) There exists $T(x)$ with $\mathbb{P}_x(T(x) < \infty) = 1$ such that

$$c_3 t^{1/3}(\log\log t)^{-12} \leq E_\omega^x[d(x,Y_t)] \leq c_4 t^{1/3}\log t \quad \text{for all } t \geq T(x).$$

We also have off-diagonal bounds for $q_t^\omega(x,y)$. For the quenched case, our theorem is the following shape.

Theorem 6.5 *(1) Let $x, y \in \mathcal{G}$, $t > 0$ be such that $N := [\sqrt{d(x,y)^3/t}] \geq 8$. Then, there exists an event $F_* = F_*(x,y,t)$ that satisfies*

$$\mathbb{P}_{x_0,y_0,b}(F_*(x,y,t)) \geq 1 - c_1 \exp(-c_2 N),$$

so that the following holds:

$$q_t^\omega(x,y) \leq c_3 t^{-2/3}\exp(-c_4 N), \qquad \forall\omega \in F_*.$$

(2) Let $x, y \in \mathcal{G}$, $m \geq 1$, $\kappa \geq 1$ and let $T = d(x,y)^3\kappa/m^2$. Then, there exists an event $G_* = G_*(x,y,m,\kappa)$ that satisfies

$$\mathbb{P}_{x,y,b}(\ G_*(x,y,m,\kappa) \text{ holds }) \geq 1 - c_1\kappa^{-1},$$

so that the following holds:

$$q_{2T}(x,y) \geq c_2 T^{-2/3} e^{-c_3(\kappa+c_4)m}, \qquad \forall\omega \in G_*.$$

For the annealed case, the off-diagonal bounds for $q_t^\omega(x,y)$ are of the same form as the bounds

$$ct^{-d_f/d_w}\exp(-c'(d(x,y)^{d_w}/t)^{1/(d_w-1)})$$

obtained for regular fractal graphs.

Theorem 6.6 *(a) Let $x, y \in \mathbb{B}$. Then*

$$\mathbb{E}_{x,y} q_t^{\omega}(x,y) \le c_1 t^{-2/3} \exp\big(- c_2 (\frac{d(x,y)^3}{t})^{1/2}\big).$$

(b) Let $x, y \in \mathbb{B}$, with $d(x,y) = R$, and $c_3 R \le t$. Then

$$\mathbb{E}_{x,y} q_t^{\omega}(x,y) \ge c_4 t^{-2/3} \exp(-c_5 (R^3/t)^{1/2}).$$

Define the continuous time rescaled height process

$$\widetilde{Z}_t^{(n)} = n^{-1/3} d(0, Y_{nt}), \quad t \ge 0.$$

By Theorem 6.4 (a) the processes $(\widetilde{Z}^{(n)}, n \ge 1)$ are tight with respect to the annealed law given by the semi-direct product $\mathbb{P}^* = \mathbb{P} \times P_{\omega}^0$. (This is much easier to prove than the full convergence given in [55].) However, the large scale fluctuations in $\mathcal{G}$ mean that we do not have quenched tightness.

Theorem 6.7 $\mathbb{P}$*-a.s., the processes $(\widetilde{Z}^{(n)}, n \ge 1)$ are not tight with respect to P_{ω}^0.*

6.3 Ideas of the proof

The proof consists of the analytic part and the probabilistic part. We would emphasize that we cannot expect (VD) for this kind of random object, so we need estimates without assuming (VD).

Definition 6.8 *Let $x \in \mathcal{G}$, $r \ge 1$. Let $M(x,r)$ be the smallest number m such that there exists a set $A = \{z_1, \dots, z_m\}$ with $d(x, z_i) \in [r/4, 3r/4]$ for each i, such that any path γ from x to $B(x,r)^c$ must pass through the set A.*

Analytic estimates For fixed $r \ge 1$ and $x_0 \in G$, we denote $B = B(x_0, r)$, $M = M(x_0, r)$, $V = V(x_0, r)$.

Proposition 6.9 *(a) Let (G, μ) be a weighted graph and suppose that the edge weights satisfy $\mu_{xy} \ge 1$ for all x and y. Then*

$$q_{2rV(x,r)}(x,x) \le \frac{2}{V(x,r)}, \quad x \in G, \, r > 0.$$

(b) Assume further that G is a tree. Let $V_1 = V_1(x_0, r) = V(x_0, r/(32M(x_0,r)))$. Then if $x \in B(x_0, r/(32M))$,

$$P^x(\tau_B \le t) \le \Big(1 - \frac{V_1}{64MV}\Big) + \frac{t}{2rV},$$

and

$$q_{2t}(x,x) \ge \frac{c_1 V_1(x_0,r)^2}{V(x_0,r)^3 M(x_0,r)^2} \qquad \textit{for } t \le \frac{r V_1(x_0,r)}{64 M(x_0,r)}.$$

(a) can be proved by carefully chasing Step A in subsection 5.2 and modifying to the current situation. For (b), first, similar argument as in Step B in subsection 5.2 (using the tree property and $M(x,r)$ instead of (VD)) gives the estimate of $E_{\omega}^x[\tau_{B(x,r)}]$. Then the argument in Step 3 in the proof of Proposition 4.1 gives the desired result. See [16] for details.

Probabilistic estimates By the above analytic estimates, we see that the information of $V(x,r)$ and $M(x,r)$ are necessary for the on-diagonal estimates. We will show that the probability that $V(x,r)$ and $M(x,r)$ behave badly is 'small'.

Proposition 6.10 *(a) Let $\lambda > 0$, $r \geq 1$ and $x, y \in \mathbb{B}$, and b be a possible backbone. Then*

$$\mathbb{P}_{x,y,b}(V(x,r) > \lambda r^2) \leq c_0 \exp(-c_1 \lambda),$$

and

$$\mathbb{P}_{x,y,b}(V(x,r) < \lambda r^2) \leq c_2 \exp(-c_3/\sqrt{\lambda}).$$

(b) For any $\varepsilon > 0$

$$\limsup_{n\to\infty} \frac{V(0,n)}{n^2(\log\log n)^{1-\varepsilon}} = \infty, \quad \mathbb{P} - a.s.$$

(c) There exist $c_4, c_5 > 0$ such that for each $r \geq 1$ and each $x, y \in \mathbb{B}$, and possible backbone b

$$\mathbb{P}_{x,y,b}(M(x,r) \geq m) \leq c_4 e^{-c_5 m}.$$

These can be obtained, basically through large deviation estimates of the total population size of the critical branching process. See [16] for details.

We now define a 'good' random set.

Definition 6.11 *Let $x \in \mathbb{B}$, $r \geq 1$, $\lambda \geq 64$. We say that $B(x,r)$ is λ–good if it satisfies the following:*

$$x \in \mathcal{G},\ r^2\lambda^{-2} \leq V(x,r) \leq r^2\lambda,\ M(x,r) \leq \frac{1}{64}\lambda,\ V(x,r/\lambda) \geq r^2\lambda^{-4}, \ \ \textit{and}\ \ V(x,r/\lambda^2) \geq r^2\lambda^{-6}.$$

By Proposition 6.10, we have the following.

Corollary 6.12 *For $x \in \mathbb{B}$ and any possible backbone b*

$$\mathbb{P}_{x,b}(B(x,r) \text{ is not } \lambda\text{–good}) \leq c_1 e^{-c_2\lambda}.$$

Combining these analytic and probabilistic estimates, we can obtain Theorem 6.2. To get off-diagonal estimates, we need to take more refined 'good' random sets. See [16] for details.

7 Some open problems

Finally, we would mention several important open problems.

- Simpler stable equivalence conditions for (PHI(Ψ)): It is not easy to check (CS(Ψ)) in concrete examples. Quite recently, Barlow-Bass ([8]) proved (PHI(β)) $\Leftrightarrow$ (VD) + (PI(β)) + (E(β)) for weighted graphs. But we do not know if (E(β)) is stable under perturbations or not. There is a conjecture that (PHI(β)) $\Leftrightarrow$ (VD) + (PI(β)) + (RES(β)).

- Stability of (EHI): We do not know if (EHI) is stable under perturbations (especially under rough isometries). This is one of the big open problems of this area.

- Stability of $(UHK(\Psi))$: As in subsection 8.2, there are various equivalence conditions for $(UHK(\Psi))$, but so far we do not know if either of those is stable under perturbations. There is a related conjecture by Grigor'yan that $(UHK(\beta))$ is equivalent to $(FK(\beta))$ plus so called the anti Faber-Krahn inequality, which guarantees the optimality of $(FK(\beta))$ for balls.

- RW on IIC on $\mathbb{Z}^d$: It will be very interesting to obtain similar results as those in Section 6 for RW on infinite incipient clusters on $\mathbb{Z}^d$. It is known (at least believed) that for the case of $d = 2$ and d large enough, RW on such IIC is in the framework of resistance forms discussed in Section 5, so we have reasonable analytic estimates. It is hard to obtain probabilistic estimates in these cases though.

8 Appendix: Upper bounds

8.1 Local ultracontractivity

In this subsection, we give a generalized version of Theorem 2.1. It is a localized version as we will treat the operator on $B(x_0, r^*)$ with Dirichlet boundary condition, but the global version can be recovered by taking $r^* = \infty$. This subsection is from [29], where the original ideas came from [27, 28] etc.

Let $r^* > 0$, let $m : X \times (0, r^*] \to \mathbb{R}_+$ be a Borel function so that for each $x \in X$, $m(x, \cdot)$ is monotone decreasing and differentiable. In this subsection, Ψ is not necessarily of the form (3.2). We simply let $\Psi : \mathbb{R}_+ \to \mathbb{R}_+$ be a monotone increasing function with the following growth condition; there exists $C_1, C_2 \geq 1$ such that

$$\Psi(2r) \leq C_1 \Psi(r) \leq \Psi(C_2 r), \qquad \forall r \in \mathbb{R}_+. \tag{8.1}$$

Denote $m_x(t) := m(x, \Psi^{-1}(t))$ and define $M_x(t) = -\log m_x(t)$. Throughout the paper, we assume that there exists $\alpha > 0$ such that

$$M_x'(u) \geq \alpha M_x'(t), \qquad \forall\, t > 0,\, u \in [t, 2t],\, x \in M, \qquad (\delta).$$

This means that the logarithmic derivative of m_x has polynomial growth.

Let $m_i : \mathbb{R}_+ \to \mathbb{R}_+$, $i = 1, 2$. We shall say that $m_1 \preccurlyeq m_2$ if there exists $C, C' > 0$ such that $m_1(t) \preccurlyeq C m_2(C' t)$. We say that m_1 and m_2 are equivalent if $m_1 \preccurlyeq m_2$ and $m_2 \preccurlyeq m_1$. In this subsection, the inequalities will be written modulo equivalence of functions. Note that we suppose m differential to have a neat theory, but this assumption can be relaxed by regularising m and get an equivalence function.

Define the spectral gap of an open set $\Omega \subset X$ by

$$\lambda_{\min}(\Omega) := \inf_{f \in \mathcal{F}_\Omega \setminus \{0\}} \frac{\mathcal{E}(f)}{\|f\|_2^2},$$

where $\mathcal{F}_\Omega := \{f \in \mathcal{F} : f = 0 \text{ in } X \setminus \Omega\}$. We will fix $r^* > 0$ and denote $B^{x_0} := B(x_0, r^*)$ for each $x_0 \in X$. When the dependency of x_0 is clear, we sometimes denote it by B.

$$\|T_t^{B^{x_0}}\|_{1 \to \infty} \leq m(x_0, \Psi^{-1}(t)), \qquad \forall t \leq \Psi(r^*), \forall x_0 \in X. \qquad (UC(\Psi))$$

$$\mathcal{E}_{B^{x_0}}(u) \geq \frac{\|u\|_2^2}{2\Psi(r)} \log \frac{\|u\|_2^2}{m(x_0, r)\|u\|_1^2}, \qquad \forall u \in \mathcal{F}_{x_0, r^*}, \forall r \leq r^*, \text{ and } \forall x_0 \in X. \qquad (logLN(\Psi))$$

$$\theta_{x_0}(\|u\|_2^2) \leq \mathcal{E}_{B^{x_0}}(u), \qquad \forall u \in \mathcal{F}_{x_0, r^*} \text{ s.t. } \|u\|_1 \leq 1 \text{ with } \|u\|_2^2 \geq m(x_0, r^*), \forall x_0 \in X. \qquad (Nash(\Psi))$$

$$\lambda_{\min}(\Omega) \geq \frac{1}{\varphi_{x_0}^2(\mu(\Omega))} \qquad \forall \Omega \subset B^{x_0} \text{ with } \mu(\Omega) \leq \frac{1}{m(x_0, r^*)}, \forall x_0 \in X. \qquad (FK(\Psi))$$

Here,

$$\begin{aligned}
\mathcal{F}_{x_0, r^*} &:= \{f \in \mathcal{F} : f = 0 \text{ in } X \setminus B(x_0, r^*)\}, \\
\theta_{x_0}(y) &= -\frac{\alpha}{4} m_{x_0}'(m_{x_0}^{-1}(y)) \text{ and } \varphi_{x_0}(y) = \frac{1}{\sqrt{y \theta_{x_0}(1/y)}}, \text{ so } \theta_{x_0}(y) = \frac{y}{\varphi_{x_0}^2(1/y)}.
\end{aligned}$$

$(FK(\Psi))$ is called the Faber-Krahn inequality.

Theorem 8.1 *Assume* (δ). *Then*

$$(UC(\Psi)) \Leftrightarrow (logLN(\Psi)) \Leftrightarrow (Nash(\Psi)) \Leftrightarrow (FK(\Psi)).$$

This theorem includes two typical cases.

Case 1: Uniform case Let $m(x,r) = m(r)$ and $\Psi(t) = t$. (So m_x does not depend on x.) This case corresponds to the work in [28].

Case 2: Volume doubling case Let $m(x,r) = 1/V(x,r)$ and assume (VD). Especially, the case $\Psi(t) = t^\beta$ for some $\beta \geq 2$ was treated in [36, 57].

Remarks. 1) We can prove the long time version of Theorem 8.1 in the same way. Namely, $(UC(\Psi))$ with $t \geq \Psi(r^*)$ is equivalent to $(logLN(\Psi))$ with $\frac{\|u\|_2^2}{\|u\|_1^2} \leq m(x_0, 2\Psi(r^*))$ and so on. The proof is the same as that of Theorem 8.1.
2) In [57], Kigami introduced the following local Nash inequality.

$$\mathcal{E}_{B^{x_0}}(u) + \alpha\frac{m(x_0,r)\|u\|_1^2}{\Psi(r)} \geq \beta\frac{\|u\|_2^2}{\Psi(r)}, \qquad \forall u \in \mathcal{F}_{x_0,r^*}, \forall r \leq r^*, \forall x_0 \in X. \qquad (KgLN(\Psi)).$$

$(logLN(\Psi)) \Rightarrow (KgLN(\Psi))$, but in general the converse is not true. If we assume the doubling condition for m (typically, Case 2 above), then it holds that $(logLN(\Psi)) \Leftrightarrow (KgLN(\Psi))$.

We will need $(FK(\Psi)) \Rightarrow (UC(\Psi))$ in subsection 4.1, so we give the proof below.

Proof of $(FK(\Psi)) \Rightarrow (Nash(\Psi))$. We adopt the argument originated in [38]. For each $\lambda > 0$, since $u < 2(u-\lambda)$ on $\{u > \lambda\}$, we have

$$\int u^2 \leq 4\int_{\{u>2\lambda\}} (u-\lambda)^2 + 2\lambda\int_{\{u\leq 2\lambda\}} \leq 4\int (u-\lambda)_+^2 + 2\lambda\|u\|_1.$$

Applying $(FK(\Psi))$ to $(u-\lambda)_+$ gives

$$\int (u-\lambda)_+^2 \leq \varphi_{x_0}(\mu(\{u>\lambda\}))^2\mathcal{E}_B((u-\lambda)_+) \leq \varphi_{x_0}(\frac{\|u\|_1}{\lambda})^2\mathcal{E}_B(u),$$

where we used the fact $\mu(\{u > \lambda\}) \leq \|u\|_1/\lambda$ and φ_{x_0} is non-decreasing in the second inequality. Therefore,

$$\|u\|_2^2 \leq 4\varphi_{x_0}(\frac{\|u\|_1}{\lambda})^2\mathcal{E}_B(u) + 2\lambda\|u\|_1.$$

Take $\lambda = \|u\|_2^2/(4\|u\|_1)$. We then obtain the following Sobolev-type inequality.

$$\|u\|_2^2 \leq 8\varphi_{x_0}^2\Big(\frac{\|u\|_1^2}{\|u\|_2^2}\Big)\mathcal{E}(u), \qquad \forall u \in \mathcal{F}_{x_0,r^*} \text{ with } \frac{\|u\|_1^2}{\|u\|_2^2} \leq \frac{1}{m(x_0,r^*)}, \forall x_0 \in X. \qquad (Sob(\Psi))$$

It is easy to see that $(Sob(\Psi))$ implies $(Nash(\Psi))$. □

Proof of $(Nash(\Psi)) \Rightarrow (UC(\Psi))$ Since $\|T^B u\|_1 \leq \|u\|_1$, replacing u by $T^B u$ in $(Nash(\Psi))$ gives

$$\theta_{x_0}(\|T^B u\|_2^2) \leq \mathcal{E}_B(T_t^B u), \qquad \forall u \in \mathcal{F}_{x_0,r^*}, \|u\|_1 = 1. \qquad (8.2)$$

Let $I(t) = \|T_t^B u\|_2^2$, then $I'(t) = 2(\frac{d}{dt}T_t^B u, T_t^B u) = -2\mathcal{E}_B(T_t^B u)$. It follows from (8.2) that

$$I'(t) \leq -2\theta_{x_0}(I(t)).$$

By integration, we have

$$-\int_0^t \frac{I'(s)}{\theta_{x_0}(I(s))}ds = \int_{I(t)}^{I(0)} \frac{dx}{\theta_{x_0}(x)} \geq 2t.$$

By definition, we have

$$2t = \int_{m_{x_0}(2t)}^{\infty} \frac{dx}{\theta_{x_0}(x)},$$

so

$$\int_{I(t)}^{\infty} \frac{dx}{\theta_{x_0}(x)} \geq \int_{I(t)}^{I(0)} \frac{dx}{\theta_{x_0}(x)} \geq \int_{m_{x_0}(2t)}^{\infty} \frac{dx}{\theta_{x_0}(x)}.$$

Thus, we obtain $I(t) \leq m_{x_0}(2t)$. It follows that $\|T_t^B\|_{1\to 2}^2 \leq m_{x_0}(2t)$. Since T_t^B is symmetric, we have

$$\|T_t^B\|_{1\to\infty} \leq \|T_{t/2}^B\|_{1\to 2}\|T_{t/2}^B\|_{2\to\infty} = \|T_{t/2}^B\|_{1\to 2}^2 \leq m_{x_0}(t),$$

which is the desired inequality. □

8.2 Equivalence to $(UHK(\beta))$

In [36], A. Grigor'yan proved various equivalence conditions for $(UHK(\beta))$ under (VD). To state his main theorem, we prepare two more notions.

• X satisfies $\bar{E}(\beta)$ if there exist $C, \nu > 0$ such that for any ball $B(x_0, r)$ in X and for any non-empty open set $\Omega \subset B(x_0, r)$,

$$\text{ess sup}_{x\in\Omega} E^x[\tau_\Omega] \leq C r^\beta \Big(\frac{\mu(\Omega)}{V(x_0, r)}\Big)^\nu. \qquad (\bar{E}(\beta))$$

• X satisfies (P_β) if there exist $\varepsilon \in (0, 1)$ and $\delta > 0$ such that

$$P^x(\tau_{B(x,r)} \leq \delta r^\beta) \leq \varepsilon, \qquad \forall x \in X, \forall r > 0. \qquad (P_\beta)$$

Clearly, $(\bar{E}(\beta)) \Rightarrow (\mathrm{E}(\beta)_\leq)$. As mentioned in the Step 1 of the proof of Proposition 4.1, $(\mathrm{E}(\beta)) \Rightarrow (P_\beta)$.

Theorem 8.2 ([36] Theorem 12.1) *Assume (VD). Then.*

$$\begin{aligned}(UHK(\beta)) \quad &\Leftrightarrow \quad (DUHK(\beta)) + (P_\beta) \Leftrightarrow (\bar{E}(\beta)) + (P_\beta) \Leftrightarrow (FK(\beta)) + (P_\beta)\\ &\Leftrightarrow \quad (DUHK(\beta)) + (E(\beta)) \Leftrightarrow (\bar{E}(\beta)) + (E(\beta)) \Leftrightarrow (FK(\beta)) + (E(\beta)).\end{aligned}$$

We believe that Theorem 8.2 can be exptended to our time scaling Ψ without any difficulties.

It will be interesting to compare Theorem 8.2 to the following ($\beta = 2$ case), which was proved in the setting of Riemannian manifolds in [38] Proposition 5.2.

$$(UHK(2) \Leftrightarrow (DUHK(2)) \Leftrightarrow (FK(2)).$$

9 Appendix 2: Miscellaneous proof

9.1 Consequences of (VD)

First, it is easy to deduce from (VD) that there exist $c_1, \alpha > 0$ such that if $x, y \in X$ and $0 < r < R$ then

$$\frac{V(x,R)}{V(y,r)} \leq c_1 \Big(\frac{d(x,y) + R}{r}\Big)^\alpha. \qquad (9.1)$$

Lemma 9.1 *Assume that X satisfies (VD). Then, there exists $\delta \in (0, 1)$ such that $V(x, r/2) \leq \delta V(x, r)$ for all $r > 0$ and $x \in X$.*

Proof. SInce X has infinite diameter and since it is connected, there exists $y \in X$ such that $d(x,y) = 3r/4$. Note that $B(x,r/2) \cap B(y,r/4) = \emptyset$ and $B(x,r/2) \cup B(y,r/4) \subset B(x,r)$, so that $V(x,r/2) + V(y,r/4) \leq V(x,r)$. Since $B(x,r/2) \subset B(y,5r/4)$, (VD) implies $V(x,r/2) \leq V(y,5r/4) \leq cV(y,r/4)$ where $c > 0$ is independent of r, x and y. Combining these facts, we obtain $(1+c^{-1})V(x,r/2) \leq V(x,r)$. □

Finally, we give the following covering lemma.

Lemma 9.2 *Assume that X satisfies (VD). For $x_0 \in X$ and $0 < s \leq R \leq \infty$, there exists a cover of $B(x_0,R)$ by balls $B(x_i,s)$ with $x_i \in B(x_0,R)$ such that no point in X is in more than L_0 of the $B(x_i,2s)$. Here L_0 depends only on X.*

Proof. Since X is a locally compact separable metric space, there is an increasing sequence of compact sets $\{K_n\}_{n\geq 1}$ such that $\cup_{n\geq 1} K_n = B(x_0,R)$. Now, take $x_1^1 \in K_1$ and choose $x_2^1, x_3^1, \cdots \in K_1$ by letting x_{i+1}^1 be any point in $K_1 \setminus \cup_{j=1}^i B(x_j^1, s)$. We do this until we can no longer proceed. Since K_1 is compact, there is a finite subset $\{x_i\}_{i=1}^{l_1} \subset \{x_i^1\}_i$ such that $K_1 \subset \cup_{i=1}^{l_1} B(x_i,s)$. We next choose $x_1^2, x_2^2, \cdots \in K_2$ by letting x_{i+1}^2 be any point in $K_2 \setminus (\cup_{i=1}^{l_1} B(x_i,s) \cup \cup_{j=1}^i B(x_j^2,s))$. Again we do this until we can no longer proceed. By doing this procedure iteratively, we obtain a desired open covering of $B(x_0,R)$. Note that the x_i must be at least s distance apart, so that the balls $\{B(x_i,s/2)\}_i$ are disjoint. Now suppose y is in N of the balls $B(x_i,2s)$, $i \in \mathbb{N}$ (N may be infinite at this stage). Using (9.1), there exists such that for each of these we have $V(y,3s)/V(x_i,s/2) \leq N_0$. Since $B(y,3s)$ contains N disjoint balls $B(x_i,s/2)$,

$$V(y,3s) \geq \sum_{i: y \in B(x_i,2s)} V(x_i,s/2) \geq N N_0^{-1} V(y,3s),$$

which implies $N \leq N_0$, independent of y and s. □

9.2 Proof of (VD) + $(DUHK(\Psi)) \Rightarrow (\mathbf{E}(\Psi)_{\leq})$

Let $c_0 \geq 1$. By (9.1) and $(DUHK(\Psi))$, we have

$$\begin{aligned} P^y(\tau_{B(x,r)} > \Psi(c_0 r)) &\leq P^y(Y_{\Psi(c_0 r)} \in B(x,r)) \leq \int_{B(x,r)} p_{\Psi(c_0 r)}(y,z)d\mu(z) \\ &\leq \int_{B(x,r)} \frac{c_1}{V(z,c_0 r)} d\mu(z) \leq \int_{B(x,r)} \frac{2^\alpha c_1}{V(x,c_0 r)} d\mu(z) = \frac{2^\alpha c_1 V(x,r)}{V(x,c_0 r)}. \end{aligned}$$

By Lemma 9.1, we may choose c_0 so that the last value of the above inequality is less than 1/2. So, by the Markov property of $\{Y_t\}$, we conclude

$$P^y(\tau_{B(x,r)} > k\Psi(c_0 r)) \leq 2^{-k}, \qquad \forall k \geq 1.$$

Hence,

$$E^y[\tau_{B(x,r)}] \leq \sum_{k\geq 0} P^y\Big((k+1)\Psi(c_0 r) \geq \tau_{B(x,r)} \geq k\Psi(c_0 r)\Big)(k+1)\Psi(c_0 r) \leq 4\Psi(c_0 r),$$

for all $r > 0$ and $x, y \in X$. We thus obtain $(\mathrm{E}(\Psi)_{\leq})$ □

9.3 Oscillation inequalities and the Hölder continuity

In this subsection, we will assume (EHI) and deduce various Oscillation inequalities and Hölder continuity of harmonic functions.

Let u be nonnegative and harmonic in $B(x_0,R)$. To be precise, the definition of (EHI) in subsection 3.2 (III) should have been,

$$\text{ess sup}_{B(x_0,R/2)} u \leq c_1 \text{ess inf}_{B(x_0,R/2)} u. \tag{9.2}$$

x_0 here is x in the definition of (EHI). We will show here that (9.2) implies the continuity of u inside the ball $B(x_0,R)$, so that (EHI) holds. Indeed, take x_1 and r such that $B(x_1,3r)\subset B(x_0,R)$. By looking at $Cu+D$ for suitable constants C and D, we may suppose that ess $\sup_{B(x_1,2r)}u=1$ and ess $\inf_{B(x_1,2r)}u=0$. Hence by (9.2), we have

$$\text{ess}\sup_{B(x_1,r)}u-\text{ess}\inf_{B(x_1,r)}u\le(1-c_1^{-1})\text{ess}\sup_{B(x_1,r)}u\le(1-c_1^{-1}).$$

So if $\rho=1-c_1^{-1}$ then

$$\text{ess}\sup_{B(x_1,r)}u-\text{ess}\inf_{B(x_1,r)}u\le\rho[\text{ess}\sup_{B(x_1,2r)}u-\text{ess}\inf_{B(x_1,2r)}].$$

It follows easily that

$$\text{ess}\sup_{B(x_1,r)}u-\text{ess}\inf_{B(x_1,r)}u\le c_2r^\gamma \tag{9.3}$$

for some $\gamma>0$. Define $\widehat{u}(x_1)=\lim_{r\to0}\text{ess}\sup_{B(x_1,r)}u$. If one takes a countable basis $\{B_i\}$ for X and excludes those points $x\in B_i$ such that $u(x)\notin[\text{ess}\inf_{B_i}u,\text{ess}\sup_{B_i}u]$, then for every other x it is easy to see, using (9.3), that $u(x)=\widehat{u}(x)$. Thus, $\widehat{u}$ is equal to u for μ-almost every x. Moreover, from (9.3) we see that $\widehat{u}$ is continuous. Recall that in our definition of harmonic function we take a quasi-continuous modification as defined in [35]. We conclude $u=\widehat{u}$ quasi-everywhere, and so u has a quasi-continuous modification that is continuous. Using this modification and (9.2), we have

$$\sup_{B(x_0,R/2)}u\le c_1\inf_{B(x_0,R/2)}u,$$

which is the desired inequality.

Let $\mathcal{H}_{B(x_0,r)}$ be a space of harmonic functions on $B(x_0,r)$. Define the oscillation of a function f over B by $\text{Osc}_Bf:=\text{ess}\sup_Bf-\text{ess}\inf_Bf$. Then, the above arguments also show the following.

Lemma 9.3 *Assume (EHI).*
1) For any $\varepsilon>0$, there exists $\delta\in(0,1)$ such that

$$Osc_{B(x_0,\delta r)}u\le\varepsilon Osc_{B(x_0,r)}u,\qquad\forall u\in\mathcal{H}_{B(x_0,r)}.$$

2) There exist $c_1,\gamma>0$ such that

$$\sup_{x,y\in B(x_0,\rho r)}|u(x)-u(y)|\le c_1\rho^\gamma\sup_{x\in B(x_0,r)}|u(x)|,\qquad\forall\rho\in(0,1),\forall u\in\mathcal{H}_{B(x_0,r)}. \tag{9.4}$$

We can now prove the following Hölder continuity of harmonic functions.

Proposition 9.4 *Assume (EHI). There exists $\gamma>0$ such that for any $\delta\in(0,1)$, there exists $C=C_\delta>0$ so that the following holds,*

$$\sup_{x,y\in B(x_0,\delta r)}\Big\{\frac{|u(x)-u(y)|}{d(x,y)^\gamma}\Big\}\le Cr^{-\gamma}\sup_{x\in B(x_0,r)}|u(x)|,\qquad\forall u\in\mathcal{H}_{B(x_0,r)}.$$

Proof. Denote $B_r:=B(x_0,r)$. For $x,y\in B_{\delta r}$, we consider two cases. first, if $d(x,y)\ge(1-\delta)r$, then

$$|u(x)-u(y)|\le2\sup_{B_r}|u|\le2\{(1-\delta)r\}^{-\gamma}d(x,y)^\gamma\sup_{B_r}|u|.$$

If $d(x,y)<(1-\delta)r$, then $B(z,(1-\delta)r)\subset B_r$ contains both x and y, where $z\in X$ is the mid point of x and y. Further $x,y\in B(z,d(x,y))$. Applying (9.4) with $\rho=d(x,y)/\{(1-\delta)r\}$ yields

$$|u(x)-u(y)|\le c_1\{(1-\delta)r\}^{-\gamma}d(x,y)^\gamma\sup_{B_r}|u|.$$

We thus obtain the result. $\square$

We next discuss about the oscillation of Green functions. Given open set $\Omega \subset X$ and $f \in \mathcal{B}(\Omega)$, define the Green operator G^Ω as

$$G^\Omega f(x) = E^x\Big[\int_0^{\tau_\Omega} f(Y_t)dt\Big].$$

Denote $\bar{E}(\Omega) := \sup_z E^z[\tau_\Omega]$. When $\Omega = B(x,r)$, we will abbreviate $\bar{E}(B(x,r))$ as $\bar{E}(x,r)$. It is easy to see

$$\|G^\Omega\|_{L^\infty\to L^\infty} \le \bar{E}(\Omega). \tag{9.5}$$

Lemma 9.5 *Assume that $\bar{E}(\Omega) < \infty$. Then, for any $f \in C_0(\Omega)$, $G^\Omega f$ is harmonic in $\Omega \setminus Supp f$. Also, for any open set $\Omega' \supset \Omega$, $G^{\Omega'} f - G^\Omega f$ is harmonic in Ω.*

Proof. Let $u_f = G^\omega f$. Since $G^\omega = (-\Delta_\Omega)^{-1}$, we see that $u_f \in \mathcal{D}(\Delta_\Omega)$. So

$$\mathcal{E}(u_f, v) = -(\Delta_\Omega u_f, v) = (f, v) = 0, \qquad \forall v \in \mathcal{F}(\Omega \setminus \mathrm{Supp} f).$$

Thus, u_f is harmonic in $\Omega \setminus \mathrm{Supp} f$. Similarly, set $w_f = G^{\Omega'} f - G^\Omega f$, then

$$\mathcal{E}(w_f, v) = \mathcal{E}(G^{\Omega'} f, v) - \mathcal{E}(G^\Omega f, v) = (f, v)_{L^2(\Omega')} - (f, v)_{L^2(\Omega)} = 0,$$

for any $v \in \mathcal{F}(\omega)$. $\square$

Proposition 9.6 *Assume (EHI). Let $f : B(x,r) \to \mathbb{R}$ be a bounded Borel function and set $u_f = G^{B(x,R)} f$. Then, for any $0 < r < R$,*

$$Osc_{B(x,\delta r)} u_f \le 2(\bar{E}(x,r) + \varepsilon \bar{E}(x,R)\|f\|_\infty,$$

where ε and δ are the same as in Lemma 9.3 1).

Proof. If $\bar{E}(x,R) = \infty$, there is nothing to prove, so assume that $\bar{E}(x,R) < \infty$. Denote $B_r := B(x,r)$ and let $v_f = G^{B_r} f$. Then, by (9.5),

$$\|u_f\|_\infty \le \bar{E}(x,R)\|f\|_\infty, \qquad \|v_f\|_\infty \le \bar{E}(x,r)\|f\|_\infty. \tag{9.6}$$

By Lemma 9.5, $w_f := u_f - v_f$ is harmonic in B_r. Using Lemma 9.3 1) and $0 \le w_f \le u_f$, we obtain

$$\mathrm{Osc}_{B_{\delta r}} w_f \le \varepsilon \mathrm{Osc}_{B_r} w_f \le \varepsilon \|w_f\|_\infty \le \varepsilon \|u_f\|_\infty.$$

Since $u_f = v_f + w_f$,

$$\mathrm{Osc}_{B_{\delta r}} u_f \le \mathrm{Osc}_{B_{\delta r}} v_f + \mathrm{Osc}_{B_{\delta r}} w_f \le \|v_f\|_\infty + \varepsilon\|u_f\|_\infty \le (\bar{E}(x,r) + \varepsilon \bar{E}(x,R)\|f\|_\infty,$$

where we used (9.6) in the last inequality. Thus we obtain the desired inequality for $f \ge 0$. For a general function f, write $f = f_+ - f_-$. Then $\mathrm{Osc}\, u_f = \mathrm{Osc}\,(u_{f_+} - u_{f_-}) \le \mathrm{Osc}\, u_{f_+} + \mathrm{Osc}\, u_{f_-}$, and the desired inequality is obtained. $\square$

9.4 Time derivative

We follow the arguments in [40, 42]. First, we show the following well-known fact in the semigroup theory.

Lemma 9.7 *For any $f \in L^2$, let $u_t = P_t f$. Then, we have*

$$\|\partial_t u_t\|_2 \le \frac{1}{t-s}\|u_s\|_2, \qquad 0 < \forall s < t.$$

Proof. Let $\{E_\lambda\}_{\lambda\geq 0}$ be spectral resolution of the operator $-\Delta$. Then we have

$$u_t = e^{t\Delta} f = \int_0^\infty e^{-t\lambda} dE_\lambda f, \quad \|u_t\|_2^2 = \int_0^\infty e^{-2t\lambda} d\|E_\lambda f\|^2.$$

Thus, we have

$$\partial_t u_t = \int_0^\infty (-\lambda) e^{-t\lambda} dE_\lambda f, \quad \|\partial_t u_t\|_2^2 = \int_0^\infty \lambda^2 e^{-2t\lambda} d\|E_\lambda f\|^2 = \int_0^\infty \lambda^2 e^{-2(t-s)\lambda} e^{-2s\lambda} d\|E_\lambda f\|^2.$$

Since $\lambda e^{-(t-s)\lambda} \leq (t-s)^{-1}$, we obtain

$$\|\partial_t u_t\|_2^2 \leq \frac{1}{(t-s)^2} \int_0^\infty e^{-2s\lambda} d\|E_\lambda f\|^2 = \frac{1}{(t-s)^2} \|u_s\|_2^2,$$

which is the desired estimate. □

Corollary 9.8 *For $t > 0$ and $z \in X$, the function $t \mapsto p_t(\cdot, z)$ is Frechet differentiable in L^2 and*

$$\|\partial_t p_t(\cdot, z)\|_2 \leq \frac{1}{t-s} \sqrt{p_{2s}(z,z)}, \qquad 0 < \forall s < t.$$

Proof. Let $f = p_\varepsilon(\cdot, z)$ for some $\varepsilon > 0$. Then, $u_t = P_t f = p_{t+\varepsilon}(\cdot, z)$. Thus, by Lemma 9.7,

$$\|\partial_t p_{t+\varepsilon}(\cdot, z)\|_2 \leq \frac{1}{t-s} \|p_{s+\varepsilon}(\cdot, z)\|_2 = \frac{1}{t-s} \sqrt{p_{2(s+\varepsilon)}(z,z)}.$$

Replacing $t+\varepsilon, s+\varepsilon$ by t, s respectively, we obtain the result. □

Proposition 9.9 *For any $x, y \in X$, the function $t \mapsto p_t(x,y)$ is differentiable in $t > 0$ and*

$$|\frac{\partial_t}{\partial t} p_t(x,y)| \leq \frac{2}{t} \sqrt{p_{t/2}(x,x) p_{t/2}(y,y)}.$$

Proof. By the Chapman-Kolmogorov equation, $p_t(x,y) = (p_{t-s}(\cdot, x), p_s(\cdot, y))$ for any $s \in (0,t)$, so that $\partial_t p_t(x,y) = (\partial_t p_{t-s}(\cdot, x), p_s(\cdot, y))$. Thus, applying Corollary 9.8,

$$|\frac{\partial_t}{\partial t} p_t(x,y)| \leq \| \partial_t p_{t-s}(\cdot, x)\|_2 \|p_s(\cdot, y)\|_2 \leq \frac{1}{t-s-r} \sqrt{p_{2r}(x,x) p_{2s}(y,y)}, \qquad 0 < \forall r < t-s.$$

Taking $s = r = t/4$, we obtain the result. □

9.5 Proof of Theorem 3.1: $(d) \Rightarrow (e)$

Recall from [35] Section 1.6 the definition of invariant sets and an irreducible Dirichlet form.

Lemma 9.10 *Let X satisfy (EHI). Then $\mathcal{E}$ is irreducible.*

Proof. Let A be an irreducible set, and suppose both $\mu(A) > 0$ and $\mu(A^c) > 0$. Then there exists a ball $B = B(x,R)$ with $\mu(A \cap B') > 0$ and $\mu(A^c \cap B') > 0$, where $B' = B(x, R/2)$. Since $P_t 1_A = 1_A$ it follows that $u = 1_A$ and $v = 1_{A^c}$ are harmonic on B. So by (EHI) we have

$$\tilde{u}(x) \leq C \tilde{u}(y), \quad x, y \in B'.$$

Since $u > 0$ on a set of positive measure, we have that there exists $x \in B'$ with $\tilde{u}(x) > 0$; hence by the (EHI), $\tilde{u} > 0$ on B'. But as $\tilde{u} = 1_A$ μ-a.e., we deduce that $\mu(A^c \cap B') = 0$, a contradiction. □

Proposition 9.11 *Let X satisfy (EHI), and $B = B(x, R)$. Then $Gg < \infty$ on B if $g \in L^1_+(B)$.*

PROOF. (sketch). Consider the Dirichlet form $\mathcal{E}_B$ with domain $\mathcal{F}_B = \{f \in \mathcal{F} : f|_{B^c} = 0\}$. Let $A = B(x, R/2)$ and $h(x) = P^x(T_A < \tau_B)$. Then h is excessive with respect to $\mathcal{E}_B$. If h were constant on B then we would have $h = 1$ on B, and the set B would be an invariant set for $\mathcal{E}$. Thus h is non-constant.

So by Ex. (4.22), p. 89 in [21], we deduce that the killed semigroup P^B_t is transient. Hence (see [35] Section 1.6) we have $Gg < \infty$ for any $g \in L^1_+(B, \mu)$. □

Lemma 9.12 *Let D be a bounded domain in X. Then (EHI) implies that there exists the Green density $g^D(\cdot, \cdot)$ which is continuous on $(X \times X) \setminus \Delta_g$ and $g_D(x, y) = g_D(y, x)$ for all $x, y \in (X \times X) \setminus \Delta_g$, where Δ_g is the diagonal. Further, there exists $C > 0$ such that for any $r > 0$, if $y_0, y_1 \in X$ satisfy $d(y_0, y_1) \geq 2r$, then*

$$g_D(y_0, x) \leq C g_D(y_0, y) \qquad \forall x, y \in B(y_1, r). \tag{9.7}$$

PROOF. Let $x_0, x_1 \in D$, Choose $r > 0$ such that $B(x_i, 2r) \subset D$, $B(x_0, 2r) \cap B(x_1, 2r) = \emptyset$. Write $B_i = B(x_i, 2r)$, $B'_i = B(x_i, r)$. Let $f, g \in \mathcal{F}$ with supports in B'_0 and B'_1, and $\int f = \int g = 1$. Let G_D be the Green operator for the process Y killed on exiting D. By Proposition 9.11 we have $G_D f < \infty$, $G_D g < \infty$.

Then if $u \in \mathcal{F}$ with $\operatorname{Supp} u \subset B(x_1, 2r)$,

$$\mathcal{E}(G_D f, u) = (f, u) = 0, \tag{9.8}$$

so $G_D f$ is harmonic on B_1. Similarly $G_D g$ is harmonic on B_0. By the (EHI) if $x \in B'_1$ then

$$G_D f(x) \leq C G_D f(y), \quad y \in B'_1. \tag{9.9}$$

Similarly

$$G_D g(x) \leq C G_D g(x_0), \quad x \in B'_0.$$

So

$$G_D f(x_1) \leq C(g, G_D f) = C(G_D g, f) \leq C^2 G_D g(x_0).$$

Now fix g such that $C_1 = G_D g(x_0) < \infty$ – such a g exists by choosing $g \leq c h_0$. Then we have $G_D f(x_1) \leq c' ||f||_1$ for all f with support in B'_0. Therefore the kernel $G_D(x_1, dx)$ has a density $g_D(x_1, y)$ on B'_0. Since $(f, G_D g) = (G_D f, g)$ for $f, g \in L^2$, it follows that $g_D(x, y) = g_D(y, x)$ $\mu \times \mu$–a.e.

Now, take $y_0, y_1 \in X$ that satisfy $d(y_0, y_1) \geq 2r$. For any $\epsilon > 0$ and $f \in L^2$ with support in $B(y_0, \epsilon r)$, similarly to (9.8) we can show that $G_D f$ is harmonic on $B(y_1, (2 - \epsilon)r)$. Thus, by the same way as (9.9), we have

$$G_D f(x) \leq C G_D f(y), \quad x, y \in B(y_1, r). \tag{9.10}$$

Now let $f_n(z) = V(y_0, r_n)^{-1} 1_{B(y_0, r_n)}(z)$ where $\epsilon r \geq r_n \downarrow 0$. Applying (9.10) to f_n and take $n \to \infty$, we obtain (9.7) for μ-a.e. y_0. By the usual oscillation argument, we can deduce that $g_D(x, y)$ is continuous on $(X \times X) \setminus \Delta_g$. Especially, $g_D(x, y) = g_D(y, x)$ for all $x, y \in (X \times X) \setminus \Delta_g$. We thus obtain (9.7) for all $y_0 \in X$. □

Now let $M \geq 2$ be fixed. (In fact, we can take M=2.)

Definition 9.13 $(\mathcal{E}, \mathcal{F})$ *satisfies (HG) if there exists a constant $c_1 > 0$ such that for any ball $B(x_0, R)$, there exists the Green kernel $g^{B_R}(x_0, y)$ and for any $0 < r \leq R/M$, we have*

$$\sup_{y \notin B(x_0, r)} g^{B_R}(x_0, y) \leq c_1 \inf_{y \in B(x_0, r)} g^{B_R}(x_0, y). \tag{HG}$$

Lemma 9.14 *(EHI) $\Rightarrow$ (HG).*

PROOF. We prove that if $d(x_0, x) = d(x_0, y) = R$, and $B(x_0, 2R) \subset D$ then

$$C_1^{-1} g_D(x_0, y) \le g_D(x_0, x) \le C_1 g_D(x_0, y). \tag{9.11}$$

Once (9.11) is proved, then (HG) holds by the maximum principle (which holds for $G_D f$ and so for g_D as well). By symmetry it is enough to prove the right hand inequality of (9.11).

Let x', y' be the midpoints of $\gamma(x_0, x)$, and $\gamma(x_0, y)$. Thus $d(x_0, x') = d(x_0, y') = R/2$. Clearly we have $d(x', y) \ge R/2$ and $d(x, y') \ge R/2$.

We now consider two cases.

Case 1. $d(x', y') \le R/3$. Let z be the midpoint of $\gamma(x', y')$. Then $d(z, x') \le R/6 \le R/4$. So applying (9.7) to $g_D(x_0, \cdot)$ in $B(x', R/4) \subset B(x', R/2)$, we deduce that

$$C_2^{-1} g_D(x_0, x') \le g_D(x_0, z) \le C_2 g_D(x_0, x').$$

Now apply (9.7) to $g_D(x_0, \cdot)$ in $B(x, R/2) \subset B(x, R)$, to deduce that

$$C_2^{-1} g_D(x_0, x) \le g_D(x_0, x') \le C_2 g_D(x_0, x).$$

Combining these inequalities we deduce that

$$C_2^{-2} g_D(x_0, x) \le g_D(x_0, z) \le C_2^2 g_D(x_0, x),$$

and this, with a similar inequality for $g_D(x_0, y)$, proves (9.11).

Case 2. $d(x', y') > R/3$. Apply (9.7) to $g_D(y, \cdot)$ in $B(x_0, R/2) \subset B(x_0, R)$, to deduce that

$$C_2^{-1} g_D(y, x') \le g_D(y, x_0) \le C_2 g_D(y, x'). \tag{9.12}$$

Now look at $g_D(x', \cdot)$. If z' is on $\gamma(y', y)$ with $d(y', z') = s \in [0, R/2]$ then as $d(x', y') > R/3$ and $d(x', y) \ge R/2$ we have $d(x', z') \ge \max(R/3 - s, s)$. Hence we deduce $d(x', z') \ge R/6$. So applying (9.7) repeatedly to $g_D(x', \cdot)$ for a chain of balls $B(z', R/12) \subset B(z', R/6)$ we deduce that

$$C_2^{-6} g_D(x', y') \le g_D(x', y) \le C_2^6 g_D(x', y'). \tag{9.13}$$

So, we obtain from (9.12) and (9.13),

$$g_D(y, x_0) \le C_2 g_D(y, x') \le C_2^7 g_D(x', y'), \quad g_D(x', y') \le C_2^6 g_D(y, x') \le C_2^7 g_D(y, x_0).$$

We have similar inequalities relating $g_D(x, x_0)$ and $g_D(x', y')$, which proves (9.11). □

Lemma 9.15 *Assume that $(\mathcal{E}, \mathcal{F})$ satisfies (HG).*
1) For any ball $B(x_0, R)$ and for any $0 < r \le R/M$, we have

$$\sup_{y \notin B(x_0, r)} g^{B_R}(x_0, y) \asymp R(B_r, B_R^c) \asymp \inf_{y \in B(x_0, r)} g^{B_R}(x_0, y). \tag{9.14}$$

2) Let $B_k = B(x_0, M^k r)$ for $k = 0, 1, \cdots$. Then, for any integers $0 \le m < n$,

$$\sup_{y \notin B_m} g^{B_n}(x_0, y) \asymp \sum_{k=m}^{n-1} R(B_k, B_{k+1}^c) \asymp \inf_{y \in B_m} g^{B_n}(x_0, y). \tag{9.15}$$

PROOF. For 1), first the following is standard (see for example (4.7) in [41]).

$$\sup_{y \notin B(x_0, r)} g^{B_R}(x_0, y) \ge R(B_r, B_R^c) \ge \inf_{y \in B(x_0, r)} g^{B_R}(x_0, y).$$

Thus, using (HG), we obtain (9.14).

For 2), note first that the following holds by the definition of resistance

$$\sum_{k=m}^{n-1} R(B_k, B_{k+1}^c) \le R(B_m, B_n^c).$$

This and (9.14) implies the lower bound for $\inf g^{B_n}$ in (9.15). Next, by the reproducing property of g^{B_k}, we know that $g^{B_{k+1}}(x,\cdot) - g^{B_k}(x,\cdot)$ is a harmonic function in B_k. Thus,

$$g^{B_{k+1}}(x,y) - g^{B_k}(x,y) \le \sup_{z\notin B_k} g^{B_{k+1}}(x,z) \le cR(B_k, B_{k+1}), \quad \forall y \in X, \tag{9.16}$$

where the first inequality is by the maximum principle and the second inequality is by (9.14). For $y \notin B_m$, by (9.14)

$$g^{B_{m+1}}(x,y) \le c' R(B_m, B_{m+1}). \tag{9.17}$$

For such y, adding up (9.17) with (9.16) for $m < k < n$, we obtain the upper bound of $\sup g^{B_n}$ in (9.15). $\square$

Proof of (VD) + (EHI) + (RES(Ψ)) $\Rightarrow$ (E(Ψ)).

$$E^{x_0}[\tau_{B_R}] = \int g^{B_R}(x_0,y)d\mu(y) \ge \int_{B(x_0,r)} g^{B_R}(x_0,y)d\mu(y) \ge cR(B_r, B_R^c)V(x_0,r) \ge c\Psi(R),$$

where we used Lemma 9.15 1) in the second inequality and (VD) + (RES(Ψ)) in the last inequality.

Now, for each $k \in \mathbb{Z}$, let $r_k = M^k$, $B_k = B(x_0, r_k)$ and let n_0 be the minimum number such that $R < r_{n_0}$. Then

$$\begin{aligned}
E^{x_0}[\tau_{B_R}] &\le E^{x_0}[\tau_{B(x_0,r_{n_0})}] = \int_{B_{n_0}} g^{B_{n_0}}(x_0,y)d\mu(y) \\
&= \sum_{m=-\infty}^{n_0-1} \int_{B_{m+1}\setminus B_m} g^{B_m}(x_0,y)d\mu(y) \le c \sum_{m=-\infty}^{n_0-1} \Big(\sum_{k=m}^{n_0-1} R(B_k, B_{k+1}^c)\Big)\mu(B_{m+1}\setminus B_m) \\
&= c \sum_{k=-\infty}^{n_0-1} \Big(\sum_{m=-\infty}^{k} \mu(B_{m+1}\setminus B_m)\Big) R(B_k, B_{k+1}^c) = c \sum_{k=-\infty}^{n_0-1} \mu(B_{k+1})R(B_k, B_{k+1}^c) \\
&\le c' \sum_{k=-\infty}^{n_0-1} \Psi(r_{k+1}) \le c''\Psi(R),
\end{aligned}$$

where we used Lemma 9.15 2) in the second inequality and (VD) + (RES(Ψ)) in the third inequality. We thus obtain (E(Ψ)). $\square$

9.6 Proof of Theorem 3.1: *(b)* $\Rightarrow$ *(a)*

Fix $x_0 \in X$ and for $R > 0$, let $B_R := B(x_0, R)$. Let $\mathcal{F}_{B_R} = \{u \in L^2(X,\mu) : u = 0$ μ-a.e. on $B_R^c\}$ and consider the part of the Dirichlet form $(\mathcal{E}, \mathcal{F}_{B_R})$ (see [35] Section 4.4). Let $\{P_t^{B_R}\}$ be the corresponding semigroup.

Lemma 9.16 *There exists a version of the heat kernel $p_t^{B_R}(x,y)$ for $\{P_t^{B_R}\}$ and, for each $\varepsilon_1, \varepsilon_2 \in (0,1)$, there exists $c_{\varepsilon_1,\varepsilon_2} > 0$ such that*

$$p_t^{B_R}(x,y) \ge \frac{c_{\varepsilon_1,\varepsilon_2}}{V(x_0, \varepsilon_1 R)},$$

for all $x, y \in B(x_0, \varepsilon_1 R)$ and $\varepsilon_2\Psi(R) < t < \Psi(R)$.

Proof. First, define
$$p_t^{B_R}(x,y) := p_t(x,y) - E^x[p_{t-\tau_{B_R}}(Y_{\tau_{B_R}},y), \tau_{B_R} \le t], \tag{9.18}$$
where Y_t is the diffusion process corresponding to $(\mathcal{E},\mathcal{F})$ and $\tau_{B_R} = \inf\{t \ge 0 : Y_t \notin B(x_0,R)\}$. Then, it is easy to check, using the strong Markov property, that $p_t^{B_R}(x,y)$ is a version of the heat kernel for $\{P_t^{B_R}\}$. The proof of (9.18) is now a standard argument (see, for example, Lemma 5.1 in [34]). □

Let $d\nu = dt \otimes d\mu$, $\mathcal{H} = L^2(\mathbb{R}^1 \times X, d\nu)$ and $\tilde{\mathcal{F}} = \{u : \mathbb{R}^1 \to \mathcal{F} : \mathcal{A}(u,u) + \|u\|_{\mathcal{H}}^2 < \infty\}$ where $\mathcal{A}(u,u) = \int_{\mathbb{R}^1} \mathcal{E}(u(t,\cdot),u(t,\cdot))dt$. Let $\tilde{\mathcal{F}}^* = \{u : \mathbb{R}^1 \to \mathcal{F}^* : \int_{\mathbb{R}^1} \|u(t,\cdot)\|_{\mathcal{F}^*}^2 dt + \|u\|_{\mathcal{H}}^2 < \infty\}$, where $\mathcal{F}^*$ is the dual of $\mathcal{F}$ in the sense $\mathcal{F} \subset L^2(X,\mu) \subset \mathcal{F}^*$. Note that $\tilde{\mathcal{F}} \subset \mathcal{H} = \mathcal{H}^* \subset \tilde{\mathcal{F}}^*$. Let

$$\begin{aligned}
\tilde{\mathcal{W}} &= \{u \in \tilde{\mathcal{F}} : \frac{\partial u}{\partial t} \in \tilde{\mathcal{F}}^*\} \\
\tilde{\mathcal{E}}(u,v) &= (u, \frac{\partial v}{\partial t})_\nu + \mathcal{A}(u,v) \quad \text{if} \;\; u \in \tilde{\mathcal{F}}, v \in \tilde{\mathcal{W}},
\end{aligned}$$

where $(u,v)_\nu = \int_{\mathbb{R}^1}\int_X uv\, d\mu\, dt$. Let $\{Y_t(x)\}$ be the diffusion process corresponding to $(\mathcal{E},\mathcal{F})$. Then the semigroup corresponding to $\tilde{\mathcal{E}}$ can be written as $P_t u(t_0,x_0) = E[u(t_0+t, Y_t(x_0))]$ so that the corresponding generator is $\frac{\partial}{\partial t} + \mathcal{L}$ (the corresponding diffusion is $Z_t = (t, Y_t)$), whereas the dual semigroup $\{\hat{P}_t\}$ can be written as $\hat{P}_t u(t_0,x_0) = E[u(t_0 - t, Y_t(x_0))]$ and the corresponding generator is $-\frac{\partial}{\partial t} + \mathcal{L}$. (See [71] for details.)

Lemma 9.17 *Let u be a non-negative solution of the heat equation on $Q := I \times G$, where $I = (a,b)$ and G is an open connected subset of X. Then $u(t,x) \ge \int p_{t-s}^B(x,y)u(s,y)d\mu(y)$ μ-a.e. x and all $0 < s < t$ where $B \subset\subset G$.*

Proof. The claim is equivalent to $(u - \hat{P}_{t-s}^Q u)(t,x) \ge 0$ for all $(t,x) \in Q$ and all $0 < s < t$.

Let $\alpha > 0$. Then, $\tilde{\mathcal{E}}_\alpha(u,g) \ge 0$ for all non-negative $g \in \tilde{\mathcal{F}}_Q$. So, for any non-negative α-excessive function (w.r.t. $(\tilde{\mathcal{E}},\tilde{\mathcal{F}}_Q)$ –see [71] Section 4.3, for a discussion of excessive functions in the parabolic case) $v \in \mathcal{F}_Q$, we have

$$\begin{aligned}
(u - e^{-\alpha s}\hat{P}_s^Q u, v)_\nu &= (u, v - e^{-\alpha s}P_s^Q v)_\nu \\
&= (u_Q, v - e^{-\alpha s}P_s^Q v)_\nu + (H_Q^\alpha u, v - e^{-\alpha s}P_s^Q v)_\nu \\
&\ge (u_Q, v - e^{-\alpha s}P_s^Q v)_\nu = \tilde{\mathcal{E}}_\alpha(u, G_\alpha^Q v - e^{-\alpha s}P_s^Q G_\alpha^Q v)_\nu =: I_1,
\end{aligned}$$

where $u = u_Q + H_Q^\alpha u$ is the orthogonal decomposition of u into $\mathcal{F}_Q \oplus \mathcal{H}_{Q^c}^\alpha$ (see p. 149 of [35] –the same proof works for the parabolic case). Here the inequality in the third line is because $H_Q^\alpha u(x) = E^x(e^{-\alpha\sigma_{Q^c}}u(Z_{\sigma_{Q^c}})) \ge 0$ (due to Lemma 5.1.3 in p. 105 of [71]) and the fact that v is α-excessive (the definition of excessive functions in [71] is different from that in [35], but the proof of Theorem 2.2.1 in [35] also establishes equivalent conditions for the parabolic case, too). Since $G_\alpha^Q v - e^{-\alpha s}P_s^Q G_\alpha^Q v = \int_0^s e^{-\alpha l} P_l^Q v dl \in \mathcal{F}_Q$ is non-negative on Q, $I_1 \ge 0$. Thus $u - e^{-\alpha s}\hat{P}_s^Q u \ge 0$ on Q. Since this holds for all $\alpha > 0$, we have $u \ge \hat{P}_s^Q u$ on Q. □

Once these properties are established, then proving (a) is standard; prove the oscillation inequality first and then use the inequality to establish (PHI(Ψ)). Indeed, the proof of Lemma 5.2 and Theorem 5.4 in [34] work line by line, with suitable changes of the scaling exponents.

9.7 Proof of Theorem 3.1: $(a) \Rightarrow (b)$

There is a standard argument, given in [73] and Section 5.5 of [72] which proves that (PHI(Ψ)) implies (VD), (PI(Ψ)), and (HK(Ψ)). See also [46] for the case $\Psi(s) \neq s^2$. However, as this argument uses existence and regularity of caloric and harmonic functions, we will give more complete details of the initial stages of this argument.

First, if $f \in L^2(X,\mu)$ we have that $P_t f \in \mathcal{D}(\Delta)$, and $v(t,x) = P_t f(x)$ is a solution to the heat equation in $X \times (0,\infty)$. Let $x \in X$, $t > 0$, $r = \Psi(t)$ and $f \geq 0$ with $\int f = 1$. Then applying $(\Psi(\Psi))$ in $Q = (0,4t) \times B(x,2r)$ we obtain

$$\sup_{Q_-} \tilde{v} \leq C \inf_{Q_+} \tilde{v}.$$

Hence if $B = B(x,r)$ then since $\int P_s f = 1$

$$\mu(B) \sup_{Q_-} \tilde{v} \leq C \int_B v(2t,y)\mu(dy) \leq C.$$

Thus for each $x \in X$ we have

$$\widetilde{P_t f}(x) \leq c(t)||f||_1.$$

Given this inequality, we can use the same arguments as in p. 52 of [7] (using the results in [79]) to deduce the existence of a transition density $p_t(x,y)$.

Lemma 9.18 *There exist an exceptional set N and a jointly measurable transition density $p_t(x,y)$, $t > 0$, $x, y \in (X \setminus N) \times (X \setminus N)$, such that*

$$\begin{aligned}
P_t(x,A) &= \int_A p_t(x,y)\mu(dy) \qquad \forall x \in X \setminus N,\ \forall t > 0,\ \forall A \in \mathcal{B}(X \setminus N), \\
p_t(x,y) &= p_t(y,x) \qquad \forall x,y,t, \\
p_{t+s}(x,z) &= \int p_s(x,y)p_t(y,z)\mu(dy) \qquad \forall x,z,t,s.
\end{aligned}$$

Since $p_t(x,y) = P_{t/2}p_{t/2}(\cdot,y)(x)$ it follows that $p_t(\cdot,y)$ is a solution of the heat equation. Now take a quasi continuous modification $\tilde{p}_t(x,y)$ w.r.t. x and use it in the procedure of (4) in [79]. Then, by Theorem 1 in [79], there exists $p_t(x,y)$ which is quasi continuous and satisfies the three equalities in Lemma 9.18. (In fact, the uniqueness criteria in Theorem 1 in [79] shows that this $p_t(x,y)$ is the same as the original one.) Thus it satisfies the (PHI(Ψ)), and so can be extended to $(0,\infty) \times X \times X$ as a jointly continuous function.

We now sketch the argument that (PHI(Ψ)) implies (VD), (PI(Ψ)), and (HK(Ψ)). We begin with (VD), which also gives a key lower bound on the transition density for the killed process. Applying the (PHI(Ψ)) to the function $u(t,x) = p_t(x_0,x)$ in the region $Q(x_0,0,R)$ we obtain (writing $T = \Psi(R)$)

$$p_{2T}(x_0,x_0) \leq c p_{4T}(x_0,y), \quad y \in B(x_0,R).$$

Integrating over $B(x_0,R)$ gives

$$p_{2T}(x_0,x_0)V(x_0,R) \leq c \int_{B(x_0,R)} u(4T,y) \leq c, \tag{9.19}$$

which gives an upper bound on $p_{2T}(x_0,x_0)$ in terms of the volume of balls.

To obtain a lower bound, write $B_\lambda = B(x_0,\lambda R)$, and let $\varphi \in \mathcal{F}$ be a cut-off function for $B_{5/2} \subset B_3$. Let $p_t^0(x,y)$ be the heat kernel for the process Y killed on exiting B_4. Define

$$u(t,x) = \begin{cases} \varphi(x), & x \in B_2,\ 0 < t \leq 2T, \\ \int_{B_3} p^0_{t-2T}(x,y)\varphi(y)\mu(dy), & x \in B_2,\ 2T < t \leq 4T. \end{cases}$$

Lemma 9.19 *u is a solution of the heat equation in $Q(x_0,T,R)$.*

Proof. The function $u_t(x,t) = \frac{\partial u}{\partial t}$ exists for $t > 2T$, and is zero for $t < 2T$. Since $u(x,t)$ is continuous at $t = 2T$ for $x \in B$, it is straightforward to check that u_t is the derivative of u in the Schwartz' distribution sense.

Since we have $u(t,\cdot) \in \mathcal{D}(\Delta)$ for all $t > 2T$, we have for $f \in \mathcal{F} \cap C(X)$ with support in B_2 that

$$\int f u_t \, d\mu = -\mathcal{E}(f, u(t,\cdot)), \qquad t > 2T. \tag{9.20}$$

If $t < 2T$ then since $u = 1$ on B_2 (9.20) also holds for $t < 2T$. Thus it follows that (3.3) holds. □

We can now, as in [73, 72, 46], use (PHI(Ψ)) in $Q(x_0, 0, R)$ to obtain

$$1 = u(y, 2T) \leq cu(x_0, 4T) \leq c \int_{B_3} p^0_{2T}(x_0, y), \quad y \in B(x_0, R). \tag{9.21}$$

Using (PHI(Ψ)) in a chain of regions $Q(y_i, t_i, r) \subset [0, 4T] \times B(x_0, 4R)$ we obtain

$$p^0_{2T}(x_0, y') \leq cp^0_{4T}(x_0, y), \quad y' \in B(x_0, 3R),\, y \in B(x_0, R). \tag{9.22}$$

Integrating (9.22) over $y' \in B_3$ gives

$$\int_{B_3} p^0_{2T}(x_0, y')\mu(dy') \leq cp^0_{4T}(x_0, x_0)V(x_0, 3R),$$

and combining this with (9.21), we deduce that

$$V(x_0, 3R)^{-1} \leq cp^0_{4T}(x_0, y), \quad y \in B(x_0, R). \tag{9.23}$$

The inequalities (9.19) and (9.23) control $p_t(x_0, x_0)$ from above and below in terms of the volume of balls, and since $t \to p_t(x_0, x_0)$ is decreasing one easily deduces, by the same arguments as in [72], that volume doubling holds.

Given the lower bound (9.23), the proof of (HK(Ψ)) now follows as in Section 5 of [46] and in the proof of Proposition 4.1. For the global lower bound one uses (9.23) and a standard chaining argument (Step 5 of the proof of Proposition 4.1). (9.23) gives uniform control of the probability that Y exits a ball radius r before time $t = \Psi(r)$, and using this the upper bounds on $p_t(x, y)$ follow as in p. 1472–1475 of [46].

We remark that (9.23) also gives a lower bound on the transition density of the process Y reflected at ∂B (see [26]). Using this the argument of [73] can be used to obtain (PI(Ψ)).

Remark. The equivalence $(a) \Leftrightarrow (b)$ is well-known for manifolds when $\Psi(s) = s^2$. For MMD with $\Psi(s) = s^2$, it is indirectly proved in [75]. (There it is proved that each condition is equivalent to (VD) + (PI(2)).) For MMD with general time scaling, [46] proves the equivalence assuming apriori that solutions to the heat equation are sufficiently regular. (See also [41] for the case of an infinite connected weighted graph.) We have proved the equivalence without assuming any apriori condition for solutions to the heat equation.

9.8 Proof of Proposition 4.5

This first step is to use (CS(Ψ)) to obtain the following weighted Poincaré and Sobolev inequalities, which will replace (2.5) in the iteration argument in subsection 2.4.

Proposition 9.20 *(Weighted Poincaré inequalities) Let $I = B(x, s)$ with $s \leq R$. Suppose f and its gradient are square integrable over $I^* = B(x, 2s)$. Let $f_A = \mu(A)^{-1} \int_A f d\mu$.*
(a) We have

$$\int_I f^2 d\gamma \leq c_1 (s/R)^{2\theta} \Psi(R) \Big(\int_{I^*} d\Gamma(f, f) + \Psi(s)^{-1} \int_{I^*} f^2 d\mu \Big). \tag{9.24}$$

(b) We have

$$\int_I (f - f_{I^*})^2 d\gamma \le c_2 (s/R)^{2\theta} \Psi(R) \int_{I^*} d\Gamma(f,f). \tag{9.25}$$

(c) If $J \subset I$, then

$$\int_J f^2 d\gamma \le c_3 (s/R)^{2\theta} \Psi(R) \int_{I^*} d\Gamma(f,f) + \mu(J)^{-1} \Big(\int_J |f| d\gamma \Big)^2.$$

(d) We have

$$\int_{B(x_0,R)} d\gamma \le c_4 V(x_0,R).$$

Proof. (a) Using the definition of γ and (3.5),

$$\begin{aligned} \int_I f^2 d\gamma &= \int_I f^2 d\mu + \Psi(R) \int_I f^2 d\Gamma(f,f) \\ &\le \int_I f^2 d\mu + c_5 (s/R)^{2\theta} \Psi(R) \int_{I^*} d\Gamma(f,f) + c_5 (s/R)^{2\theta} \Psi(R) \Psi(s)^{-1} \int_{I^*} f^2 d\mu. \end{aligned}$$

Since $\beta \ge \bar\beta \ge 2\theta$, and $s \le R$ this implies (a).

For (b), applying (9.24) to $f - f_{I^*}$ we have

$$\int_I (f - f_{I^*})^2 d\gamma \le c_6 (s/R)^{2\theta} \Psi(R) \Big(\int_{I^*} d\Gamma(f,f) + \Psi(s)^{-1} \int_{I^*} (f - f_{I^*})^2 d\mu \Big). \tag{9.26}$$

Using (PI(Ψ)) applied to the ball I^* we have

$$\int_{I^*} (f - f_{I^*})^2 d\mu \le c_7 \Psi(s) \int_{I^*} d\Gamma(f,f).$$

Substituting this into (9.26) gives (9.25).
(c) Now let $b = \int_J f d\gamma / \int_J d\gamma$. Then

$$\begin{aligned} \int_J f^2 d\gamma &= \int_J (f-b)^2 d\gamma + b^2 \int_J d\gamma \\ &\le \int_J (f - f_{I^*})^2 d\gamma + \Big(\int_J d\gamma \Big)^{-1} \Big(\int_J f d\gamma \Big)^2. \\ &\le \int_I (f - f_{I^*})^2 d\gamma + \mu(J)^{-1} \Big(\int_J |f| d\gamma \Big)^2. \end{aligned}$$

Using (9.25) to bound the first term of the above inequalities completes the proof of (c).
(d) follows from (a) by taking $s = R$ and $f = 1$, and using (VD). □

Our next result is a weighted Nash inequality. Recall that for any set $J \subset X$, $J^s := \{y : d(y,J) \le s\}$.

Proposition 9.21 *(Weighted Nash inequality) Let $s \le R$ and $J \subset B(x_0,R)$ be a finite union of balls of radius s. Suppose the gradient of f is square integrable over J^s and $\int_{J^s} f^2 d\gamma < \infty$. There exist $c_1 < \infty$ and $\alpha_1 \in (0,1)$ such that*

$$\mu(J)^{-1} \int_J f^2 d\gamma \le c_1 \Big[\Psi(R) \mu(J)^{-1} \int_{J^s} d\Gamma(f,f) + (s/R)^{-2\theta} \mu(J)^{-1} \int_J f^2 d\gamma \Big]^{1-\alpha_1} \Big[\mu(J)^{-1} \int_J |f| d\gamma \Big]^{2\alpha_1}.$$

PROOF. Suppose that $0 < t < s$. Using Lemma 9.2, we can cover J by balls $B(x_i, t)$ with $x_i \in J$ so that any point of J^s is in at most L_0 of the balls $B(x_i, 2t)$. Set $B_i = B(x_i, t) \cap J$ and $B_i^* = B(x_i, 2t)$. Then $\cup_i B_i = J$, $\cup_i B^* \subset J^s$, and $\sum \mu(B_i^*) \le L_0 \mu(J^s)$.

As J is a union of balls, for each i there exists y_i so that $d(x_i, y_i) = t/2$ and $B(y_i, t/2) \subset J$. Then by (9.1),

$$\frac{\mu(J)}{\mu(B_i)} \le \frac{\mu(J)}{\mu(B(y_i, t/2))} \le c_2 \left(\frac{R}{t}\right)^\alpha. \tag{9.27}$$

By Proposition 9.20 (c), and (9.27)

$$\begin{aligned}
\int_J f^2 d\gamma &\le \sum_i \int_{B_i} f^2 d\gamma \\
&\le c_3 (t/R)^{2\theta} \Psi(R) \sum_i \int_{B_i^*} d\Gamma(f,f) + \sum_i \frac{1}{\mu(B_i)} \Big(\int_{B_i} |f| d\gamma \Big)^2 \\
&\le c_4 (t/R)^{2\theta} \Psi(R) L_0 \int_{J^s} d\Gamma(f,f) + c_5 (R/t)^\alpha \mu(J)^{-1} \Big(\sum_i \int_{B_i} |f| d\gamma \Big)^2 \\
&\le c_6 (t/R)^{2\theta} \Psi(R) \int_{J^s} d\Gamma(f,f) + c_7 (R/t)^\alpha \mu(J)^{-1} \Big(\int_J |f| d\gamma \Big)^2.
\end{aligned}$$

Hence

$$\mu(J)^{-1} \int_J f^2 d\gamma \le c_8 [(t/R)^{2\theta} A + (R/t)^\alpha B], \tag{9.28}$$

where

$$A = \Big[\Psi(R) \mu(J)^{-1} \int_{J^s} d\Gamma(f,f) + (s/R)^{-2\theta} \mu(J)^{-1} \int_J f^2 d\gamma \Big], \quad B = \Big[\mu(J)^{-1} \int_J |f| d\gamma \Big]^2.$$

If $t \ge s$, (9.28) is obvious.

We choose t so that the two terms on the right hand side of (9.28) are equal. Thus $(t/R)^{2\theta+\alpha} = B/A$ (so $(t/R)^{2\theta} A = A^{1-2\theta/(2\theta+\alpha)} B^{2\theta/(2\theta+\alpha)}$), and substituting this into (9.28) completes the proof, with $\alpha_1 = 2\theta/(2\theta + \alpha)$. Note that if $\theta = 1$ and $\alpha = d$ we obtain the powers in the standard Nash inequality. $\square$

It is known that the Nash inequality is equivalent to the Sobolev inequality ([78, 25]). Using the fact, we obtain Proposition 4.5.

9.9 Proof of (4.27)

Without loss of generality, we multiply u by a constant so that $V(x_0, R)^{-1} \int_{B(x_0,R)} \log v = \overline{w} = 0$. Recall that v is either u or u^{-1} and define $\Phi(t) = \text{ess sup}_{\overline{Q}(t)} \log v$.

Lemma 9.22 *Let $1 \ge s > t > 0$. Then*

$$\Phi(s) \le \tfrac{3}{4} \Phi(t) + c_1 (s-t)^{-\zeta_1}. \tag{9.29}$$

PROOF. Fix t and write Φ for $\Phi(t)$. Let $c_2 > e$ satisfy $c_2 = 6 \log c_2$. If $\Phi(t) \le c_2$, then

$$\Phi(s) \le \Phi(t) \le \frac{3}{4} \Phi(t) + \frac{1}{4} c_2,$$

so that (9.29) holds provided $c_1 \ge c_2/4$.

Now suppose $\Phi > c_2$. From Proposition 9.20 (d) we have $\int_{Q(t)} d\gamma \le c_3 V(x_0,R)$. By Proposition 4.9 (b) and the fact that $v^p \le e^{p\Phi}$ on $Q(t)$,

$$\begin{aligned}
\int_{Q(t)} v^{2p} d\gamma &= \int_{Q(t)\cap\{\log v \ge \Phi/2\}} v^{2p} d\gamma + \int_{Q(t)\cap\{\log v < \Phi/2\}} v^{2p} d\gamma \\
&\le e^{2p\Phi} \int_{Q(t)\cap\{\log v \ge \Phi/2\}} d\gamma + e^{p\Phi} \int_{Q(t)\cap\{\log v < \Phi/2\}} d\gamma \\
&\le \frac{4c_4 e^{2p\Phi}}{\Phi^2} V(x_0,R) + e^{p\Phi} \int_{Q(t)} d\gamma \le c_5 \Big(\frac{e^{2p\Phi}}{\Phi^2} + e^{p\Phi} \Big) V(x_0,R).
\end{aligned}$$

Let $p = \frac{2}{\Phi} \log \Phi$, so that $e^{p\Phi} = \Phi^2$. As $\Phi > c_2$ we have $p < (2/c_2)\log c_2 = \frac{1}{3}$. So

$$V(x_0,R)^{-1} \int_{Q(t)} v^{2p} d\gamma \le c_5 e^{p\Phi} \Big(1 + \frac{e^{p\Phi}}{\Phi^2} \Big) = 2c_5 e^{p\Phi}.$$

Therefore by Corollary 4.8,

$$\begin{aligned}
\Phi(s) &= \frac{1}{2p} \log[\text{ess sup}_{Q(s)} v^{2p}] \le \frac{1}{2p} \log \Big[c_6 (s-t)^{-\zeta_1} V(x_0,R)^{-1} \int_{Q(t)} v^{2p} d\gamma \Big] \\
&\le \frac{1}{2p} \log \Big[c_7 (s-t)^{-\zeta_1} e^{p\Phi} \Big] = \Big[1 + \frac{\log(c_7 (s-t)^{-\zeta_1})}{2\log\Phi} \Big] \frac{\Phi}{2}. \qquad (9.30)
\end{aligned}$$

Without loss of generality we may take c_7 larger than c_2. If $\Phi(t) \ge c_7(s-t)^{-\zeta_1}$, then by (9.30) $\Phi(s) \le \frac{3}{4}\Phi(t)$, and (9.29) is satisfied. If, on the other hand, $\Phi(t) \le c_7(s-t)^{-\zeta_1}$, then since $\Phi(s) \le \Phi(t)$, we have (9.29) satisfied with $c_1 = c_7$. □

Proof of (4.27). Multiplying u by a constant we can assume $\int_{B(x_0,R)} \log u d\mu = 0$ as before. Choose $t_j = 1/(j+1)$, so that $t_0 = 1$ and $t_i \downarrow 0$. Then by Lemma 9.22,

$$\begin{aligned}
\Phi(t_0) &\le \tfrac{3}{4}\Phi(t_1) + c_2(t_0 - t_1)^{-\zeta_1} \\
&\le (\tfrac{3}{4})^2 \Phi(t_2) + c_2(t_0-t_1)^{-\zeta_1} + \tfrac{3}{4} c_2 (t_1 - t_2)^{-\zeta_1} \\
&\le \cdots \le (\tfrac{3}{4})^n \Phi(t_n) + \textstyle\sum_{i=1}^n (\tfrac{3}{4})^{i-1} c_2 (t_{i-1} - t_i)^{-\zeta_1},
\end{aligned}$$

for any $n \ge 0$. Since $\Phi(t_n) \le \text{ess sup}_{B(x_0,R)} \log v < \infty$, and

$$\sum_{i=1}^{\infty} (\tfrac{3}{4})^{i-1} c_2 (t_{i-1} - t_i)^{-\zeta_1} = c_3 < \infty,$$

we obtain (4.27). □

References

[1] S. Alexander and R. Orbach, *Density of states on fractals: "fractons"*, J. Physique (Paris) Lett., **43**, L625–L631 (1982).

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

[3] M.T. Barlow, *Anomalous diffusion and stability of Harnack inequalities*, Surveys in Differential Geometry IX, to appear.

[4] M.T. Barlow, *Some remarks on the elliptic Harnack inequality*, Bull. London Math. Soc., **37** (2005), 200–208.

[5] M.T. Barlow, *Random walks on supercritical percolation clusters*, Ann. Probab., **32** (2004), 3024-3084.

[6] M.T. Barlow, *Which values of the volume growth and escape time exponent are possible for a graph?*, Rev. Math. Iberoamericana, **20** (2004), 1–31.

[7] M.T. Barlow, *Diffusions on fractals*, L.N.M. **1690**, Springer, 1998.

[8] M.T. Barlow and R.F. Bass, Personal communications.

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

[10] M.T. Barlow and R.F. Bass, *Divergence form operators on fractal-like domains*, J. Funct. Analysis, **175** (2000), 214–247.

[11] M.T. Barlow and R.F. Bass, *Brownian motion and harmonic analysis on Sierpiński carpets*, Canad. J. Math., **51** (1999), 673–744.

[12] M.T. Barlow and R.F. Bass, *Construction of Brownian motion on the Sierpiński carpet,* Ann. Inst. Henri Poincaré, **25** (1989), 225–257.

[13] M.T. Barlow, R.F. Bass, T. Kumagai. *Note on the equivalence of parabolic Harnack inequalities and heat kernel estimates*, Unpublished note.
Available at: `http://www.kurims.kyoto-u.ac.jp/~kumagai/kumpre.html`

[14] M.T. Barlow and R.F. Bass and T. Kumagai, *Stability of parabolic Harnack inequalities on measure metric spaces*, J. Math. Soc. Japan, to appear.

[15] M. T. Barlow, T. Coulhon and T. Kumagai, *Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs*, Comm. Pure Appl. Math., to appear.

[16] M.T. Barlow and T. Kumagai, *Random walk on the incipient infinite cluster on trees*, preprint 2005.

[17] M.T. Barlow and E.A. Perkins, *Brownian Motion on the Sierpiński gasket*, Probab. Theory Relat. Fields, **79** (1988), 543–623.

[18] R.F. Bass, *Diffusions on the Sierpinski carpet*, Trends in probability and related analysis, Taipei, 1996, World Sci. Publishing, River Edge, NJ, 1997, 1–34.

[19] D. Ben-Avraham and S. Havlin, *Diffusion and reactions in fractals and disordered systems*, Cambridge Univ. Press, 2000.

[20] M. Biroli and U.A. Mosco, *Saint-Venant type principle for Dirichlet forms on discontinuous media*, Ann. Mat. Pura Appl., **169** (1995), 125–181.

[21] R.M. Blumenthal and R.K. Getoor, *Markov processes and potential theory.* Pure and Applied Mathematics, Vol. 29 Academic Press, New York-London 1968.

[22] E. Bombieri, *Theory of minimal surfaces and a counter-example to the Bernstein conjecture in high dimensions*, Mimeographed Notes of Lectures held at Courant Inst., New York Univ., 1970.

[23] E. Bombieri and E. Giusti, *Harnack's inequality for elliptic differential equations on minimal surfaces*, Invent. Math., **15** (1972), 24–46.

[24] C. Borgs, J.T. Chayes, H. Kesten and J. Spencer, *The birth of the infinite cluster: finite-size scaling in percolation*, Comm. Math. Phys., **224** (2001), no. 1, 153–204.

[25] E.A. Carlen, S. Kusuoka, and D.W. Stroock, *Upper bounds for symmetric Markov transition functions*, Ann. L'IHP, Sup au No2: 245–287, 1987.

[26] Z.Q. Chen, *On reflected Dirichlet spaces*, Probab. Theory Relat. Fields, **94** (1992), 135–162.

[27] T. Coulhon, *Heat kernel and isoperimetry on non-compact Riemannian manifolds*, Contemporary Mathematics, **338**, pp. 65–99, Amer. Math. Soc. 2003.

[28] T. Coulhon, *Ultracontractivity and Nash type inequalities*, J. Funct. Anal., **141** (1996), 510-539.

[29] T. Coulhon and T. Kumagai, in preparation.

[30] E.B. Davies, *Heat kernels and spectral theory,* (1989), Cambridge Univ. Press, Cambridge.

[31] E.B. Davies, *Explicit constants for Gaussian upper bounds on heat kernels,* Amer. J. Math., **109** (1987), 319–333.

[32] T. Delmotte, *Parabolic Harnack inequality and estimates of Markov chains on graphs*, Rev. Math. Iberoamericana, **15** (1999), 181–232.

[33] A. De Masi, P.A. Ferrari, S. Goldstein and W.D. Wick, *An invariance principle for reversible Markov processes. Applications to random motions in random environments*, J. Stat. Phys., **55** (1989), 787–855.

[34] E.B. Fabes and D.W. Stroock, *A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash*, Arch. Rational Mech. Anal., **96** (1986), 327–338.

[35] M. Fukushima, Y. Oshima and M.Takeda, *Dirichlet forms and symmetric Markov processes*, de Gruyter, Berlin, 1994.

[36] A. Grigoryan, *Heat kernel upper bounds on fractal spaces*, J. London Math. Soc., to appear.

[37] A. Grigor'yan, *Heat kernels and function theory on metric measure spaces,* Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), 143–172, Contemp. Math. **338**, A.M.S., Providence, RI, 2003.

[38] A. Grigoryan, *Heat kernel upper bounds on a complete non-compact manifold*, Rev. Mat. Iberoamericana, **10** (1994), 395–452.

[39] A. Grigor'yan, *The heat equation on non-compact Riemannian manifolds*, (in Russian) Matem. Sbornik., **182** (1991), 55–87. (English transl.) Math. USSR Sb., **72** (1992), 47–77.

[40] A. Grigor'yan and A. Telcs, *Two-sided estimates of heat kernels in metric measure spaces*, in preparation.

[41] A. Grigor'yan and A. Telcs, *Harnack inequalities and sub-Gaussian estimates for random walks*, Math. Annalen, **324** (2002), 521–556.

[42] A. Grigor'yan and A. Telcs, *Sub-Gaussian estimates of heat kernels on infinite graphs*, Duke Math. J., **109** (2001), 451–510.

[43] G.R. Grimmett, *Percolation.* (2nd edition), Springer, 1999.

[44] B.M. Hambly and T. Kumagai, *Heat kernel estimates for symmetric random walks on a class of fractal graphs and stability under rough isometries*, In: Fractal geometry and applications: A Jubilee of B. Mandelbrot (M.L. Lapidus and M. van Frankenhuijsen (eds.)), Proc. of Sympos. Pure Math. **72**, Part 2, pp. 233–260, Amer. Math. Soc. 2004.

[45] S. Havlin and D. Ben-Avraham, *Diffusion in disordered media*, Adv. Phys., **36** (1987), 695–798.

[46] W. Hebisch, L. Saloff-Coste, *On the relation between elliptic and parabolic Harnack inequalities*, Ann. Inst. Fourier (Grenoble), **51** (2001), 1437–1481.

[47] M. Hino, *On the singularity of energy measures on self-similar sets*, Probab. Theory Relat. Fields, **132** (2005), 265–290.

[48] R. van der Hofstad, *Infinite canonical super-Brownian motion and scaling limits*, Preprint 2004.

[49] R. van der Hofstad and A.A. Járai, *The incipient infinite cluster for high-dimensional unoriented percolation*, J. Stat. Phys., **114** (2004), 625-663.

[50] R. van der Hofstad, F. den Hollander and G. Slade, *Construction of the incipient infinite cluster for spread-out oriented percolation above* $4+1$ *dimensionals*, Comm. Math. Phys., **231** (2002), 435-461.

[51] R. van der Hofstad and G. Slade, *Convergence of critical oriented percolation to super-Brownian motion above* $4+1$ *dimensions*, Ann. Inst. Henri Poincaré Probab. Statist, **39** (2003), no. 3, 413–485.

[52] M. Kanai, *Analytic inequalities, and rough isometries between non-compact riemannian manifolds*, In: Curvature and topology of Riemannian manifolds (Katata, 1985), 122–137, Lect. Notes Math. **1201**, Springer, Berlin, 1986.

[53] M. Kanai, *Rough isometries and combinatorial approximations of geometries of non-compact riemannian manifolds*, J. Math. Soc. Japan, **37** (1985), 391–413.

[54] H. Kesten, *The incipient infinite cluster in two-dimensional percolation*, Probab. Theory Related Fields, **73** (1986), 369–394.

[55] H. Kesten, *Subdiffusive behavior of random walk on a random cluster*, Ann. Inst. Henri Poincaré, **22** (1986), 425–487.

[56] H. Kesten, *Subadditive behavior of random walk on a random cluster*, Unpublished notes.

[57] J. Kigami, *Local Nash inequality and inhomogeneity of heat kernels*, Proc. London Math. Soc., **89** (2004), 525–544.

[58] J. Kigami, *Analysis on Fractals*, Cambridge Univ. Press, 2001.

[59] T. Kumagai, *Heat kernel estimates and parabolic Harnack inequalities on graphs and resistance forms*, Publ. RIMS, Kyoto Univ., **40** (2004), 793–818.

[60] T. Kumagai, *Brownian motion penetrating fractals -An application of the trace theorem of Besov spaces-*, J. Funct. Anal., **170** (2000), 69–92.

[61] S. Kusuoka, *Dirichlet forms on fractals and products of random matrices*, Publ. RIMS. Kyoto Univ., **25** (1989), 659–680.

[62] S. Kusuoka, *A diffusion process on a fractal*, Probabilistic methods in mathematical physics (Katata/Kyoto, 1985), 251–274, Academic Press, Boston, MA, 1987.

[63] S. Kusuoka and D.W. Stroock, *Applications of the Malliavin calculus. III*, J. Fac. Sci. Univ. Tokyo Sect. IA Math., **34** (1987), 391–442.

[64] Y. Le Jan. *Mesures associees a une forme de Dirichlet. Applications*, Bull. Soc. Math. France, **106** (1978), no. 1, 61–112.

[65] P. Li and S.T. Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math., **156** (1986), 153–201.

[66] U. Mosco. *Composite media and asymptotic Dirichlet forms.* J. Funct. Anal. **123** (1994), no. 2, 368–421.

[67] J. Moser, *On a pointwise estimate for parabolic differential equations*, Comm. Pure Appl. Math., **24** (1971), 727–740.

[68] J. Moser, *On Harnack's inequality for parabolic differential equations*, Comm. Pure Appl. Math., **17** (1964), 101–134.

[69] J. Moser, *On Harnack's inequality for elliptic differential equations*, Comm. Pure Appl. Math., **14** (1961), 577–591.

[70] J. Nash, *Continuity of solutions of parabolic and elliptic equations*, Amer. Math. J., **80** (1958), 931–954.

[71] Y. Oshima, *Lecture notes on Dirichlet forms*, Unpublished lecture notes.

[72] L. Saloff-Coste, *Aspects of Sobolev-type inequalities*, (2002), Cambridge Univ. Press, Cambridge.

[73] L. Saloff-Coste, *A note on Poincaré, Sobolev and Harnack inequality*, Internat. Math. Res. Notices, **2** (1992), 27–38.

[74] V. Sidoravicius and A.-S. Sznitman, *Quenched invariance principles for walks on clusters of percolation or among random conductances*, Probab. Theory Related Fields, **124** (2004), 219–244.

[75] K.T. Sturm, *Analysis on local Dirichlet spaces -III. the parabolic Harnack inequality*, J. Math. Pure Appl., **75** (1996), 273–297.

[76] K.T. Sturm, *Analysis on local Dirichlet spaces -II. Gaussian upper bounds for the fundamental solutions of parabolic Harnack equations*, Osaka J. Math., **32** (1995), 275–312.

[77] M. Tomisaki, *Comparison theorems on Dirichlet norms and their applications*, Forum Math., **2** (1990), 277–295.

[78] N. Th. Varopoulos, *Hardy-Littlewood theory for semigroups*, J. Funct. Anal., **63** (1985), 240–260.

[79] J.-A. Yan, *A formula for densities of transition functions*, Sem. Prob. XXII, pp. 92-100. Lect Notes Math. **1321**, Springer, 1988.