Random walks on disordered media and their scaling limits

Takashi Kumagai *
Department of Mathematics
Kyoto University, Kyoto 606-8502, Japan
Email: kumagai@math.kyoto-u.ac.jp

(June 27, 2010)
(Version for St. Flour Lectures)

Abstract

The main theme of these lectures is to analyze heat conduction on disordered media such as fractals and percolation clusters using both probabilistic and analytic methods, and to study the scaling limits of Markov chains on the media.

The problem of random walk on a percolation cluster 'the ant in the labyrinth' has received much attention both in the physics and the mathematics literature. In 1986, H. Kesten showed an anomalous behavior of a random walk on a percolation cluster at critical probability for trees and for $\mathbb{Z}^2$. (To be precise, the critical percolation cluster is finite, so the random walk is considered on an incipient infinite cluster (IIC), namely a critical percolation cluster conditioned to be infinite.) Partly motivated by this work, analysis and diffusion processes on fractals have been developed since the late eighties. As a result, various new methods have been produced to estimate heat kernels on disordered media, and these turn out to be useful to establish quenched estimates on random media. Recently, it has been proved that random walks on IICs are sub-diffusive on $\mathbb{Z}^d$ when d is high enough, on trees, and on the spread-out oriented percolation for $d > 6$.

Throughout the lectures, I will survey the above mentioned developments in a compact way. In the first part of the lectures, I will summarize some classical and non-classical estimates for heat kernels, and discuss stability of the estimates under perturbations of operators and spaces. Here Nash inequalities and equivalent inequalities will play a central role. In the latter part of the lectures, I will give various examples of disordered media and obtain heat kernel estimates for Markov chains on them. In some models, I will also discuss scaling limits of the Markov chains. Examples of disordered media include fractals, percolation clusters, random conductance models and random graphs.

*Research partially supported by the Grant-in-Aid for Scientific Research (B) 22340017.

Contents

0 Plan of the lectures and remark

A rough plan of lectures at St. Flour is as follows.

Lecture 1–3: In the first lecture, I will discuss general potential theory for symmetric Markov chains on weighted graphs (Section 1). Then I will show various equivalent conditions for the heat kernel upper bounds (the Nash inequality (Section 2)). In the third lecture, I will use effective resistance to estimate Green functions, exit times from balls etc.. On-diagonal heat kernel bounds are also obtained using the effective resistance (Section 3).

Lecture 4–$6\frac{1}{2}$: I will discuss random walk on an incipient infinite cluster (IIC) for a critical percolation. I will give some sufficient condition for the sharp on-diagonal heat kernel bounds for random walk on random graphs (Section 4). I then prove the Alexander-Orbach conjecture for IICs when two-point functions behave nicely, especially for IICs of high dimensional critical bond percolations on $\mathbb{Z}^d$ (Section 5). I also discuss heat kernel bounds and scaling limits on related random models (Section 6).

Lecture $6\frac{1}{2}$–8: The last 2 (and half) lectures will be devoted to the quenched invariance principle for the random conductance model on $\mathbb{Z}^d$ (Section 7). Put i.i.d. conductance μ_e on each bond in $\mathbb{Z}^d$ and consider the Markov chain associated with the (random) weighted graph. I consider two cases, namely $0 \le \mu_e \le c$ $\mathbb{P}$-a.e. and $c \le \mu_e < \infty$ $\mathbb{P}$-a.e.. Although the behavior of heat kernels are quite different, for both cases the scaling limit is Brownian motion in general. I will discuss some technical details about correctors of the Markov chains, which play a key role to prove the invariance principle.

This is the version for St. Flour Lectures. There are several ingredients (which I was planning to include) missing in this version. Especially, I could not explain enough about techniques on heat kernel estimates; for example off-diagonal heat kernel estimates, isoperimetric profiles, relations to Harnack inequalities etc. are either missing or mentioned only briefly. (Because of that, I could not give proof to most of the heat kernel estimates in Section 7.) This was because I was much busier than I had expected while preparing for the lectures. However, even if I could have included them, most likely there was not enough time to discuss them during the 8 lectures. Anyway, my plan is to revise these notes and include the missing ingredients in the version for publication from Springer.

I referred many papers and books during the preparation of the lecture notes. Especially, I owe a lot to the lecture notes by Barlow [10] and by Coulhon [40] for Section 1–2. Section 3–4 (and part of Section 6) are mainly from [18, 87]. Section 5 is mainly from the beautiful paper by Kozma and Nachmias [86] (with some simplification in [101]). In Section 7, I follow the arguments of the papers by Barlow, Biskup, Deuschel, Mathieu and their co-authors [16, 30, 31, 32, 90, 91].

1 Weighted graphs and the associated Markov chains

In this section, we discuss general potential theory for symmetric (reversible) Markov chains on weighted graphs. Note that there are many nice books and lecture notes that treat potential theory and/or Markov chains on graphs, for example [6, 10, 51, 59, 88, 107, 109]. While writing this section, we are largely influenced by the lecture notes by Barlow [10].

1.1 Weighted graphs

Let X be a finite or countably infinite set, and E is a subset of $\{\{x,y\} : x,y \in X, x \neq y\}$. A graph is a pair (X,E). For $x,y \in X$, we write $x \sim y$ if $\{x,y\} \in E$. A sequence $x_0, x_2, \cdots, x_n$ is called a path with length n if $x_i \in X$ for $i = 0,1,2,\cdots,n$ and $x_j \sim x_{j+1}$ for $j = 0,1,2,\cdots,n-1$. For $x \neq y$, define $d(x,y)$ to be the length of the shortest path from x to y. If there is no such path, we set $d(x,y) = \infty$ and we set $d(x,x) = 0$. $d(\cdot,\cdot)$ is a metric on X and it is called a graph distance. (X,E) is connected if $d(x,y) < \infty$ for all $x,y \in X$, and it is locally finite if $\sharp\{y : \{x,y\} \in E\} < \infty$ for all $x \in X$. Throughout the lectures, we will consider connected locally finite graphs (except when we consider the trace of them in Subsection 1.3).

Assume that the graph (X,E) is endowed with a weight (conductance) μ_{xy}, which is a symmetric nonnegative function on $X \times X$ such that $\mu_{xy} > 0$ if and only if $x \sim y$. We call the pair (X,μ) a weighted graph.

Let $\mu_x = \mu(x) = \sum_{y \in X} \mu_{xy}$ and define a measure μ on X by setting $\mu(A) = \sum_{x \in A} \mu_x$ for $A \subset X$. Also, we define $B(x,r) = \{y \in X : d(x,y) < r\}$ for each $x \in X$ and $r \geq 1$.

Definition 1.1 *We say that (X,μ) has controlled weights (or (X,μ) satisfies p_0-condition) if there exists $p_0 > 0$ such that*

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

If (X,μ) has controlled weights, then clearly $\sharp\{y \in X : x \sim y\} \leq p_0^{-1}$.

Once the weighted graph (X,μ) is given, we can define the corresponding quadratic form, Markov chain and the discrete Laplace operator.

Quadratic form We define a quadratic form on (X,μ) as follows.

$$\begin{aligned} H^2(X,\mu) = H^2 &= \{f : X \to \mathbb{R} : \mathcal{E}(f,f) = \frac{1}{2} \sum_{\substack{x,y \in X \\ x \sim y}} (f(x) - f(y))^2 \mu_{xy} < \infty\}, \\ \mathcal{E}(f,g) &= \frac{1}{2} \sum_{\substack{x,y \in X \\ x \sim y}} (f(x) - f(y))(g(x) - g(y)) \mu_{xy} \qquad \forall f,g \in H^2. \end{aligned}$$

Physically, $\mathcal{E}(f,f)$ is the energy of the electrical network for an (electric) potential f.

Since the graph is connected, one can easily see that $\mathcal{E}(f,f) = 0$ if and only if f is a constant function. We fix a base point $0 \in X$ and define

$$\|f\|_{H^2}^2 = \mathcal{E}(f,f) + f(0)^2 \qquad \forall f \in H^2.$$

Note that

$$\mathcal{E}(f,f)=\frac{1}{2}\sum_{x\sim y}(f(x)-f(y))^2\mu_{xy}\le\sum_x\sum_y(f(x)^2+f(y)^2)\mu_{xy}=2\|f\|_2^2\quad\forall f\in\mathbb{L}^2,\tag{1.1}$$

where $\|f\|_2$ is the $\mathbb{L}^2$-norm of f. So $\mathbb{L}^2\subset H^2$. We give basic facts in the next lemma.

Lemma 1.2 *(i) Convergence in H^2 implies the pointwise convergence.*
(ii) H^2 is a Hilbert space.

Proof. (i) Suppose $f_n\to f$ in H^2 and let $g_n=f_n-f$. Then $\mathcal{E}(g_n,g_n)+g_n(0)^2\to 0$ so $g_n(0)\to 0$. For any $x\in X$, there is a sequence $\{x_i\}_{i=0}^l\subset X$ such that $x_0=0, x_l=x$ and $x_i\sim x_{i+1}$ for $i=0,1,\cdots,l-1$. Then

$$|g_n(x)-g_n(0)|^2\le l\sum_{i=0}^{l-1}|g_n(x_i)-g_n(x_{i+1})|^2\le 2l(\min_{i=0}^{l-1}\mu_{x_ix_{i+1}})^{-1}\mathcal{E}(g_n,g_n)\to 0\tag{1.2}$$

as $n\to\infty$ so we have $g_n(x)\to 0$.
(ii) Assume that $\{f_n\}_n\subset H^2$ is a Cauchy sequence in H^2. Then $f_n(0)$ is a Cauchy sequence in $\mathbb{R}$ so converges. Thus, similarly to (1.2) f_n converges pointwise to f, say. Now using Fatou's lemma, we have $\|f_n-f\|_{H^2}^2\le\liminf_m\|f_n-f_m\|_{H^2}^2$, so that $\|f_n-f\|_{H^2}^2\to 0$. □

Markov chain Let $Y=\{Y_n\}$ be a Markov chain on X whose transition probabilities are given by

$$\mathbb{P}(Y_{n+1}=y|Y_n=x)=\frac{\mu_{xy}}{\mu_x}=:P(x,y)\qquad\forall x,y\in X.$$

We write $\mathbb{P}^x$ when the initial distribution of Y is concentrated on x (i.e. $Y_0=x$, $\mathbb{P}$-a.s.). $(P(x,y))_{x,y\in X}$ is the transition matrix for Y. Y is called a simple random walk when $\mu_{xy}=1$ whenever $x\sim y$. Y is μ-symmetric since for each $x,y\in X$,

$$\mu_xP(x,y)=\mu_{xy}=\mu_{yx}=\mu_yP(y,x).$$

We define the *heat kernel* of Y by

$$p_n(x,y)=\mathbb{P}^x(Y_n=y)/\mu_y\qquad\forall x,y\in X.$$

Using the Markov property, we can easily show the Chapman-Kolmogorov equation:

$$p_{n+m}(x,y)=\sum_z p_n(x,z)p_m(z,y)\mu_z,\qquad\forall x,y\in X.\tag{1.3}$$

Using this and the fact $p_1(x,y)=\mu_{xy}/(\mu_x\mu_y)=p_1(y,x)$, one can verify the following inductively

$$p_n(x,y)=p_n(y,x),\qquad\forall x,y\in X.$$

For $n \geq 1$, let

$$P_n f(x) = \sum_y p_n(x,y) f(y)\mu_y = \sum_y \mathbb{P}^x(Y_n = y) f(y) = \mathbb{E}^x[f(Y_n)], \qquad \forall f : X \to \mathbb{R}.$$

We sometimes consider a continuous time Markov chain $\{Y_t\}_{t\geq 0}$ w.r.t. μ which is defined as follows: each particle stays at a point, say x for (independent) exponential time with parameter 1, and then jumps to another point, say y with probability $P(x,y)$. The heat kernel for the continuous time Markov chain can be expressed as follows.

$$p_t(x,y) = \mathbb{P}^x(Y_t = y)/\mu_y = \sum_{n=0}^{\infty} e^{-t}\frac{t^n}{n!} p_n(x,y), \qquad \forall x, y \in X.$$

Discrete Laplace operator   For $f : X \to \mathbb{R}$, the discrete Laplace operator is defined by

$$\mathcal{L}f(x) = \sum_y P(x,y)f(y) - f(x) = \frac{1}{\mu_x}\sum_y (f(y) - f(x))\mu_{xy} = \mathbb{E}^x[f(Y_1)] - f(x) = (P_1 - I)f(x). \quad (1.4)$$

Note that according to the Ohm's law '$I = V/R$', $\sum_y (f(y) - f(x))\mu_{xy}$ is the total flux flowing into x, given the potential f.

Definition 1.3 *Let $A \subset X$. A function $f : X \to \mathbb{R}$ is harmonic on A if*

$$\mathcal{L}f(x) = 0, \qquad \forall x \in A.$$

h is sub-harmonic (resp. super-harmonic) on A if $\mathcal{L}f(x) \geq 0$ (resp. $\mathcal{L}f(x) \leq 0$) for $x \in A$.

$\mathcal{L}f(x) = 0$ means that the total flux flowing into x is 0 for the given a potential f. This is the behavior of the currents in a network called Kirchoff's (first) law.

For $A \subset X$, we define the (exterior) boundary of A by

$$\partial A = \{x \in A^c : \exists z \in A \text{ such that } z \sim x\}.$$

Proposition 1.4 (Maximum principle) *Let A be a connected subset of X and $h : A \cup \partial A \to \mathbb{R}$ be sub-harmonic on A. If the maximum of h over $A \cup \partial A$ is attained in A, then h is constant on $A \cup \partial A$.*

Proof. Let $x_0 \in A$ be the point where h attains the maximum and let $H = \{z \in A \cup \partial A : h(z) = h(x_0)\}$. If $y \in H \cap A$, then since $h(y) \geq h(x)$ for all $x \in A \cup \partial A$, we have

$$0 \leq \mu_y \mathcal{L}h(y) = \sum_y (h(x) - h(y))\mu_{xy} \leq 0.$$

Thus, $h(x) = h(y)$ (i.e. $x \in H$) for all $x \sim y$. Since A is connected, this implies $H = A \cup \partial A$. □

We can prove the minimum principle for a super-harmonic function h by applying the maximum principle to $-h$.

For $f, g \in \mathbb{L}^2$, denote their $\mathbb{L}^2$-inner product as (f, g), namely $(f, g) = \sum_x f(x)g(x)\mu_x$.

Lemma 1.5 *(i) $\mathcal{L} : H^2 \to \mathbb{L}^2$ and $\|\mathcal{L}f\|_2^2 \le 2\|f\|_{H^2}^2$.*
(ii) For $f \in H^2$ and $g \in \mathbb{L}^2$, we have $(-\mathcal{L}f, g) = \mathcal{E}(f, g)$.
(iii) $\mathcal{L}$ is a self-adjoint operator on $\mathbb{L}^2(X, \mu)$ and the following holds:

$$(-\mathcal{L}f, g) = (f, -\mathcal{L}g) = \mathcal{E}(f, g), \qquad \forall f, g \in \mathbb{L}^2. \tag{1.5}$$

Proof. (i) Using Schwarz's inequality, we have

$$\begin{aligned} \|\mathcal{L}f\|_2^2 &= \sum_x \frac{1}{\mu_x} (\sum_y (f(y) - f(x))\mu_{xy})^2 \\ &\le \sum_x \frac{1}{\mu_x} (\sum_y (f(y) - f(x))^2 \mu_{xy})(\sum_y \mu_{xy}) = 2\mathcal{E}(f, f) \le 2\|f\|_{H^2}^2. \end{aligned}$$

(ii) Using (i), both sides of the equality are well-defined. Further, using Schwarz's inequality,

$$\sum_{x,y} |\mu_{xy}(f(y) - f(x))g(x)| \le (\sum_{x,y} \mu_{xy}(f(y) - f(x))^2)^{1/2} (\sum_{x,y} \mu_{xy} g(y)^2)^{1/2} = \mathcal{E}(f, f)^{1/2} \|g\|_2 < \infty.$$

So we can use Fubini's theorem, and we have

$$(-\mathcal{L}f, g) = -\sum_x (\sum_y \mu_{xy}(f(y) - f(x)))g(x) = \frac{1}{2} \sum_x \sum_y \mu_{xy}(f(y) - f(x))(g(y) - g(x)) = \mathcal{E}(f, g).$$

(iii) We can prove $(f, -\mathcal{L}g) = \mathcal{E}(f, g)$ similarly and obtain (1.5). □

(1.5) is the discrete Gauss-Green formula.

Lemma 1.6 *Set $p_n^x(\cdot) = p_n(x, \cdot)$. Then, the following hold for all $x, y \in X$.*

$$p_{n+m}(x, y) = (p_n^x, p_m^y), \qquad P_1 p_n^x(y) = p_{n+1}(x, y), \tag{1.6}$$

$$\mathcal{L}p_n^x(y) = p_{n+1}(x, y) - p_n(x, y), \qquad \mathcal{E}(p_n^x, p_m^y) = p_{n+m}(x, y) - p_{n+m+1}(x, y), \tag{1.7}$$

$$p_{2n}(x, y) \le \sqrt{p_{2n}(x, x) p_{2n}(y, y)}. \tag{1.8}$$

Proof. The two equations in (1.6) are due to the Chapman-Kolmogorov equation (1.3). The first equation in (1.7) is then clear since $\mathcal{L} = P_1 - I$. The last equation can be obtained by these equations and (1.5). Using (1.6) and the Schwarz inequality, we have

$$p_{2n}(x, y)^2 = (p_n^x, p_n^y)^2 \le (p_n^x, p_n^x)(p_n^y, p_n^y) = p_{2n}(x, x) p_{2n}(y, y),$$

which gives (1.8). □

It can be easily shown that $(\mathcal{E}, \mathbb{L}^2)$ is a regular Dirichlet form on $\mathbb{L}^2(X, \mu)$ (c.f. [55]). Then the corresponding Hunt process is the continuous time Markov chain $\{Y_t\}_{t \ge 0}$ w.r.t. μ and the corresponding self-adjoint operator on $\mathbb{L}^2$ is $\mathcal{L}$ in (1.4).

Remark 1.7 *Note that* $\{Y_t\}_{t\geq 0}$ *has the transition probability* $P(x,y) = \mu_{xy}/\mu_x$ *and it waits at* x *for an exponential time with mean 1 for each* $x \in X$. *Since the 'speed' of* $\{Y_t\}_{t\geq 0}$ *is independent of the location, it is sometimes called* constant speed random walk *(CSRW for short). We can also consider a continuous time Markov chain with the same transition probability* $P(x,y)$ *and wait at* x *for an exponential time with mean* μ_x^{-1} *for each* $x \in X$. *This Markov chain is called* variable speed random walk *(VSRW for short). We will discuss VSRW in Section 7. The corresponding discrete Laplace operator is*

$$\mathcal{L}_V f(x) = \sum_y (f(y) - f(x))\mu_{xy}. \tag{1.9}$$

For each f, g *that have finite support, we have*

$$\mathcal{E}(f,g) = -(\mathcal{L}_V f, g)_\nu = -(\mathcal{L}_C f, g)_\mu,$$

where ν *is a measure on* X *such that* $\nu(A) = |A|$ *for all* $A \subset X$. *So VSRW is the Markov process associated with the Dirichlet form* $(\mathcal{E},\mathcal{F})$ *on* $\mathbb{L}^2(X,\nu)$ *and CSRW is the Markov process associated with the Dirichlet form* $(\mathcal{E},\mathcal{F})$ *on* $\mathbb{L}^2(X,\mu)$. *VSRW is a time changed process of CSRW and vice versa.*

We now introduce the notion of rough isometry.

Definition 1.8 *Let* $(X_1,\mu_1),(X_2,\mu_2)$ *be weighted graphs that have controlled weights.*
(i) A map $T : X_1 \to X_2$ *is called a rough isometry if the following holds.*
There exist constants $c_1, c_2, c_3 > 0$ *such that*

$$c_1^{-1} d_1(x,y) - c_2 \leq d_2(T(x),T(y)) \leq c_1 d_1(x,y) + c_2 \quad \forall x,y \in X_1, \tag{1.10}$$

$$d_2(T(X_1),y') \leq c_2 \quad \forall y' \in X_2, \tag{1.11}$$

$$c_3^{-1}\mu_1(x) \leq \mu_2(T(x)) \leq c\mu_1(x) \quad \forall x \in X_1, \tag{1.12}$$

where $d_i(\cdot,\cdot)$ *is the the graph distance of* (X_i,μ_i), *for* $i = 1,2$.
(ii) $(X_1,\mu_1),(X_2,\mu_2)$ *are said to be rough isometric if there is a rough isometry between them.*

It is easy to see that rough isometry is an equivalence relation. One can easily prove that $\mathbb{Z}^2$ and the triangular lattice, the hexagon lattice are all roughly isometric. It can be proved that $\mathbb{Z}^1$ and $\mathbb{Z}^2$ are not roughly isometric.

The notion of rough isometry was first introduced by M. Kanai ([73, 74]). As this work was mainly concerned with Riemannian manifolds, definition of rough isometry included only (1.10), (1.11). The definition equivalent to Definition 1.8 is given in [42] (see also [65]). Note that rough isometry corresponds to quasi-isometry in the field of geometric group theory.

While discussing various properties of Markov chains/Laplace operators, it is important to think about their 'stability'. In the following, we introduce two types of stability.

Definition 1.9 *(i) We say a property is stable under bounded perturbation if whenever* (X,μ) *satisfies the property and* (X,μ') *satisfies* $c^{-1}\mu_{xy} \leq \mu'_{xy} \leq c\mu_{xy}$ *for all* $x,y \in X$, *then* (X,μ') *satisfies*

the property.
(ii) We say a property is stable under rough isometry if whenever (X,μ) satisfies the property and (X',μ') is rough isometric to (X,μ), then (X',μ') satisfies the property.

If a property P is stable under rough isometry, then it is clearly stable under bounded perturbation.

It is known that the following properties of weighted graphs are stable under rough isometry.

(i) Transience and recurrence

(ii) The Nash inequality, i.e. $p_n(x,y) \le c_1 n^{-\alpha}$ for all $n \ge 1, x \in X$ (for some $\alpha > 0$)

(iii) Parabolic Harnack inequality

We will see (i) later in this section and (ii) in Section 2. One of the important open problem is to show if the elliptic Harnack inequality is stable under these perturbations or not.

Definition 1.10 (X,μ) *has the Liouville property if there is no bounded non-constant harmonic functions.* (X,μ) *has the strong Liouville property if there is no positive non-constant harmonic functions.*

It is known that both Liouville and strong Liouville properties are not stable under bounded perturbation (see [89]).

1.2 Harmonic functions and effective resistances

For $A \subset X$, define

$$\sigma_A = \inf\{n \ge 0 : Y_n \in A\}, \quad \sigma_A^+ = \inf\{n > 0 : Y_n \in A\}, \quad \tau_A = \inf\{n \ge 0 : Y_n \notin A\}.$$

For $A \subset X$ and $f : A \to \mathbb{R}$, consider the following *Dirichlet problem.*

$$\begin{cases} \mathcal{L}v(x) = 0 & \forall x \in A^c, \\ v|_A = f. \end{cases} \tag{1.13}$$

Proposition 1.11 *Assume that $f : A \to \mathbb{R}$ is bounded and set*

$$\varphi(x) = \mathbb{E}^x[f(Y_{\sigma_A}) : \sigma_A < \infty].$$

(i) φ is a solution of (1.13).
(ii) If $\mathbb{P}^x(\sigma_A < \infty) = 1$ for all $x \in X$, then φ is the unique solution of (1.13).

Proof. (i) $\varphi|_A = f$ is clear. For $x \in A^c$, using the Markov property of Y, we have

$$\varphi(x) = \sum_y P(x,y)\varphi(y),$$

so $\mathcal{L}\varphi(x) = 0$.
(ii) Let φ' be another solution and let $H_n = \varphi(Y_n) - \varphi'(Y_n)$. Then H_n is a bounded martingale up to σ_A, so using the optional stopping theorem, we have

$$\varphi(x) - \varphi'(x) = \mathbb{E}^x H_0 = \mathbb{E}^x H_{\sigma_A} = \mathbb{E}^x[\varphi(Y_{\sigma_A}) - \varphi'(Y_{\sigma_A})] = 0$$

since $\sigma_A < \infty$ a.s. and $\varphi(x) = \varphi'(x)$ for $x \in A$. □

Remark 1.12 *(i) In particular, we see that φ is the unique solution of* (1.13) *when A^c is finite. In this case, here is another proof of the uniqueness of the solution of* (1.13)*: let $u(x) = \varphi(x) - \varphi'(x)$, then $u|_A = 0$ and $\mathcal{L}u(x) = 0$ for $x \in A^c$. So, noting $u \in \mathbb{L}^2$ and using Lemma 1.5, $\mathcal{E}(u,u) = (-\mathcal{L}u, u) = 0$ which implies that u is constant on X (so it is 0 since $u|_A = 0$).*
(ii) If $h_A(x) := \mathbb{P}^x(\sigma_A = \infty) > 0$ for some $x \in X$, then the function $\varphi + \lambda h_A$ is also a solution of (1.13) *for all $\lambda \in \mathbb{R}$, so the uniqueness of the Dirichlet problem fails.*

For $A, B \subset X$ such that $A \cap B = \emptyset$, define

$$R_{\text{eff}}(A,B)^{-1} = \inf\{\mathcal{E}(f,f) : f \in H^2, f|_A = 1, f|_B = 0\}. \tag{1.14}$$

(We define $R_{\text{eff}}(A,B) = \infty$ when the right hand side is 0.) We call $R_{\text{eff}}(A,B)$ the *effective resistance* between A and B. It is easy to see that $R_{\text{eff}}(A,B) = R_{\text{eff}}(B,A)$. If $A \subset A'$, $B \subset B'$ with $A' \cap B' = \emptyset$, then $R_{\text{eff}}(A', B') \le R_{\text{eff}}(A,B)$.

Take a bond $e = \{x,y\}$, $x \sim y$ in a weighted graph (X, μ). We say *cutting the bond e* when we take the conductance μ_{xy} to be 0, and we say *shorting the bond e* when we identify $x = y$ and take the conductance μ_{xy} to be ∞. Clearly, shorting decreases the effective resistance (*shorting law*), and cutting increases the effective resistance (*cutting law*).

The following proposition shows that among feasible potentials whose voltage is 1 on A and 0 on B, it is a harmonic function on $(A \cup B)^c$ that minimizes the energy.

Proposition 1.13 *(i) The right hand side of (1.14) is attained by a unique minimizer φ.*
(ii) φ in (1) is a solution of the following Dirichlet problem

$$\begin{cases} \mathcal{L}\varphi(x) = 0 & \forall x \in X \setminus (A \cup B), \\ \varphi|_A = 1, \quad \varphi|_B = 0. \end{cases} \tag{1.15}$$

Proof. (i) We fix a based point $x_0 \in B$ and recall that H^2 is a Hilbert space with $\|f\|_{H^2} = \mathcal{E}(f,f) + f(x_0)^2$ (Lemma 1.2 (ii)). Since $\mathcal{V} := \{f \in H^2 : f|_A = 1, f|_B = 0\}$ is a closed convex subset of H^2, a general theorem shows that $\mathcal{V}$ has a unique minimizer for $\|\cdot\|_{H^2}$ (which is equal to $\mathcal{E}(\cdot,\cdot)$ on $\mathcal{V}$).
(ii) Let g be a function on X whose support is finite and is contained in $X \setminus (A \cup B)$. Then, for any $\lambda \in \mathbb{R}$, $\varphi + \lambda g \in \mathcal{V}$, so $\mathcal{E}(\varphi + \lambda g, \varphi + \lambda g) \ge \mathcal{E}(\varphi, \varphi)$. Thus $\mathcal{E}(\varphi, g) = 0$. Applying Lemma 1.5(ii), we have $(\mathcal{L}\varphi, g) = 0$. For each $x \in X \setminus (A \cup B)$, by choosing $g(z) = \delta_x(z)$, we obtain $\mathcal{L}\varphi(x)\mu_x = 0$. □

As we mentioned in Remark 1.12 (ii), we do not have uniqueness of the Dirichlet problem in general. So in the following of this section, we will assume that A^c is finite in order to guarantee uniqueness of the Dirichlet problem.

The next theorem gives a probabilistic interpretation of the effective resistance.

Theorem 1.14 *If A^c is finite, then for each $x_0 \in A^c$,*

$$R_{\mathrm{eff}}(x_0, A)^{-1} = \mu_{x_0} \mathbb{P}^{x_0}(\sigma_A < \sigma_{x_0}^+). \tag{1.16}$$

Proof. Let $v(x) = \mathbb{P}^x(\sigma_A < \sigma_{x_0})$. Then, by Proposition 1.11, v is the unique solution of Dirichlet problem with $v(x_0) = 0$, $v|_A = 1$. By Proposition 1.13 and Lemma 1.5 (noting that $1 - v \in \mathbb{L}^2$),

$$R_{\mathrm{eff}}(x_0, A)^{-1} = \mathcal{E}(v, v) = \mathcal{E}(-v, 1 - v) = (\mathcal{L}v, 1 - v) = \mathcal{L}v(x_0)\mu_{x_0} = \mathbb{E}^{x_0}[v(Y_1)]\mu_{x_0}.$$

By definition of v, one can see $\mathbb{E}^{x_0}[v(Y_1)] = \mathbb{P}^{x_0}(\sigma_A < \sigma_{x_0}^+)$ so the result follows. □

Similarly, if A^c is finite one can prove

$$R_{\mathrm{eff}}(B, A)^{-1} = \sum_{x \in B} \mu_x \mathbb{P}^x(\sigma_A < \sigma_B^+).$$

Note that by Ohm's law, the right hand side of (1.16) is the current flowing from x_0 to A^c.

The following lemma is useful and will be used later in Proposition 3.18.

Lemma 1.15 *Let $A, B \subset X$ and assume that both A^c, B^c are finite. Then the following holds.*

$$\frac{R_{\mathrm{eff}}(x, A \cup B)}{R_{\mathrm{eff}}(x, A)^{-1} - R_{\mathrm{eff}}(x, B)^{-1}} \le \mathbb{P}^x(\sigma_A < \sigma_B) \le \frac{R_{\mathrm{eff}}(x, A \cup B)}{R_{\mathrm{eff}}(x, A)}, \quad \forall x \notin A \cup B.$$

Proof. Using the strong Markov property, we have

$$\begin{aligned} \mathbb{P}^x(\sigma_A < \sigma_B) &= \mathbb{P}^x(\sigma_A < \sigma_B, \sigma_{A\cup B} < \sigma_x^+) + \mathbb{P}^x(\sigma_A < \sigma_B, \sigma_{A\cup B} > \sigma_x^+) \\ &= \mathbb{P}^x(\sigma_A < \sigma_B, \sigma_{A\cup B} < \sigma_x^+) + \mathbb{P}^x(\sigma_{A\cup B} > \sigma_x^+)\mathbb{P}^x(\sigma_A < \sigma_B). \end{aligned}$$

So

$$\mathbb{P}^x(\sigma_A < \sigma_B) = \frac{\mathbb{P}^x(\sigma_A < \sigma_B, \sigma_{A\cup B} < \sigma_x^+)}{\mathbb{P}^x(\sigma_{A\cup B} < \sigma_x^+)} \le \frac{\mathbb{P}^x(\sigma_A < \sigma_x^+)}{\mathbb{P}^x(\sigma_{A\cup B} < \sigma_x^+)}.$$

Using (1.16), the upper bound is obtained. For the lower bound,

$$\mathbb{P}^x(\sigma_A < \sigma_B, \sigma_{A\cup B} < \sigma_x^+) \ge \mathbb{P}^x(\sigma_A < \sigma_x^+ < \sigma_B) \ge \mathbb{P}^x(\sigma_A < \sigma_x^+) - \mathbb{P}^x(\sigma_B < \sigma_x^+),$$

so using (1.16) again, the proof is complete. □

As we see in the proof, we only need to assume that A^c is finite for the upper bound.

Now let (X, μ) be an infinite weighted graph. Let $\{A_n\}_{n=1}^{\infty}$ be a family of finite sets such that $A_n \subset A_{n+1}$ for $n \in \mathbb{N}$ and $\cup_{n\geq 1} A_n = X$. Let $x_0 \in A_1$. By the short law, $R_{\rm eff}(x_0, A_n^c) \leq R_{\rm eff}(x_0, A_{n+1}^c)$, so the following limit exists.

$$R_{\rm eff}(x_0) := \lim_{n\to\infty} R_{\rm eff}(x_0, A_n^c). \tag{1.17}$$

Further, the limit $R_{\rm eff}(x_0)$ is independent of the choice of the sequence $\{A_n\}$ mentioned above. (Indeed, if $\{B_n\}$ is another such family, then for each n there exists N_n such that $A_n \subset B_{N_n}$, so $\lim_{n\to\infty} R_{\rm eff}(x_0, A_n^c) \leq \lim_{n\to\infty} R_{\rm eff}(x_0, B_n^c)$. By changing the role of A_n and B_n, we have the opposite inequality.)

Theorem 1.16 *Let (X, μ) be an infinite weighted graph. For each $x \in X$, the following holds*

$$\mathbb{P}^x(\sigma_x^+ = \infty) = (\mu_x R_{\rm eff}(x))^{-1}.$$

Proof. By Theorem 1.14, we have

$$\mathbb{P}^x(\sigma_{A_n^c} < \sigma_x^+) = (\mu_x R_{\rm eff}(x, A_n^c)^{-1})^{-1}.$$

Taking $n \to \infty$ and using (1.17), we have the desired equality. □

Definition 1.17 *We say a Markov chain is recurrent at $x \in X$ if $\mathbb{P}^x(\sigma_x^+ = \infty) = 0$. We say a Markov chain is transient at $x \in X$ if $\mathbb{P}^x(\sigma_x^+ = \infty) > 0$.*

The following is well-known for irreducible Markov chains (so in particular it holds for Markov chains corresponding to weighted graphs). See, for example [97].

Proposition 1.18 *(1) $\{Y_n\}_n$ is recurrent at $x \in X$ if and only if $m := \sum_{n=0}^{\infty} \mathbb{P}^x(Y_n = x) = \infty$. Further, $m^{-1} = \mathbb{P}^x(\sigma_x^+ = \infty)$.*
(2) If $\{Y_n\}_n$ is recurrent (resp. transient) at some $x \in X$, then it is recurrent (resp. transient) for all $x \in X$.
(3) $\{Y_n\}_n$ is recurrent if and only if $\mathbb{P}^x(\{Y \text{ hits } y \text{ infinitely often}\}) = 1$ for all $x, y \in X$. $\{Y_n\}_n$ is transient if and only if $\mathbb{P}^x(\{Y \text{ hits } y \text{ finitely often}\}) = 1$ for all $x, y \in X$.

From Theorem 1.16 and Proposition 1.18, we have the following.

$$\begin{aligned} \{Y_n\} \text{ is transient (resp. recurrent)} \quad &\Leftrightarrow \quad R_{\rm eff}(x) < \infty \text{ (resp. } R_{\rm eff}(x) = \infty\text{)}, \ \exists x \in X \\ &\Leftrightarrow \quad R_{\rm eff}(x) < \infty \text{ (resp. } R_{\rm eff}(x) = \infty\text{)}, \ \forall x \in X. \end{aligned} \tag{1.18}$$

Example 1.19 *Consider $\mathbb{Z}^2$ with weight 1 on each nearest neighbor bond. Let $\partial B_n = \{(x, y) \in \mathbb{Z}^2 :$ either $|x|$ or $|y|$ is $n\}$. By shorting ∂B_n for all $n \in \mathbb{N}$, one can obtain*

$$R_{\rm eff}(0) \geq \sum_{n=0}^{\infty} \frac{1}{4(2n+1)} = \infty.$$

So the simple random walk on $\mathbb{Z}^2$ is recurrent.

Let us recall the following fact.

Theorem 1.20 (Pólya 1921) *Simple random walk on* $\mathbb{Z}^d$ *is recurrent if* $d = 1, 2$ *and transient if* $d \geq 3$.

The combinatorial proof of this theorem is well-known. For example, for $d = 1$, by counting the total number of paths of length $2n$ that moves both right and left n times,

$$\mathbb{P}^0(Y_{2n} = 0) = 2^{-2n} \binom{2n}{n} = \frac{(2n)!}{2^{2n} n! n!} \sim (\pi n)^{-1/2},$$

where Stirling's formula is used in the end. Thus

$$m = \sum_{n=0}^{\infty} \mathbb{P}^0(Y_n = 0) \sim \sum_{n=1}^{\infty} (\pi n)^{-1/2} + 1 = \infty,$$

so $\{Y_n\}$ is recurrent.

This argument is not robust. For example, if one changes the weight on $\mathbb{Z}^d$ so that $c_1 \leq \mu_{xy} \leq c_2$ for $x \sim y$, one cannot apply the argument at all. The advantage of the characterization of transience/recurrence using the effective resistance is that one can make a robust argument. Indeed, by (1.18) we can see that transience/recurrence is stable under bounded perturbation. This is because, if $c_1 \mu'_{xy} \leq \mu_{xy} \leq c_2 \mu'_{xy}$ for all $x, y \in X$, then $c_1 R_{\text{eff}}(x) \leq R'_{\text{eff}}(x) \leq c_2 R_{\text{eff}}(x)$. We can further prove that transience/recurrence is stable under rough isometry.

Finally in this subsection, we will give more equivalence condition for the transience and discuss some decomposition of H^2. Let H_0^2 be the closure of $C_0(X)$ in H^2, where $C_0(X)$ is the space of compactly supported function on X. For a finite set $B \subset X$, define the capacity of B by

$$\text{Cap}\,(B) = \inf\{\mathcal{E}(f, f) : f \in H_0^2, f|_B = 1\}.$$

We first give a lemma.

Lemma 1.21 *If a sequence of non-negative functions* $v_n \in H^2$, $n \in \mathbb{N}$ *satisfies* $\lim_{n\to\infty} v_n(x) = \infty$ *for all* $x \in X$ *and* $\lim_{n\to\infty} \mathcal{E}(v_n, v_n) = 0$, *then*

$$\lim_{n\to\infty} \|u - (u \wedge v_n)\|_{H^2} = 0, \qquad \forall u \in H^2, u \geq 0.$$

Proof. Let $u_n = u \wedge v_n$ and define $U_n = \{x \in X : u(x) > v_n(x)\}$. By the assumption, for each $N \in \mathbb{N}$, there exists $N_0 = N_0(N)$ such that $U_n \subset B(0, N)^c$ for all $n \geq N_0$. For $A \subset X$, denote $\mathcal{E}_A(u) = \frac{1}{2}\sum_{x,y\in A}(u(x) - u(y))^2 \mu_{xy}$. Since $\mathcal{E}_{U_n^c}(u - u_n) = 0$, we have

$$\begin{aligned} \mathcal{E}(u - u_n, u - u_n) &\leq 2 \cdot \frac{1}{2} \sum_{x\in U_n} \sum_{y:\, y\sim x} \Big(u(x) - u_n(x) - (u(y) - u_n(y))\Big)^2 \mu_{xy} \\ &\leq 2\mathcal{E}_{B(0,N-1)^c}(u - u_n) \leq 2\Big(\mathcal{E}_{B(0,N-1)^c}(u) + \mathcal{E}_{B(0,N-1)^c}(u_n)\Big) \qquad (1.19) \end{aligned}$$

for all $n \geq N_0$. As $u_n = (u + v_n - |u - v_n|)/2$, we have

$$\begin{aligned}\mathcal{E}_{B(0,N-1)^c}(u_n) &\leq c_1\Big(\mathcal{E}_{B(0,N-1)^c}(u) + \mathcal{E}_{B(0,N-1)^c}(v_n) + \mathcal{E}_{B(0,N-1)^c}(|u - v_n|)\Big) \\ &\leq c_2\Big(\mathcal{E}_{B(0,N-1)^c}(u) + \mathcal{E}_{B(0,N-1)^c}(v_n)\Big).\end{aligned}$$

Thus, together with (1.19), we have

$$\mathcal{E}(u - u_n, u - u_n) \leq c_3\Big(\mathcal{E}_{B(0,N-1)^c}(u) + \mathcal{E}_{B(0,N-1)^c}(v_n)\Big) \leq c_3\Big(\mathcal{E}_{B(0,N-1)^c}(u) + \mathcal{E}(v_n, v_n)\Big).$$

Since $u \in H^2$, $\mathcal{E}_{B(0,N-1)^c}(u) \to 0$ as $N \to \infty$ and by the assumption, $\mathcal{E}(v_n, v_n) \to 0$ as $n \to \infty$. So we obtain $\mathcal{E}(u - u_n, u - u_n) \to 0$ as $n \to \infty$. By the assumption, it is clear that $u - u_n \to 0$ pointwise, so we obtain $\|u - u_n\|_{H^2} \to 0$. □

We say that a quadratic form $(\mathcal{E}, \mathcal{F})$ is Markovian if $u \in \mathcal{F}$ and $v = (0 \vee u) \wedge 1$, then $v \in \mathcal{F}$ and $\mathcal{E}(v, v) \leq \mathcal{E}(u, u)$. It is easy to see that quadratic forms determined by weighted graphs are Markovian.

Proposition 1.22 *The following are equivalent.*
(i) The Markov chain corresponding to (X, μ) is transient.
(ii) $1 \notin H_0^2$
(iii) Cap $(\{x\}) > 0$ for some $x \in X$.
(iii)′ Cap $(\{x\}) > 0$ for all $x \in X$.
(iv) $H_0^2 \neq H^2$
(v) There exists a non-negative super-harmonic function which is not a constant function.
(vi) For each $x \in X$, there exists $c_1(x) > 0$ such that

$$|f(x)|^2 \leq c_1(x)\mathcal{E}(f, f) \qquad \forall f \in C_0(X). \tag{1.20}$$

Proof. For fixed $x \in X$, define $\varphi(z) = \mathbb{P}^z(\sigma_x < \infty)$. We first show the following: $\varphi \in H_0^2$ and

$$\mathcal{E}(\varphi,\varphi) = (-\mathcal{L}\varphi, 1_{\{x\}}) = R_{\text{eff}}(x)^{-1} = \text{Cap}\,(\{x\}). \tag{1.21}$$

Indeed, let $\{A_n\}_{n=1}^{\infty}$ be a family of finite sets such that $A_n \subset A_{n+1}$ for $n \in \mathbb{N}$, $x \in A_1$, and $\cup_{n\geq 1} A_n = X$. Then $R_{\text{eff}}(x, A_n^c)^{-1} \downarrow R_{\text{eff}}(x)^{-1}$. Let $\varphi_n(z) = \mathbb{P}^z(\sigma_x < \tau_{A_n})$. Using Lemma 1.5 (ii), and noting $\varphi_n \in C_0(X)$, we have, for $m \leq n$,

$$\mathcal{E}(\varphi_m, \varphi_n) = (\varphi_m, -\mathcal{L}\varphi_n) = (1_{\{x\}}, -\mathcal{L}\varphi_n) = \mathcal{E}(\varphi_n, \varphi_n) = R_{\text{eff}}(x, A_n^c)^{-1}. \tag{1.22}$$

This implies

$$\mathcal{E}(\varphi_m - \varphi_n, \varphi_m - \varphi_n) = R_{\text{eff}}(x, A_m^c)^{-1} - R_{\text{eff}}(x, A_n^c)^{-1}.$$

Hence $\{\varphi_m\}$ is a $\mathcal{E}$-Cauchy sequence. Noting that $\varphi_n \to \varphi$ pointwise, we see that $\varphi_n \to \varphi$ in H^2 as well and $\varphi \in H_0^2$. Taking $n = m$ and $n \to \infty$ in (1.22), we obtain (1.21) except the last equality. To prove the last equality of (1.21), take any $f \in H_0^2$ with $f(x) = 1$. Then $g := f - \varphi \in H_0^2$ and

$g(x) = 0$. Let $g_n \in C_0(X)$ with $g_n \to g$ in H_0^2. Then, by Lemma 1.5 (ii), $\mathcal{E}(\varphi, g_n) = (-\mathcal{L}\varphi, g_n)$. Noting that φ is harmonic except at x, we see that $\mathcal{L}\varphi \in C_0(X)$. so, letting $n \to \infty$, we have

$$\mathcal{E}(\varphi, g) = (-\mathcal{L}\varphi, g) = -\mathcal{L}\varphi(x)g(x)\mu_x = 0.$$

Thus,

$$\mathcal{E}(f, f) = \mathcal{E}(\varphi + g, \varphi + g) = \mathcal{E}(\varphi, \varphi) + \mathcal{E}(g, g) \geq \mathcal{E}(\varphi, \varphi),$$

which means that φ is the unique minimizer in the definition of Cap $(\{x\})$. So the last equality of (1.21) is obtained.

Given (1.21), we now prove the equivalence.

$(i) \Longrightarrow (iii)'$: This is a direct consequence of (1.18) and (1.21).

$(iii) \Longleftrightarrow (ii) \Longleftrightarrow (iii)'$: This is easy. Indeed, Cap $(\{x\}) = 0$ if and only if there is $f \in H_0^2$ with $f(x) = 1$ and $\mathcal{E}(f, f) = 0$, that is f is identically 1.

$(iii)' \Longrightarrow (vi)$: Let $f \in C_0(X) \subset H_0^2$ with $f(x) \neq 0$, and define $g = f/f(x)$. Then

$$\text{Cap}(\{x\}) \leq \mathcal{E}(g, g) = \mathcal{E}(f, f)/f(x)^2.$$

So, letting $c_1(x) = 1/\text{Cap}(\{x\}) > 0$, we obtain (vi).

$(vi) \Longrightarrow (i)$: As before, let $\varphi_n(z) = \mathbb{P}^z(\sigma_x < \tau_{A_n})$. Then by (1.20), $\mathcal{E}(\varphi_n, \varphi_n) \geq c_1(x)^{-1}$. So, using the fact $\varphi_n \to \varphi$ in H^2 and (1.21), $R_{\text{eff}}(x)^{-1} = \mathcal{E}(\varphi, \varphi) = \lim_n \mathcal{E}(\varphi_n, \varphi_n) \geq c_1(x)^{-1}$. This means the transience by (1.18).

$(ii) \Longleftrightarrow (iv)$: $(ii) \Longrightarrow (iv)$ is clear since $1 \in H^2$, so we will prove the opposite direction. Suppose $1 \in H_0^2$. Then there exists $\{f_n\}_n \subset C_0(X)$ such that $\|1 - f_n\|_{H^2} < n^{-2}$. Since $\mathcal{E}$ is Markovian, we have $\|1 - f_n|_{H^2} \geq \|1 - (f_n \vee 0) \wedge 1\|_{H^2}$, so without loss of generality we may assume $f_n \geq 0$. Let $v_n = nf_n \geq 0$. Then $\lim_n v_n(x) = \infty$ for all $x \in X$ and $\mathcal{E}(v_n, v_n) = n^2\mathcal{E}(f_n, f_n) \leq n^{-2} \to 0$ so by Lemma 1.21, $\|u - (u \wedge v_n)\|_{H^2} \to 0$ for all $u \in H^2$ with $u \geq 0$. Since $u \wedge v_n \in C_0(X)$, this implies $u \in H_0^2$. For general $u \in H^2$, we can decompose it into $u_+ - u_-$ where $u_+, u_- \geq 0$ are in H^2. So applying the above, we have $u_+, u_- \in H_0^2$ and conclude $u \in H_0^2$.

$(i) \Longrightarrow (v)$: If the corresponding Markov chain is transient, then $\psi(z) = \mathbb{P}^z(\sigma_x^+ < \infty)$ is the non-constant super-harmonic function.

$(i) \Longleftarrow (v)$: Suppose the corresponding Markov chain $\{Y_n\}_n$ is recurrent. For a super-harmonic function $\psi \geq 0$, $M_n = \psi(Y_n) \geq 0$ is a supermartingale, so it converges $\mathbb{P}^x$-a.s. Let M_∞ be the limiting random variable. Since the set $\{n \in \mathbb{N} : Y_n = y\}$ is unbounded $\mathbb{P}^x$-a.s. for all $y \in X$ (due to the recurrence), we have $\mathbb{P}^x(\psi(y) = M_\infty) = 1$ for all $y \in X$, so ψ is constant. □

Remark 1.23 $(v) \Longrightarrow (i)$ *implies that if the Markov chain corresponding to* (X, μ) *is recurrent, then it has the strong Liouville property.*

For A, B which are subspaces of H^2, we write $A \oplus B = \{f + g : f \in A, g \in B\}$ if $\mathcal{E}(f, g) = 0$ for all $f \in A$ and $g \in B$.

As we see above, the Markov chain corresponding to (X, μ) is recurrent if and only if $H^2 = H_0^2$. When the Markov chain is transient, we have the following decomposition of H^2, which is called the Royden decomposition (see [107, Theorem 3.69]).

Proposition 1.24 *Suppose that the Markov chain corresponding to (X, μ) is transient. Then*

$$H^2 = \mathcal{H} \oplus H_0^2,$$

where $\mathcal{H} := \{h \in H^2 : h$ is a harmonic functions on $X.\}$. Further the decomposition is unique.

Proof. For each $f \in H^2$, let $a_f = \inf_{h \in H_0^2} \mathcal{E}(f-h, f-h)$. Then, similarly to the proof of Proposition 1.13, we can show that there is a unique minimizer $v_f \in H_0^2$ such that $a_f = \mathcal{E}(f - v_f, f - v_f)$, $\mathcal{E}(f - v_f, g) = 0$ for all $g \in H_0^2$, and in particular $f - v_f$ is harmonic on X. For the uniqueness of the decomposition, suppose $f = u + v = u' + v'$ where $u, u' \in \mathcal{H}$ and $v, v' \in H_0^2$. Then, $w := u - u' = v' - v \in \mathcal{H} \cap H_0^2$, so $\mathcal{E}(w, w) = 0$, which implies w is constant. Since $w \in H_0^2$ and the Markov chain is transient, by Proposition 1.22 we have $w \equiv 0$. □

1.3 Trace of weighted graphs

Finally in this section, we briefly mention the trace of weighted graphs, which will be used in Section 3 and Section 7. Note that there is a general theory on traces for Dirichlet forms (see [55]). Also note that a trace to infinite subset of X may not satisfy locally finiteness, but one can consider quadratic forms on them similarly.

Proposition 1.25 (Trace of the weighted graph) *Let $V \subset X$ be a non-void set such that $\mathbb{P}(\sigma_V < \infty) = 1$ and let f be a function on V. Then there exists a unique $u \in H^2$ which attains the following infimum:*

$$\inf\{\mathcal{E}(v, v) : v \in H^2, v|_V = f\}. \tag{1.23}$$

Moreover, the map $f \mapsto u =: H_V f$ is a linear map and there exist weights $\{\hat{\mu}_{xy}\}_{x,y \in V}$ such that the corresponding quadratic form $\hat{\mathcal{E}}_V(\cdot, \cdot)$ satisfies the following:

$$\hat{\mathcal{E}}_V(f, f) = \mathcal{E}(H_V f, H_V f) \qquad \forall f : V \to \mathbb{R}.$$

Proof. The first part can be proved similarly to Proposition 1.11 and Proposition 1.13. It is clear that $H_V(cf) = cH_V f$, so we will show $H_V(f_1 + f_2) = H_V(f_1) + H_V(f_2)$. Let $\varphi = H_V(f_1 + f_2)$, $\varphi_i = H_V(f_i)$ for $i = 1, 2$, and for $f : V \to \mathbb{R}$, define $Ef : X \to \mathbb{R}$ by $(Ef)|_V = f$ and $Ef(x) = 0$ when $x \in V^c$. As in the proof of Proposition 1.13 (ii), we see that $\mathcal{E}(H_V f, g) = 0$ whenever Supp $g \subset V^c$. So

$$\mathcal{E}(\varphi_1 + \varphi_2, \varphi_1 + \varphi_2) = \mathcal{E}(\varphi_1 + \varphi_2, Ef_1 + Ef_2) = \mathcal{E}(\varphi_1 + \varphi_2, \varphi) = \mathcal{E}(E(f_1 + f_2), \varphi) = \mathcal{E}(\varphi, \varphi).$$

Using the uniqueness of (1.23), we obtain $\varphi_1 + \varphi_2 = \varphi$. This establishes the linearity of H_V. Set $\hat{\mathcal{E}}(f, f) = \mathcal{E}(H_V f, H_V f)$. Clearly, $\hat{\mathcal{E}}$ is a non-negative definite symmetric bilinear form and $\hat{\mathcal{E}}(f, f) = 0$ if any only if f is a constant function. So, there exists $\{a_{xy}\}_{x,y \in V}$ with $a_{xy} = a_{yx}$ such that $\hat{\mathcal{E}}(f, f) = \frac{1}{2} \sum_{x,y \in V} a_{xy}(f(x) - f(y))^2$.

Next, we show that $\hat{\mathcal{E}}$ is Markovian. Indeed, writing $\bar{u} = (0 \vee u) \wedge 1$ for a function u, since $\overline{H_V u}|_V = \bar{u}$, we have

$$\hat{\mathcal{E}}(\bar{u},\bar{u}) = \mathcal{E}(H_V\bar{u}, H_V\bar{u}) \le \mathcal{E}(\overline{H_V u}, \overline{H_V u}) \le \mathcal{E}(H_V u, H_V u) = \hat{\mathcal{E}}(u,u), \qquad \forall u : V \to \mathbb{R},$$

where the fact that $\mathcal{E}$ is Markovian is used in the second inequality. Now take $p, q \in V$ with $p \neq q$ arbitrary, and consider a function h such that $h(p) = 1, h(q) = -\alpha < 0$ and $h(z) = 0$ for $z \in V \backslash \{p, q\}$. Then, there exist c_1, c_2 such that

$$\begin{aligned}\hat{\mathcal{E}}(h,h) &= a_{pq}(h(p)-h(q))^2 + c_1 h(p)^2 + c_2 h(q)^2 = a_{pq}(1+\alpha)^2 + c_1 + c_2\alpha^2 \\ &\ge \hat{\mathcal{E}}(\bar{h},\bar{h}) = a_{pq}(\bar{h}(p)-\bar{h}(q))^2 + c_1\bar{h}(p)^2 + c_2\bar{h}(q)^2 = a_{pq} + c_1.\end{aligned}$$

So $(a_{pq}+c_2)\alpha^2 + 2a_{pq}\alpha \ge 0$. Since this holds for all $\alpha > 0$, we have $a_{pq} \ge 0$. Putting $\hat{\mu}_{pq} = a_{pq}$ for each $p, q \in V$ with $p \neq q$, we have $\hat{\mathcal{E}}_V = \hat{\mathcal{E}}$, that is $\hat{\mathcal{E}}$ is associated with the weighted graph $(V, \hat{\mu})$. □

We call the induced weights $\{\hat{\mu}_{xy}\}_{x,y\in V}$ as the *trace of* $\{\mu_{xy}\}_{x,y\in X}$ *to* V. From this proposition, we see that for $x, y \in V$, $R_{\text{eff}}(x,y) = R^V_{\text{eff}}(x,y)$ where $R^V_{\text{eff}}(\cdot,\cdot)$ is the effective resistance for $(V,\hat{\mu})$.

2 Heat kernel upper bounds (The Nash inequality)

In this section, we will consider various equivalent inequalities to the Nash-type heat kernel upper bound, i.e. $p_t(x,y) \le c_1 t^{-\theta/2}$ for some $\theta > 0$. We would prefer to discuss them under a general framework including weighted graphs. However, some arguments here are rather sketchy to apply for the general framework. (The whole arguments are fine for weighted graphs, so readers may only consider them.) This section is strongly motivated by Coulhon's paper [40].

Let X be a locally compact separable metric space and μ be a Radon measure on X such that $\mu(B) > 0$ for any non-void ball. $(\mathcal{E}, \mathcal{F})$ is called a Dirichlet form on $\mathbb{L}^2(X,\mu)$ if it is a symmetric Markovian closed bilinear form on $\mathbb{L}^2$. It is well-known that given a Dirichlet form, there is a corresponding symmetric strongly continuous Markovian semigroup $\{P_t\}_{t\ge 0}$ on $\mathbb{L}^2(X,\mu)$ (see [55, Section 1.3, 1.4]). Here Markovian means if $u \in \mathbb{L}^2$ satisfies $0 \le u \le 1$ μ-a.s., then $0 \le P_t u \le 1$ μ-a.s. for all $t \ge 0$. We denote the corresponding non-negative definite $\mathbb{L}^2$-generator by $-\mathcal{L}$.

We denote the inner product of $\mathbb{L}^2$ by $(\cdot,\cdot)$ and for $p \ge 1$ denote $\|f\|_p$ for the $\mathbb{L}^p$-norm of $f \in \mathbb{L}^2(X,\mu)$. For each $\alpha > 0$, define

$$\mathcal{E}_\alpha(\cdot,\cdot) = \mathcal{E}(\cdot,\cdot) + \alpha(\cdot,\cdot).$$

$(\mathcal{E}_1, \mathcal{F})$ is then a Hilbert space.

2.1 The Nash inequality

We first give a preliminary lemma.

Lemma 2.1 *(i)* $\|P_t f\|_1 \le \|f\|_1$ *for all* $f \in \mathbb{L}^1 \cap \mathbb{L}^2$.
(ii) For $f \in \mathbb{L}^2$*, define* $u(t) = (P_t f, P_t f)$*. Then* $u'(t) = -2\mathcal{E}(P_t f, P_t f)$.
(iii) For $f \in \mathcal{F}$ *and* $t \ge 0$, $\exp(-\mathcal{E}(f,f)t/\|f\|_2^2) \le \|P_t f\|_2/\|f\|_2$.

Proof. (i) We first show that if $0 \le f \in \mathbb{L}^2$, 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 $\mathbb{L}^2$. Since $0 \le f_n \le n$, the Markovian property of $\{P_t\}$ implies that $0 \le P_t f_n \le n$. Taking $n \to \infty$, we obtain $0 \le P_t f$. So we have $P_t|f| \ge |P_t f|$, since $-|f| \le f \le |f|$. Using this and the Markovian property, we have for all $f \in \mathbb{L}^2 \cap \mathbb{L}^1$ and all Borel set $A \subset X$,

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

Hence we have $P_t f \in \mathbb{L}^1$ and $\|P_t f\|_1 \le \|f\|_1$.
(ii) Since $P_t f \in Dom(\mathcal{L})$, we have

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

Hence $u'(t) = -2\mathcal{E}(P_t f, P_t f)$.
(iii) We will prove the inequality for $f \in Dom(\mathcal{L})$; then one can obtain the result for $f \in \mathcal{F}$ by approximation. Let $-\mathcal{L} = \int_0^\infty \lambda dE_\lambda$ be the spectral decomposition of $-\mathcal{L}$. Then $P_t = e^{\mathcal{L}t} = \int_0^\infty e^{-\lambda t} dE_\lambda$ and $\|f\|_2^2 = \int_0^\infty (dE_\lambda f, f)$. Since $\lambda \mapsto e^{-2\lambda t}$ is convex, by Jensen's inequality,

$$\exp\Big(-2\int_0^\infty \lambda t \frac{(dE_\lambda f, f)}{\|f\|_2^2}\Big) \le \int_0^\infty e^{-2\lambda t}\frac{(dE_\lambda f, f)}{\|f\|_2^2} = \frac{(P_{2t}f, f)}{\|f\|_2^2} = \frac{\|P_t f\|_2^2}{\|f\|_2^2}.$$

Taking the square root in each term, we obtain the desired inequality. □

Remark 2.2 *An alternative proof of (iii) is to use the logarithmic convexity of* $\|P_t f\|_2^2$. *Indeed,*

$$\|P_{(t+s)/2}f\|_2^2 = (P_{t+s}f, f) = (P_t f, P_s f) \le \|P_t f\|_2 \|P_s f\|_2, \quad \forall s, t > 0$$

so $\|P_t f\|_2^2$ *is logarithmic convex. Thus,*

$$t \mapsto \frac{d}{dt}\log\|P_t f\|_2^2 = \frac{\frac{d}{dt}(\|P_t f\|_2^2)}{\|P_t f\|_2^2} = -\frac{2\mathcal{E}(P_t f, P_t f)}{\|P_t f\|_2^2} \tag{2.1}$$

is non-decreasing. (The last equality is due to Lemma 2.1(ii).) The right hand side of (2.1) is $-\frac{2\mathcal{E}(f,f)}{\|f\|_2^2}$ *when* $t = 0$, *so integrating (2.1) over* $[0, t]$, *we have*

$$\log\frac{\|P_t f\|_2^2}{\|f\|_2^2} = \int_0^t \frac{d}{ds}\log\|P_s f\|_2^2 ds \ge -\frac{2t\mathcal{E}(f,f)}{\|f\|_2^2}.$$

The following is easy to see. (Note that we only need the first assertion.)

Lemma 2.3 *Let* $(\mathcal{E}, \mathcal{F})$ *be a symmetric closed bilinear form on* $\mathbb{L}^2(X, \mu)$, *and let* $\{P_t\}_{t\ge 0}$, $-\mathcal{L}$ *be the corresponding semigroup and the self-adjoint operator respectively. Then, for each* $\delta > 0$, $(\mathcal{E}_\delta, \mathcal{F})$ *is also a symmetric closed bilinear form and the corresponding semigroup and the self-adjoint operator are* $\{e^{-\delta t}P_t\}_{t\ge 0}$, $\delta I - \mathcal{L}$, *respectively. Further, if* $(\mathcal{E}, \mathcal{F})$ *is the regular Dirichlet form on* $\mathbb{L}^2(X, \mu)$ *and the corresponding Hunt process is* $\{Y_t\}_{t\ge 0}$, *then* $(\mathcal{E}_\delta, \mathcal{F})$ *is also the regular Dirichlet form and the corresponding hunt process is* $\{Y_{t\wedge\zeta}\}_{t\ge 0}$ *where* ζ *is the independent exponential random variable with parameter* δ. *(*ζ *is the killing time; i.e. the process goes to the cemetery point at* ζ.*)*

The next theorem was proved by Carlen-Kusuoka-Stroock ([38]), where the original idea of the proof of $(i) \Rightarrow (ii)$ was due to Nash [96].

Theorem 2.4 (The Nash inequality, [38])
The following are equivalent for any $\delta \geq 0$.
(i) There exist $c_1, \theta > 0$ such that for all $f \in \mathcal{F} \cap L^1$,

$$\|f\|_2^{2+4/\theta} \leq c_1(\mathcal{E}(f,f) + \delta\|f\|_2^2)\|f\|_1^{4/\theta}. \tag{2.2}$$

(ii) 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

$$\|P_t\|_{1\to\infty} \leq c_2 e^{\delta t} t^{-\theta/2}, \qquad \forall t > 0. \tag{2.3}$$

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

When $\delta = 0$, we cite (2.2) as (N_θ) and (2.3) as (UC_θ).

Proof. First, note that using Lemma 2.1, it is enough to prove the theorem when $\delta = 0$.
$(i) \Rightarrow (ii)$: Let $f \in L^2 \cap L^1$ with $\|f\|_1 = 1$ and $u(t) := (P_t f, P_t f)_2$. Then, by Lemma 2.1 (ii), $u'(t) = -2\mathcal{E}(P_t f, P_t f)$. Now by (i) and Lemma 2.1 (i),

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

so $u'(t) \leq -c_2 u(t)^{1+2/\theta}$. Set $v(t) = u(t)^{-2/\theta}$, then we obtain $v'(t) \geq 2c_2/\theta$. Since $\lim_{t\downarrow 0} v(t) = u(0)^{-2/\theta} = \|f\|_2^{-4/\theta} > 0$, it follows that $v(t) \geq 2c_2 t/\theta$. This means $u(t) \leq c_3 t^{-\theta/2}$, whence $\|P_t f\|_2 \leq c_3 t^{-\theta/4}\|f\|_1$ for all $f \in \mathbb{L}^2 \cap \mathbb{L}^1$, which implies $\|P_t\|_{1\to 2} \leq c_3 t^{-\theta/4}$. 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 (ii).
$(ii) \Rightarrow (i)$: Let $f \in L^2 \cap L^1$ with $\|f\|_1 = 1$. Using (ii) and Lemma 2.1 (iii), we have

$$\exp(-2\frac{\mathcal{E}(f,f)t}{\|f\|_2^2}) \leq \frac{c_4 t^{-\theta/2}}{\|f\|_2^2}.$$

Rewriting, we have $\mathcal{E}(f,f)/\|f\|_2^2 \geq (2t)^{-1}\log(t^{\theta/2}\|f\|_2^2) - (2t)^{-1}A$, where $A = \log c_4$ and we may take $A > 0$. Set $\Psi(x) = \sup_{t>0}\{\frac{x}{2t}\log(xt^{\theta/2}) - Ax/(2t)\}$. By elementary computations, we have $\Psi(x) \geq c_5 x^{1+2/\theta}$. So

$$\mathcal{E}(f,f) \geq \Psi(\|f\|_2^2) \geq c_5\|f\|_2^{2(1+2/\theta)} = c_5\|f\|^{2+4/\theta}.$$

Since this holds for all $f \in L^2 \cap L^1$ with $\|f\|_1 = 1$, we obtain (i). □

Remark 2.5 *(1) When one is only concerned about $t \geq 1$ (for example on graphs), we have the following equivalence under the assumption of $\|P_t\|_{1\to\infty} \leq C$ for all $t \geq 0$ (C is independent of t).*

(i) (N_θ) *with $\delta = 0$ holds for $\mathcal{E}(f,f) \leq \|f\|_1^2$.*

(ii) (UC_θ) *with $\delta = 0$ holds for $t \geq 1$.*

(2) We have the following generalization of the theorem due to [41]. Let $m : \mathbb{R}_+ \to \mathbb{R}_+$ be a decreasing C^1 bijection which satisfies the following: $M(t) := -\log m(t)$ satisfies $M'(u) \ge c_0 M'(t)$ for all $t \ge 0$ and all $u \in [t, 2t]$. (Roughly, this condition means that the logarithmic derivative of $m(t)$ is polynomial growth. So exponential growth functions may satisfy the condition, but double exponential growth functions do not.) Let $\Psi(x) = -m'(m^{-1}(x))$. Then the following are equivalent.

(i) $c_1 \Psi(\|f\|_2^2) \le \mathcal{E}(f,f)$ *for all* $f \in \mathcal{F}, \|f\|_1 \le 1$.

(ii) $\|P_t\|_{1\to\infty} \le c_2 m(t)$ *for all* $t > 0$.

The above theorem corresponds to the case $\Psi(y) = c_4 y^{1+2/\theta}$ *and* $m(t) = t^{-\theta/2}$.

Corollary 2.6 *Suppose the Nash inequality (Theorem 2.4) holds. Let φ be an eigenfunction of $-\mathcal{L}$ with eigenvalue $\lambda \ge 1$. Then*

$$\|\varphi\|_\infty \le c_1 \lambda^{\theta/4} \|\varphi\|_2,$$

where $c_1 > 0$ is a constant independent of φ and λ.

Proof. Since $-\mathcal{L}\varphi = \lambda\varphi$, $P_t\varphi = e^{t\mathcal{L}}\varphi = e^{-\lambda t}\varphi$. By Theorem 2.4, $\|P_t\|_{2\to\infty} = \|P_t\|_{1\to\infty}^{1/2} \le ct^{-\theta/4}$ for $t \le 1$. Thus

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

Taking $t = \lambda^{-1}$ and $c_1 = ce$, we obtain the result. □

Example 2.7 *Consider $\mathbb{Z}^d$, $d \ge 2$ and put weight 1 for each edge $\{x,y\}$, $x, y \in \mathbb{Z}^d$ with $\|x-y\| = 1$. Then it is known that the corresponding simple random walk enjoys the following heat kernel estimate: $p_n(x,y) \le c_1 n^{-d/2}$ for all $x, y \in \mathbb{Z}^d$ and all $n \ge 1$. Now, let*

$$H = \{\{(2n_1, 2n_2, \cdots, 2n_d), (2n_1+1, 2n_2, \cdots, 2n_d)\} : n_1, \cdots, n_d \in \mathbb{Z}\}$$

and consider a random subgraph $\mathcal{C}(\omega)$ by removing each $e \in H$ with probability $p \in [0,1]$ independently. (So the set of vertices of $\mathcal{C}(\omega)$ is $\mathbb{Z}^d$. Here ω is the randomness of the environments.) If we define the quadratic forms for the original graph and $\mathcal{C}(\omega)$ by $\mathcal{E}(\cdot,\cdot)$ and $\mathcal{E}^\omega(\cdot,\cdot)$ respectively, then it is easy to see that $\mathcal{E}^\omega(f,f) \le \mathcal{E}(f,f) \le 4\mathcal{E}^\omega(f,f)$ for all $f \in \mathbb{L}^2$. Thus, by Remark 2.5(1), the heat kernel of the simple random walk on $\mathcal{C}(\omega)$ still enjoys the estimate $p_n^\omega(x,y) \le c_1 n^{-d/2}$ for all $x, y \in \mathcal{C}(\omega)$ and all $n \ge 1$, for almost every ω.

2.2 The Faber-Krahn, Sobolev and isoperimetric inequalities

In this subsection, we denote$\Omega \subset\subset X$ when Ω is an open relative compact subset of X. (For weighted graphs, it simply means that Ω is a finite subset of X.) Let $C_0(X)$ be the space of continuous, compactly supported functions on X. Define

$$\lambda_1(\Omega) = \inf_{\substack{f \in \mathcal{F}\cap C_0(X),\\ \mathrm{Supp}\, f \subset Cl(\Omega)}} \frac{\mathcal{E}(f,f)}{\|f\|_2^2}, \qquad \forall \Omega \subset\subset X.$$

By the min-max principle, this is the first eigenvalue for the corresponding Laplace operator which is zero outside Ω.

Definition 2.8 (The Faber-Krahn inequality)
Let $\theta > 0$. We say $(\mathcal{E}, \mathcal{F})$ satisfies the Faber-Krahn inequality of order θ if the following holds:

$$\lambda_1(\Omega) \geq c\mu(\Omega)^{-2/\theta}, \qquad \forall \Omega \subset\subset X. \tag{$FK(\theta)$}$$

Theorem 2.9 $(N_\theta) \Leftrightarrow (FK(\theta))$

Proof. $(N_\theta) \Rightarrow (FK(\theta))$: This is an easy direction. From (N_θ), we have

$$\|f\|_2^2 \leq \Big(\frac{\|f\|_1^2}{\|f\|_2^2}\Big)^{2/\theta} \mathcal{E}(f,f), \qquad \forall f \in \mathcal{F} \cap \mathbb{L}^1. \tag{2.4}$$

On the other hand, if Supp $f \subset Cl(\Omega)$, then by Schwarz's inequality, $\|f\|_1^2 \leq \mu(\Omega)\|f\|_2^2$, so $(\|f\|_1^2/\|f\|_2^2)^{2/\theta}$ $\leq \mu(\Omega)^{2/\theta}$. Putting this into (2.4), we obtain $(FK(\theta))$.
$(FK(\theta)) \Rightarrow (N_\theta)$: We adopt the argument originated in [60]. Let $u \in \mathcal{F} \cap C_0(X)$ be a non-negative function. For each $\lambda > 0$, since $u < 2(u - \lambda)$ on $\{u > 2\lambda\}$, we have

$$\begin{aligned} \int u^2 d\mu &= \int_{\{u>2\lambda\}} u^2 d\mu + \int_{\{u\leq 2\lambda\}} u^2 d\mu \\ &\leq 4\int_{\{u>2\lambda\}} (u-\lambda)^2 d\mu + 2\lambda \int_{\{u\leq 2\lambda\}} u d\mu \leq 4\int (u-\lambda)_+^2 d\mu + 2\lambda\|u\|_1. \end{aligned} \tag{2.5}$$

Note that $(u - \lambda)_+ \in \mathcal{F}$ since $(\mathcal{E}, \mathcal{F})$ is Markovian (cf. [55, Theorem 1.4.1]). Set $\Omega = \{u > \lambda\}$; then Ω is an open relative compact set since u is compactly supported, and Supp $(u-\lambda)_+ \subset \Omega$. So, applying $(FK(\theta))$ to $(u - \lambda)_+$ gives

$$\int (u-\lambda)_+^2 d\mu \leq \mu(\Omega)^{2/\theta}\mathcal{E}((u-\lambda)_+, (u-\lambda)_+) \leq \Big(\frac{\|u\|_1}{\lambda}\Big)^{2/\theta}\mathcal{E}(u,u),$$

where we used the Chebyshev inequality in the second inequality. Putting this into (2.5),

$$\|u\|_2^2 \leq 4\Big(\frac{\|u\|_1}{\lambda}\Big)^{2/\theta}\mathcal{E}(u,u) + 2\lambda\|u\|_1.$$

Optimizing the right hand side by taking $\lambda = c_1\mathcal{E}(u,u)^{\theta/(\theta+2)}\|u\|_1^{(2-\theta)/(2+\theta)}$, we obtain

$$\|u\|_2^2 \leq c_2\mathcal{E}(u,u)^{\frac{\theta}{\theta+2}}\|u\|_1^{\frac{4}{2+\theta}},$$

and thus obtain (N_θ). For general compactly supported $u \in \mathcal{F}$, we can obtain (N_θ) for u_+ and u_-, so for u as well. For general $u \in \mathcal{F} \cap \mathbb{L}^1$, approximation by compactly supported functions gives the desired result. □

Remark 2.10 *We can generalize Theorem 2.9 as follows. Let*

$$\lambda_1(\Omega) \geq \frac{c}{\varphi(\mu(\Omega))^2}, \qquad \forall \Omega \subset\subset X, \tag{2.6}$$

where $\varphi : (0, \infty) \to (0, \infty)$ *is a non-decreasing function. Then, it is equivalent to Remark 2.5(2)(i), where* $\Psi(x) = x/\varphi(1/x)^2$*, or* $\varphi(x) = (x\Psi(1/x))^{-1/2}$*. Theorem 2.9 is the case* $\varphi(x) = x^{1/\theta}$*.*

In the following, we define $\|\nabla f\|_1$ for the two cases. Case 1: When one can define the gradient on the space and $\mathcal{E}(f, f) = \frac{1}{2}\int_X |\nabla f(x)|^2 d\mu(x)$, then $\|\nabla f\|_1 := \int_X |\nabla f(x)| d\mu(x)$. Case 2: When (X, μ) is a weighted graph, then $\|\nabla f\|_1 := \frac{1}{2}\sum_{x,y\in X} |f(y) - f(x)|\mu_{xy}$. Whenever $\|\nabla f\|_1$ appears, we consider that we are either of the two cases.

Definition 2.11 (Sobolev inequalities)
(i) Let $\theta > 2$*. We say* $(\mathcal{E}, \mathcal{F})$ *satisfies* (S^2_θ) *if*

$$\|f\|^2_{2\theta/(\theta-2)} \leq c_1 \mathcal{E}(f, f), \qquad \forall f \in \mathcal{F} \cap C_0(X). \tag{S^2_θ}$$

(ii) Let $\theta > 1$*. We say* $(\mathcal{E}, \mathcal{F})$ *satisfies* (S^1_θ) *if*

$$\|f\|_{\theta/(\theta-1)} \leq c_2 \|\nabla f\|_1, \qquad \forall f \in \mathcal{F} \cap C_0(X). \tag{S^1_θ}$$

In the following, we define $|\partial\Omega|$ for the two cases. Case 1: When X is a d-dimensional Riemannian manifold and Ω is a smooth domain, then $|\partial\Omega|$ is the surface measure of Ω. Case 2: When (X, μ) is a weighted graph, then $|\partial\Omega| = \sum_{x\in\Omega}\sum_{y\in\Omega^c} \mu_{xy}$. Whenever $|\partial\Omega|$ appears, we consider that we are either of the two cases.

Definition 2.12 (The isoperimetric inequality)
Let $\theta > 1$*. We say* $(\mathcal{E}, \mathcal{F})$ *satisfies the isoperimetric inequality of order* θ *if*

$$\mu(\Omega)^{(\theta-1)/\theta} \leq c_1 |\partial\Omega|, \qquad \forall \Omega \subset\subset X. \tag{I_θ}$$

We write (I_∞) *when* $\theta = \infty$*, namely when*

$$\mu(\Omega) \leq c_1 |\partial\Omega|, \qquad \forall \Omega \subset\subset X. \tag{I_∞}$$

Remark 2.13 *(i) For the weighted graph* (X, μ) *with* $\mu_x \geq 1$ *for all* $x \in X$*, if* (I_β) *holds, then* (I_α) *holds for any* $\alpha \leq \beta$*. So* (I_∞) *is the strongest inequality among all the isoperimetric inequalities.*
(ii) $\mathbb{Z}^d$ *satisfies* (I_d)*. The binary tree satisfies* (I_∞)*.*

Theorem 2.14 *The following holds for* $\theta > 0$*.*

$$(I_\theta) \overset{\theta>1}{\Longleftrightarrow} (S^1_\theta) \overset{\theta>2}{\Longrightarrow} (S^2_\theta) \overset{\theta>2}{\Longleftrightarrow} (N_\theta) \Longleftrightarrow (UC_\theta) \Longleftrightarrow (FK(\theta))$$

Proof. Note that the last two equivalence relations are already proved in Theorem 2.4 and 2.9.
$(I_\theta) \overset{\theta>1}{\Longleftarrow} (S^1_\theta)$: When (X,μ) is the weighted graph, simply apply (S^1_θ) to $f = 1_\Omega$ and we can obtain (I_θ). When X is the Riemannian manifold, by using a Lipschitz function which approximates $f = 1_\Omega$ nicely, we can obtain (I_θ).
$(I_\theta) \overset{\theta>1}{\Longrightarrow} (S^1_\theta)$: We will use the co-area formula given in Lemma 2.16 below. Let f be a support compact non-negative function on X (when X is the Riemannian manifold, $f \in C^1_0(X)$ and $f \geq 0$). Let $H_t(f) = \{x \in X : f(x) > t\}$ and set $p = \theta/(\theta-1)$. Applying (I_θ) to f and using Lemma 2.16 below, we have

$$\|\nabla f\|_1 = \int_0^\infty |\partial H_t(f)| dt \geq c_1 \int_0^\infty \mu(H_t(f))^{1/p} dt = c_1 \int_0^\infty \|1_{H_t(f)}\|_p dt. \tag{2.7}$$

Next take any $g \in \mathbb{L}^q$ such that $g \geq 0$ and $\|g\|_q = 1$ where q is the value that satisfies $p^{-1} + q^{-1} = 1$. Then, by the Hölder inequality,

$$\int_0^\infty \|1_{H_t(f)}\|_p dt \geq \int_0^\infty \|g \cdot 1_{H_t(f)}\|_1 dt = \int_X g(x) \int_0^\infty 1_{H_t(f)}(x) dt d\mu(x) = \|fg\|_1,$$

since $\int_0^\infty 1_{H_t(f)}(x)dt = f(x)$. Putting this into (2.7), we obtain

$$\|f\|_p = \sup_{g \in \mathbb{L}^q : \|g\|_q = 1} \|fg\|_1 \leq c_1^{-1} \|\nabla f\|_1,$$

so we have (S^1_θ). We can obtain (S^1_θ) for general $f \in \mathcal{F} \cap C_0(X)$ by approximations.
$(S^1_\theta) \overset{\theta>2}{\Longrightarrow} (S^2_\theta)$: Set $\hat{\theta} = 2(\theta-1)/(\theta-2)$ and let $f \in \mathcal{F} \cap C_0(X)$. Applying (S^1_θ) to $f^{\hat{\theta}}$ and using Schwarz's inequality,

$$\begin{aligned}\Big(\int f^{\frac{2\theta}{\theta-2}} d\mu\Big)^{\frac{\theta-1}{\theta}} &= \|f^{\hat{\theta}}\|_{\frac{\theta}{\theta-1}} \leq c_1 \|\nabla f^{\hat{\theta}}\|_1 \\ &\leq c_2 \|f^{\hat{\theta}-1} \nabla f\|_1 \leq c_2 \|\nabla f\|_2 \|f^{\hat{\theta}-1}\|_2 = c_2 \|\nabla f\|_2 \Big(\int f^{\frac{2\theta}{\theta-2}} d\mu\Big)^{1/2}.\end{aligned}$$

rearranging, we obtain (S^2_θ).
$(S^2_\theta) \overset{\theta>2}{\Longrightarrow} (N_\theta)$: For $f \in \mathcal{F} \cap C_0(X)$, applying the Hölder inequality (with $p^{-1} = 4/(\theta+2)$, $q^{-1} = (\theta-2)/(\theta+2)$) and using (S^2_θ), we have

$$\|f\|_2^{2+\frac{4}{\theta}} \leq \|f\|_1^{\frac{4}{\theta}} \|f\|^2_{\frac{2\theta}{\theta-2}} \leq c_1 \|f\|_1^{\frac{4}{\theta}} \mathcal{E}(f,f),$$

so we have (N_θ) in this case. Usual approximation arguments give the desired fact for $f \in \mathcal{F} \cap \mathbb{L}^1$.
$(S^2_\theta) \overset{\theta>2}{\Longleftarrow} (N_\theta)$: For $f \in \mathcal{F} \cap C_0(X)$ such that $f \geq 0$, define

$$f_k = (f - 2^k)_+ \wedge 2^k = 2^k 1_{A_k} + (f - 2^k) 1_{B_k}, \qquad k \in \mathbb{Z},$$

where $A_k = \{f \geq 2^{k+1}\}, B_k = \{2^k \leq f < 2^{k+1}\}$. Then $f = \sum_{k \in \mathbb{Z}} f_k$ and $f_k \in \mathcal{F} \cap C_0(X)$. So

$$\mathcal{E}(f,f) = \sum_{k \in \mathbb{Z}} \mathcal{E}(f_k, f_k) + \sum_{k \in \mathbb{Z}, k \neq k'} \sum_{k' \in \mathbb{Z}} \mathcal{E}(f_k, f_{k'}) \geq \sum_{k \in \mathbb{Z}} \mathcal{E}(f_k, f_k), \tag{2.8}$$

where the last inequality is due to $\sum_{k\neq k'}\mathcal{E}(f_k,f_{k'})\geq 0$. (This can be verified in a elementary way for the case of weighted graphs. When $\mathcal{E}$ is strongly local, $\sum_{k\neq k'}\mathcal{E}(f_k,f_{k'})=0$.)

Next, we have

$$\begin{aligned}\Big(2^{2k}\mu(A_k)\Big)^{1+2/\theta} &= \Big(\int_{A_k} f_k^2 d\mu\Big)^{1+2/\theta}\\ &\leq \|f_k\|_2^{2+4/\theta}\leq c_1\|f_k\|_1^{4/\theta}\mathcal{E}(f_k,f_k)\leq c_1\Big(2^k\mu(A_{k-1})\Big)^{4/\theta}\mathcal{E}(f_k,f_k), \qquad (2.9)\end{aligned}$$

where we used (N_θ) for f_k in the second inequality. Let $\alpha=2\theta/(\theta-2)$, $\beta=\theta/(\theta+2)\in(1/2,1)$, and define $a_k=2^{\alpha k}\mu(A_k)$, $b_k=\mathcal{E}(f_k,f_k)$. Then (2.9) can be rewritten as $a_k\leq c_2a_{k-1}^{2(1-\beta)}b_k^\beta$. Summing over $k\in\mathbb{Z}$ and using the Hölder inequality (with $p^{-1}=1-\beta, q^{-1}=\beta$), we have

$$\sum_k a_k\leq c_2\sum_k a_{k-1}^{2(1-\beta)}b_k^\beta\leq c_2(\sum_k a_{k-1}^2)^{1-\beta}(\sum_k b_k)^\beta\leq c_2(\sum_k a_k)^{2(1-\beta)}(\sum_k b_k)^\beta.$$

Putting (2.8) into this, we have

$$\sum_k a_k\leq c_2\mathcal{E}(f,f)^{\beta/(2\beta-1)}. \qquad (2.10)$$

On the other hand, we have

$$\|f\|_\alpha^\alpha=\sum_k\int_{B_k}f^\alpha d\mu\leq\sum_k 2^{\alpha(k+1)}\mu(A_{k-1})=2^{2\alpha}\sum_k a_k.$$

Plugging (2.10) into this, we obtain (S_θ^2). □

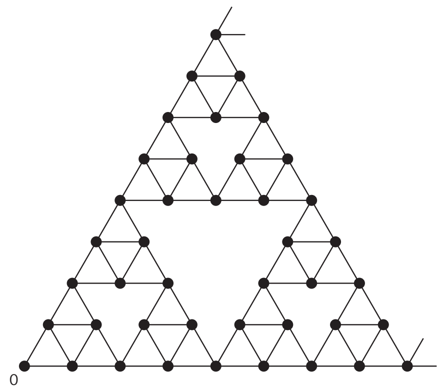


Figure 1: 2-dimensional pre-Sierpinski gasket

Remark 2.15 *(i) Generalizations of (S_θ^p) and (I_θ) are the following:*

$$\begin{aligned}&\|f\|_p\leq c_1\varphi(\mu(\Omega))\|\nabla f\|_p, && \forall\Omega\subset\subset X,\ \forall f\in\mathcal{F}\cap C_0(X)\ \textit{such that}\ \ Supp\,f\subset Cl(\Omega),\\ &\frac{1}{\varphi(\mu(\Omega))}\leq c_2\frac{|\partial\Omega|}{\mu(\Omega)}, && \forall\Omega\subset\subset X,\end{aligned}$$

where $\varphi:(0,\infty)\to(0,\infty)$ is a non-decreasing function as in Remark 2.10. Note that we defined (S_θ^p) for $p=1,2$, but the above generalization makes sense for all $p\in[1,\infty]$ (at least for Riemannian

manifolds). Theorem 2.14 is the case $\varphi(x) = x^{1/\theta}$ – see [40] for details.
(ii) As we see in the proof, in addition to the equivalence $(I_\theta) \iff (S^1_\theta)$, one can see that the best constant for c_1 in (I_θ) is equal to the best constant for c_2 in (S^1_θ). This is sometimes referred as the Federer-Fleming theorem.
(iii) It is easy to see that the pre-Sierpinski gasket (Figure 1) does not satisfy (I_θ) for any $\theta > 1$. On the other hand, we will prove in (3.28) *that the heat kernel of the simple random walk enjoys the following estimate $p_{2n}(0,0) \asymp n^{-\log 3/\log 5}$, $\forall n \geq 1$. This gives an example that (N_θ) cannot imply (I_θ) in general. In other word, the best exponent for isoperimetric inequalities is not necessarily the best exponent for Nash inequalities.*
(iv) In [93], there is an interesting approach to prove $(I_\theta) \Rightarrow (N_\theta)$ directly using the evolution of random sets.

The following lemma was used in the proof of Theorem 2.14.

Lemma 2.16 (Co-area formula) *Let f be a non-negative function on X (when X is the Riemannian manifold, $f \in C^1_0(X)$ and $f \geq 0$), and define $H_t(f) = \{x \in X : f(x) > t\}$. Then*

$$\|\nabla f\|_1 = \int_0^\infty |\partial H_t(f)| dt.$$

Proof. For simplicity we will prove it only when (X,μ) is a weighted graph. Then,

$$\begin{aligned} \|\nabla f\|_1 &= \frac{1}{2}\sum_{x,y\in X} |f(y) - f(x)|\mu_{xy} = \sum_{x\in X}\sum_{y\in X: f(y)>f(x)} (f(y) - f(x))\mu_{xy} \\ &= \sum_{x\in X}\sum_{y\in X: f(y)>f(x)} (\int_0^\infty 1_{\{f(y)>t\geq f(x)\}} dt)\mu_{xy} = \int_0^\infty dt \sum_{x,y\in X} 1_{\{f(y)>t\geq f(x)\}}\mu_{xy} \\ &= \int_0^\infty dt \sum_{x,\in H_t(f)}\sum_{y\in H_t(f)^c} \mu_{xy} = \int_0^\infty |\partial H_t(f)| dt. \end{aligned}$$

□

The next fact is an immediate corollary to Theorem 2.14.

Corollary 2.17 *Let $\theta > 2$. If $(\mathcal{E},\mathcal{F})$ satisfies (I_θ), then the following holds.*

$$p_t(x,y) \leq ct^{-\theta/2} \qquad \forall t > 0,\ \mu - a.e.\ x,y.$$

3 Heat kernel estimates using effective resistance

In this section, we will consider the weighted graph (X,μ). We say that (X,μ) is *loop-free* if for any $l \geq 3$, there is no set of distinct points $\{x_i\}_{i=1}^l \subset X$ such that $x_i \sim x_{i+1}$ for $1 \leq i \leq l$ where we set $x_{l+1} := x_1$.

Set $R_{\mathrm{eff}}(x,x) = 0$ for all $x \in X$. We now give an important lemma on the effective resistance.

Lemma 3.1 *(i) If* $c_1 := \inf_{x,y\in X: x\sim y}\mu_{xy} > 0$ *then* $R_{\rm eff}(x,y) \le c_1^{-1} d(x,y)$ *for all* $x,y \in X$.
(ii) If (X,μ) *is loop-free and* $c_2 := \sup_{x,y\in X:x\sim y}\mu_{xy} < \infty$, *then* $R_{\rm eff}(x,y) \ge c_2^{-1}d(x,y)$ *for all* $x,y\in X$.
(iii) $|f(x)-f(y)|^2 \le R_{\rm eff}(x,y)\mathcal{E}(f,f)$ *for all* $x,y\in X$ *and* $f \in H^2$.
(iv) $R_{\rm eff}(\cdot,\cdot)$ *and* $R_{\rm eff}(\cdot,\cdot)^{1/2}$ *are both metrics on* X.

Proof. (i) Take a shortest path between x and y and cut all the bonds that are not along the path. Then we have the inequality by the cutting law.
(ii) Suppose $d(x,y) = n$. Take a shortest path $(x_0, x_1, \cdots, x_n)$ between x and y so that $x_0 = x, x_n = y$. Now take $f : X \to \mathbb{R}$ so that $f(x_i) = (n-i)/n$ for $0 \le i \le n$, and $f(y) = f(x_i)$ if y is in the branch from x_i, i.e. if y can be connected to x_i without crossing $\{x_k\}_{k=0}^n \setminus \{x_i\}$. This f is well-defined because (X,μ) is loop-free, and $f(x) = 1, f(y) = 0$. So $R_{\rm eff}(x,y)^{-1} \le \sum_{i=0}^{n-1}(1/n)^2\mu_{x_ix_{i+1}} \le c_2/n = c_2/d(x,y)$, and the result follows.
(iii) For any non-constant function $u \in H^2$ and any $x \ne y \in X$, we can construct $f \in H^2$ such that $f(x) = 1, f(y) = 0$ by a linear transform $f(z) = au(z) + b$ (where a, b are chosen suitably). So

$$\sup\Big\{\frac{|u(x)-u(y)|^2}{\mathcal{E}(u,u)} : u \in H^2, \mathcal{E}(u,u) > 0\Big\} = \sup\Big\{\frac{1}{\mathcal{E}(f,f)} : f \in H^2, f(x) = 1, f(y) = 0\Big\} = R_{\rm eff}(x,y), \tag{3.1}$$

and we have the desired inequality.
(iv) It is easy to see $R_{\rm eff}(x,y) = R_{\rm eff}(y,x)$ and $R_{\rm eff}(x,y) = 0$ if and only if $x = y$. So we only need to check the triangle inequality.

Let $\tilde{H}^2 = \{u \in H^2 : \mathcal{E}(u,u) > 0\}$. Then, for $x,y,z \in X$ that are distinct, we have by (3.1)

$$\begin{aligned} R_{\rm eff}(x,y)^{1/2} &= \sup\Big\{\frac{|u(x)-u(y)|}{\mathcal{E}(u,u)^{1/2}} : u \in \tilde{H}^2\Big\} \\ \le \quad & \sup\Big\{\frac{|u(x)-u(z)|}{\mathcal{E}(u,u)^{1/2}} : u \in \tilde{H}^2\Big\} + \sup\Big\{\frac{|u(z)-u(y)|}{\mathcal{E}(u,u)^{1/2}} : u \in \tilde{H}^2\Big\} = R_{\rm eff}(x,z)^{1/2} + R_{\rm eff}(z,y)^{1/2}. \end{aligned}$$

So $R_{\rm eff}(\cdot,\cdot)^{1/2}$ is a metric on X.

Next, let $V = \{x,y,z\} \subset X$ and let $\{\hat\mu_{xy}, \hat\mu_{yz}, \hat\mu_{zx}\}$ be the trace of $\{\mu_{xy}\}_{x,y\in X}$ to V. Define $R_{xy}^{-1} = \hat\mu_{xy}$, $R_{yz}^{-1} = \hat\mu_{yz}$, $R_{zx}^{-1} = \hat\mu_{zx}$. Then, using Proposition 1.25 and the resistance formula of series and parallel circuits, we have

$$R_{\rm eff}(z,x) = \frac{1}{R_{zx}^{-1} + (R_{xy}+R_{yz})^{-1}} = \frac{R_{zx}(R_{xy}+R_{yz})}{R_{xy}+R_{yz}+R_{zx}}, \tag{3.2}$$

and similarly $R_{\rm eff}(x,y) = \frac{R_{xy}(R_{yz}+R_{zx})}{R_{xy}+R_{yz}+R_{zx}}$ and $R_{\rm eff}(y,z) = \frac{R_{yz}(R_{zx}+R_{xy})}{R_{xy}+R_{yz}+R_{zx}}$. Hence

$$\frac{1}{2}\{R_{\rm eff}(x,z) + R_{\rm eff}(z,y) - R_{\rm eff}(x,y)\} = \frac{R_{yz}R_{zx}}{R_{xy}+R_{yz}+R_{zx}} \ge 0, \tag{3.3}$$

which shows that $R_{\rm eff}(\cdot,\cdot)$ is a metric on X. □

Remark 3.2 *(i) Different proofs of the triangle inequality of* $R_{\text{eff}}(\cdot,\cdot)$ *in Lemma 3.1(iv) can be found in [10] and [75, Theorem 1.12].*
(ii) Weighted graphs are resistance forms. (See [79] for definition and properties of the resistance form.) In fact, most of the results in this section (including this lemma) hold for resistance forms.

3.1 Green density killed on a finite set

For $y \in X$ and $n \in \mathbb{N}$, let

$$L(y,n) = \sum_{k=0}^{n-1} 1_{\{Y_k = y\}}$$

be the local time at y up to time $n-1$. For a finite set $B \subset X$ and $x, y \in X$, define the Green density by

$$g_B(x,y) = \frac{1}{\mu_y}\mathbb{E}^x[L(y,\tau_B)] = \frac{1}{\mu_y}\sum_k \mathbb{P}^x(Y_k = y, k <\tau_B). \tag{3.4}$$

Clearly $g_B(x,y) = 0$ when either x or y is outside B. Since $\mu_y^{-1}\mathbb{P}^x(Y_k = y, k <\tau_B) = \mu_x^{-1}\mathbb{P}^y(Y_k = x, k <\tau_B)$, we have

$$g_B(x,y) = g_B(y,x) \qquad \forall x, y \in X. \tag{3.5}$$

Using the strong Markov property of Y,

$$g_B(x,y) = \mathbb{P}^x(\sigma_y < \tau_B) g_B(y,y) \le g_B(y,y). \tag{3.6}$$

Below are further properties of the Green density.

Lemma 3.3 *Let $B \subset X$ be a finite set. Then the following hold.*
(i) For $x \in B$, $g_B(x,\cdot)$ is harmonic on $B \setminus \{x\}$ and $= 0$ outside B.
(ii) (Reproducing property of the Green density) *For all $f \in H^2$ with $\mathrm{Supp}\, f \subset B$, it holds that $\mathcal{E}(g_B(x,\cdot), f) = f(x)$ for each $x \in X$.*
(iii) $E^x[\tau_B] = \sum_{y\in B} g_B(x,y)\mu_y$ for each $x \in X$.
(iv) $R_{\text{eff}}(x, B^c) = g_B(x,x)$ for each $x \in X$.
(v) $E^x[\tau_B] \le R_{\text{eff}}(x, B^c)\mu(B)$ for each $x \in X$.

Proof. (i) Let $\nu(z) = g_B(x,z)$. Then, for each $y \in B \setminus \{x\}$, noting that $Y_0, Y_{\tau_B} \notin B$, we have

$$\nu(y)\mu_y = \mathbb{E}^x[\sum_{i=0}^{\tau_B - 1} 1_y(Y_{i+1})] = \mathbb{E}^x[\sum_{i=0}^{\tau_B-1}\sum_z 1_z(Y_i)P(z,y)] = \sum_z \nu(z)\mu_z\frac{\mu_{zy}}{\mu_z} = \sum_z \nu(z)\mu_{yz}$$

Dividing both sides by μ_y, we have $\nu(y) = \sum_z P(y,z)\nu(z)$, so ν is harmonic on $B \setminus \{x\}$.
(ii) Let $u(y) = \nu(y)\mu_y$. Since ν is harmonic on $B \setminus \{x\}$ and $\mathrm{Supp}\, f \subset B$, noting that we can apply

Lemma 1.5 (ii) because $f \in \mathbb{L}^2$ when B is finite, we have

$$\begin{aligned}
\mathcal{E}(\nu, f) &= -\Delta\nu(x)f(x)\mu_x = \{\nu(x) - \sum_y P(x,y)\nu(y)\}f(x)\mu_x \\
&= \{\nu(x)\mu_x - \sum_y \frac{\mu_{xy}}{\mu_x}\mu_x\nu(y)\}f(x) = \{\nu(x)\mu_x - \sum_y \frac{\mu_{xy}}{\mu_y}\mu_y\nu(y)\}f(x) \\
&= \{u(x) - \sum_y P(y,x)u(y)\}f(x). \qquad (3.7)
\end{aligned}$$

Since $u(y) = E^x[\sum_{k=0}^{\tau_B-1} 1_{\{Y_k=y\}}]$ and $\sum_y 1_{\{Y_k=y\}}P(y,x) = 1_{\{Y_{k+1}=x\}}$, we have

$$\begin{aligned}
u(x) - \sum_y P(y,x)u(y) &= E^x\Big[\sum_{k=0}^{\tau_B-1} 1_{\{Y_k=x\}}\Big] - E^x\Big[\sum_{k=0}^{\tau_B-1}\sum_y 1_{\{Y_k=y\}}P(y,x)\Big] \\
&= 1 + E^x\Big[\sum_{k=1}^{\tau_B-1} 1_{\{Y_k=x\}}\Big] - E^x\Big[\sum_{k=0}^{\tau_B-1} 1_{\{Y_{k+1}=x\}}\Big] \\
&= 1 + E^x\Big[\sum_{k=1}^{\tau_B-1} 1_{\{Y_k=x\}}\Big] - E^x\Big[\sum_{k=1}^{\tau_B} 1_{\{Y_k=x\}}\Big] = 1.
\end{aligned}$$

Putting this into (3.7), we obtain $\mathcal{E}(\nu, f) = f(x)$.
(iii) Multiplying both sides of (3.4) by μ_y and summing over $y \in B$, we obtain the result.
(iv) If $x \notin B$, both sides are 0, so let $x \in B$. Let $p_B^x(z) = g_B(x,z)/g_B(x,x)$. Then, by Proposition 1.13 and (i) above, we see that p_B^x attains the minimum in the definition of the effective resistance. (Note that the assumption that B is finite is used to guarantee the uniqueness of the minimum.) Thus, using (ii) above,

$$R_{\text{eff}}(x, B^c)^{-1} = \mathcal{E}(p_B^x, p_B^x) = \frac{\mathcal{E}(g_B(x,\cdot), g_B(x,\cdot))}{g_B(x,x)^2} = \frac{1}{g_B(x,x)}. \qquad (3.8)$$

(v) If $x \notin B$, both sides are 0, so let $x \in B$. Using (iii), (iv) and (3.6), we have

$$E^x[\tau_B] = \sum_{y\in B} g_B(x,y)\mu_y \le \sum_{y\in B} g_B(x,x)\mu_y = R_{\text{eff}}(x, B^c)\mu(B).$$

We thus obtain the desired inequality. □

Remark 3.4 *As mentioned before Definition 1.3, $-(u(x) - \sum_y P(y,x)u(y))$ in (3.7) is the total flux flowing into x, given the potential ν. So, we see that the total flux flowing out from x is 1 when the potential $g_B(x,\cdot)$ is given at x.*

The next example shows that $R_{\text{eff}}(x, B^c) = g_B(x,x)$ does not hold in general when B is not finite.

Example 3.5 *Consider $\mathbb{Z}^3$ with weight 1 on each nearest neighbor bond. Let p be an additional point and put a bond with weight 1 between the origin of $\mathbb{Z}^3$ and p; $X = \mathbb{Z}^3 \cup \{p\}$ with the above* *mentioned weights is the weighted graph in this example. Let $B = \mathbb{Z}^3$ and let $B_n = B \cap B(0,n)$. By Lemma 3.3(iv), $R_{\rm eff}(0, B_n^c) = g_{B_n}(0,0)$. If we set $c_n^{-1} := R_{\rm eff}^{\mathbb{Z}^3}(0, B(0,n))$, then it is easy to compute $R_{\rm eff}(0, B_n^c) = (1+c_n)^{-1}$. Since the simple random walk on $\mathbb{Z}^3$ is transient, $\lim_{n\to\infty} c_n =: c_0 > 0$. As will be proved in the proof of Lemma 3.9, $g_B(0,0) = \lim_{n\to\infty} g_{B_n}(0,0)$, so $g_B(0,0) = (1+c_0)^{-1}$. On the other hand, it is easy to see $R_{\rm eff}(0, B^c) = 1 > (1+c_0)^{-1}$, so $R_{\rm eff}(0,B^c) > g_B(0,0)$. This also shows that in general the resistance between two sets cannot be approximated by the resistance of finite approximation graphs.*

For any $A \subset X$, and $A_1, A_2 \subset X$ with $A_1 \cap A_2 = \emptyset$, $A \cap A_i = \emptyset$ (for either $i = 1$ or 2) define

$$R_{\rm eff}^A(A_1, A_2)^{-1} = \inf\{\mathcal{E}(f,f) : f \in H^2, f|_{A_1} = 1, f|_{A_2} = 0,\ f \text{ is a constant on } A\}.$$

In other word, $R_{\rm eff}^A(\cdot,\cdot)$ is the effective resistance for the network where the set A is shorted and reduced to one point. Clearly $R_{\rm eff}^A(x, A) = R_{\rm eff}(x,A)$ for $x \in X \setminus A$. We then have the following.

Proposition 3.6 *Let $B \subset X$ be a finite set. Then*

$$g_B(x,y) = \frac{1}{2}(R_{\rm eff}(x,B^c) + R_{\rm eff}(y,B^c) - R_{\rm eff}^{B^c}(x,y)), \qquad \forall x,y \in B.$$

Proof. Since the set B^c is reduced to one point, it is enough to prove this when B^c is a point, say z. Noting that $R_{\rm eff}^{\{z\}}(x,y) = R_{\rm eff}(x,y)$, we will prove the following.

$$g_{X\setminus\{z\}}(x,y) = \frac{1}{2}(R_{\rm eff}(x,z) + R_{\rm eff}(y,z) - R_{\rm eff}(x,y)), \qquad \forall x,y \in B. \tag{3.9}$$

By Lemma 3.3 (iv) and (3.6), we have

$$g_{X\setminus\{z\}}(x,y) = \mathbb{P}^y(\sigma_x < \tau_{X\setminus\{z\}}) g_{X\setminus\{z\}}(x,x) = \mathbb{P}^y(\sigma_x < \sigma_z) R_{\rm eff}(x,z). \tag{3.10}$$

Now $V = \{x,y,z\} \subset X$ and consider the trace of the network to V, and consider the function u on V such that $u(x) = 1$, $u(z) = 0$ and u is harmonic on y. Using the same notation as in the proof of Lemma 3.1 (iv), we have

$$\mathbb{P}^y(\sigma_x < \sigma_z) = u(y) = \frac{R_{xy}^{-1}}{R_{xy}^{-1} + R_{yz}^{-1}} \times 1 + \frac{R_{yz}^{-1}}{R_{xy}^{-1} + R_{yz}^{-1}} \times 0 = \frac{R_{yz}}{R_{xy} + R_{yz}}.$$

Putting this into (3.10) and using (3.2), we have

$$g_B(x,y) = \frac{R_{yz}}{R_{xy}+R_{yz}} \cdot \frac{R_{zx}(R_{xy}+R_{yz})}{R_{xy}+R_{yz}+R_{zx}} = \frac{R_{zx}R_{yz}}{R_{xy}+R_{yz}+R_{zx}}. \tag{3.11}$$

By (3.3), we obtain (3.9). □

Remark 3.7 *Take an additional point p_0 and consider the Δ-Y transform between $V = \{x,y,z\}$ and $W = \{p_0,x,y,z\}$. Namely, $W = \{p_0,x,y,z\}$ is the network such that $\tilde\mu_{p_0x}, \tilde\mu_{p_0y}, \tilde\mu_{p_0z} > 0$ and other weights are 0, and the trace of $(W,\tilde\mu)$ to V is $(V,\hat\mu)$. (See, for example, [79, Lemma 2.1.15].) Then $\frac{R_{zx}R_{yz}}{R_{xy}+R_{yz}+R_{zx}}$ in (3.11) is equal to $\tilde\mu_{p_0z}^{-1}$, so $g_B(x,y) = \tilde\mu_{p_0z}^{-1}$.*

Corollary 3.8 *Let $B \subset X$ be a finite set. Then*

$$|g_B(x,y) - g_B(x,z)| \le R_{\rm eff}(y,z), \qquad \forall x,y,z \in X.$$

Proof. By Proposition 3.6, we have

$$\begin{aligned} |g_B(x,y) - g_B(x,z)| &\le \frac{|R_{\rm eff}(y,B^c) - R_{\rm eff}(z,B^c)| + |R_{\rm eff}^{B^c}(x,y) - R_{\rm eff}^{B^c}(x,z)|}{2} \\ &\le \frac{R_{\rm eff}(y,z) + R_{\rm eff}^{B^c}(y,z)}{2} \le R_{\rm eff}(y,z), \end{aligned}$$

which gives the desired result. □

3.2 Green density on a general set

In this subsection, will discuss the Green density when B can be infinite. Since we do not use results in this subsection later, readers may skip this subsection.

Let $B \subset X$. We define the Green density by (3.4). By Proposition 1.18, $g_B(x,y) < \infty$ for all $x,y \in X$ when $\{Y_n\}$ is transient, whereas $g_X(x,y) = \infty$ for all $x,y \in X$ when $\{Y_n\}$ is recurrent. Since there is nothing interesting when $g_X(x,y) = \infty$, throughout this subsection we will only consider the case

$$B \ne X \text{ when } \{Y_n\} \text{ is recurrent.}$$

Then we can easily see that the process $\{Y_n^B\}$ killed on exiting B is transient, so $g_B(x,y) < \infty$ for all $x,y \in X$. It is easy to see that (3.5), (3.6), and Lemma 3.3 (i), (iii) hold without any change of the proof.

Recall that H_0^2 is the closure of $C_0(X)$ in H^2. We can generalize Lemma 3.3 (ii) as follows.

Lemma 3.9 (Reproducing property of Green density) *For each $x \in X$, $g_B(x,\cdot) \in H_0^2$. Further, for all $f \in H_0^2$ with Supp $f \subset B$, it holds that $\mathcal{E}(g_B(x,\cdot), f) = f(x)$ for each $x \in X$.*

Proof. When B is finite, this is already proved in Lemma 3.3 (ii), so let B be finite (and $B \ne X$ if $\{Y_n\}$ is recurrent). Fix $x_0 \in X$ and let $B_n = B(x_0,n) \cap B$ and write $\nu_n(z) = g_{B_n}(x,z)$. Then $\tau_{B_n} \uparrow \tau_B$ so that $\nu_n(z) \uparrow \nu(z) < \infty$ for all $z \in X$. Using the reproducing property, for $m \le n$, we have

$$\mathcal{E}(\nu_n - \nu_m, \nu_n - \nu_m) = g_{B_n}(x,x) - g_{B_m}(x,x),$$

which implies that $\{\nu_n\}$ is the Cauchy sequence in H^2. It follows that $\nu_n \to \nu$ in H^2 and $\nu \in H_0^2$. Now for each $f \in H_0^2$ with Supp $f \subset B$, choose $f_n \in C_0(X)$ so that Supp $f \subset B_n$ and $f_n \to f$ in H^2. Then, as we proved above, $\mathcal{E}(f_n, \nu_n) = f_n(x)$. Taking $n \to \infty$ and using Lemma 1.2 (i), we obtain $\mathcal{E}(f,\nu) = f(x)$. □

As we see in Example 3.5, $R_{\rm eff}(0,B^c) = \lim_{n\to\infty} R_{\rm eff}(0,(B \cap B(0,n)^c)$ does not hold in general. So we introduce another resistance metric as follows.

$$R_*(x,y) := \sup\Big\{\frac{|u(x)-u(y)|^2}{\mathcal{E}(u,u)} : u \in H_0^2 \oplus 1, \mathcal{E}(u,u) > 0\Big\}, \qquad \forall x,y \in X, x \ne y,$$

where $H_0^2 \oplus 1 = \{f + a : f \in H_0^2, a \in \mathbb{R}\}$. By (3.1), the difference between this and the effective resistance metric is that either supremum is taken over all $H_0^2 \oplus 1$ or over all H^2. Clearly $R_*(x,y) \le R_{\rm eff}(x,y)$.

Here and in the following of this subsection, we will consider $R_*(x, B^c)$, so we assume $B \neq X$. (For $R_*(x, B^c)$, we consider B^c as one point by shorting.) Using $R_*(\cdot,\cdot)$, we can generalize Lemma 3.3 (iv),(v) as follows.

Lemma 3.10 *Let $B \subset X$ be a set such that $B \neq X$. Then the following hold.*
(i) $R_*(x, B^c) = g_B(x,x)$ for each $x \in X$.
(ii) $E^x[\tau_B] \le R_*(x, B^c)\mu(B)$ for each $x \in X$.

Proof. (i) If $x \notin B$, both sides are 0, so let $x \in B$. Let $p_B^x(z) = g_B(x,z)/g_B(x,x)$. Note that we cannot follow the proof Lemma 3.3 (iv) directly because we do not have uniqueness for the solution of the Dirichlet problem in general. Instead, we discuss as follows. Rewriting the definition, we have

$$R_*(x,B^c)^{-1} = \inf\{\mathcal{E}(f,f) : f \in H_{0,x}^2(B)\} \quad \text{where} \quad H_{0,x}^2(B) := \{f \in H_0^2 \oplus 1 : \ \mathrm{Supp}\, f \subset B, f(x) = 1\}. \tag{3.12}$$

Take any $v \in H_{0,x}^2(B)$. (Note that $H_{0,x}^2(B) \subset H_0^2$ since $B \neq X$.) Then, by Lemma 3.9,

$$\mathcal{E}(v - p_B^x, p_B^x) = \frac{\mathcal{E}(v - p_B^x, g_B(x,\cdot))}{g_B(x,x)} = \frac{v(x) - p_B^x(x)}{g_B(x,x)} = 0.$$

So we have

$$\mathcal{E}(v,v) = \mathcal{E}(v - p_B^x, v - p_B^x) + \mathcal{E}(p_B^x, p_B^x) \ge \mathcal{E}(p_B^x, p_B^x),$$

which shows that the infimum in (3.12) is attained by p_B^x. Thus by (3.8) with $R_*(\cdot,\cdot)$ instead of $R_{\rm eff}(\cdot,\cdot)$, we obtain (i).

Given (i), (ii) can be proved exactly in the same way as the proof of Lemma 3.3 (v). □

Remark 3.11 *As mentioned in [78, Section 2], when (X,μ) is transient, one can show that $(\mathcal{E}, H_0^2 \oplus 1)$ is the resistance form on $X \cup \{\Delta\}$, where $\{\Delta\}$ is a new point that can be regarded as a point of infinity. Note that there is no weighted graph $(X \cup \{\Delta\}, \bar{\mu})$ whose associated resistance form is $(\mathcal{E}, H_0^2 \oplus 1)$. Indeed, if there is, then $1_{\{\Delta\}} \in H_0^2 \oplus 1$, which contradicts the fact $1 \notin H_0^2$ (Proposition 1.22).*

Given Lemma 3.10, Proposition 3.6 and Corollary 3.8 holds in general by changing $R_{\rm eff}(\cdot,\cdot)$ to $R_*(\cdot,\cdot)$ without any change of the proof.

Finally, note that by Proposition 1.24, we see that $H^2 = H_0^2 \oplus 1$ if and only if there is no non-constant harmonic functions of finite energy. One can see that it is also equivalent to $R_{\rm eff}(x,y) = R_*(x,y)$ for all $x, y \in X$. (The necessity can be shown by the fact that the resistance metric determines the resistance form: see [79, Section 2.3].)

3.3 General heat kernel estimates

In this subsection, we give general on-diagonal upper and lower heat kernel estimates. Define

$$B(x,r) = \{y \in X : d(x,y) < r\}, \quad V(x,R) = \mu(B(x,R)).$$

For$\Omega \subset X$, let $r(\Omega)$ be the *inradius*, that is

$$r(\Omega) = \max\{r \in \mathbb{N} : \exists x_0 \in \Omega \text{ such that } B(x_0,r) \subset \Omega\}.$$

Lemma 3.12 *Assume that* $\inf_{x,y\in X:x\sim y} \mu_{xy} > 0$.
(i) For any non-void finite set $\Omega \subset X$,

$$\lambda_1(\Omega) \geq \frac{c_1}{r(\Omega)\mu(\Omega)}. \tag{3.13}$$

(ii) Suppose that there exists a strictly increasing function v *on* $\mathbb{N}$ *such that*

$$V(x,r) \geq v(r), \qquad \forall x \in X, \forall r \in \mathbb{N}. \tag{3.14}$$

Then, for any non-void finite set $\Omega \subset X$,

$$\lambda_1(\Omega) \geq \frac{c_1}{v^{-1}(\mu(\Omega))\mu(\Omega)}. \tag{3.15}$$

Proof. (i) Let f be any function with $\mathrm{Supp}\, f \subset \Omega$ normalized as $\|f\|_\infty = 1$. Then $\|f\|_2^2 \leq \mu(\Omega)$. Now consider a point $x_0 \in X$ such that $|f(x_0)| = 1$ and the largest integer n such that $B(x_0,n) \subset \Omega$. Then $n \leq r(\Omega)$ and there exists a sequence $\{x_i\}_{i=0}^n \subset X$ such that $x_i \sim x_{i+1}$ for $i = 0, 1, \cdots, n-1$, $x_j \in \Omega$ for $j = 0, 1, \cdots, n-1$ and $x_n \notin \Omega$. So we have

$$\mathcal{E}(f,f) \geq \frac{1}{2}\sum_{i=0}^{n-1}(f(x_i) - f(x_{i+1}))^2\mu_{x_i x_{i+1}} \geq \frac{c_1}{n}\Big(\sum_{i=0}^{n-1}|f(x_i) - f(x_{i+1})|\Big)^2 \geq \frac{c_1}{n},$$

where the last inequality is due to $\sum_{i=0}^{n-1}|f(x_i) - f(x_{i+1})| \geq |f(x_0) - f(x_n)| = 1$. Combining these,

$$\frac{\mathcal{E}(f,f)}{\|f\|_2^2} \geq \frac{c_1}{n\mu(\Omega)} \geq \frac{c_1}{r(\Omega)\mu(\Omega)}.$$

Taking infimum over all such f, we obtain the result.
(ii) Denote $r = r(\Omega)$. Then, there exists $x_0 \in \Omega$ such that $B(x_0,r) \subset \Omega$, so (3.14) implies $v(r) \leq \mu(\Omega)$. Thus $r \leq v^{-1}(\mu(\Omega))$, and (3.15) follows from (3.13). □

Proposition 3.13 (Upper bound: slow decay) *Assume that* $\inf_{x,y\in X:x\sim y} \mu_{xy} > 0$ *and*

$$V(x,r) \geq c_1 r^D, \qquad \forall x \in X, \forall r \in \mathbb{N}, \tag{3.16}$$

for some $D \geq 1$. *Then the following holds.*

$$\sup_{x\in X} p_t(x,x) \leq c_2 t^{-\frac{D}{D+1}}, \qquad \forall t \geq 1.$$

Proof. By (3.15), we have $\lambda_1(\Omega) \geq c_1\mu(\Omega)^{-1-1/D}$. Thus the result is obtained by Theorem 2.14 by taking $\theta = 2D/(D+1)$. □

Remark 3.14 *We can generalize Proposition 3.13 as follows (see [14]):*
Assume that $\inf_{x,y\in X:x\sim y}\mu_{xy} > 0$ *and (3.14) holds for all* $r \geq r_0$*. Then the following holds.*

$$\sup_{x\in X} p_t(x,x) \leq c_1 m(t), \qquad \forall t \geq r_0^2,$$

where m *is defined by*

$$t - r_0^2 = \int_{v(r_0)}^{1/m(t)} v^{-1}(s)ds.$$

Indeed, by (3.14), we see that (2.6) holds with $\varphi(s)^2 = cv^{-1}(s)s$*. Thus the result can be obtained by applying Remark 2.5 and Remark 2.10. In fact, the above generalized version of Proposition 3.13 also holds for geodetically complete non-compact Riemannian manifolds with bounded geometry (see [14]).*

Below is the table of the slow heat kernel decay $m(t)$*, given the information of the volume growth* $v(r)$*.*

$V(x,r) \geq$	$\exp(cr)$	$c\exp(cr^\alpha)$	cr^D	cr
$\sup_{x\in X} p_t(x,x) \leq$	$ct^{-1}\log t$	$ct^{-1}(\log t)^{1/\alpha}$	$ct^{-D/(D+1)}$	$ct^{-1/2}$

Next we discuss general form of the on-diagonal heat kernel estimate. This lower bound is quite robust, and the argument works as long as there is a Hunt process and the heat kernel exists. Note that the estimate is independent of the upper bound, so in general the two estimates may not coincide.

Proposition 3.15 (Lower bound) *Let* $B \subset X$ *and* $x \in B$*. Then*

$$p_{2t}(x,x) \geq \frac{\mathbb{P}^x(\tau_B > t)^2}{\mu(B)}, \qquad \forall t > 0.$$

Proof. Using the Chapman-Kolmogorov equation and the Schwarz inequality, we have

$$\mathbb{P}^x(\tau_B > t)^2 \leq \mathbb{P}^x(Y_t \in B)^2 = (\int_B p_t(x,y)d\mu(y))^2 \leq \mu(B)\int_B p_t(x,y)^2 d\mu(y) \leq \mu(B)p_{2t}(x,x),$$

which gives the desired inequality. □

3.4 Strongly recurrent case

In this subsection, we will restrict ourselves to the 'strongly recurrent' case and give sufficient conditions for precise on-diagonal upper and lower estimates of the heat kernel. (To be precise, Proposition 3.16 holds for general weighted graphs, but the assumption $F_{R,\lambda}$ given in other propositions holds only for the strongly recurrent case.)

Throughout this subsection, we fix a based point $0 \in X$ and let $D \geq 1, 0 < \alpha \leq 1$. As before $d(\cdot,\cdot)$ is a graph distance, but we do not use the property of the graph distance except in Remark 3.17. In fact all the results in this subsection hold for any metric (not necessarily a geodesic metric) on X without any change of the proof.

Proposition 3.16 *For $n \in \mathbb{N}$, let $f_n(x) = p_n(0,x) + p_{n+1}(0,x)$. Assume that $R_{\rm eff}(0,y) \leq c_* d(0,y)^\alpha$ holds for all $y \in X$. Let $r \in \mathbb{N}$ and $n = 2[r]^{D+\alpha}$. Then*

$$f_n(0) \leq c_1 n^{-\frac{D}{D+\alpha}} \Big(c_* \vee \frac{r^D}{V(0,r)} \Big). \tag{3.17}$$

Especially, if $c_2 r^D \leq V(0,r)$, then $f_n(0) \leq c_3 n^{-D/(D+\alpha)}$.

Proof. First, note that similarly to (1.7), we can easily check that

$$\mathcal{E}(f_n, f_n) = f_{2n}(0) - f_{2n+2}(0). \tag{3.18}$$

Choose $x_* \in B(0,r)$ such that $f_n(x_*) = \min_{x \in B(0,r)} f_n(x)$. Then

$$f_n(x_*) V(0,r) \leq \sum_{x \in B(0,r)} f_n(x)\mu_x \leq \sum_{x \in G} p_n(0,x)\mu_x + \sum_{x \in G} p_{n+1}(0,x)\mu_x \leq 2,$$

so that $f_n(x_*) \leq 2/V(0,r)$. Using Lemma 3.1 (iii), $R_{\rm eff}(0,y) \leq c_* d(0,y)^\alpha$, and (3.18), we have

$$\begin{aligned} f_n(0)^2 &\leq 2\left(f_n(x_*)^2 + |f_n(0) - f_n(x_*)|^2\right) \leq \frac{8}{V(0,r)^2} + 2R_{\rm eff}(0,x_*)\mathcal{E}(f_n,f_n) \\ &\leq \frac{8}{V(0,r)^2} + 2c_* d(0,x_*)^\alpha \mathcal{E}(f_n,f_n) \leq \frac{8}{V(0,r)^2} + 2c_* d(0,x_*)^\alpha (f_{2n}(0) - f_{2n+2}(0)). \end{aligned} \tag{3.19}$$

The spectral decomposition gives that $k \to f_{2k}(0) - f_{2k+2}(0)$ is non-increasing. Thus

$$\begin{aligned} n\left(f_{2n}(0) - f_{2n+2}(0)\right) &\leq (2[n/2]+1)\left(f_{4[n/2]}(0) - f_{4[n/2]+2}(0)\right) \\ &\leq 2\sum_{i=[n/2]}^{2[n/2]} (f_{2i}(0) - f_{2i+2}(0)) \leq 2 f_{2[n/2]}(0). \end{aligned}$$

Since $n = 2[r]^{D+\alpha}$ is even, putting this into (3.19), we have $f_n(0)^2 \leq \frac{8}{V(0,r)^2} + \frac{4c_* r^\alpha f_n(0)}{n}$. Using $a + b \leq 2(a \vee b)$, we have

$$f_n(0) \leq c_1 \Big(\frac{1}{V(0,r)} \vee \frac{c_* r^\alpha}{n}\Big). \tag{3.20}$$

Rewriting, we obtain (3.17). □

Remark 3.17 *(i) Putting $n = 2[r^\alpha V(0,r)]$ in (3.20), we have the following estimate:*

$$f_{2[r^\alpha V(0,r)]}(0) \leq \frac{c_1(1 \vee c_*)}{V(0,r)}. \tag{3.21}$$

(ii) When $\alpha = 1$, using Lemma 3.1(i), we see that the assumption of Proposition 3.16 holds if $\inf_{x,y \in X: x \sim y} \mu_{xy} > 0$. So (3.17) gives another proof of Proposition 3.13.

In the following, we write $B_R = B(0,R)$. Let $\varepsilon_\lambda = (3c_*\lambda)^{-1/\alpha}$. We say the weighted graph (X,μ) satisfies $F_{R,\lambda}$ (or simply say that $F_{R,\lambda}$ holds) if it satisfies the following estimates:

$$V(0,R) \le \lambda R^D,\ R_{\rm eff}(0,z) \le c_* d(0,z)^\alpha,\ \forall z \in B_R,\ R_{\rm eff}(0,B_R^c) \ge \frac{R^\alpha}{\lambda},\ V(0,\varepsilon_\lambda R) \ge \frac{(\varepsilon_\lambda R)^D}{\lambda}. \quad (3.22)$$

Proposition 3.18 *In the following, we fix $R \ge 1$ and $\lambda > 0$.*
(i) If $V(0,R) \le \lambda R^D, R_{\rm eff}(0,z) \le c_* d(0,z)^\alpha$ for all $z \in B_R$, then the following holds.

$$E^x[\tau_{B_R}] \le 2c_*\lambda R^{D+\alpha} \qquad \forall x \in B_R. \quad (3.23)$$

(ii) If $F_{R,\lambda}$ holds, then there exist $c_1 = c_1(c_*), q_0 > 0$ such that the following holds for all $x \in B(0,\varepsilon_\lambda R)$.

$$\begin{aligned} E^x[\tau_{B_R}] &\ge c_1\lambda^{-q_0}R^{D+\alpha}, & (3.24)\\ P^x(\tau_{B_R} > n) &\ge \frac{c_1\lambda^{-q_0}R^{D+\alpha} - n}{2c_*\lambda R^{D+\alpha}} \quad \forall n \ge 0. & (3.25)\end{aligned}$$

Proof. (i) Using Lemma 3.3 (v) and the assumption, we have

$$E^x[\tau_{B_R}] \le R_{\rm eff}(x,B_R^c)\mu(B_R) \le (R_{\rm eff}(0,x) + R_{\rm eff}(0,B_R^c))\mu(B_R) \le 2c_*\lambda R^{D+\alpha}, \qquad \forall x \in B_R. \quad (3.26)$$

(ii) Denote $B' = B(0,\varepsilon_\lambda R)$. By Lemma 1.15 and the assumption, we have for $y \in B'$,

$$\begin{aligned}\mathbb{P}^y(\tau_{B_R} < \sigma_{\{0\}}) &= \mathbb{P}^y(\sigma_{B_R^c} < \sigma_{\{0\}}) \le \frac{R_{\rm eff}(y, B_R^c \cup \{0\})}{R_{\rm eff}(y,B_R^c)} \le \frac{R_{\rm eff}(y,0)}{R_{\rm eff}(0,B_R^c) - R_{\rm eff}(0,y)}\\ &\le \frac{c_* d(y,0)^\alpha}{R_{\rm eff}(0,B_R^c) - c_* d(y,0)^\alpha} \le \frac{R^\alpha/(3\lambda)}{R^\alpha/\lambda - R^\alpha/(3\lambda)} = \frac{1}{2}.\end{aligned}$$

Applying this into (3.6) and using Lemma 3.3 (v), we have for $y \in B'$,

$$g_{B_R}(0,y) = g_{B_R}(0,0)\mathbb{P}^y(\sigma_{\{0\}} < \tau_{B_R}) \ge \frac{1}{2}R_{\rm eff}(0,B_R^c) \ge \frac{R^\alpha}{2\lambda}.$$

Thus,

$$\mathbb{E}^0[\tau_{B_R}] \ge \sum_{y \in B'} g_B(0,y)\mu_y \ge \frac{R^\alpha}{2\lambda}\mu(B') \ge \frac{\varepsilon_\lambda^D R^{D+\alpha}}{2\lambda^2}.$$

Further, for $x \in B'$,

$$\mathbb{E}^x[\tau_{B_R}] \ge \mathbb{P}^x(\sigma_{\{0\}} < \sigma_{B_R})\mathbb{E}^0[\tau_{B_R}] \ge \frac{1}{2}\mathbb{E}^0[\tau_{B_R}] \ge \frac{\varepsilon_\lambda^D R^{D+\alpha}}{4\lambda^2},$$

so (3.24) is obtained.

Next, by (3.23), (3.24), and the Markov property of Y, we have

$$c_1\lambda^{-q_0}R^{D+\alpha} \le E^x[\tau_{B_R}] \le n + \mathbb{E}^x[1_{\{\tau_{B_R} > n\}}\mathbb{E}^{Y_n}[\tau_{B_R}]] \le n + 2c_*\lambda R^{D+\alpha}\mathbb{P}^x(\tau_{B_R} > n).$$

So (3.25) is obtained. □

Proposition 3.19 *If $F_{R,\lambda}$ holds, then there exist $c_1 = c_1(c_*), q_0, q_1 > 0$ such that the following holds for all $x \in B(0, \varepsilon_\lambda R)$.*

$$p_{2n}(x,x) \geq c_1 \lambda^{-q_1} n^{-D/(D+\alpha)} \quad \textit{for} \quad \frac{c_{3.18.1}}{4\lambda^{q_0}} R^{D+\alpha} \leq n \leq \frac{c_{3.18.1}}{2\lambda^{q_0}} R^{D+\alpha}.$$

Proof. Using Proposition 3.15 and (3.25), we have

$$p_{2n}(x,x) \geq \frac{\mathbb{P}^x(\tau_{B_R} > n)^2}{\mu(B_R)} \geq \frac{(c_{3.18.1}/(2c_* \lambda^{q_0+1}))^2}{\lambda R^D} \geq c_1 \lambda^{-q_1} n^{-D/(D+\alpha)},$$

for some $c_1, q_1 > 0$. □

Remark 3.20 *The above results can be generalized as follows. Let $v, r : \mathbb{N} \to [0, \infty)$ be strictly increasing functions with $v(1) = r(1) = 1$ which satisfy*

$$C_1^{-1}\Big(\frac{R}{R'}\Big)^{d_1} \leq \frac{v(R)}{v(R')} \leq C_1\Big(\frac{R}{R'}\Big)^{d_2}, \quad C_2^{-1}\Big(\frac{R}{R'}\Big)^{\alpha_1} \leq \frac{r(R)}{r(R')} \leq C_2\Big(\frac{R}{R'}\Big)^{\alpha_2}$$

for all $1 \leq R' \leq R < \infty$, where $C_1, C_2 \geq 1$, $1 \leq d_1 \leq d_2$ and $0 < \alpha_1 \leq \alpha_2 \leq 1$. Assume that the following holds instead of (3.22) for a suitably chosen $\varepsilon_\lambda > 0$:

$$V(0,R) \leq \lambda v(R),\ R_{\text{eff}}(0,z) \leq c_* r(d(0,z)),\ \forall z \in B_R,\ R_{\text{eff}}(0,B_R^c) \geq \frac{r(R)}{\lambda},\ V(0,\varepsilon_\lambda R) \geq \frac{v(\varepsilon_\lambda R)}{\lambda}.$$

$x \in B(0, \varepsilon_\lambda R)$. Then, the following estimates hold.

$$\frac{c_1}{\lambda^{q_1} v(\mathcal{I}(n))} \leq p_{2n}(x,x) \leq \frac{c_2}{\lambda^{q_2} v(\mathcal{I}(n))} \qquad \textit{for} \quad \frac{c_0}{4\lambda^{q_0}} v(R) r(R) \leq n \leq \frac{c_0}{2\lambda^{q_0}} v(R) r(R),$$

where $\mathcal{I}(\cdot)$ is the inverse function of $(v \cdot r)(\cdot)$ (see [87] for details).

3.5 Applications to fractal graphs

In this subsection, we will apply the estimates obtained in the previous subsection to fractal graphs.

2-dimensional pre-Sierpinski gasket

Let V_0 be the vertices of the pre-Sierpinski gasket (Figure 1) and define $V_{-n} = 2^n V_0$. Let $a_n = (2^n, 0), b_n = (2^{n-1}, 2^{n-1}\sqrt{3})$ be the vertices in V_{-n}.

Lemma 3.21 *It holds that $R_{\text{eff}}(0, \{a_n, b_n\}) = \frac{1}{2}\left(\frac{5}{3}\right)^n$.*

Proof. Let $p_n = \mathbb{P}^0(\sigma_{\{a_n,b_n\}} < \sigma_{\{0\}}^+)$. Let $z = (3/2, \sqrt{3}/2)$ and define $q_1 = \mathbb{P}^z(\sigma_{\{a_1,b_1\}} < \sigma_{\{0\}}^+)$. Then, by the Markov property, $p_1 = 4^{-1}(p_1 + q_1 + 1)$ and $q_1 = 2^{-1}(p_1 + 1)$. Solving them, we have $p_1 = 3/5$. Next, for a simple random walk $\{Y_k\}$ on the pre-Sierpinski gasket, define the induced random walk $\{Y_k^{(n)}\}$ on V_{-n} as follows.

$$\eta_0 = \min\{k \geq 0 : Y_k \in V_{-n}\}, \quad \eta_i = \min\{k > \eta_{i-1} : Y_k \in V_{-n} \setminus Y_{\eta_{i-1}}\},\ Y_i^{(n)} = Y_{\eta_i}, \quad \text{for } i \in \mathbb{N}.$$

Then it can be easily seen that $\{Y_k^{(n)}\}$ is a simple random walk on V_{-n}. Using this fact, we can inductively show $p_{n+1} = p_n \cdot p_1 = p_n \cdot (3/5) = \cdots = (3/5)^{n+1}$. We thus obtain the result by using Theorem 1.14. □

Let $d_f = \log 3/\log 2,\ d_w = \log 5/\log 2$.

Proposition 3.22 *The following hold for all $R \geq 1$.*
(i) $c_1 R^{d_f} \leq V(0,R) \leq c_2 R^{d_f}$.
(ii) $R_{\rm eff}(0,z) \leq c_3 d(0,z)^{d_w-d_f}, \quad \forall z \in X$.
(iii) $R_{\rm eff}(0,B(0,R)^c) \geq c_4 R^{d_w-d_f}$.

Proof. (i) Since there are 3^n triangles with length 1 in $B(0,2^n)$, we see that $c_1 3^n \leq V(0,2^n) \leq c_2 3^n$. Next, for each R, take $n \in \mathbb{N}$ such that $2^{n-1} \leq R < 2^n$. Then $c_1 3^{n-1} \leq V(0,2^{n-1}) \leq V(0,R) \leq V(0,2^n) \leq c_2 3^n$. Since $3 = 2^{d_f}$, we have the desired estimate.
(ii) We first prove the following:

$$c_3(5/3)^n \leq R_{\rm eff}(0,a_n) \leq c_4(5/3)^n. \tag{3.27}$$

Indeed, by the shorting law and Lemma 3.21, $2^{-1}(5/3)^n = R_{\rm eff}(0,\{a_n,b_n\}) \leq R_{\rm eff}(0,a_n)$. On the other hand, let $\varphi_n^{(1)}$ and $\varphi_n^{(2)}$ be such that

$$\mathcal{E}(\varphi_n^{(1)},\varphi_n^{(1)}) = R_{\rm eff}(0,a_n)^{-1} = R_{\rm eff}(0,b_n)^{-1} = \mathcal{E}(\varphi_n^{(2)},\varphi_n^{(2)}).$$

Then, by symmetry, $\varphi_n^{(1)}(b_n) = \varphi_n^{(2)}(a_n) =: C$. So

$$\begin{aligned}
\frac{1}{2}\Big(\frac{5}{3}\Big)^n = R_{\rm eff}(0,\{a_n,b_n\}) &\geq \mathcal{E}\Big(\frac{\varphi_n^{(1)}+\varphi_n^{(2)}}{1+C},\frac{\varphi_n^{(1)}+\varphi_n^{(2)}}{1+C}\Big)^{-1}\\
&\geq \Big\{\frac{2}{(1+C)^2}\Big(\mathcal{E}(\varphi_n^{(1)},\varphi_n^{(1)})+\mathcal{E}(\varphi_n^{(2)},\varphi_n^{(2)})\Big)\Big\}^{-1} = \frac{(1+C)^2}{4}R_{\rm eff}(0,a_n).
\end{aligned}$$

We thus obtain (3.27). Now for each $z \in X$, choose $n \geq 0$ such that $2^n \leq d(0,z) < 2^{n+1}$. We can then take a sequence $z = z_0, z_1, \cdots, z_n$ such that $z_i \in V_{-i}$, $d(z_i,z_{i+1}) \leq 2^i$ for $i = 0,1,\cdots,n-1$ ($z_i = z_{i+1}$ is allowed) and $d(0,z_n) = 2^n$. Similarly to (3.27), we can show that $R_{\rm eff}(z_i,z_{i+1}) \leq c_5(5/3)^i$. Thus, using the triangle inequality, we have

$$R_{\rm eff}(0,z) \leq \sum_{i=0}^{n-1} R_{\rm eff}(z_i,z_{i+1}) + R_{\rm eff}(0,z_n) \leq c_5\Big(\sum_{i=0}^{n-1}(5/3)^i + (5/3)^n\Big) \leq c_6(5/3)^n \leq c_7 d(0,z)^{d_w-d_f}.$$

So the desired estimate is obtained.
(iii) For each R, take $n \in \mathbb{N}$ such that $2^{n-1} \leq R < 2^n$. Then, by the shorting law and Lemma 3.21,

$$R_{\rm eff}(0,B(0,R)^c) \geq R_{\rm eff}(0,\{a_n,b_n\}) = \frac{1}{2}\Big(\frac{5}{3}\Big)^n = \frac{1}{2}2^{n(d_w-d_f)} \geq c_8 R^{d_w-d_f},$$

so we have the desired estimate. □

By Proposition 3.16, 3.19, 3.22, we see that the following heat kernel estimate holds for simple random walk on the 2-dimensional Sierpinski gasket.

$$c_1 n^{-d_f/d_w} \le p_{2n}(0,0) \le c_2 n^{-d_f/d_w}, \quad \forall n \ge 1. \tag{3.28}$$

Vicsek sets

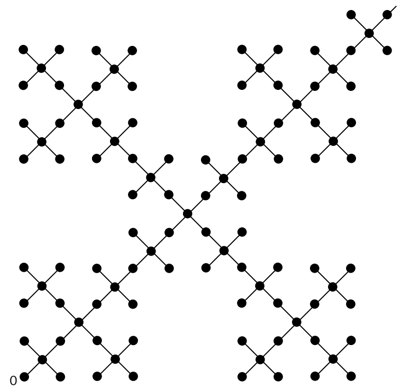


Figure 2: 2-dimensional Vicsek set

We next consider the Vicsek set (Figure 2). Noting that this graph is loop-free, we can apply Lemma 3.1 (ii), and Lemma 3.1 (i) trivially holds. So by a similar proof we can obtain Proposition 3.22 with $d_f = \log 5/\log 3$ and $d_w = d_f + 1$, so (3.28) holds. This shows that the estimate in Proposition 3.13 is in general the best possible estimate when $D = \log 5/\log 3$. By considering the generalization of Vicsek set in $\mathbb{R}^d$, we can see that the same is true when $D = \log(1+2^d)/\log 3$ (which is a sequence that goes to infinity as $d \to \infty$). In fact it is proved in [14, Theorem 5.1] that the estimate in Proposition 3.13 is in general the best possible estimate for any $D \ge 1$.

Remark 3.23 *It is known that the following heat kernel estimate holds for simple random walk on the pre-Sierpinski gasket, on Vicsek sets, and in general on the so-called nested fractal graphs (see [65, 72]):*

$$c_1 n^{-d_f/d_w} \exp\Big(-c_2\Big(\frac{d(x,y)^{d_w}}{n}\Big)^{\frac{1}{d_w-1}}\Big) \le p_n(x,y)+p_{n+1}(x,y) \le c_3 n^{-d_f/d_w} \exp\Big(-c_4\Big(\frac{d(x,y)^{d_w}}{n}\Big)^{\frac{1}{d_w-1}}\Big),$$

for all $x, y \in X$, $n \ge d(x,y)$. (Note that $p_n(x,y) = 0$ if $n < d(x,y)$.)

4 Heat kernel estimates for random weighted graphs

From this section, we consider the situation where we have a random weighted graph $\{(X(\omega), \mu^\omega) : \omega \in \Omega\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that, for each $\omega \in \Omega$, the graph $X(\omega)$ is infinite, locally finite and connected, and contains a marked vertex $0 \in G$. We denote balls in $X(\omega)$ by $B_\omega(x,r)$, their volume by $V_\omega(x,r)$, and write

$$B(R) = B_\omega(R) = B_\omega(0,R), \qquad V(R) = V_\omega(R) = V_\omega(0,R).$$

We write $Y = (Y_n, n \geq 0, P_\omega^x, x \in X(\omega))$ for the Markov chain on $X(\omega)$, and denote by $p_n^\omega(x,y)$ its transition density with respect to μ^ω. Let

$$\tau_R = \tau_{B(0,R)} = \min\{n \geq 0 : Y_n \notin B(0,R)\}.$$

4.1 Random walk on a random graph

Recall that $F_{R,\lambda}$ is defined in (3.22). We have the following quenched estimates.

Theorem 4.1 *Let $R_0, \lambda_0 \geq 1$. Assume that there exist $p(\lambda) \geq 0$ with $p(\lambda) \leq c_1\lambda^{-q_0}$ for some $q_0, c_1 > 0$ such that for each $R \geq R_0$ and $\lambda \geq \lambda_0$,*

$$\mathbb{P}(\{\omega : (X(\omega), \mu^\omega) \text{ satisfies } F_{R,\lambda}\}) \geq 1 - p(\lambda). \tag{4.1}$$

Then there exist $\alpha_1, \alpha_2 > 0$ and $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that the following holds: For all $\omega \in \Omega_0$ and $x \in X(\omega)$, there exist $N_x(\omega), R_x(\omega) \in \mathbb{N}$ such that

$$(\log n)^{-\alpha_1} n^{-\frac{D}{D+\alpha}} \leq p_{2n}^\omega(x,x) \leq (\log n)^{\alpha_1} n^{-\frac{D}{D+\alpha}}, \qquad \forall n \geq N_x(\omega), \tag{4.2}$$

$$(\log R)^{-\alpha_2} R^{D+\alpha} \leq E_\omega^x \tau_R \leq (\log R)^{\alpha_2} R^{D+\alpha}, \qquad \forall R \geq R_x(\omega). \tag{4.3}$$

Further, if (4.1) holds with $p(\lambda) \leq \exp(-c_2\lambda^{q_0})$ for some $q_0, c_2 > 0$, then (4.2), (4.3) hold with $\log\log n$ instead of $\log n$.

Proof. We will take$\Omega_0 = \Omega_1 \cap \Omega_2$ whereΩ_1 andΩ_2 are defined below.

First we prove (4.2). Write $w(n) = p_{2n}^\omega(0,0)$. By Proposition 3.16 and 3.19, we have, taking $n = [c_1(\lambda)R^{D+\alpha}]$,

$$\mathbb{P}((c_1\lambda^{q_1})^{-1} \leq n^{D/(D+\alpha)} w(n) \leq c_1\lambda^{q_1}) \geq 1 - 2p(\lambda). \tag{4.4}$$

Let $n_k = [e^k]$ and $\lambda_k = k^{2/q_0}$ (by choosing R suitably). Then, since $\sum p(\lambda_k) < \infty$, by the Borel–Cantelli lemma there exists $K_0(\omega)$ with $\mathbb{P}(K_0 < \infty) = 1$ such that $c_1^{-1}k^{-2q_1/q_0} \leq n_k^{D/(D+\alpha)} w(n_k) \leq c_1 k^{2q_1/q_0}$ for all $k \geq K_0(\omega)$. Let$\Omega_1 = \{K_0 < \infty\}$. For $k \geq K_0$ we therefore have

$$c_2^{-1}(\log n_k)^{-2q_1/q_0} n_k^{-D/(D+\alpha)} \leq w(n_k) \leq c_2(\log n_k)^{2q_1/q_0} n_k^{-D/(D+\alpha)},$$

so that (4.2) holds for the subsequence n_k. The spectral decomposition gives that $p_{2n}^\omega(0,0)$ is monotone decreasing in n. So, if $n > N_0 = e^{K_0} + 1$, let $k \geq K_0$ be such that $n_k \leq n < n_{k+1}$. Then

$$w(n) \leq w(n_k) \leq c_2(\log n_k)^{2q_1/q_0} n_k^{-D/(D+\alpha)} \leq 2e^{D/(D+\alpha)} c_2 (\log n)^{2q_1/q_0} n^{-D/(D+\alpha)}.$$

Similarly $w(n) \geq w(n_{k+1}) \geq c_3 n^{-D/(D+\alpha)} (\log n)^{-2q_1/q_0}$. Taking $q_2 > 2q_1/q_0$, so that the constants c_2, c_3 can be absorbed into the $\log n$ term, we obtain (4.2) for $x = 0$.

If $x, y \in X(\omega)$ and $k = d_\omega(x,y)$, then using the Chapman–Kolmogorov equation

$$p_{2n}^\omega(x,x)(p_k^\omega(x,y)\mu_x(\omega))^2 \leq p_{2n+2k}^\omega(y,y).$$

Let $\omega \in \Omega_1$, $x \in X(\omega)$, write $k = d_\omega(0,x)$, $h^\omega(0,x) = (p^\omega_k(x,0)\mu_x(\omega))^{-2}$, and let $n \geq N_0(\omega) + 2k$. Then

$$p^\omega_{2n}(x,x) \leq h^\omega(0,x) p^\omega_{2n+2k}(0,0) \leq h^\omega(0,x) \frac{(\log(n+k))^{q_2}}{(n+k)^{D/(D+\alpha)}} \leq h^\omega(0,x) \frac{(\log(2n))^{q_2}}{n^{D/(D+\alpha)}} \leq \frac{(\log n)^{1+q_2}}{n^{D/(D+\alpha)}}$$

provided $\log n \geq 2^{q_2} h^\omega(0,x)$. Taking

$$N_x(\omega) = \exp(2^{q_2} h^\omega(0,x)) + 2d_\omega(0,x) + N_0(\omega), \tag{4.5}$$

and $\alpha_1 = 1 + q_2$, this gives the upper bound in (4.2). The lower bound is obtained in the same way.

Next we prove (4.3). Write $F(R) = E^x_\omega \tau_R$. By (3.23) and (3.24), we have

$$\mathbb{P}(c_1 \lambda^{-q_0} R^{D+\alpha} \leq F(R) \leq 2c_2 \lambda R^{D+\alpha}, \quad \forall x \in B(0, \varepsilon_\lambda R)) \geq 1 - p(\lambda). \tag{4.6}$$

Let $R_n = e^n$ and $\lambda_n = n^{2/q_0}$, and let F_n be the event of the left hand side of (4.6) when $R = R_n$, $\lambda = \lambda_n$. Then we have $\mathbb{P}(F_n^c) \leq p(\lambda_n) \leq n^{-2}$, so by Borel–Cantelli, if$\Omega_2 = \liminf F_n$, then $\mathbb{P}(\Omega_2) = 1$. Hence there exist M_0 with $M_0(\omega) < \infty$ onΩ_2, and $c_3, q_3 > 0$ such that for $\omega \in \Omega_2$ and $x \in X(\omega)$,

$$(c_3 \lambda_n^{q_3})^{-1} \leq \frac{F(R_n)}{R_n^{D+\alpha}} \leq c_3 \lambda_n^{q_3}, \tag{4.7}$$

provided $n \geq M_0(\omega)$ and n is also large enough so that $x \in B(\varepsilon_{\lambda_n} R_n)$. Writing $M_x(\omega)$ for the smallest such n,

$$c_3^{-1} (\log R_n)^{-2q_3/q_0} R_n^{D+\alpha} \leq F(R_n) \leq c_3 (\log R_n)^{2q_3/q_0} R_n^{D+\alpha}, \qquad \text{for all } n \geq M_x(\omega).$$

As $F(R)$ is monotone increasing, the same argument as in the proof of (4.2) above enables us to replace $F(R_n)$ by $F(R)$, for all $R \geq R_x = 1 + e^{M_x}$. Taking $\alpha_2 > 2q_3/q_0$, we obtain (4.3).

The case $p(\lambda) \leq \exp(-c_2 \lambda^{q_0})$ can be proved similarly by the following changes; take $\lambda_k = (e + (2/c_2) \log k)^{1/q_0}$ instead of $\lambda_k = k^{2/q_0}$, and take $N_x(\omega) = \exp(\exp(C h^\omega(0,x))) + 2d_\omega(0,x) + N_0(\omega)$ in (4.5). Then, $\log n$ (resp. $\log n_k$, $\log R_n$) in the above proof are changed to $\log\log n$ (resp. $\log\log n_k$, $\log\log R_n$) and the proof goes through. □

We also have the following annealed estimates. Note that there are no log terms in them.

Proposition 4.2 *Suppose* (4.1) *holds for some* $p(\lambda) \geq 0$ *which goes to* 0 *as* $\lambda \to \infty$*, and suppose the following hold,*

$$\mathbb{E}[R_{\mathrm{eff}}(0, B(R)^c) V(R)] \leq c_1 R^{D+\alpha}, \qquad \forall R \geq 1. \tag{4.8}$$

Then

$$c_2 R^{D+\alpha} \leq \mathbb{E}[E^0_\omega \tau_R] \leq c_3 R^{D+\alpha}, \qquad \forall R \geq 1, \tag{4.9}$$

$$c_4 n^{-D/(D+\alpha)} \leq \mathbb{E}[p^\omega_{2n}(0,0)], \qquad \forall n \geq 1. \tag{4.10}$$

Suppose in addition that there exist $c_5 > 0, \lambda_0 > 1$ *and* $q_0' > 2$ *such that*

$$\mathbb{P}(\lambda^{-1}R^D \le V(R),\ R_{\rm eff}(0,y) \le \lambda d(0,y)^\alpha,\ \forall y \in B(R)) \ge 1 - \frac{c_5}{\lambda^{q_0'}}, \tag{4.11}$$

for each $R \ge 1, \lambda \ge \lambda_0$. *Then*

$$\mathbb{E}[p_{2n}^\omega(0,0)] \le c_6 n^{-D/(D+\alpha)}, \qquad \forall n \ge 1. \tag{4.12}$$

Proof. We begin with the upper bounds in (4.9). By (3.26) and (4.8),

$$\mathbb{E}[E_\omega^0 \tau_R] \le \mathbb{E}[R_{\rm eff}(0, B(R)^c)V(R)] \le c_1 R^{D+\alpha}.$$

For the lower bounds, it is sufficient to find a set $F \subset \Omega$ of 'good' graphs with $\mathbb{P}(F) \ge c_2 > 0$ such that, for all $\omega \in F$ we have suitable lower bounds on $E_\omega^0 \tau_R$ or $p_{2n}^\omega(0,0)$. We assume that $R \ge 1$ is large enough so that $\varepsilon_{\lambda_0} R \ge 1$, where λ_0 is chosen large enough so that $p(\lambda_0) < 1/4$. We can then obtain the results for all n (chosen below to depend on R) and R by adjusting the constants c_2, c_4 in (4.9) and (4.10).

Let F be the event of the left hand side of (4.6) when $\lambda = \lambda_0$. Then $\mathbb{P}(F) \ge \frac{3}{4}$, and for $\omega \in F$, $E_\omega^0 \tau_R \ge c_3 \lambda_0^{-q_0} R^{D+\alpha}$. So,

$$\mathbb{E}[E_\cdot^0 \tau_R] \ge \mathbb{E}[E_\cdot^0 \tau_R : F] \ge c_3 \lambda_0^{-q_0} R^{D+\alpha} \mathbb{P}(F) \ge \frac{3c_3}{4} \lambda_0^{-q_0} R^{D+\alpha}.$$

Also, by (4.4), if $n = [c_4(\lambda_0)R^{D+\alpha}]$, then $p_{2n}^\omega(0,0) \ge c_5 \lambda_0^{-q_1} n^{-D/(D+\alpha)}$. So, given $n \in \mathbb{N}$, choose R so that $n = [c_4(\lambda_0)R^{D+\alpha}]$ and let F be the event of the left hand side of (4.4). Then

$$\mathbb{E}p_{2n}^\cdot(0,0) \ge \mathbb{P}(F) c_5 \lambda_0^{-q_1} n^{-D/(D+\alpha)} \ge \frac{3c_5}{4} \lambda_0^{-q_1} n^{-D/(D+\alpha)},$$

giving the lower bound in (4.10).

Finally we prove (4.12). Let H_k be the event of the left hand side of (4.11) with $\lambda = k$. By (3.17), we see that $p_{2n}^\omega(0,0) \le c_6 k n^{-D/(D+\alpha)}$ if $\omega \in H_k$, where R is chosen to satisfy $n = 2[R]^{D+\alpha}$. Since $\mathbb{P}(\cup_k H_k) = 1$, using (4.11), we have

$$\begin{aligned} \mathbb{E}p_{2n}^\omega(0,0) &\le \sum_k c_6(k+1)n^{-D/(D+\alpha)}\mathbb{P}(H_{k+1} \setminus H_k) \le \sum_k c_6(k+1)n^{-D/(D+\alpha)}\mathbb{P}(H_k^c) \\ &\le c_7 n^{-D/(D+\alpha)} \sum_k (k+1)k^{-q_0'} < \infty, \end{aligned}$$

since $q_0' > 2$. We thus obtain (4.12). □

Remark 4.3 *With some extra efforts, one can obtain quenched estimates of* $\max_{0 \le k \le n} d(0, Y_k)$ *and* $\mu(W_n)$ *where* $W_n = \{Y_0, Y_1, \cdots, Y_n\}$ *(range of the random walk), and annealed lower bound of* $E_\omega^0 d(0, Y_n)$. *See [87, (1.23), (1.28), (1.31)] and [18, (1.16), (1.20), (1.23)].*

4.2 The IIC and the Alexander-Orbach conjecture

The problem of random walk on a percolation cluster — the 'ant in the labyrinth' [57] — has received much attention both in the physics and the mathematics literature. From the next section, we will consider random walk on a percolation cluster, so $X(\omega)$ will be a percolation cluster.

Let us first 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 [61]) 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 mentioned below) it is known (see for example [61, 106]) that $\theta(p_c) = 0$, and it is conjectured that this holds for $d \geq 2$. 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. 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 $\mathcal{G}$, called the incipient infinite cluster (IIC for short). This IIC $\mathcal{G} = \mathcal{G}(\omega)$ is our random weighted graph $X(\omega)$.

IIC was constructed when $d = 2$ in [76], 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 [69], using the lace expansion. It is believed that the results there will hold for any $d > 6$. [69] 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. We write $\mathcal{G}_d$ for the IIC in $\mathbb{Z}^d$. It is believed that the global properties of $\mathcal{G}_d$ are the same for all $d > d_c$, both for nearest neighbor and spread-out models. (Here d_c is the *critical dimension* which is 6 for the percolation model.) In [69] it is proved for spread-out models that $\mathcal{G}_d$ has one end – i.e. any two paths from 0 to infinity intersect infinitely often. See [106] for a summary of the high-dimensional results.

Let $Y = \{Y_t^\omega\}_{t\geq 0}$ be the (continuous time) simple random walk on $\mathcal{G}_d = \mathcal{G}_d(\omega)$, and $q_t^\omega(x,y)$ be its heat kernel. Define the *spectral dimension* of $\mathcal{G}_d$ by

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

(if this limit exists). Alexander and Orbach [7] conjectured that, for any $d \geq 2$, $d_s(\mathcal{G}_d) = 4/3$. While it is now thought that this is unlikely to be true for small d (see [71, Section 7.4]), the results on the geometry of $\mathcal{G}_d$ for spread-out models in [69] are consistent with this holding for d above the critical dimension.

Recently, it is proved that the Alexander-Orbach conjecture holds for random walk for the IIC on a tree ([19]), on a high dimensional oriented percolation cluster ([18]), and on a high dimensional percolation cluster ([86]). In all cases, we apply Theorem 4.1, namely we verify (4.1) with $D = 2, \alpha = 1$ to prove the Alexander-Orbach conjecture. We will discuss details in the next two sections.

5 The Alexander-Orbach conjecture holds when two-point functions behave nicely

This section is based on the paper by Kozma-Nachmias ([86]) with some simplification by [101].

5.1 The model and main theorems

We write $x \leftrightarrow y$ if x is connected to y by an open path. We write $x \overset{r}{\leftrightarrow} y$ if there is an open path of length less than or equal to r that connects x and y.

Definition 5.1 *Let (X, E) be a infinite graph.*
(i) A bijection map $f : X \to X$ is called a graph automorphism for X if $\{f(u), f(v)\} \in E$ if and only if $\{u, v\} \in E$. Denote the set of all the automorphisms of X by $\mathrm{Aut}(X)$
(ii) (X, E) is said to be transitive if for any $u, v \in X$, there exists $\phi \in \mathrm{Aut}(X)$ such that $\phi(u) = v$.
(iii) For each $x \in X$, define the stabilizer of x by

$$S(x) = \{\phi \in \mathrm{Aut}(X) : \phi(x) = x\}.$$

(iv) A transitive graph X is unimodular if for each $x, y \in X$,

$$|\{\phi(y) : \phi \in S(x)\}| = |\{\phi(x) : \phi \in S(y)\}|.$$

One of the important property of the unimodular transitive graph is that it satisfies the following

$$\sum_{\substack{x,y,z\in X\\ d(0,z)=K}} \mathbb{P}(x\leftrightarrow 0 \not\leftrightarrow z\leftrightarrow y) = \sum_{\substack{x,y,z\in X\\ d(x,z)=K}} \mathbb{P}(0\leftrightarrow x \not\leftrightarrow z\leftrightarrow y), \tag{5.1}$$

for each $K \in \mathbb{N}$. This can be verified on $\mathbb{Z}^d$ using the property of group translations, and it can be extended to unimodular transitive graphs using the mass transport technique (see for example, [102, (3.14)]).

Let (X, μ) be a unimodular transitive weighted graph that has controlled weights, with weight 1 on each bond, i.e. $\mu_{xy} = 1$ for $x \sim y$. In this section, allowing some abuse of notation, we denote $B(x, r) = \{y \in X : d(x, y) \le r\}$ and set $\partial A = \{x \in A : \exists z \in A^c \text{ such that } z \sim x\}$.

We fix a root $0 \in X$ as before. (Note that since the graph is unimodular and transitive, the results in this section is independent of the choice of $0 \in X$.) Consider a bond percolation on X and let $p_c = p_c(X)$ be its critical probability, namely

$$\begin{aligned} &\mathbb{P}_p(\text{There exists an open } \infty\text{-cluster.}) > 0 \quad \text{for} \quad p > p_c(X),\\ &\mathbb{P}_p(\text{There exists an open } \infty\text{-cluster.}) = 0 \quad \text{for} \quad p < p_c(X). \end{aligned}$$

In this section, we will consider the case $p = p_c$ and denote $\mathbb{P} := \mathbb{P}_{p_c}$. Throughout this section, we assume that the following limit exists (independently on how $d(0, x) \to \infty$).

$$\mathbb{P}_{\mathrm{IIC}}(F) = \lim_{d(0,x)\to\infty} \mathbb{P}(F|0 \leftrightarrow x) \tag{5.2}$$

for any cylinder event F (i.e., an event that depends only on the status of a finite number of edges). We denote the realization of IIC by $\mathcal{G}(\omega)$. For the critical bond percolations on $\mathbb{Z}^d$, this is proved in [69] for d large using the lace expansion.

We consider the following two conditions. The first is the triangle condition by Aizenman-Newman ([3]), which indicates mean-field behavior of the model:

$$\sum_{x,y\in X} \mathbb{P}(0 \leftrightarrow x)\mathbb{P}(x \leftrightarrow y)\mathbb{P}(y \leftrightarrow 0) < \infty. \tag{5.3}$$

Note that the value of the left hand side of (5.3) does not change if 0 is changed to any $v \in X$ because X is unimodular and transitive.

The second is the following condition for two-point functions: There exist $c_0, c_1, c_2 > 0$ and a decreasing function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\psi(r) \le c_0\psi(2r)$ for all $r > 0$ such that

$$c_1\psi(d(0,x)) \le \mathbb{P}(0 \leftrightarrow x) \le c_2\psi(d(0,x)) \qquad \forall x \in X \text{ such that } x, \tag{5.4}$$

where $d(\cdot,\cdot)$ is the original graph distance on X. Because X is transitive, we have $\mathbb{P}(y \leftrightarrow z) = \mathbb{P}(0 \leftrightarrow x)$ when $d(0,x) = d(y,z)$.

Since X is unimodular and transitive, we can deduce the following from the above mentioned two conditions.

Lemma 5.2 *(i) (5.3) implies the following open triangle condition:*

$$\lim_{K\to\infty} \sup_{w:d(0,w)\ge K} \sum_{x,y\in X} \mathbb{P}(0{\leftrightarrow}x)\mathbb{P}(x{\leftrightarrow}y)\mathbb{P}(y{\leftrightarrow}w) = 0\,. \tag{5.5}$$

(ii) (5.5) implies the following estimates: There exist $C_1, C_2, C_3 > 0$ such that

$$\mathbb{P}\big(|\mathcal{C}(0)| > n\big) \le C_1 n^{-1/2}, \qquad \forall n \ge 1, \tag{5.6}$$

$$C_2(p_c - p)^{-1} \le \mathbb{E}[|\mathcal{C}(0)|] \le C_3(p_c - p)^{-1}, \qquad \forall p < p_c. \tag{5.7}$$

where $\mathcal{C}(0)$ is the connected component containing 0.

Proof. (i) This is proved in [83] (see [23, Lemma 2.1] for $\mathbb{Z}^d$).
(ii) (5.5) implies (5.7) (see [3, Proposition 3.2] for $\mathbb{Z}^d$ and [102, Page 291] for general unimodular transitive graphs), and the following for $h > 0$ small (see [23, (1.13)] for $\mathbb{Z}^d$ and [102, Page 292] for general unimodular transitive graphs).

$$c_1 h^{1/2} \le \sum_{j=1}^{\infty} \mathbb{P}(|\mathcal{C}(0)| = j)(1 - e^{-jh}) \le c_2 h^{1/2}.$$

Taking $h = 1/n$, we have $\mathbb{P}\big(|\mathcal{C}(0)| > n\big) \le C_1 n^{-1/2}$. □

For the critical bond percolations on $\mathbb{Z}^d$, (5.4) with $\psi(x) = x^{2-d}$ was obtained by Hara, van der Hofstad and Slade [67] for the spread-out model and $d > 6$, and by Hara [66] for the nearest-neighbor model with $d \geq 19$ using the lace expansion. (They obtained the right asymptotic behavior including the constant.) Given (5.4) with $\psi(x) = x^{2-d}$ for $d > 6$, it is easy to check that (5.3) holds as well.

For each subgraph G of X, write G_p for the result of p-bond percolation. Define $B(x, r; G) = \{z : d_{G_{p_c}}(x, z) \leq r\}$ where $d_{G_{p_c}}(x, z)$ is the length of the shortest path between x and z in G_{p_c} (it is ∞ if there is no such path). Note that $p_c = p_c(X)$. Now define

$$H(r; G) = \{\partial B(0, r; G) \neq \emptyset\}, \qquad \Gamma(r) = \sup_{G \subset X} \mathbb{P}(H(r; G)).$$

Note that

$$\Gamma(r) = \sup_{G \subset X} \mathbb{P}(\{\partial B(v, r; G) \neq \emptyset\}), \quad \mathbb{E}[|B(0, r)|] = \mathbb{E}[|B(v, r)|], \qquad \forall v \in X, \tag{5.8}$$

since X is transitive. The following two propositions play a key role.

Proposition 5.3 *Assume that the triangle condition (5.3) holds. Then there exists a constant $C > 0$ such that the following hold for all $r \geq 1$.*

$$\text{i) } \mathbb{E}[|B(0, r)|] \leq Cr, \qquad \text{ii)} \Gamma(\ r) \leq C/r.$$

Proposition 5.4 *Assume that (5.4) and i), ii) in Proposition 5.3 hold. Then (4.1) in Theorem 4.1 holds for $\mathbb{P}_{\mathrm{IIC}}$ with $p(\lambda) = \lambda^{-1/2}$, $D = 2$ and $\alpha = 1$.*

Consider simple random walk on the IIC and let $p_n^\omega(\cdot, \cdot)$ be its heat kernel. Combining the above two propositions with Theorem 4.1, we can derive the following theorem.

Theorem 5.5 *Assume that (5.3) and (5.4) hold. Then there exist $\alpha_1 > 0$ and $N_0(\omega) \in \mathbb{N}$ with $\mathbb{P}(N_0(\omega) < \infty) = 1$ such that*

$$(\log n)^{-\alpha_1} n^{-\frac{2}{3}} \leq p_{2n}^\omega(0, 0) \leq (\log n)^{\alpha_1} n^{-\frac{2}{3}}, \qquad \forall n \geq N_0(\omega), \quad \mathbb{P}_{\mathrm{IIC}} - a.e.\ \omega.$$

In particular, the Alexander-Orbach conjecture holds for the IIC.

By the above mentioned reason, for the critical bond percolations on $\mathbb{Z}^d$, the Alexander-Orbach conjecture holds for the IIC for the spread-out model with $d > 6$, and for the nearest-neighbor model with $d \geq 19$.

Remark 5.6 (i) *In fact, the existence of the limit in (5.2) is not relevant in the arguments. Indeed, even if the limit does not exist, subsequential limits exist due to compactness, and the above results hold for each limit. So the conclusions of Theorem 5.5 hold for any IIC measure (i.e. any subsequential limit).*
(ii) *The opposite inequalities of Theorem 5.3 (i.e. $\mathbb{E}[|B(0, r)|] \geq C'r$ and $\Gamma(r) \geq C'/r$ for all $r \geq 1$) hold under weaker assumption. See Proposition 5.12.*

In the following subsections, we prove Proposition 5.3 and 5.4.

5.2 Proof of Proposition 5.4

The proof splits into three lemmas.

Lemma 5.7 *Assume that (5.4) and i) in Proposition 5.3 hold. Then there exists a constant $C>0$ such that for $r\geq 1$ and $x\in X$ with $d(0,x)\geq 2r$, we have*

$$\mathbb{P}(|B(0,r)|\geq \lambda r^2|0\leftrightarrow x)\leq C\lambda^{-1},\qquad \forall\lambda\geq 1. \tag{5.9}$$

Proof. It is enough to prove the following for $r\geq 1$ and $x\in X$ with $d(0,x)\geq 2r$:

$$\mathbb{E}[|B(0,r)|\cdot \mathbf{1}_{\{0\leftrightarrow x\}}]\leq c_1 r^2\psi(d(0,x)). \tag{5.10}$$

Indeed, we then have, using (5.4) and (5.10),

$$\mathbb{P}(|B(0,r)|\geq \lambda r^2|0\leftrightarrow x)\leq \frac{\mathbb{E}[|B(0,r)||0\leftrightarrow x]}{\lambda r^2}=\frac{\mathbb{E}[|B(0,r)|\cdot \mathbf{1}_{\{0\leftrightarrow x\}}]}{\lambda r^2\mathbb{P}(0\leftrightarrow x)}\leq \frac{c_1 r^2\psi(d(0,x))}{\lambda r^2 c_2\psi(d(0,x))}\leq c_3\lambda^{-1},$$

which gives (5.9). So we will prove (5.10).

We have

$$\begin{aligned}\mathbb{E}[|B(0,r)|\cdot \mathbf{1}_{\{0\leftrightarrow x\}}] &= \sum_z \mathbb{P}(0\overset{r}{\leftrightarrow}z\,,\,0\leftrightarrow x)\leq \sum_{z,y}\mathbb{P}(\{0\overset{r}{\leftrightarrow}y\}\circ\{y\overset{r}{\leftrightarrow}z\}\circ\{y\leftrightarrow x\})\\ &\leq \sum_{z,y}\mathbb{P}(0\overset{r}{\leftrightarrow}y)\mathbb{P}(y\overset{r}{\leftrightarrow}z)\mathbb{P}(y\leftrightarrow x). \end{aligned}\tag{5.11}$$

Here the first inequality is because, if $\{0\overset{r}{\leftrightarrow}z, 0\leftrightarrow x\}$ occurs, then there must exist some y such that $\{0\overset{r}{\leftrightarrow}y\}$, $\{y\overset{r}{\leftrightarrow}z\}$ and $\{y\leftrightarrow x\}$ occur disjointly. The second inequality uses the BK inequality twice.

For $d(0,y)\leq r$ and $d(0,x)\geq 2r$, we have $d(x,y)\geq d(0,x)-d(0,y)\geq d(0,x)/2$, so that

$$\mathbb{P}(y\leftrightarrow x)\leq c_4\psi(d(x,y))\leq c_4\psi(d(0,x)/2)\leq c_5\psi(d(0,x)).$$

Thus,

$$(\text{RHS of (5.11)})\leq c_5\psi(d(0,x))\sum_{z,y}\mathbb{P}(0\overset{r}{\leftrightarrow}y)\mathbb{P}(y\overset{r}{\leftrightarrow}z)\leq c_5 r\psi(d(0,x))\sum_{z,y}\mathbb{P}(0\overset{r}{\leftrightarrow}y)\leq c_5 r^2\psi(d(0,x)).$$

Here we use i) in Proposition 5.3 to sum, first over z (note that (5.8) is used here) and then over y in the second and the third inequality. We thus obtain (5.10). □

Lemma 5.8 *Assume that (5.4) and ii) in Proposition 5.3 hold. Then there exists a constant $C>0$ such that for $r\geq 1$ and $x\in X$ with $d(0,x)\geq 2r$, we have*

$$\mathbb{P}(|B(0,r)|\leq r^2\lambda^{-1}|0\leftrightarrow x)\leq C\lambda^{-1},\qquad \forall\lambda\geq 1. \tag{5.12}$$

Proof. It is enough to prove the following for $r \geq 1$, $\varepsilon < 1$ and $x \in X$ with $d(0,x) \geq 2r$:

$$\mathbb{P}\Big(|B(0,r)| \leq \varepsilon r^2 \,,\, 0 \leftrightarrow x\Big) \leq c_1 \varepsilon \psi(d(0,x)). \tag{5.13}$$

Indeed, we then have, using (5.4) and (5.13),

$$\mathbb{P}(|B(0,r)| \leq r^2\lambda^{-1}|0 \leftrightarrow x) = \frac{\mathbb{P}(|B(0,r)| \leq r^2\lambda^{-1}, 0 \leftrightarrow x)}{\mathbb{P}(0 \leftrightarrow x)} \leq \frac{c_1\lambda^{-1}\psi(d(0,x))}{c_2\psi(d(0,x))} \leq c_3\lambda^{-1},$$

which gives (5.12). So we will prove (5.13).

If $|B(0,r)| \leq \varepsilon r^2$, there must exist $j \in [r/2, r]$ such that $|\partial B(0,j)| \leq 2\varepsilon r$. We fix the smallest such j. Now, if $0 \leftrightarrow x$, there exists a vertex $y \in \partial B(0,j)$ which is connected to x by a path that does not use any of the vertices in $B(0,j-1)$. We say this "$x \leftrightarrow y$ off $B(0,j-1)$". Let A be a subgraph of X such that $\mathbb{P}(B(0,j) = A) > 0$. It is clear that, for any A and any $y \in \partial A$, $\{y \leftrightarrow x \text{ off } A \setminus \partial A\}$ is independent of $\{B(0,j) = A\}$. Thus,

$$\begin{aligned}\mathbb{P}(0 \leftrightarrow x \mid B(0,j) = A) &\leq \sum_{y \in \partial A} \mathbb{P}(y \leftrightarrow x \text{ off } A \setminus \partial A \mid B(0,j) = A) \\ &= \sum_{y \in \partial A} \mathbb{P}(y \leftrightarrow x \text{ off } A \setminus \partial A) \leq \sum_{y \in \partial A} \mathbb{P}(y \leftrightarrow x) \leq C|\partial A|\psi(d(0,x)),\end{aligned} \tag{5.14}$$

where we used $d(x,y) \geq d(0,x) - d(0,y) \geq d(0,x)/2$ in the last inequality. By the definition of j we have $|\partial A| \leq 2\varepsilon r$ and summing over all A with $\mathbb{P}(B(0,j) = A) > 0$ and $\partial B(0, r/2) \neq \emptyset$ (because $0 \leftrightarrow x$) gives

$$\mathbb{P}(|B(0,r)| \leq \varepsilon r^2, 0 \leftrightarrow x) \leq C\varepsilon r\psi(d(0,x)) \cdot \sum_A \mathbb{P}(B(0,j) = A).$$

Since the events $\{B(0,j) = A_1\}$ and $\{B(0,j) = A_2\}$ are disjoint for $A_1 \neq A_2$, we have

$$\sum_A \mathbb{P}(B(0,j) = A) = \mathbb{P}(\partial B(0,r/2) \neq \emptyset) \leq c/r,$$

where ii) in Proposition 5.3 is used in the last inequality. We thus obtain (5.13). □

Lemma 5.9 *Assume that (5.4) and i), ii) in Proposition 5.3 hold. Then there exists a constant $C > 0$ such that for $r \geq 1$ and $x \in X$ with $d(0,x) \geq 2r$, we have*

$$\mathbb{P}(R_{\text{eff}}(0, \partial B(0,r)) \leq r\lambda^{-1}|0 \leftrightarrow x) \leq C\lambda^{-1/2}, \qquad \forall \lambda \geq 1. \tag{5.15}$$

Proof of Proposition 5.4. Note that $R_{\text{eff}}(0,z) \leq d(0,z)$ holds for all $z \in B(0,R)$, and $|B(0,R)| \leq V(0,R) \leq c_1|B(0,R)|$ for all $R \geq 1$. By Lemma 5.7, 5.8 and 5.9, for each $R \geq 1$ and $x \in X$ with $d(0,x) \geq 2R$, we have

$$\mathbb{P}(\{\omega : B(0,R) \text{ satisfies } F_{R,\lambda}\}|0 \leftrightarrow x) \geq 1 - 3C\lambda^{-1/2} \qquad \forall \lambda \geq 1.$$

Using (5.2), we obtain the desired estimate. □

So, all we need is to prove Lemma 5.9. We first give definition of *lane* introduced in [95] (similar notion was also given in [19]).

Definition 5.10 *(i) An edge e between $\partial B(0,j-1)$ and $\partial B(0,j)$ is called a lane for r if there exists a path with initial edge e from $\partial B(0,j-1)$ to $\partial B(0,r)$ that does not return to $\partial B(0,j-1)$.*
(ii) Let $0<j<r$ and $\lambda \geq 1$. We say that a level j has λ-lanes for r if there exist at least λ edges between $\partial B(0,j-1)$ and $\partial B(0,j)$ which are lanes for r.
(iii) We say that 0 is λ-lane rich for r if more than half of $j \in [r/4, r/2]$ have λ-lanes for r.

Note that if 0 is *not* λ-lane rich for r, then

$$R_{\rm eff}(0, \partial B(0,r)) \geq \frac{r}{8\lambda}. \tag{5.16}$$

Indeed, since 0 is not λ-lane rich, there exist $j_1, j_2, \cdots, j_l \in [r/4, r/2]$, $l \geq r/8$ that do not have λ-lanes. For $j \in [r/4, r/2]$, let

$$J_j = \{e : e \text{ is a lane for } r \text{ that is between } \partial B(0,j-1) \text{ and } \partial B(0,j)\}. \tag{5.17}$$

Then $\{J_{j_k}\}_{k=1}^l$ are disjoint cut-sets separating 0 from $\partial B(0,r)$. Since $|J_{j_k}| \leq \lambda$, by the shorting law we have $R_{\rm eff}(0, \partial B(0,r)) \geq \sum_{k=1}^l |J_{j_k}|^{-1} \geq l/\lambda \geq r/(8\lambda)$, so that (5.16) holds.

Proof of Lemma 5.9. It is enough to prove the following for $r \geq 1$, $\lambda > 1$ and $x \in X$ with $d(0,x) \geq 2r$:

$$\mathbb{P}\Big(R_{\rm eff}(0, \partial B(0,r)) \leq \lambda^{-1} r\,,\, 0 \leftrightarrow x\Big) \leq c_1 \lambda^{-1/2} \psi(d(0,x)). \tag{5.18}$$

Indeed, we then have, using (5.4) and (5.18),

$$\begin{aligned}
\mathbb{P}(R_{\rm eff}(0, \partial B(0,r)) \leq r\lambda^{-1} | 0 \leftrightarrow x) &= \frac{\mathbb{P}(R_{\rm eff}(0, \partial B(0,r)) \leq r\lambda^{-1}, 0 \leftrightarrow x)}{\mathbb{P}(0 \leftrightarrow x)} \\
&\leq \frac{c_1 \lambda^{-1/2} \psi(d(0,x))}{c_2 \psi(d(0,x))} \leq c_3 \lambda^{-1/2},
\end{aligned}$$

which gives (5.15). So we will prove (5.18). The proof consists of two steps.

Step 1 We will prove the following; There exists a constant $C>0$ such that for any $r \geq 1$, for any event E measurable with respect to $B(0,r)$ and for any $x \in X$ with $d(0,x) \geq 2r$,

$$\mathbb{P}(E \cap \{0 \leftrightarrow x\}) \leq C \sqrt{r \mathbb{P}(E)} \psi(d(0,x)). \tag{5.19}$$

We first note that by (5.10), there exists some $j \in [r/2, r]$ such that

$$\mathbb{E}\big[|\partial B(0,j)| \cdot \mathbf{1}_{\{0 \leftrightarrow x\}}\big] \leq C r \psi(d(0,x)).$$

Using this and the Chebyshev inequality, for each $M>0$ we have

$$\begin{aligned}
\mathbb{P}(E \cap \{0 \leftrightarrow x\}) &\leq \mathbb{P}(|\partial B(0,j)| > M,\, 0 \leftrightarrow x) + \mathbb{P}(E \cap \{|\,\partial B(0,j)| \leq M, 0 \leftrightarrow x\}) \\
&\leq \frac{C r \psi(d(0,x))}{M} + \mathbb{P}(E \cap \{|\,\partial B(0,j)| \leq M, 0 \leftrightarrow x\}).
\end{aligned}$$

For the second term, using (5.14) we have, for each A,

$$\mathbb{P}(\{B(0,j)=A\}\cap\{0\leftrightarrow x\})\leq C|\partial A|\psi(d(0,x))\mathbb{P}(B(0,j)=A).$$

Summing over all subgraphs A which satisfy E (measurability of E with respect to $B(0,r)$ is used here) and $|\partial A|\leq M$ gives $\mathbb{P}(E\cap\{|\,\partial B(0,j)|\leq M, 0\leftrightarrow x\})\leq CM\psi(d(0,x))\mathbb{P}(E)$. Thus

$$\mathbb{P}(E\cap\{0\leftrightarrow x\})\leq\frac{Cr\psi(d(0,x))}{M}+CM\psi(d(0,x))\mathbb{P}(E).$$

Taking $M=\sqrt{r/\mathbb{P}(E)}$, we obtain (5.19).

Step 2 For $j\in[r/4,r/2]$, let us condition on $B(0,j)$, take an edge e between $\partial B(0,j-1)$ and $\partial B(0,j)$, and denote the end vertex of e in $\partial B(0,j)$ by v_e. Let G_j be a graph that one gets by removing all the edges with at least one vertex in $B(0,j-1)$. Then, $\{e \text{ is a lane for } r\}\subset\{\partial B(v_e,r/2;G_j)\neq\emptyset\}$ in the graph G_j. By the definition of Γ and ii) in Proposition 5.3 (note that (5.8) is used here), we have

$$\mathbb{P}(\partial B(v_e,r/2;G_j)\neq\emptyset|B(0,j))\leq\Gamma(r/2)\leq C/r.$$

Recall the definition of J_j in (5.17). By the above argument, we obtain

$$\begin{aligned}\mathbb{E}[|J_j|\mid B(0,j)] &= P(\sum_{v_e\in\partial B(0,j)}1_{\{e\text{ is a lane for }r\}}|B(0,j))\\ &\leq P(\sum_{v_e\in\partial B(0,j)}1_{\{\partial B(v_e,r/2;G_j)\neq\emptyset\}}|B(0,j))\leq\frac{C}{r}|\partial B(0,j)|.\end{aligned}$$

This together with i) in Proposition 5.3 implies

$$\begin{aligned}\mathbb{E}[\sum_{j=r/4}^{r/2}|J_j|] &= \sum_{j=r/4}^{r/2}\sum_A\mathbb{E}[|J_j|1_{\{B(0,j)=A\}}]\\ &\leq \frac{C}{r}\sum_{j=r/4}^{r/2}\sum_A|\partial A|\mathbb{P}(B(0,j)=A)=\frac{C}{r}\sum_{j=r/4}^{r/2}\mathbb{E}[|\partial B(0,j)|]\leq\frac{C}{r}\mathbb{E}[|B(0,r)|]\leq C'.\end{aligned}$$

So, $\mathbb{P}(0\text{ is }\lambda\text{-lane rich for }r)\leq C/(\lambda r)$. Combining this with (5.16), we obtain

$$\mathbb{P}\Big(R_{\text{eff}}(0,\partial B(0,r))\leq\lambda^{-1}r\Big)\leq\frac{C}{\lambda r}.$$

(Here $R_{\text{eff}}(0,\partial B(0,r))=\infty$ if $\partial B(0,r)=\emptyset$.) This together with (5.19) in Step 1 implies (5.18). □

5.3 Proof of Proposition 5.3 i)

The original proof of Proposition 5.3 i) by Kozma-Nachmias ([86]) (with some simplification in [82]) used an induction scheme which is new and nice, but it requires several pages. Very recently (in fact, 2 weeks before the deadline of the lecture notes), I learned a nice short proof by Sapozhnikov [101] which we will follow. The proof is also robust in the sense we do not need the graph to be transitive (nor unimodular).

Proposition 5.11 *If there exists $C_1 > 0$ such that $\mathbb{E}_p[|\mathcal{C}(0)|] < C_1(p_c - p)^{-1}$ for all $p < p_c$, then there exists $C_2 > 0$ such that $\mathbb{E}[|B(0,r)|] \le C_2 r$ for all $r \ge 1$.*

Proof. It is sufficient to prove for $r \ge 2/p_c$. For $p < p_c$, we will consider the coupling of percolation with parameter p and p_c as follows. First, each edge is open with probability p_c and closed with $1-p_c$ independently. Then, for each open edge, the edge is kept open with probability p/p_c and gets closed with $1 - p/p_c$ independently. By the construction, for each $r \in \mathbb{N}$, we have

$$\mathbb{P}_p(x \overset{r}{\leftrightarrow} y) \ge \left(\frac{p}{p_c}\right)^r \mathbb{P}_{p_c}(x \overset{r}{\leftrightarrow} y), \qquad \forall x, y \in X.$$

Summing over $y \in X$ and using $\mathbb{P}_p(x \overset{r}{\leftrightarrow} y) \le \mathbb{P}_p(x \leftrightarrow y)$ and the assumption, we have

$$\mathbb{E}[|B(0,r)|] \le \left(\frac{p}{p_c}\right)^r \mathbb{E}_p[|\mathcal{C}(0)|] \le \left(\frac{p}{p_c}\right)^r (p_c - p)^{-1}.$$

Taking $p = p_c - 1/r$, we obtain the result. □

Using Lemma 5.2, we see that (5.3) implies (5.7) for unimodular transitive graphs. So the proof of Proposition 5.3 i) is completed.

As mentioned in Remark 5.6, the opposite inequalities of Proposition 5.3 hold under weaker assumption. Let X be a connected, locally finite graph with $0 \in X$.

Proposition 5.12
(i) If X is transitive, then there exists $c_1 > 0$ such that $\mathbb{E}[|B(0,r)|] \ge c_1 r$ for all $r \ge 1$.
(ii) If X is unimodular transitive and satisfies (5.3), then there exists $c_2 > 0$ such that $\Gamma(r) \le c_2/r$ for all $r \ge 1$.

Proof. (i) It is enough to prove $\mathbb{E}[|\{x : d(0,x) = r\}|] \ge 1$ for $r \ge 1$. Assume by contradiction that $\mathbb{E}(|\{x : d(0,x) = r_0\}|) \le 1 - c$ for some $r_0 \in \mathbb{N}$ and $c > 0$. Let $G(r) = \mathbb{E}(|\{x : d(0,x) \le r\}|)$. Then

$$\begin{aligned} G(2r_0) - G(r_0) &= \mathbb{E}[|\{x : d(0,x) \in (r_0, 2r_0]\}|] \\ &\le \mathbb{E}[|\{(y,x) : d(0,y) = r_0, d(y,x) \in (1, r_0] \text{ and } y \text{ is on a geodesic from 0 to } x\}|] \\ &\le \mathbb{E}[|\{y : d(0,y) = r_0\}|] \cdot G(r_0) \le (1-c)G(r_0). \end{aligned}$$

where we used Reimer's inequality and the transitivity of X in the second inequality. (Note that we can use Reimer's inequality here because $H := \{y : d(0,y) = r_0\}$ can be verified by examining all the open edges of $B(0,r_0)$ and the closed edges in its boundary, while $\{x : d(y,x) \le r_0\}$ for $y \in H$ can be verified using open edges outside $B(0,r_0)$.) Similarly $G(nr_0) \le (1-c)G(n-1,r_0) + G(r_0)$. A simple calculation shows that this implies that $G(nr_0) \not\to \infty$ as $n \to \infty$, which contradicts criticality.
(ii) We use a second moment argument. By (i) and Proposition 5.3 i), we have $\mathbb{E}|B(0,\lambda r)| \ge c_1 \lambda r$ and $\mathbb{E}|B(0,r)| \le c_2 r$ for each $r \ge 1$, $\lambda \ge 1$. Putting $\lambda = 2c_2/c_1$, we get

$$\mathbb{E}|B(0,\lambda r) \setminus B(0,r)| \ge c_1 \lambda r - c_2 r = c_2 r\,.$$

Next, noting that $\{0\overset{\lambda r}{\leftrightarrow}x, 0\overset{\lambda r}{\leftrightarrow}y\} \subset \{0\overset{\lambda r}{\leftrightarrow}z\}\circ\{z\overset{\lambda r}{\leftrightarrow}x\}\circ\{z\overset{\lambda r}{\leftrightarrow}y\}$ for some $z \in \mathbb{Z}^d$, the BK inequality gives

$$\mathbb{E}[|B(0,\lambda r)|^2] \le \sum_{x,y,z} \mathbb{P}(0\overset{\lambda r}{\leftrightarrow}z)\mathbb{P}(z\overset{\lambda r}{\leftrightarrow}x)\mathbb{P}(z\overset{\lambda r}{\leftrightarrow}y) = \Big[\sum_{x\in\mathbb{Z}^d} \mathbb{P}(0\overset{\lambda r}{\leftrightarrow}x)\Big]^3 \le c_3 r^3\,,$$

where the last inequality is due to Proposition 5.3 i). The 'inverse Chebyshev' inequality $\mathbb{P}(Z>0) \ge (\mathbb{E}Z)^2/\mathbb{E}Z^2$ valid for any non-negative random variable Z yields that

$$\mathbb{P}\big(|B(0,\lambda r)\setminus B(0,r)|>0\big) \ge \frac{c_2^2 r^2}{c_3 r^3} \ge \frac{c_4}{r}\,,$$

which completes the proof since $\{|B(0,\lambda r)\setminus B(0,r)|>0\} \subset H(r)$. □

Note that the idea behind the above proof of (i) is that if $\mathbb{E}[|\{x : d(0,x) = r\}|] < 1$ then the percolation process is dominated above by a subcritical branching process which has finite mean, and this contradicts the fact that the mean is infinity. This argument works not only for the boundary of balls for the graph distance, but also for the boundary of balls for any reasonable geodesic metric. I learned the above proof of (i), which is based on that of [85, Lemma 3.1], from Kozma and Nachmias ([84]). The proof of (ii) is from that of [86, Theorem 1.3 (ii)].

5.4 Proof of Proposition 5.3 ii)

In order to prove Proposition 5.3 ii), we will only need (5.6). First, note that (5.6) implies the following estimate: There exists $C>0$ such that

$$\mathbb{P}\big(|\mathcal{C}_G(0)|>n\big) \le \frac{C_1}{n^{1/2}} \qquad \forall G\subset X,\ \forall n\ge 1, \tag{5.20}$$

because $|\mathcal{C}(0)| \ge |\mathcal{C}_G(0)|$. Here $\mathcal{C}_G(0)$ is the connected component containing 0 for G_{p_c}, where $p_c = p_c(X)$.

The key idea of the proof of Proposition 5.3 ii) is to make a regeneration argument, which is similar to the one given in the proof of Lemma 5.9.

Proof of Proposition 5.3 ii). Let $A\ge 1$ be a large number that satisfies $3^3A^{2/3}+C_1A^{2/3}\le A$, where C_1 is from (5.20). We will prove that$\Gamma(\ r) \le 3Ar^{-1}$. For this, it suffices to prove

$$\Gamma(3^k) \le \frac{A}{3^k}, \tag{5.21}$$

for all $k\in\mathbb{N}\cup\{0\}$. Indeed, for any r, by choosing k such that $3^{k-1}\le r<3^k$, we have

$$\Gamma(r) \le \Gamma(3^{k-1}) \le \frac{A}{3^{k-1}} < \frac{3A}{r}.$$

We will show (5.21) by induction – it is trivial for $k=0$ since $A\ge 1$. Assume that (5.21) holds for all $j<k$ and we prove it for k. Let $\varepsilon>0$ be a (small) constant to be chosen later. For any $G\subset X$, we have

$$\begin{aligned}\mathbb{P}(H(3^k;G)) &\le \mathbb{P}\Big(\partial B(0,3^k;G)\ne\emptyset,\ |\mathcal{C}_G(0)|\le\varepsilon 9^k\Big) + \mathbb{P}\big(|\mathcal{C}_G(0)|>\varepsilon 9^k\big)\\ &\le \mathbb{P}\Big(\partial B(0,3^k;G)\ne\emptyset,\ |\mathcal{C}_G(0)|\le\varepsilon 9^k\Big) + \frac{C_1}{\sqrt{\varepsilon}3^k}\,, \end{aligned}\tag{5.22}$$

where the last inequality is due to (5.20). We claim that

$$\mathbb{P}\Big(\partial B(0,3^k;G)\neq\emptyset,\ |\mathcal{C}_G(0)|\le \varepsilon 9^k\Big)\le \varepsilon 3^{k+1}(\Gamma(3^{k-1}))^2\,. \tag{5.23}$$

If (5.23) holds, then by (5.22) and the induction hypothesis, we have

$$\mathbb{P}(H(3^k;G))\le \varepsilon 3^{k+1}(\Gamma(3^{k-1}))^2+\frac{C_1}{\sqrt{\varepsilon}3^k}\le \frac{\varepsilon 3^3A^2+C_1\varepsilon^{-1/2}}{3^k}. \tag{5.24}$$

Put $\varepsilon = A^{-4/3}$. Since (5.24) holds for any $G\subset X$, we have

$$\Gamma(3^k)\le \frac{3^3A^{2/3}+C_1A^{2/3}}{3^k}\le \frac{A}{3^k}\,,$$

where the last inequality is by the choice of A. This completes the inductive proof of (5.21).

So, we will prove (5.23). Observe that if $|\mathcal{C}_G(0)|\le\varepsilon 9^k$ then there exists $j\in[3^{k-1},2\cdot 3^{k-1}]$ such that $|\partial B(0,j;G)|\le \varepsilon 3^{k+1}$. We fix the smallest such j. If, in addition, $\partial B(0,3^k;G)\neq\emptyset$ then at least one vertex y of the $\varepsilon 3^{k+1}$ vertices of level j satisfies $\partial B(y,3^{k-1};G_2)\neq\emptyset$, where $G_2\subset G$ is determined from G by removing all edges needed to calculate $B(0,j;G)$. By (5.8) and definition of Γ (with G_2), this has probability $\le\Gamma(3^{k-1})$. Summarizing, we have

$$\mathbb{P}\Big(\partial B(0,3^k;G)\neq\emptyset,\ |\mathcal{C}_G(0)|\le\varepsilon 9^k\ \Big|\ B(0,j;G)\Big)\le \varepsilon 3^{k+1}\Gamma(3^{k-1})\,.$$

We now sum over possible values of $B(0,j;G)$ and get an extra term of $\mathbb{P}(H(3^{k-1};G))$ because we need to reach level 3^{k-1} from v. Since $\mathbb{P}(H(3^{k-1};G))\le\Gamma(3^{k-1})$, we obtain (5.23). □

6 Further results for random walk on IIC

In this section, we will summarize results for random walks on various IICs.

6.1 Random walk on IIC for critical percolation on trees

Let $n_0\ge 2$ and let $\mathbb{B}$ be the n_0-ary homogeneous tree with a root 0. We consider the critical bond percolation on $\mathbb{B}$, i.e. let $\{\eta_e : e \text{ is a bond on } \mathbb{B}\}$ be i.i.d. such that $P(\eta_e=1)=\frac{1}{n_0}$, $P(\eta_e=0)=1-\frac{1}{n_0}$ for each e. Set

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

Let $\mathbb{B}_n=\{x\in\mathbb{B}: d(0,x)=n\}$ where $d(\cdot,\cdot)$ is the graph distance. Then, it is easy to see that $Z_n:=|\mathcal{C}(0)\cap\mathbb{B}_n|$ is a branching process with offspring distribution $Bin(n_0,\frac{1}{n_0})$. Since $E[Z_1]=1$, $\{Z_n\}$ dies out with probability 1, so $\mathcal{C}(0)$ is a finite cluster P-a.s.. In this case, we can construct the incipient infinite cluster easily as follows.

Lemma 6.1 (Kesten [77]) *Let $A\subset\mathbb{B}_{\le k}:=\{x\in\mathbb{B}: d(0,x)\le k\}$. Then*

$$\exists\lim_{n\to\infty}P(\mathcal{C}(0)\cap\mathbb{B}_{\le k}=A|Z_n\neq 0)=|A\cap\mathbb{B}_k|P(\mathcal{C}(0)\cap\mathbb{B}_{\le k}=A)=:\mathbb{P}_0(A).$$

Further, there exists a unique probability measure $\mathbb{P}$ which is an extension of $\mathbb{P}_0$ to a probability on the set of ∞-connected subsets of $\mathbb{B}$ containing 0.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space given above and for each $\omega \in \Omega$, let $\mathcal{G}(\omega)$ be the rooted labeled tree with distribution $\mathbb{P}$. So, $(\Omega, \mathcal{F}, \mathbb{P})$ governs the randomness of the media and for each $\omega \in \Omega$, $\mathcal{G}(\omega)$ is the incipient infinite cluster (IIC) on $\mathbb{B}$.

For each $\mathcal{G} = \mathcal{G}(\omega)$, let $\{Y_n\}$ be a simple random walk on $\mathcal{G}$. Let μ be a measure on $\mathcal{G}$ given by $\mu(A) = \sum_{x\in\mathcal{G}} \mu_x$, where μ_x is the number of open bonds connected to $x \in \mathcal{G}$. Define $p_n^\omega(x,y) := \mathbb{P}^x(Y_n = y)/\mu_y$.

In this example, (4.1) in Theorem 4.1 holds for $\mathbb{P}$ with $p(\lambda) = \exp(-c\lambda)$, $D = 2$ and $\alpha = 1$, so we can obtain the following (6.1). In [19], further results are obtained for this example. Let $\mathbb{P}_x(\cdot) := \mathbb{P}(\cdot | x \in \mathcal{G})$.

Theorem 6.2 *(i) There exist $c_0, c_1, c_2, \alpha_1 > 0$ and a positive random variable $S(x)$ with $\mathbb{P}_x(S(x) \geq m) \leq \frac{c_0}{\log m}$ for all $x \in \mathbb{B}$ such that the following holds*

$$c_1 n^{-2/3}(\log\log n)^{-\alpha_1} \leq p_{2n}^\omega(x,x) \leq c_2 n^{-2/3}(\log\log n)^{\alpha_1} \qquad \text{for all} \quad n \geq S(x), x \in \mathbb{B}. \tag{6.1}$$

(ii) There exists $C \in (0,\infty)$ such that for each $\varepsilon \in (0,1)$, the following holds for $\mathbb{P}$-a.e. ω

$$\liminf_{n\to\infty} (\log\log n)^{(1-\varepsilon)/3} n^{2/3} p_{2n}^\omega(0,0) \leq C. \tag{6.2}$$

We will give sketch of the proof in the next subsection.

(6.2) together with (4.10) show that one cannot take $\alpha_1 = 0$ in (6.1). Namely, there is a oscillation of order $\log\log$ for the quenched heat kernel estimates.

Remark 6.3 *(i) For $N \in \mathbb{N}$, let $\widetilde{Z}_n^{(N)} = N^{-1/3} d(0, Y_{Nn})$, $n \geq 0$. In [77], Kesten proved that $\mathbb{P}$-distribution of $\widetilde{Z}_n^{(N)}$ converges as $N \to \infty$. Especially, $\{\widetilde{Z}_\cdot^{(N)}\}$ is tight with respect to the annealed law $\mathbb{P}^* := \mathbb{P} \times P_\omega^0$. On the other hand, by (6.2), it can be shown that $\{\widetilde{Z}_\cdot^{(N)}\}$ is NOT tight with respect to the quenched law.*
(ii) In [45], it is proved that the $\mathbb{P}$-distribution of the rescaled simple random walk on IIC converges to Brownian motion on the Aldous tree ([5]).
(iii) (4.2) is proved with $D = 2$ and $\alpha = 1$ for simple random walk on IIC of the family tree for the critical Galton-Watson branching process with finite variance offspring distribution ([54]), and for invasion percolation on regular trees ([8]).
(iv) The behavior of simple random walk on IIC is different for the family tree of the critical Galton-Watson branching process with infinite variance offspring distribution. In [48], it is proved that when the offspring distribution is in the domain of attraction of a stable law with index $\beta \in (1,2)$, then (4.2) holds with $D = \beta/(\beta-1)$ and $\alpha = 1$. In particular, the spectral dimension of the random walk is $2\beta/(2\beta-1)$. It is further proved that there is an oscillation of order log *for the quenched heat kernel estimates. Namely, it is proved that there exists $\alpha_2 > 0$ such that*

$$\liminf_{n\to\infty} n^{\frac{\beta}{2\beta-1}} (\log n)^{\alpha_2} p_{2n}^\omega(0,0) = 0, \ \mathbb{P}-a.e.\ \omega.$$

Note that for this case convergence of $\mathbb{P}$-distribution of $N^{-(\beta-1)/(2\beta-1)} d(0, Y_{Nn})$ as $N \to \infty$ is already proved in [77].

One may wonder if the off-diagonal heat kernel estimates enjoy sub-Gaussian estimates given in Remark 3.23. As we have seen, there is an oscillation already in the on-diagonal estimates for quenched heat kernel, so one cannot expect estimates precisely the same one as in Remark 3.23. However, the following theorem shows that such sub-Gaussian estimates holds with high probability for the quenched heat kernel, and the precise sub-Gaussian estimates holds for annealed heat kernel.

Let $\{Y_t\}_{t\geq 0}$ be the continuous time Markov chain with based measure μ. Set $q_t^\omega(x,y) = \mathbb{P}^x(Y_t = y)/\mu_y$, and let $\mathbb{P}_{x,y}(\cdot) = \mathbb{P}(\cdot|x, y \in \mathcal{G})$, $\mathbb{E}_{x,y}(\cdot) = \mathbb{E}(\cdot|x, y \in \mathcal{G})$. Then the following holds.

Theorem 6.4 ([19]) *(i) Quenched heat kernel bounds:*
a) Let $x, y \in \mathcal{G} = \mathcal{G}(\omega)$, $t > 0$ be such that $N := [\sqrt{d(x,y)^3/t}] \geq 8$ and $t \geq c_1 d(x,y)$. Then, there exists $F_* = F_*(x,y,t)$ with $\mathbb{P}_{x_0,y_0}\big(F_*(x,y,t)\big) \geq 1 - c_2 \exp(-c_3 N)$, such that

$$q_t^\omega(x,y) \leq c_4 t^{-2/3} \exp(-c_5 N), \qquad \text{for all} \quad \omega \in F_*.$$

b) Let $x, y \in \mathcal{G}$ with $x \neq y$, $m \geq 1$ and $\kappa \geq 1$. Then there exists $G_* = G_*(x,y,m,\kappa)$ with $\mathbb{P}_{x_0,y_0}\big(G_*(x,y,m,\kappa)\big) \geq 1 - c_6 \kappa^{-1}$ such that

$$q_{2T}^\omega(x,y) \geq c_7 T^{-2/3} e^{-c_8(\kappa + c_9)m}, \qquad \text{for all} \quad \omega \in G_* \quad \text{where} \quad T = d(x,y)^3 \kappa/m^2.$$

(ii) Annealed heat kernel bounds: Let $x, y \in \mathbb{B}$. Then

$$c_1 t^{-2/3} \exp\big(-c_2(\frac{d(x,y)^3}{t})^{1/2}\big) \leq \mathbb{E}_{x,y} q_t^\cdot(x,y) \leq c_3 t^{-2/3} \exp\big(-c_4(\frac{d(x,y)^3}{t})^{1/2}\big), \tag{6.3}$$

where the upper bound is for $c_5 d(x,y) \leq t$ and the lower bound is for $c_5(d(x,y) \vee 1) \leq t$.

Note that (6.3) coincides the estimate in Remark 3.23 with $d_f = 2, d_w = 3$.

6.2 Sketch of the Proof of Theorem 6.2

For $x \in \mathcal{G}$ and $r \geq 1$, let $M(x,r)$ be the smallest number $m \in \mathbb{N}$ such that there exists $A = \{z_1, \ldots, z_m\}$ with $d(x, z_i) \in [r/4, 3r/4], 1 \leq i \leq m$, so that any path γ from x to $B(x,r)^c$ must pass through the set A. Similarly to (5.16), we have

$$R_{\text{eff}}(0, B(0,R)^c) \geq R/(2M(0,R)). \tag{6.4}$$

So, in order to prove (4.1) with $p(\lambda) = \exp(-c\lambda)$ (which implies Theorem 6.2 (i) due to the last assertion of Theorem 4.1), it is enough to prove the following.

Proposition 6.5 *(i) Let $\lambda > 0$, $r \geq 1$. Then*

$$\mathbb{P}(V(0,r) > \lambda r^2) \leq c_0 \exp(-c_1\lambda), \quad \mathbb{P}(V(0,r) < \lambda r^2) \leq c_2 \exp(-c_3/\sqrt{\lambda}).$$

(ii) Let $r, m \geq 1$. Then

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

Proof. (Sketch) Basically, all the estimates can be obtained through large deviation estimates of the total population size of the critical branching process.

(i) Since $\mathcal{G}$ is a tree, $|B(x,r)| \le V(x,r) \le 2|B(x,r)|$, so we consider $|B(x,r)|$. Let η be the offspring distribution and $p_i := P(\eta = i)$. (Note that $E[\eta] = 1$, $\mathrm{Var}[\eta] < \infty$.) Now define the size biased branching process $\{\tilde{Z}_n\}$ as follows: $\tilde{Z}_0 = 1$, $P(\tilde{Z}_1 = j) = (j+1)p_{j+1}$ for all $j \ge 0$, and $\{\tilde{Z}_n\}_{n\ge 2}$ is the usual branching process with offspring distribution η. Let $\tilde{Y}_n := \sum_{k=0}^n \tilde{Z}_k$. Then,

$$\tilde{Y}_{r/2}[r/2] \overset{(d)}{\le} |B(0,r)| \overset{(d)}{\le} \tilde{Y}_r[r].$$

Here, for random variable ξ, we denote $\xi[n] :\overset{(d)}{=} \sum_{i=1}^n \xi_i$, where $\{\xi_i\}$ are i.i.d. copies of ξ.

Let $\bar{Y}_n := \sum_{k=0}^n Z_k$, i.e. the total population size up to generation n. Then, it is easy to get

$$P(\bar{Y}_n[n] \ge \lambda n^2) \le c\exp(-c'\lambda), \quad P(\bar{Y}_n[n] \le \lambda n^2) \le c\exp(-c'/\sqrt{\lambda}).$$

We can obtain similar estimates for $\tilde{Y}_n[n]$ so (i) holds.

(ii) For $x, y \in \mathcal{G}$, let $\gamma(x,y)$ be the unique geodesic between x and y. Let $D(x)$ be the descendents of x and $D_r(x) = \{y \in D(x) : d(x,y) = r\}$. Let H be the backbone and $b = b_{r/4} \in H$ be the point where $d(0,b) = r/4$. Define

$$A := \cup_{z\in\gamma(0,b)\setminus\{b\}}(D_{r/4}(z) \setminus H), \quad A^* = \{z \in A : D_{r/4}(z) \ne \emptyset\}.$$

Then, any path from 0 to $B(0,r)^c$ must cross $A^* \cup \{b\}$, so that $M(0,r) \le |A^*| + 1$. Define $p_r := P(z \in A^* | z \in A) = P(Z_{r/4} > 0) \le c/r$. Let $\{\kappa_i\}$ be i.i.d. with distribution $\mathrm{Ber}(p_r)$ that are independent of $|A|$. Then we see that

$$|A^*| \overset{(d)}{=} \sum_{i=1}^{|A|} \kappa_i, \quad |A| \overset{(d)}{\le} \tilde{Z}_{r/4}[r/4].$$

Using these, it is easy to obtain $P(|A^*| > m) \le e^{-cm}$, so (ii) holds. □

We next give the key lemma for the proof of Theorem 6.2 (ii)

Lemma 6.6 *For any $\varepsilon \in (0,1)$,*

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

Proof. (Sketch) Let $D(x;z) = \{y \in D(x) : \gamma(x,y) \cap \gamma(x,z) = \{x\}\}$, and define

$$Z_n = |\{x : x \in D(y_i; y_{i+1}),\, d(x,y_i) \le 2^{n-2},\, 2^{n-1} \le i \le 2^{n-1} + 2^{n-2}\}|.$$

Thus Z_n is the number of descendants off the backbone, to level 2^{n-2}, of points y on the backbone between levels 2^{n-1} and $2^{n-1} + 2^{n-2}$. So $\{Z_n\}_n$ are independent, $|B(0,2^n)| \ge Z_n$, and $Z_n \overset{(d)}{=}$

$\tilde{Y}_{2^{n-2}}[2^{n-2}]$. It can be shown that $\tilde{Y}_n[n] \overset{(d)}{\geq} c_1 n^2 \mathrm{Bin}(n, p_1/n)$ for some $p_1 > 0$ so we have, if $a_n = (\log n)^{1-\varepsilon}$,

$$\begin{aligned}
\mathbb{P}_b(|B(0,2^n)| \geq a_n 4^n) &\geq \mathbb{P}_b(Z_n \geq a_n 4^n) \geq P(\tilde{Y}_{2^{n-2}}[2^{n-2}] \geq a_n 4^n) \\
&\geq P(\mathrm{Bin}\,(2^{n-2}, p_1 2^{-n+2}) \geq c_2 a_n) \geq c_3 e^{-c_2 a_n \log a_n} \geq c_3/n.
\end{aligned}$$

As Z_n are independent, the desired estimate follows by the second Borel-Cantelli Lemma. □

Proof of Theorem 6.2 (ii). Let $a_n = V(0,2^n)2^{-2n}, t_n = 2^n V(0,2^n) = a_n 2^{2n}$. Using (3.21) with $\alpha = 1$, we have $f_{rV(0,r)}(0) \leq c/V(0,r)$, so $p^{\omega}_{t_n}(0,0) \leq c/V(0,2^n) = ct_n^{-2/3}/a_n^{1/3}$. By Lemma 6.6, $a_n \geq (\log n)^{1-\varepsilon} \asymp (\log\log t_n)^{1-\varepsilon}$. We thus obtain the result. □

6.3 Random walk on IIC for the critical oriented percolation cluster in $\mathbb{Z}^d$ $(d > 6)$

We first introduce the spread-out oriented percolation model. Let $d > 4$ and $L \geq 1$. We consider an oriented graph with vertices $\mathbb{Z}^d \times \mathbb{Z}_+$ and oriented bonds $\{((x,n),(y,n+1)) : n \geq 0,\ x,y \in \mathbb{Z}^d$ with $0 \leq \|x-y\|_\infty \leq L\}$. We consider bond percolation on the graph. We write $(x,n) \to (y,m)$ if there is a sequence of open oriented bonds that connects (x,n) and (y,m). Let

$$\mathcal{C}(x,n) = \{(y,m) : (x,n) \to (y,m)\}$$

and define $\theta(p) = \mathbb{P}_p(|C(0,0)| = \infty)$. Then, there exists $p_c = p_c(d,L) \in (0,1)$ such that $\theta(p) > 0$ for $p > p_c$ and $\theta(p) = 0$ for $p \leq p_c$. In particular, there is no infinite cluster when $p = p_c$ (see [61, Page 369], [62]).

For this example, the construction of incipient infinite cluster is given by van der Hofstad, den Hollander and Slade [68] for $d > 4$. Let $\tau_n(x) = \mathbb{P}_{p_c}((0,0) \to (x,n))$ and $\tau_n = \sum_x \tau_n(x)$.

Proposition 6.7 ([68]) *(i) There exists $L_0(d)$ such that if $L \geq L_0(d)$, then*

$$\exists \lim_{n\to\infty} \mathbb{P}_{p_c}(E|(0,0) \to n) =: \mathbb{Q}_{\mathrm{IIC}}(E) \qquad \textit{for any cylindrical events } E,$$

where $(0,0) \to n$ means $(0,0) \to (x,n)$ for some $x \in \mathbb{Z}^d$. Further, $\mathbb{Q}_{\mathrm{IIC}}$ can be extended uniquely to a probability measure on the Borel σ-field and $\mathcal{C}(0,0)$ is $\mathbb{Q}_{\mathrm{IIC}}$–a.s. an infinite cluster.
(ii) There exists $L_0(d)$ such that if $L \geq L_0(d)$, then

$$\exists \lim_{n\to\infty} \frac{1}{\tau_n} \sum_x \mathbb{P}_{p_c}(E \cap \{(0,0) \to (x,n)\}) =: \mathbb{P}_{\mathrm{IIC}}(E) \qquad \textit{for any cylindrical events } E.$$

Further, $\mathbb{P}_{\mathrm{IIC}}$ can be extended uniquely to a probability measure on the Borel σ-field and $\mathbb{P}_{\mathrm{IIC}} = \mathbb{Q}_{\mathrm{IIC}}$.

Let us consider simple random walk on the IIC. It is proved in [18] that (4.1) in Theorem 4.1 holds for $\mathbb{P}_{\mathrm{IIC}}$ with $p(\lambda) = \lambda^{-1}$, $D = 2$ and $\alpha = 1$, so we have (4.2).

Note that although many of the arguments in Section 5 can be generalized to this oriented model straightforwardly, some are not, due to the orientedness. For example, it is not clear how to adapt

Proposition 5.3 ii) to the oriented model. (The definition of$\Gamma(\,r)$ needs to be modified in order to take orientedness into account.) For the reference, we will briefly explain how the proof of (4.1) goes in [18], and explain why $d > 6$ is needed.

Volume estimates Let $Z_R := cV(0,R)/R^2$ and $Z = \int_0^1 dt W_t(\mathbb{R}^d)$, where W_t is the canonical measure of super-Brownian motion conditioned to survive for all time.

Proposition 6.8 *The following holds for $d > 4$ and $L \geq L_0(d)$.*
(i) $\lim_{R\to\infty} \mathbb{E}_{\rm IIC} Z_R^l = \mathbb{E}Z^l \leq 2^{-l}(l+1)!$ for all $l \in \mathbb{N}$. In particular, $c_1R^2 \leq \mathbb{E}_{\rm IIC}V(0,R) \leq c_2R^2$ for all $R \geq 1$.
(ii) $\mathbb{P}_{\rm IIC}(V(0,R)/R^2 < \lambda) \leq c_1 \exp\{-c_2/\sqrt{\lambda}\}$ for all $R,\lambda \geq 1$.

Proof. (Sketch) (i) First, note that for $l \geq 1$ and R large,

$$\mathbb{E}_{\rm IIC} Z_R^l \sim \mathbb{E}_{\rm IIC}[(c'R^{-2}|B(0,R)|)^l] = (c'R^{-2})^l \mathbb{E}_{\rm IIC}[(\sum_{n=0}^{R-1}\sum_{y\in\mathbb{Z}^d} I_{\{(0,0)\to(y,n)\}})^l]. \tag{6.5}$$

For $l \geq 1$ and $\mathbf{m} = (m_1, \cdots, m_l)$, define the IIC $(l+1)$-points function as

$$\hat{\rho}_{\mathbf{m}}^{(l+1)} = \sum_{y_1,\cdots,y_l\in\mathbb{Z}^d} \mathbb{P}_{\rm IIC}((0,0) \to (y_i, m_i),\ \forall i = 1, \cdots, l).$$

In [68], it is proved that for $\mathbf{t} = (t_1, \cdots, t_l) \in (0,1]^l$,

$$\lim_{s\to\infty} (\frac{c'}{s})^l \hat{\rho}_{s\mathbf{t}}^{(l+1)} = \hat{M}_{1,\mathbf{t}}^{(l+1)} := \mathbb{N}(X_1(\mathbb{R}^d), X_{t_1}(\mathbb{R}^d), \cdots, X_{t_l}(\mathbb{R}^d)),$$

where $\hat{M}_{1,\mathbf{t}}^{(l+1)}$ is the $(l+1)$-st moment of the canonical measure $\mathbb{N}$ of super-BM X_t. So taking $R \to \infty$

$$\begin{aligned}(\text{RHS of (6.5)}) &= \frac{c'^l}{R^{2l}}\sum_{n_1=0}^{R-1}\cdots\sum_{n_l=0}^{R-1}\hat{\rho}_{n_1,\cdots,n_l}^{(l+1)} = \frac{1}{R}\sum_{n_1=0}^{R-1}\cdots\frac{1}{R}\sum_{n_l=0}^{R-1}\frac{c'^l}{R^l}\hat{\rho}_{R\vec{t}}^{(l+1)} \\ &\to \int_0^1 dt_1\cdots\int_0^1 dt_l \hat{M}_{1,\mathbf{t}}^{(l+1)} = \mathbb{E}Z^l \quad \text{where } \vec{t} = (n_1R^{-1}, \cdots n_lR^{-1}).\end{aligned}$$

(ii) By (i), we see that for any $\epsilon > 0$, there exists $\delta > 0$ such that $\mathbb{Q}_{\rm IIC}(V(0,R)/R^2 < \delta) < \epsilon$. Now the chaining argument gives the desired estimate. □

From Proposition 6.8, we can verify the volume estimates in (4.1). As we see, we can obtain it for all $d > 4$.

Resistance estimates Following [18], define cut set at level $n \in [1,R]$ as follows:

$$D(n) = \left\{ e = ((w,n-1),(x,n)) \subset \mathcal{G} : \begin{array}{l}(x,n) \text{ is RW-connected to level } R \text{ by a path} \\ \text{in } \mathcal{G} \cap \{(z,l) : l \geq n\}\end{array} \right\}.$$

Here '(x,n) is RW-connected to level R' means there exists a sequence of non-oriented bonds connecting (x,n) to (y,R) for some $y \in \mathbb{Z}^d$.

Proposition 6.9 *For $d > 6$, there exists $L_1(d) \geq L_0(d)$ such that for $L \geq L_1(d)$,*

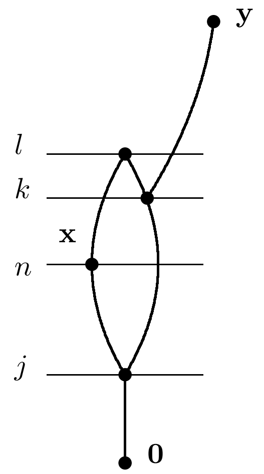
$$\mathbb{E}_{\rm IIC}(|D(n)|) \leq c_1(a,d), \quad 0 < \forall n < aR, \quad 0 < \forall a < 1, \ \forall R \geq 1.$$

Note that similarly to (5.16), we have $R_{\rm eff}(0, \partial B(0,R)) \geq \sum_{n=1}^{R} \frac{1}{|D(n)|}$. So, once Proposition 6.9 is proved, we can prove $\mathbb{Q}_{\rm IIC}(R_{\rm eff}(0, \partial B(R)) \leq \varepsilon R) \leq c\varepsilon$. Thus we can verify the resistance estimates in (4.1) and the proof is complete.

Proof of Proposition 6.9 is quite involved and use $d > 6$. Let us indicate why $d > 6$ is needed.

$$\mathbb{E}_{\rm IIC}|D(n)| = \sum_{w,x \in \mathbb{Z}^d} \mathbb{Q}_{\rm IIC}\left[(\mathbf{w},\mathbf{x}) \in D(n)\right] = \frac{1}{A} \lim_{N\to\infty} \sum_{w,x,y \in \mathbb{Z}^d} \mathbb{P}_{p_c}\left[(\mathbf{w},\mathbf{x}) \in D(n), \, \mathbf{0} \to \mathbf{y}\right],$$

where $\mathbf{w} = (w, n-1), \mathbf{x} = (x, n), \mathbf{y} = (y, N)$ and $A = \lim_{N\to\infty} \tau_N$. Let us consider one configuration in the event $\{(\mathbf{w},\mathbf{x}) \in D(n), \, \mathbf{0} \to \mathbf{y}\}$ which is indicated by the following figure.



By [70], we have $\sup_{x\in\mathbb{Z}^d} \tau_n(x) \leq K\beta(n+1)^{-d/2}$ for $n \geq 1$, and $\tau_n = A(1 + O(n^{(4-d)/2}))$ as $n \to \infty$. Using these and the BK inequality, the configuration in the above figure can be bounded above by

$$c \sum_{l=n}^{\infty} \sum_{k=n}^{l} \sum_{j=0}^{n} (l-j+1)^{-d/2} \leq c \sum_{l=n}^{\infty} \sum_{k=n}^{l} (l-n+1)^{(2-d)/2}$$
$$\leq \ c \sum_{l=n}^{\infty} (l-n+1)^{(4-d)/2} = c \sum_{m=1}^{\infty} m^{(4-d)/2} < \infty \quad \text{for } d > 6.$$

In order to prove Proposition 6.9, one needs to estimate more complicated zigzag paths efficiently.

(Open problem) The critical dimension for oriented percolation is 4. How does the random walk on IIC for oriented percolation behaves when $d = 5, 6$?

6.4 Below critical dimension

We have seen various IIC models where simple random walk on the IIC enjoys the estimate (4.2) with $D = 2$ and $\alpha = 1$. This is a typical mean field behavior that may hold for high dimensions (above the critical dimension). It is natural to ask for the behavior of simple random walk on low dimensions. Very few is known. Let us list up rigorous results that are known so far.

(i) Random walk on the IIC for 2-dimensional critical percolation ([76])

In [76], Kesten shows the existence of IIC for 2-dimensional critical percolation cluster. Further, he proves subdiffusive behavior of simple random walk on IIC in the following sense. Let $\{Y_n\}_n$ be a simple random walk on the IIC, then there exists $\epsilon > 0$ such that the $\mathbb{P}$-distribution of $n^{-\frac{1}{2}+\epsilon}d(0, Y_n)$ is tight.

(ii) Random walk on the 2-dimensional uniform spanning tree ([20])

It is proved that (4.2) holds with $D = 8/5$ and $\alpha = 1$. Especially, it is shown that the spectral dimension of the random walk is 16/13.

(iii) Brownian motion on the critical percolation cluster for the diamond lattice ([64])

Brownian motion is constructed on the critical percolation cluster for the diamond lattice. Further, it is proved that the heat kernel enjoys continuous version of (4.2) with $\alpha = 1$ and some non-trivial D that is determined by the maximum eigenvalue of the matrix for the corresponding multi-dimensional branching process.

It is believed that the critical dimension for percolation is 6. It would be very interesting to know the spectral dimension of simple random walk on IIC for the critical percolation cluster for $d < 6$. (Note that in this case even the existence of IIC is not proved except for $d = 2$.) The following numerical simulations (which we borrow from [27]) suggest that the Alexander-Orbach conjecture does not hold for $d \le 5$.

$$d = 5 \Rightarrow d_s = 1.34 \pm 0.02, \quad d = 4 \Rightarrow d_s = 1.30 \pm 0.04,$$
$$d = 3 \Rightarrow d_s = 1.32 \pm 0.01, \quad d = 2 \Rightarrow d_s = 1.318 \pm 0.001.$$

6.5 Random walk on random walk traces and on the Erdös-Rényi random graphs

In this subsection, we discuss the behavior of random walk in two random environments, namely on random walk traces and on the Erdös-Rényi random graphs. They are not IIC, but the technique discussed in Section 3 and Subsection 4.1 can be applied to some extent. We give a brief overview of the results.

Random walk on random walk traces

Let $X(\omega)$ be the trace of simple random walk on $\mathbb{Z}^d$, $d \ge 3$, started at 0. Let $\{Y_t^\omega\}_{t\ge 0}$ be the simple random walk on $X(\omega)$ and $p_n^\omega(\cdot,\cdot)$ be its heat kernel. It is known in general that if a Markov chain corresponding to a weighted graph is transient, then the simple random walk on the trace of the Markov chain is recurrent $\mathbb{P}$-a.s. (see [28]). The question is to have more detailed properties of the random walk when the initial graph is $\mathbb{Z}^d$. The following results show that it behaves like 1-dimensional simple random walk when $d \ge 5$.

Theorem 6.10 ([44]) *Let $d \ge 5$, and let $\{B_t\}_{t\ge 0}, \{W_t^{(d)}\}_{t\ge 0}$ be independent standard Brownian motions on $\mathbb{R}$ and $\mathbb{R}^d$ respectively, both started at* 0.
(i) There exist $c_1, c_2 > 0$ such that

$$c_1 n^{-1/2} \le p_{2n}^\omega(0,0) \le c_2 n^{-1/2} \quad \textit{for large } n, \quad \mathbb{P}\textit{-a.e. } \omega.$$

(ii) There exists $\sigma_1 = \sigma_1(d) > 0$ such that $\{n^{-1/2} d^\omega(0, Y^\omega_{[tn]})\}_{t\ge 0}$ converges weakly to $\{|B_{\sigma_1 t}|\}_{t\ge 0}$ $\mathbb{P}$-a.e. ω, where $d^\omega(\cdot,\cdot)$ is the graph distance on $X(\omega)$. Also, there exists $\sigma_2 = \sigma_2(d) > 0$ such that $\{n^{-1/4} Y^\omega_{[tn]}\}_{t\ge 0}$ converges weakly to $\{W^{(d)}_{|B_{\sigma_2 t}|}\}_{t\ge 0}$ $\mathbb{P}$-a.e. ω.

On the other hand, the behavior is different for $d = 3, 4$.

Theorem 6.11 ([103, 104])
(i) Let $d = 4$. Then there exist $c_1, c_2 > 0$ and a slowly varying function ψ such that

$$c_1 n^{-\frac{1}{2}} \big(\psi(n)\big)^{\frac{1}{2}} \le p^\omega_{2n}(0,0) \le c_2 n^{-\frac{1}{2}} \big(\psi(n)\big)^{\frac{1}{2}} \qquad \text{for large } n, \quad \mathbb{P}\text{-a.e. } \omega, \tag{6.6}$$
$$n^{\frac{1}{4}} (\log n)^{\frac{1}{24}-\delta} \le \max_{1\le k\le n} |Y^\omega_k| \le n^{\frac{1}{4}} (\log n)^{\frac{13}{12}+\delta} \qquad \text{for large } n, \quad P^0_\omega\text{-a.s. and } \mathbb{P}\text{-a.e. } \omega,$$

for any $\delta > 0$. Further, $\psi(n) \approx (\log n)^{-\frac{1}{2}}$, that is

$$\lim_{n\to\infty} \frac{\log \psi(n)}{\log\log n} = -\frac{1}{2}.$$

(ii) Let $d = 3$. Then there exists $\alpha > 0$ such that

$$p^\omega_{2n}(0,0) \le n^{-\frac{10}{19}} (\log n)^\alpha \quad \text{for large } n, \quad \mathbb{P}\text{-a.s. } \omega.$$

These estimates suggest that the 'critical dimension' for random walk on random walk trace for $\mathbb{Z}^d$ is 4. Note that some annealed estimates for the heat kernel (which are weaker than the above) are obtained in [44] for $d = 4$.

One of the key estimates to establish (6.6) is to obtain a sharp estimate for $\mathbb{E}[R_{\rm eff}(0, S_n)]$ where $\{S_n\}_n$ is the simple random walk on $\mathbb{Z}^4$ started at 0. In [37], Burdzy and Lawler obtained

$$c_1 (\log n)^{-\frac{1}{2}} \le \frac{1}{n} \mathbb{E}[R_{\rm eff}(0, S_n)] \le c_2 (\log n)^{-\frac{1}{3}} \qquad \text{for } d = 4. \tag{6.7}$$

This comes from naive estimates $\mathbb{E}[L_n] \le \mathbb{E}[R_{\rm eff}(0, S_n)] \le \mathbb{E}[A_n]$, where L_n is the number of cut points for $\{S_0, \cdots, S_n\}$, and A_n is the number of points for loop-erased random walk for $\{S_0, \cdots, S_n\}$. (The logarithmic estimates in (6.7) are those of $\mathbb{E}[L_n]$ and $\mathbb{E}[A_n]$.) In [103], Shiraishi proves the following.

$$\frac{1}{n} \mathbb{E}[R_{\rm eff}(0, S_n)] \approx (\log n)^{-\frac{1}{2}} \qquad \text{for } d = 4,$$

which means the exponent that comes from the number of cut points is the right one. Intuitively the reason is as follows. Let $\{T_j\}$ be the sequence of cut times up to time n. Then the random walk trace near S_{T_j} and $S_{T_{j+1}}$ intersects typically when $T_{j+1} - T_j$ is large, i.e. there exists a 'long range intersection'. So $\mathbb{E}[R_{\rm eff}(S_{T_j}, S_{T_{j+1}})] \approx 1$, and $\mathbb{E}[R_{\rm eff}(0, S_n)] \asymp \mathbb{E}[\sum_{j=1}^{a_n} R_{\rm eff}(S_{T_j}, S_{T_{j+1}})] \approx \mathbb{E}[a_n] \asymp n(\log n)^{-1/2}$, where $a_n := \sup\{j : T_j \le n\}$.

Random walk on the Erdös-Rényi random graphs

Let $V_n := \{1, 2, \cdots, n\}$ be labeled vertices. Each bond $\{i, j\}$ $(i, j \in V_n)$ is open with probability $p \in (0, 1)$, independently of all the others. The realization $G(n, p)$ is the Erdös-Rényi random graph.

It is well-known (see [34]) that this model has a phase transition at $p = 1/n$ in the following sense. Let $\mathcal{C}_i^n$ be the i-th largest connected component. If $p \sim c/n$ with $c < 1$ then $|\mathcal{C}_1^n| = O(\log n)$, with $c > 1$ then $|\mathcal{C}_1^n| \asymp n$ and $|\mathcal{C}_j^n| = O(\log n)$ for $j \geq 2$, and if $p \sim c/n$ with $c = 1$ then $|\mathcal{C}_j^n| \asymp n^{2/3}$ for $j \geq 1$.

Now consider the finer scaling $p = 1/n + \lambda n^{-4/3}$ for fixed $\lambda \in \mathbb{R}$ – the so-called critical window. Let $|\mathcal{C}_1^n|$ and S_1^n be the size and the surplus (i.e., the minimum number of edges which would need to be removed in order to obtain a tree) of $\mathcal{C}_1^n$. Then the well-known result by Aldous ([4]) says that $(n^{-2/3}|\mathcal{C}_1^n|, S_1^n)$ converges weakly to some random variables which are determined by a length of the largest excursion of reflected Brownian motion with drift and by some Poisson point process. Recently, Addario-Berry, Broutin and Goldschmidt ([1]) prove further that there exists a (random) compact metric space $\mathcal{M}_1$ such that $n^{-1/3}\mathcal{C}_1^n$ converges weakly to $\mathcal{M}_1$ in the Gromov-Hausdorff sense. (In fact, these results hold not only for $\mathcal{C}_1^n, S_1^n$ but also for the sequences $(\mathcal{C}_1^n, \mathcal{C}_2^n, \cdots)$, $(S_1^n, S_2^n, \cdots)$.) Here $\mathcal{M}_1$ can be constructed from a random real tree (given by the excursion mentioned above) by gluing a (random) finite number of points (chosen according to the Poisson point process) – see [1] for details.

Now we consider simple random walk $\{Y_m^{\mathcal{C}_1^n}\}_m$ on $\mathcal{C}_1^n$. The following results on the scaling limit and the heat kernel estimates are obtained by by Croydon ([43]).

Theorem 6.12 ([43]) *There exists a diffusion process ('Brownian motion') $\{B_t^{\mathcal{M}_1}\}_{t\geq 0}$ on $\mathcal{M}_1$ such that $\{n^{-1/3}Y_{[nt]}^{\mathcal{C}_1^n}\}_{t\geq 0}$ converges weakly to $\{B_t^{\mathcal{M}_1}\}_{t\geq 0}$ $\mathbb{P}$-a.s.. Further, there exists a jointly continuous heat kernel $p_t^{\mathcal{M}_1}(\cdot,\cdot)$ for $\{B_t^{\mathcal{M}_1}\}_{t\geq 0}$ which enjoys the following estimates.*

$$c_1 t^{-2/3}(\ln_1 t^{-1})^{-\alpha_1} \exp\Big(-c_2\Big(\frac{d(x,y)^3}{t}\Big)^{1/2}\Big(\ln_1(\frac{d(x,y)}{t})\Big)^{\alpha_2}\Big) \leq p_t^{\mathcal{M}_1}(x,y)$$
$$\leq c_3 t^{-2/3}(\ln_1 t^{-1})^{1/3} \exp\Big(-c_4\Big(\frac{d(x,y)^3}{t}\Big)^{1/2}\Big(\ln_1(\frac{d(x,y)}{t})\Big)^{-\alpha_3}\Big), \qquad \forall x,y \in \mathcal{M}_1, t \leq 1,$$

where $\ln_1 x := 1 \vee \log x$, and $c_1,\cdots,c_4,\alpha_1,\cdots,\alpha_3$ are positive (non-random) constants.

The same results hold for $\mathcal{C}_i^n$ and $\mathcal{M}_i$ for each $i \in \mathbb{N}$ (with constants depending on i).

7 Random conductance model

Consider weighted graph (X,μ). As we have seen in Remark 1.7, there are two continuous time Markov chains with transition probability $P(x,y) = \mu_{xy}/\mu_x$. One is constant speed random walk (CSRW) for which the holding time is the exponential distribution with mean 1 for each point, and the other is variable speed random walk (VSRW) for which the holding time at x is the exponential distribution with mean μ_x^{-1} for each $x \in X$. The corresponding discrete Laplace operators are given in (1.4), (1.9), which we rewrite here.

$$\mathcal{L}_C f(x) = \frac{1}{\mu_x}\sum_y (f(y)-f(x))\mu_{xy},$$
$$\mathcal{L}_V f(x) = \sum_y (f(y)-f(x))\mu_{xy}.$$

Recall also that for each f, g that have finite support, we have

$$\mathcal{E}(f,g) = -(\mathcal{L}_V f, g)_\nu = -(\mathcal{L}_C f, g)_\mu.$$

7.1 Overview

Consider $\mathbb{Z}^d$, $d \geq 2$ and let E_d be the set of non-oriented nearest neighbor bonds and let the conductance $\{\mu_e : e \in E_d\}$ be stationary and ergodic on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For each $\omega \in \omega$, let $(\{Y_t\}_{t\geq 0}, \{P_\omega^x\}_{x\in\mathbb{Z}^d})$ be either the CSRW or VSRW and define

$$q_t^\omega(x,y) = P_\omega^x(Y_t = y)/\theta_y$$

be the transition density of $\{X_t\}_{t\geq 0}$ where θ is either ν or μ. This model is called the random conductance model (RCM for short). We are interested in the long time behavior of $\{Y_t\}_{t\geq 0}$, especially in obtaining the heat kernel estimates for $q_t^\omega(\cdot,\cdot)$ and a quenched invariance principle (to be precise, quenched functional central limit theorem) for $\{Y_t\}_{t\geq 0}$. Note that when $\mathbb{E}\mu_e < \infty$, it was proved in the 1980s that $\varepsilon Y_{t/\varepsilon^2}$ converges as $\varepsilon \to 0$ to Brownian motion on $\mathbb{R}^d$ with covariance $\sigma^2 I$ in law under $\mathbb{P} \times P_\omega^0$ with the possibility $\sigma = 0$ for some cases (see [50, 80, 81]). This is sometimes referred as the annealed (or averaged) invariance principle.

From now on, we will discuss the case when $\{\mu_e : e \in E_d\}$ are i.i.d.. If $p_+ := \mathbb{P}(\mu_e > 0) < p_c(\mathbb{Z}^d)$ where $p_c(\mathbb{Z}^d)$ is the critical probability for bond percolation on $\mathbb{Z}^d$, then $\{Y_t\}_{t\geq 0}$ is confined to a finite set $\mathbb{P} \times P_\omega^x$-a.s., so we consider the case $p_+ > p_c(\mathbb{Z}^d)$. Under the condition, there exists unique infinite connected components of edges with strictly positive conductances, which we denote by $\mathcal{C}_\infty$. Typically, we will consider the case where $0 \in \mathcal{C}_\infty$, namely we consider $\mathbb{P}(\cdot|0 \in \mathcal{C}_\infty)$.

We will consider the following cases:

- Case 0: $c^{-1} \leq \mu_e \leq c$ for some $c \geq 1$,
- Case 1: $0 \leq \mu_e \leq c$ for some $c > 0$,
- Case 2: $c \leq \mu_e < \infty$ for some $c > 0$.

(Of course, Case 0 is the special case of Case 1 and Case 2.) For Case 0, the following both sides quenched Gaussian heat kernel estimates

$$c_1 t^{-d/2} \exp(-c_2|x-y|^2/t) \leq q_t^\omega(x,y) \leq c_3 t^{-d/2} \exp(-c_4|x-y|^2/t) \tag{7.1}$$

holds $\mathbb{P}$-a.s. for $t \geq |x-y|$ by the result in [49], and the quenched invariance principle is proved in [105]. When $\mu_e \in \{0,1\}$, which is a special case of Case 1, the corresponding Markov chain is a random walk on supercritical percolation clusters. In this case, isoperimetric inequalities are proved in [92] (see also [56]), both sides quenched Gaussian long time heat kernel estimates are obtained in [11] (precisely, (7.1) holds for $1 \vee S_x(\omega) \vee |x-y| \leq t$ where $\{S_x\}_{x\in\mathbb{Z}^d}$ satisfies $\mathbb{P}_p(S_x \geq n, x \in \mathcal{C}(0)) \leq c_1 \exp(-c_2 n^{\varepsilon_d})$ for some $\varepsilon_d > 0$), and the quenched invariance principle is proved in [105] for $d \geq 4$ and later extended to all $d \geq 2$ in [30, 91].

Case 1 This case is treated in [31, 32, 53, 90] for $d \geq 2$. (Note that the papers [31, 32] consider a discrete time random walk and [53, 90] considers CSRW. In fact, one can see that there is not a big difference between CSRW and VSRW in this case, as we will see in Theorem 7.2.)

Heat kernel estimates In [53, 31], it is proved that Gaussian heat kernel bounds do not hold in general and anomalous behavior of the heat kernel is established for d large (see also [36]). In [53], Fontes and Mathieu consider VSRW on $\mathbb{Z}^d$ with conductance given by $\mu_{xy} = \omega(x) \wedge \omega(y)$ where $\{\omega(x) : x \in \mathbb{Z}^d\}$ are i.i.d. with $\omega(x) \leq 1$ for all x and

$$\mathbb{P}(\omega(0) \leq s) \asymp s^\gamma \quad \text{as} \quad s \downarrow 0,$$

for some $\gamma > 0$. They prove the following anomalous annealed heat kernel behavior.

$$\lim_{t\to\infty} \frac{\log \mathbb{E}[P^0_\omega(Y_t = 0)]}{\log t} = -(\frac{d}{2} \wedge \gamma).$$

We now state the main results in [31]. Here we consider discrete time Markov chain with transition probability $\{P(x, y) : x, y \in \mathbb{Z}^d\}$ and denote by $P^n_\omega(0, 0)$ the heat kernel for the Markov chain, which (in this case) coincides with the return probability for the Markov chain started at 0 to 0 at time n.

Theorem 7.1 *(i) For $\mathbb{P}$-a.e. ω, there exists $C_1(\omega) < \infty$ such that for each $n \geq 1$,*

$$P^n_\omega(0, 0) \leq C_1(\omega) \begin{cases} n^{-d/2}, & d = 2, 3, \\ n^{-2} \log n, & d = 4, \\ n^{-2}, & d \geq 5. \end{cases} \tag{7.2}$$

Further, for $d \geq 5$, $\lim_{n\to\infty} n^2 P^n_\omega(0, 0) = 0$ $\mathbb{P}$-a.s.
(ii) Let $d \geq 5$ and $\kappa > 1/d$. There exists an i.i.d. law $\mathbb{P}$ on bounded nearest-neighbor conductances with $p_+ > p_c(d)$ and $C_2(\omega) > 0$ such that for a.e. $\omega \in \{|\mathcal{C}(0)| = \infty\}$,

$$P^n_\omega(0, 0) \geq C_2(\omega) n^{-2} \exp(-(\log n)^\kappa), \quad \forall n \geq 1.$$

(iii) Let $d \geq 5$. For any increasing sequence $\{\lambda_n\}_{n\in\mathbb{N}}$, $\lambda_n \to \infty$, there exists an i.i.d. law $\mathbb{P}$ on bounded nearest-neighbor conductances with $p_+ > p_c(d)$ and $C_3(\omega) > 0$ such that for a.e. $\omega \in \{|\mathcal{C}(0)| = \infty\}$,

$$P^n_\omega(0, 0) \geq C_3(\omega) n^{-2} \lambda_n^{-1}$$

along a subsequence that does not depend on ω.

As we can see, Theorem 7.1 (ii), (iii) shows anomalous behavior of the Markov chain for $d \geq 5$. Once (ii) is proved, (iii) can be proved by a suitable choice of subsequence. We will give a key idea of the proof of (ii) here and give complete proof of it in Subsection 7.3.

Suppose we can show that for large n, there is a box of side length ℓ_n centered at the origin such that in the box a bond with conductance 1 ('strong' bond) is separated from other sites by bonds with conductance $1/n$ ('weak' bonds), and at least one of the 'weak' bonds is connected to the origin by a path of bonds with conductance 1 within the box. Then the probability that the walk is back to the origin at time n is bounded below by the probability that the walk goes directly towards the above place (which costs $e^{O(\ell_n)}$ of probability) then crosses the weak bond (which costs $1/n$), spends time $n - 2\ell_n$ on the strong bond (which costs only $O(1)$ of probability), then crosses a weak bond again (another $1/n$ term) and then goes back to the origin on time (another $e^{O(\ell_n)}$ term). The cost of this strategy is $O(1)e^{O(\ell_n)}n^{-2}$ so if can take $\ell_n = o(\log n)$ then we get leading order n^{-2}.

Quenched invariance principle For $t \geq 0$, let $\{Y_t\}_{t\geq 0}$ be either CSRW or VSRW and define

$$Y_t^{(\varepsilon)} := \varepsilon Y_{t/\varepsilon^2}. \tag{7.3}$$

In [32, 90], they prove the following quenched invariance principle.

Theorem 7.2 *(i) Let $\{Y_t\}_{t\geq 0}$ be the VSRW. Then $\mathbb{P}$-a.s. $Y^{(\varepsilon)}$ converges (under P_ω^0) in law to Brownian motion on $\mathbb{R}^d$ with covariance $\sigma_V^2 I$ where $\sigma_V > 0$ is non-random.*
(ii) Let $\{Y_t\}_{t\geq 0}$ be the CSRW. Then $\mathbb{P}$-a.s. $Y^{(\varepsilon)}$ converges (under P_ω^0) in law to Brownian motion on $\mathbb{R}^d$ with covariance $\sigma_C^2 I$ where $\sigma_C^2 = \sigma_V^2/(2d\mathbb{E}\mu_e)$.

Case 2 This case is treated in [16] for $d \geq 2$.

Heat kernel estimates The following heat kernel estimates for the VSRW is proved in [16]. (We do not give proof here.)

Theorem 7.3 *Let $q_t^\omega(x,y)$ be the heat kernel for the VSRW and let $\eta \in (0,1)$. Then, there exist constants $c_1, \cdots, c_{11} > 0$ (depending on d and the distribution of μ_e) and a family of random variables $\{U_x\}_{x\in\mathbb{Z}^d}$ with*

$$\mathbb{P}(U_x \geq n) \leq c_1 \exp(-c_2 n^\eta),$$

such that the following hold.
(a) For all $x, y \in \mathbb{Z}^d$ and $t > 0$,

$$q_t^\omega(x,y) \leq c_3 t^{-d/2}.$$

(b) For $x, y \in \mathbb{Z}^d$ and $t > 0$ with $|x-y| \vee t^{1/2} \geq U_x$,

$$q_t^\omega(x,y) \leq c_3 t^{-d/2} \exp(-c_4 |x-y|^2/t) \qquad \text{if } t \geq |x-y|,$$
$$q_t^\omega(x,y) \leq c_3 \exp(-c_4 |x-y|(1 \vee \log(|x-y|/t))) \qquad \text{if } t \leq |x-y|.$$

(c) For $x, y \in \mathbb{Z}^d$ and $t > 0$,

$$q_t^\omega(x,y) \geq c_5 t^{-d/2} \exp(-c_6 |x-y|^2/t) \qquad \text{if } t \geq U_x^2 \vee |x-y|^{1+\eta}.$$

(d) For $x, y \in \mathbb{Z}^d$ and $t > 0$ with $t \geq c_7 \vee |x-y|^{1+\eta}$,

$$c_8 t^{-d/2} \exp(-c_9 |x-y|^2/t) \leq \mathbb{E}[q_t^\omega(x,y)] \leq c_{10} t^{-d/2} \exp(-c_{11} |x-y|^2/t).$$

Quenched invariance principle For $t \geq 0$, define $Y_t^{(\varepsilon)}$ as in (7.3). Then the following quenched invariance principle is proved in [16].

Theorem 7.4 *(i) Let $\{Y_t\}_{t\geq 0}$ be the VSRW. Then $\mathbb{P}$-a.s. $Y^{(\varepsilon)}$ converges (under P_ω^0) in law to Brownian motion on $\mathbb{R}^d$ with covariance $\sigma_V^2 I$ where $\sigma_V > 0$ is non-random.*
(ii) Let $\{Y_t\}_{t\geq 0}$ be the CSRW. Then $\mathbb{P}$-a.s. $Y^{(\varepsilon)}$ converges (under P_ω^0) in law to Brownian motion on $\mathbb{R}^d$ with covariance $\sigma_C^2 I$ where $\sigma_C^2 = \sigma_V^2/(2d\mathbb{E}\mu_e)$ if $\mathbb{E}\mu_e < \infty$ and $\sigma_C^2 = 0$ if $\mathbb{E}\mu_e = \infty$.

Local central limit theorem In [17], a sufficient condition is given for the quenched local limit theorem to hold (see [47] for a generalization to sub-Gaussian type local CLT). Using the results, the following local CLT is proved in [16]. (We do not give proof here.)

Theorem 7.5 *Let $q_t^\omega(x,y)$ be the VSRW and write $k_t(x) = (2\pi t\sigma_V^2)^{-d/2}\exp(-|x|^2/(2\sigma_V^2 t)$ where σ_V is as in Theorem 7.4 (i). Let $T > 0$, and for $x \in \mathbb{R}^d$, write $[x] = ([x_1], \cdots, [x_d])$. Then*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}^d} \sup_{t\geq T} |n^{d/2} q_{nt}^\omega(0, [n^{1/2}x]) - k_t(x)| = 0, \quad \mathbb{P} - a.s.$$

The key idea of the proof is as follows: one can prove the parabolic Harnack inequality using Theorem 7.3. This implies the uniform Hölder continuity of $n^{d/2}q_{nt}^\omega(0,[n^{1/2}\cdot])$, which, together with Theorem 7.4 implies the pointwise convergence.

For the case of simple random walk on the supercritical percolation, this local CLT is proved in [17]. Note that in general when $\mu_e \leq c$, such local CLT does NOT hold because of the anomalous behavior of the heat kernel and the quenched invariance principle.

More about CSRW with $\mathbb{E}\mu_e = \infty$ According to Theorem 7.4 (ii), one does not have the usual central limit theorem for CSRW with $\mathbb{E}\mu_e = \infty$ in the sense the scaled process degenerates as $\varepsilon \to 0$. A natural question is what is the right scaling order and what is the scaling limit. The answers are given in [13, 22] for the case of heavy-tailed environments with $d \geq 3$. Let $\{\mu_e\}$ satisfies

$$\mathbb{P}(\mu_e \geq c_1) = 1, \quad \mathbb{P}(\mu_e \geq u) = c_2 u^{-\alpha}(1 + o(1)) \text{ as } u \to \infty, \tag{7.4}$$

for some constants $c_1, c_2 > 0$ and $\alpha \in (0, 1]$.

In order to state the result, we first introduce the Fractional-Kinetics (FK) process.

Definition 7.6 *Let $\{B_d(t)\}$ be a standard d-dimensional Brownian motion started at 0, and for $\alpha \in (0,1)$, let $\{V_\alpha(t)\}_{t\geq 0}$ be an α-stable subordinator independent of$\{B_d(t)\}$, which is determined by $\mathbb{E}[\exp(-\lambda V_\alpha(t))] = \exp(-t\lambda^\alpha)$. Let $V_\alpha^{-1}(s) := \inf\{t : V_\alpha(t) > s\}$ be the rightcontinuous inverse of $V_\alpha(t)$. We define the fractional-kinetics process $\mathbf{FK}_{d,\alpha}$ by*

$$\mathbf{FK}_{d,\alpha}(s) = B_d(V_\alpha^{-1}(s)), \quad s \in [0, \infty).$$

The FK process is non-Markovian process, which is γ-Hölder continuous for all $\gamma < \alpha\ 2\!\!\!/$ and is self-similar, i.e. $\mathbf{FK}_{d,\alpha}(\cdot) \stackrel{(d)}{=} \lambda^{-\alpha/2}\mathbf{FK}_{d,\alpha}(\lambda\cdot)$ for all $\lambda > 0$. The density of the process $p(t,x)$ started at 0 satisfies the fractional-kinetics equation

$$\frac{\partial^\alpha}{\partial t^\alpha}p(t,x) = \frac{1}{2}\Delta p(t,x) + \delta_0(x)\frac{t^{-\alpha}}{\Gamma(1-\alpha)}.$$

This process is well-known in physics literatures, see [110] for details.

Theorem 7.7 *Let $d \geq 3$ and Let $\{Y_t\}_{t\geq 0}$ be the CSRW of RCM that satisfies (7.4).*
(i) ([13]) *Let $\alpha \in (0,1)$ in (7.4) and let $Y_t^{(\varepsilon)} := \varepsilon Y_{t/\varepsilon^{2/\alpha}}$. Then $\mathbb{P}$-a.s. $Y^{(\varepsilon)}$ converges (under P_ω^0) in law to a multiple of the fractional-kinetics process $c \cdot \mathbf{FK}_{d,\alpha}$ on $D([0,\infty), \mathbb{R}^d)$ equipped with the topology of the uniform convergence on compact subsets of $[0,\infty)$.*
(ii) ([22]) *Let $\alpha = 1$ in (7.4) with $c_1 = c_2 = 1$ and let $Y_t^{(\varepsilon)} := \varepsilon Y_{t\log(1/\varepsilon)/\varepsilon^2}$. Then $\mathbb{P}$-a.s. $Y^{(\varepsilon)}$ converges (under P_ω^0) in law to Brownian motion on $\mathbb{R}^d$ with covariance $\sigma_C^2 I$ where $\sigma_C = 2^{-1/2}\sigma_V > 0$.*

Remark 7.8 *(i) In [24], a scaling limit theorem similar to Theorem 7.7 (i) was shown for symmetric Bouchaud's trap model (BTM) for $d \geq 2$. Let $\{\tau_x\}_{x\in\mathbb{Z}^d}$ be a positive i.i.d. and let $a \in [0,1]$ be a parameter. Define a random weight (conductance) by*

$$\mu_{xy} = \tau_x^a \tau_y^a \qquad \text{if } x \sim y,$$

and let $\mu_x = \tau_x$ be the measure. Then, the BTM is the CSRW with the transition probability $\mu_{xy}/\sum_y \mu_{xy}$ and the measure μ_x. If $a = 0$, then the BTM is a time change of the simple random walk on $\mathbb{Z}^d$ and it is called symmetric BMT, while non-symmetric refers to the general case $a \neq 0$. (This terminology is a bit confusing. Note that the Markov chain for the BTM is symmetric (reversible) w.r.t. μ for all $a \in [0,1]$.) According to the result in [24], one may expect that Theorem 7.7 (i) holds for $d = 2$ as well with a suitable log correction in the scaling exponent.
(ii) For $d = 1$, the scaling limit is quite different from the FK process. In [52, 25], it is proved that the scaling limit (in the sense of finite-dimensional distributions) of the BTM on $\mathbb{R}$ is a singular diffusion in a random environment, called FIN diffusion. The result can be extended to RCM for $d = 1$.

Remark 7.9 *Isoperimetric inequalities and heat kernel estimates are very useful to obtain various properties of the random walk. For the case of supercritical percolation, estimates of mixing times ([29]) and the Laplace transform of the range of a random walk ([99]) are obtained with the help of isoperimetric inequalities and heat kernel estimates. For 2-dimensional supercritical percolation ([21, 39]), and other models including IICs on trees and IIC on $\mathbb{Z}^d$ with $d \geq 19$ ([21]), it is proved that two independent random walks started at the same point collide infinitely often $\mathbb{P}$-a.s..*

Remark 7.10 *(i) RCM is a special case of random walk in random environment (RWRE). The subject of RWRE has a long history; we refer to [35, 111] for overviews of this field.*
(ii) For directed random walks in (space-time) random environments, the quenched invariance principle is obtained in [98]. Since the walk enters a new environment at every time step, one can use independence more efficiently, but the process is no longer symmetric (reversible). Because of the directed nature of the environment, one may consider distributions with a drift for which a CLT is not even expected to hold in general for the undirected setting; see for example [108, 26] for 'pathologies' that may arise.

7.2 Percolation estimates

Consider the supercritical bond percolation $p > p_c(d)$ on $\mathbb{Z}^d$ for $d \geq 2$. In this subsection, we will give some percolation estimates that are needed later. We do not give proof here, but mention the corresponding references.

Let $\mathcal{C}(0)$ be the open cluster containing 0 and for $x, y \in \mathcal{C}(0)$, let $d_\omega(x,y)$ be the graph distance for $\mathcal{C}(0)$ and $|x-y|$ be the Euclidean distance. The first statement of the next proposition gives a stretched-exponential decay of truncated connectivities due to [61, Theorem 8.65]. The second statement is a comparison of the graph distance and the Euclidean distance due to Antal and Pisztora [9, Corollary 1.3].

Proposition 7.11 *Let $p > p_c(d)$. Then the following hold.*
(i) There exists $c_1 = c_1(p)$ such that

$$\mathbb{P}_p(|\mathcal{C}(0)| = n) \leq \exp(-c_1 n^{(d-1)/d}) \qquad \forall n \in \mathbb{N}.$$

(ii) There exists $c_2 = c_2(p,d) > 0$ such that the following holds $\mathbb{P}_p$-almost surely,

$$\limsup_{|y|\to\infty} \frac{d_\omega(0,y) 1_{\{0\leftrightarrow y\}}}{|y|} \leq c_2.$$

For $\alpha > 0$, denote $\mathcal{C}_{\infty,\alpha}$ the set of sites in $\mathbb{Z}^d$ that are connected to infinity by a path whose edges satisfy $\mu_b \geq \alpha$. The following proposition is due to [32, Proposition 2.3]. Similar estimate for the size of 'holes' in $\mathcal{C}_\infty$ can be found in [90, Lemma 3.1].

Proposition 7.12 *Assume $p_+ = \mathbb{P}(\mu_b > 0) > p_c(d)$. Then there exists $c(p_+, d) > 0$ such that if α satisfies*

$$\mathbb{P}(\mu_b \geq \alpha) > p_c(d) \quad \text{and} \quad \mathbb{P}(0 < \mu_b < \alpha) < c(p_+, d), \tag{7.5}$$

then $\mathcal{C}_{\infty,\alpha} \neq \emptyset$ and $\mathcal{C}_\infty \setminus \mathcal{C}_{\infty,\alpha}$ has only finite components a.s.

Further, if $\mathcal{K}(x)$ is the (possibly empty) component of $\mathcal{C}_\infty \setminus \mathcal{C}_{\infty,\alpha}$ containing x, then

$$\mathbb{P}\big(x \in \mathcal{C}_\infty,\ \text{diam}\, \mathcal{K}(x) \geq n\big) \leq c_1 e^{-c_2 n}, \qquad n \geq 1, \tag{7.6}$$

for some $c_1, c_2 > 0$. Here 'diam' is the diameter in Euclidean distance on $\mathbb{Z}^d$.

7.3 Proof of some heat kernel estimates

Proof of Theorem 7.1(ii). For $\kappa > 1/d$ let $\epsilon > 0$ be such that $(1+4d\epsilon)/d < \kappa$. Let $\mathbb{P}$ be an i.i.d. conductance law on $\{2^{-N} : N \geq 0\}^{E_d}$ such that

$$\mathbb{P}(\mu_e = 1) > p_c(d), \quad \mathbb{P}(\mu_e = 2^{-N}) = cN^{-(1+\epsilon)}, \quad \forall N \geq 1, \tag{7.7}$$

where $c = c(\epsilon)$ is the normalized constant. Let $\mathbf{e}_1$ denote the unit vector in the first coordinate direction. Define the scale $\ell_N = N^{(1+4d\epsilon)/d}$ and for each $x \in \mathbb{Z}^d$, let $A_N(x)$ be the event that the configuration near x, $y = x + \mathbf{e}_1$ and $z = x + 2\mathbf{e}_1$ is as follows:

(1) $\mu_{yz} = 1$ and $\mu_{xy} = 2^{-N}$, while every other bond containing y or z has $\mu_e \le 2^{-N}$.

(2) x is connected to the boundary of the box of side length $(\log \ell_N)^2$ centered at x by bonds with conductance one.

Since bonds with $\mu_e = 1$ percolate and since $\mathbb{P}(\mu_e \le 2^{-N}) \asymp N^{-\epsilon}$, we have

$$\mathbb{P}\big(A_N(x)\big) \ge cN^{-[1+(4d-2)\epsilon]}. \tag{7.8}$$

Now consider a grid of vertices $\mathbb{G}_N$ in $[-\ell_N, \ell_N]^d \cap \mathbb{Z}^d$ that are located by distance $2(\log \ell_N)^2$. Since $\{A_N(x) \colon x \in \mathbb{G}_N\}$ are independent, we have

$$\mathbb{P}\Big(\bigcap_{x\in\mathbb{G}_N} A_N(x)^c\Big) \le \Big(1 - cN^{-[1+(4d-2)\epsilon]}\Big)^{|\mathbb{G}_N|} \le \exp\Big\{-c\Big(\frac{\ell_N}{(\log \ell_N)^2}\Big)^d N^{-[1+(4d-2)\epsilon]}\Big\} \le e^{-cN^{\epsilon}}, \tag{7.9}$$

so using the Borel-Cantelli, $\cap_{x\in\mathbb{G}_N} A_N(x)^c$ occurs only for finitely many N.

By Proposition 7.11 (i), every connected component of side length $(\log \ell_N)^2$ in $[-\ell_N, \ell_N]^d \cap \mathbb{Z}^d$ will eventually be connected to $\mathcal{C}_\infty$ in $[-2\ell_N, 2\ell_N]^d \cap \mathbb{Z}^d$. Summarizing, there exists $N_0 = N_0(\omega)$ with $\mathbb{P}(N_0 < \infty) = 1$ such that for $N \ge N_0$, $A_N(x)$ occurs for some $x = x_N(\omega) \in [-\ell_N, \ell_N]^d \cap \mathbb{Z}^d$ that is connected to 0 by a path (say Path_N) in $[-2\ell_N, 2\ell_N]^d$, on which only the last N_0 edges (i.e. those close to the origin) may have conductance smaller than one.

Now let $N \ge N_0$ and let n be such that $2^N \le 2n < 2^{N+1}$. Let $x_N \in [-\ell_N, \ell_N]^d \cap \mathbb{Z}^d$ be such that $A_N(x_N)$ occurs and let r_N be the length of Path_N. Let $\alpha = \alpha(\omega)$ be the minimum of μ_e for e within N_0 steps of the origin. The Markov chain moves from 0 to x_N in time r_N with probability at least $\alpha^{N_0}(2d)^{-r_N}$, and the probability of staying on the bond (y, z) for time $2n - 2r_N - 2$ is bounded independently of ω. The transitions across (x, y) cost order 2^{-N} each. Hence we have

$$P_\omega^{2n}(0, 0) \ge c\alpha^{2N_0}(2d)^{-2r_N} 2^{-2N}. \tag{7.10}$$

By Proposition 7.11 (ii), we have $r_N \le c\ell_N$ for large N. Since $n \asymp 2^N$ and $\ell_N \le (\log n)^\kappa$, we obtain the result. □

7.4 Corrector and quenched invariance principle

Our goal is to prove Theorem 7.2 and Theorem 7.4 assuming the heat kernel estimates. Let us first give overview of the proof. As usual for the functional central limit theorem, we use 'corrector'. Let $\varphi = \varphi_\omega : \mathbb{Z}^d \to \mathbb{R}^d$ be a harmonic map, so that $M_t = \varphi(Y_t)$ is a P_ω^0-martingale. Let I be the identity map on $\mathbb{Z}^d$. The corrector is

$$\chi(x) = (\varphi - I)(x) = \varphi(x) - x.$$

It is referred to as the 'corrector' because it corrects the non-harmonicity of the position function. For simplicity, let us consider CLT (instead of functional CLT) for Y. By definition, we have

$$\frac{Y_t}{t^{1/2}} = \frac{M_t}{t^{1/2}} - \frac{\chi(Y_t)}{t^{1/2}}.$$

Since we can control φ (due to the heat kernel estimates), the martingale CLT gives that $M_t/t^{1/2}$ converges weakly to the normal distribution. So all we need is to prove $\chi(Y_t)/t^{1/2} \to 0$. This can be done once we have (a) $P_\omega^0(|Y_t| \geq At^{1/2})$ is small and (b) $|\chi(x)|/|x| \to 0$ as $|x| \to \infty$. (a) holds by the heat kernel upper bound, so the key is to prove (b), namely sublinearity of the corrector. Note that there maybe many global harmonic functions, so we should chose one such that (b) holds. As we will see later, we in fact prove the sublinearity of the corrector for $\mathcal{C}_{\infty,\alpha}$.

We now discuss details. Let

$$\Omega = \begin{cases} [0,1]^{E_d} & \text{for Case 1,} \\ [1,\infty]^{E_d} & \text{for Case 2.} \end{cases}$$

(Note that one can choose $c = 1$ in Case 1 and Case 2.) The conductance $\{\mu_e : e \in E_d\}$ are defined on $(\Omega, \mathbb{P})$ and we write $\mu_{\{x,y\}}(\omega) = \omega_{x,y}$ for the coordinate maps. Let $T_x \colon \Omega \to \Omega$ denote the shift by x, namely $(T_z\omega)_{xy} := \omega_{x+z,y+z}$.

The construction of the corrector is simple and robust. Let $\{Q_{x,y}(\omega) : x, y \in \mathbb{Z}^d\}$ be a sequence of non-negative random variables such that $Q_{x,y}(T_z\omega) = Q_{x+z,y+z}(\omega)$, which is stationary and ergodic. Assume that there exists $C > 0$ such that the following hold:

$$\sum_{x\in\mathbb{Z}^d} Q_{0,x}(\omega) \leq C \quad \mathbb{P}\text{-a.e. } \omega, \quad \text{and} \quad \mathbb{E}[\sum_{x\in\mathbb{Z}^d} Q_{0,x}|x|^2] < \infty. \tag{7.11}$$

The following construction is based on Mathieu and Piatnitski [91].

Theorem 7.13 *There exists a function $\chi\colon \Omega \times \mathbb{Z}^d \to \mathbb{R}^d$ that satisfies the following (1)–(3) $\mathbb{P}$-a.e. ω.*

(1) (Cocycle property) $\chi(\omega, 0) = 0$ and, for all $x, y \in \mathbb{Z}^d$,

$$\chi(\omega, x) - \chi(\omega, y) = \chi(T_y\omega, x - y). \tag{7.12}$$

(2) (Harmonicity) $\varphi_\omega(x) := x + \chi(\omega, x)$ enjoys $\mathcal{L}_Q\varphi_\omega^j(z) = 0$ for all $1 \leq j \leq d$ and $z \in \mathbb{Z}^d$ where φ_ω^j is the j-th coordinate of φ_ω and

$$(\mathcal{L}_Q f)(x) = \sum_{y\in\mathbb{Z}^d} Q_{x,y}(\omega)(f(y) - f(x)). \tag{7.13}$$

(3) (Square integrability) There exists $C < \infty$ such that for all $x, y \in \mathbb{Z}^d$,

$$\mathbb{E}\big[|\chi(\cdot, y) - \chi(\cdot, x)|^2\, Q_{x,y}(\omega)\big] < C. \tag{7.14}$$

Remark 7.14 *(i) In fact, one can construct the corrector under milder condition $\mathbb{E}[\sum_{x\in\mathbb{Z}^d} Q_{0,x}|x|^2] < \infty$ instead of (7.11) by suitable changes of function spaces in the proof.*
(ii) In [33, Lemma 3.3], it is proved that for the uniform elliptic case (Case 0), i.e. $\mathbb{P}(c_1 \leq Q_{xy} \leq c_2) = 1$ for some $c_1, c_2 > 0$, the corrector is uniquely determined by the properties (1)–(3) in Theorem

7.13 and (3-2) in Proposition 7.16 below.
(iii) In [91, 32] the corrector is defined essentially on $\Omega \times B$ where $B = \{\mathbf{e}_1, \cdots, \mathbf{e}_d, -\mathbf{e}_1, \cdots, -\mathbf{e}_d\}$ is the set of unit vectors. However, one can easily extend the domain of the corrector to $\Omega \times \mathbb{Z}^d$ by using the cocycle property.

Given the construction of the corrector, we proceed as in Biskup and Prescott [32]. We first give sufficient condition for the (uniform) sublinearity of the corrector in this general setting and then show that the assumption given for the sufficiency of the sublinearity can be verified for Case 1 and Case 2.

We consider open clusters with respect to the (in general long range) percolation for $\{Q_{xy}\}_{x,y\in\mathbb{Z}^d}$. Assume that the distribution of $\{Q_{xy}\}$ is given so that there exists unique infinite open cluster $\mathbb{P}$-a.s., and denote it by $\mathcal{C}_\infty$. We also consider (random) one-parameter family of infinite sub-clusters which we denote by $\mathcal{C}_{\infty,\alpha}$, where we set $\mathcal{C}_{\infty,0} = \mathcal{C}_\infty$. We assume $\mathbb{P}(0 \in \mathcal{C}_{\infty,\alpha}) > 0$. (The concrete choice of $\mathcal{C}_{\infty,\alpha}$ will be given later.)

Let $Y = \{Y_t\}_{t\geq 0}$ be the VSRW on $\mathcal{C}_\infty$ (with based measure $\nu_x \equiv 1$ for $x \in \mathcal{C}_\infty$) that corresponds to $\mathcal{L}_Q$ in (7.13). We introduce the trace of Markov chain to $\mathcal{C}_{\infty,\alpha}$ (cf. Subsection 1.3). Define $\sigma_1, \sigma_2, \ldots$ the time intervals between successive visits of Y to $\mathcal{C}_{\infty,\alpha}$, namely, let

$$\sigma_{j+1} := \inf\big\{n \geq 1 \colon Y_{\sigma_0+\cdots+\sigma_j+n} \in \mathcal{C}_{\infty,\alpha}\big\}, \tag{7.15}$$

with $\sigma_0 = 0$. For each $x, y \in \mathcal{C}_{\infty,\alpha}$, let $\hat{Q}_{xy} := \hat{Q}^{(\alpha)}_{xy}(\omega) = P^x_\omega(Y_{\sigma_1} = y)$ and define the operator

$$\mathcal{L}_{\hat{Q}} f(x) := \sum_{y\in\mathcal{C}_{\infty,\alpha}} \hat{Q}_{xy}\big(f(y) - f(x)\big). \tag{7.16}$$

Let $\hat{Y} = \{\hat{Y}_t\}_{t\geq 0}$ be the continuous-time random walk corresponding to $\mathcal{L}_{\hat{Q}}$.

The following theorem gives sufficient condition for the sublinearity of the corrector ψ_ω on $\mathcal{C}_{\infty,\alpha}$. Let $\mathbb{P}_\alpha(\cdot) := \mathbb{P}(\cdot|0 \in \mathcal{C}_{\infty,\alpha})$ and let $\mathbb{E}_\alpha$ be the expectation w.r.t. $\mathbb{P}_\alpha$.

Theorem 7.15 *Fix $\alpha \geq 0$ and suppose $\psi_\omega \colon \mathcal{C}_{\infty,\alpha} \to \mathbb{R}^d$, $\theta > 0$ and $\hat{Y}$ satisfy the following (1)–(5) for $\mathbb{P}_\alpha$-a.e. ω:*

(1) (Harmonicity) If $\varphi_\omega(x) = (\varphi^1_\omega(x), \cdots, \varphi^d_\omega(x)) := x + \psi_\omega(x)$, then $\mathcal{L}_{\hat{Q}}\varphi^j_\omega = 0$ on $\mathcal{C}_{\infty,\alpha}$ for $1 \leq j \leq d$.

(2) (Sublinearity on average) For every $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{n^d} \sum_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\leq n}} \mathbf{1}_{\{|\psi_\omega(x)|\geq\epsilon n\}} = 0. \tag{7.17}$$

(3) (Polynomial growth) There exists $\theta > 0$ such that

$$\lim_{n\to\infty} \max_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\leq n}} \frac{|\psi_\omega(x)|}{n^\theta} = 0. \tag{7.18}$$

(4) (Diffusive upper bounds) For a deterministic sequence $b_n = o(n^2)$ and a.e. ω,

$$\sup_{n\geq 1} \max_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\leq n}} \sup_{t\geq b_n} \frac{E^x_\omega|\hat{Y}_t - x|}{\sqrt{t}} < \infty, \tag{7.19}$$

$$\sup_{n\geq 1} \max_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\leq n}} \sup_{t\geq b_n} t^{d/2} P^x_\omega(\hat{Y}_t = x) < \infty. \tag{7.20}$$

(5) (Control of big jumps) Let $\tau_n = \{t \geq 0 : |\hat{Y}_t - \hat{Y}_0| \geq n\}$. There exist $c_1 > 1$ and $N_\omega > 0$ which is finite for $\mathbb{P}_\alpha$-a.e. ω such that $|\hat{Y}_{\tau_n} - \hat{Y}_0| \leq c_1 n$ for all $t > 0$ and $n \geq N_\omega$.

Then for $\mathbb{P}_\alpha$-a.e. ω,

$$\lim_{n\to\infty} \max_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\leq n}} \frac{|\psi_\omega(x)|}{n} = 0. \tag{7.21}$$

Now we discuss how to apply the theorem for Case 1 and Case 2.

Case 1: In this case, we define $Q_{x,y}(\omega) = \omega_{x,y}$. Then it satisfies the conditions for $\{Q_{xy}\}$ given above including (7.11). Denote by $\mathcal{L}_\omega$ the generator of VSRW (which we denote by $\{Y_t\}_{t\geq 0}$), i.e.,

$$(\mathcal{L}_\omega f)(x) = \sum_{y:\, y\sim x} \omega_{xy}(f(y) - f(x)). \tag{7.22}$$

$\mathcal{L}_Q = \mathcal{L}_\omega$ in this case. The infinite cluster $\mathcal{C}_\infty$ is the cluster for $\{b \in E_d : \mu_b > 0\}$ (since we assumed $p_+ = \mathbb{P}(\mu_b > 0) > p_c(d)$, it is the supercritical percolation so there exists unique infinite cluster.) Fix $\alpha \geq 0$ that satisfies (7.5) and $\mathcal{C}_{\infty,\alpha}$ the the cluster for $\{b \in E_d : \mu_b \geq \alpha\}$. (Again it is the supercritical percolation so there exists unique infinite cluster.) We let $\hat{Q}_{xy} := P^x_\omega(Y_{\sigma_1} = y)$. Note that although $\hat{Y}$ may jump the 'holes' of $\mathcal{C}_{\infty,\alpha}$, Proposition 7.12 shows that all jumps are finite. Let $\varphi_\omega(x) := x + \chi(\omega, x)$ where χ is the corrector on $\mathcal{C}_\infty$. Then, by the optional stopping theorem, it is easy to see that $\mathcal{L}_{\hat{Q}}\varphi_\omega = 0$ on $\mathcal{C}_{\infty,\alpha}$ (see Lemma 7.21 (i)).

Case 2: First, note that if we define $Q_{x,y}(\omega)$ similarly as in Case 1, it does not satisfy (7.11). Especially when $\mathbb{E}\mu_e = \infty$, we cannot define correctors for $\{Y_t\}_{t\geq 0}$ in a usual manner. The idea of [16] is to discretize $\{Y_t\}_{t\geq 0}$ and construct the corrector for the discretized process $\{Y_{[t]}\}_{t\geq 0}$. Let $q^\omega_t(x, y)$ be the heat kernel of $\{Y_t\}_{t\geq 0}$. Note that $q^\omega_t(x, y) = P^x_\omega(Y_t = y) = q^\omega_t(y, x)$. We define

$$Q_{x,y}(\omega) = q^\omega_1(x, y), \qquad \forall x, y \in \mathbb{Z}^d.$$

Then $Q_{x,y} \leq 1$ and $\sum_y Q_{x,y} = 1$. Note that in this case $\mathbb{P}(Q_{xy} > 0) > 0$ for all $x, y \in \mathbb{Z}^d$, so $\mathcal{C}_\infty = \mathbb{Z}^d$. Integrating Theorem 7.3 (b), we have $\mathbb{E}[E^0_\cdot|Y_t|^2] = \mathbb{E}[\sum_x Q_{0x}|x|^2] \leq ct$ for $t \geq 1$, so (7.11) holds. One can easily check other conditions for $\{Q_{xy}\}$ given above. In this case, we do not need to consider $\mathcal{C}_{\infty,\alpha}$ and there is no need to take the trace process. So, $\alpha = 0$, $\hat{Q}_{x,y} = Q_{x,y}$, $\mathcal{L}_{\hat{Q}} = \mathcal{L}_Q$, and $\hat{Y}_t = Y_{[t]}$ (discrete time random walk) in this case.

The next proposition provides some additional properties of the corrector for Case 1 and Case 2. This together with Lemma 7.28 below verify for Case 1 and Case 2 the sufficient conditions for sublinearity of the corrector given in Theorem 7.15. As mentioned above, for Case 2, we consider only $\alpha = 0$.

Proposition 7.16 *Let $\alpha > 0$ for Case 1 and $\alpha = 0$ for Case 2, and let χ be the corrector given in Theorem 7.13. Then χ satisfies (2), (3) in Theorem 7.13 for $\mathbb{P}_\alpha$-a.e. ω if $\mathbb{P}(0 \in \mathcal{C}_{\infty,\alpha}) > 0$. Further, it satisfies the following.*
(1) (Polynomial growth) There exists $\theta > d$ such that the following holds $\mathbb{P}_\alpha$-a.e. ω:

$$\lim_{n\to\infty} \max_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\ |x|\le n}} \frac{|\chi(\omega,x)|}{n^\theta} = 0. \tag{7.23}$$

(2) (Sublinearity on average) For each $\epsilon > 0$, the following holds $\mathbb{P}_\alpha$-a.e. ω:

$$\lim_{n\to\infty} \frac{1}{n^d} \sum_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\ |x|\le n}} \mathbf{1}_{\{|\chi(\omega,x)|\ge\epsilon n\}} = 0.$$

(3-1) Case 1: (Zero mean under random shifts) Let $Z\colon \Omega \to \mathbb{Z}^d$ be a random variable such that (a) *$Z(\omega) \in \mathcal{C}_{\infty,\alpha}(\omega)$,* (b) *$\mathbb{P}_\alpha$ is preserved by $\omega \mapsto \tau_{Z(\omega)}(\omega)$, and* (c) *$\mathbb{E}_\alpha[d_\omega^{(\alpha)}(0, Z(\omega))^q] < \infty$ for some $q > 3d$. Then $\chi(\cdot, Z(\cdot)) \in L^1(\Omega, \mathcal{F}, \mathbb{P}_\alpha)$ and*

$$\mathbb{E}_\alpha\big[\chi(\cdot, Z(\cdot))\big] = 0. \tag{7.24}$$

(3-2) Case 2: (Zero mean) $\mathbb{E}\big[\chi(\cdot, x)\big] = 0$ for all $x \in \mathbb{Z}^d$.

7.5 Construction of the corrector

In this subsection, we prove Theorem 7.13. We follow the arguments in Barlow and Deuschel [16, Section 5] (see also [32, 33, 91]).

We define the process that gives the 'environment seen from the particle' by

$$Z_t = T_{Y_t}(\omega), \qquad \forall t \in [0, \infty), \tag{7.25}$$

where $\{Y_t\}_{t\ge0}$ is the Markov chain corresponding to $\mathcal{L}_Q$. Note that the process Z is ergodic under the time shift on Ω (see for example, [30, Section 3], [50, Lemma 4.9] for the proof in discrete time).

Let $\mathbb{L}^2 = \mathbb{L}^2(\Omega, \mathbb{P})$ and for $F \in \mathbb{L}^2$, write $F_x = F \circ T_x$. Then the generator of Z is

$$\hat{L}F(\omega) = \sum_{x\in\mathbb{Z}^d} Q_{0,x}(\omega)(F_x(\omega) - F(\omega)).$$

Define

$$\hat{\mathcal{E}}(F, G) = \mathbb{E}\Big[\sum_{x\in\mathbb{Z}^d} Q_{0,x}(F - F_x)(G - G_x)\Big] \qquad \forall F, G \in \mathbb{L}^2.$$

The following lemma shows that $\hat{\mathcal{E}}$ is the quadratic form on $\mathbb{L}^2$ corresponding to $\hat{L}$.

Lemma 7.17 *(i) For all $F \in \mathbb{L}^2$ and $x \in \mathbb{Z}^d$, it holds that $\mathbb{E}F = \mathbb{E}F_x$ and $\mathbb{E}[Q_{0,x}F_x] = \mathbb{E}[Q_{0,-x}F]$.*
(ii) For $F \in \mathbb{L}^2$, it holds that $\hat{\mathcal{E}}(F, F) < \infty$ and $\hat{L}F \in \mathbb{L}^2$.
(iii) For $F, G \in \mathbb{L}^2$, it holds that $\hat{\mathcal{E}}(F, G) = -\mathbb{E}[G\hat{L}F]$.

Proof. (i) The first equality is because $\mathbb{P} = \mathbb{P} \circ T_x$. Since $Q_{0,x} \circ T_{-x} = Q_{-x,0} = Q_{0,-x}$, $\mathbb{E}[Q_{0,x} F_x] = \mathbb{E}[(Q_{0,x} \circ T_{-x})F] = \mathbb{E}[Q_{0,-x} F]$ so the second equality holds.
(ii) For $F \in \mathbb{L}^2$, we have

$$\begin{aligned}\hat{\mathcal{E}}(F,F) &= \mathbb{E}[\sum_x Q_{0,x}(F-F_x)^2] \le 2\mathbb{E}[\sum_x Q_{0,x}(F^2+F_x^2)] \\ &= 2C\mathbb{E}F^2 + 2\mathbb{E}[\sum_x Q_{0,-x}F^2] = 4C\|F\|_2^2, \end{aligned} \tag{7.26}$$

where we used (i) in the second equality, and C is a constant in (7.11). Also,

$$\begin{aligned}\mathbb{E}|\hat{L}F|^2 &= \mathbb{E}[\sum_{x,y} Q_{0,x}Q_{0,y}(F_x-F)(F_y-F)] \\ &\le \mathbb{E}\Big[(\sum_{x,y} Q_{0,x}Q_{0,y}(F_x-F)^2)^{1/2}(\sum_{x,y} Q_{0,x}Q_{0,y}(F_y-F)^2)^{1/2}\Big] \le C\hat{\mathcal{E}}(F,F) \le 4C^2\|F\|_2^2\end{aligned}$$

where (7.26) is used in the last inequality. We thus obtain (ii).
(iii) Using (i), we have

$$\mathbb{E}[Q_{0,-x}G(F-F_{-x})] = \mathbb{E}[Q_{0,x}G_x(F_x-F)]. \tag{7.27}$$

So

$$\begin{aligned}-\mathbb{E}[G\hat{L}F] &= \sum_x \mathbb{E}[GQ_{0,x}(F-F_x)] = \frac{1}{2}\sum_x \mathbb{E}[GQ_{0,x}(F-F_x)] + \frac{1}{2}\sum_x \mathbb{E}[GQ_{0,-x}(F-F_{-x})] \\ &= \frac{1}{2}\sum_x \mathbb{E}[Q_{0,x}(GF - GF_x + G_xF_x - G_xF)] = \hat{\mathcal{E}}(F,G),\end{aligned}$$

where (7.27) is used in the third equation, and (iii) is proved. □

Next we look at 'vector fields'. Let M be the measure on$\Omega \times \mathbb{Z}^d$ defined as

$$\int_{\Omega\times\mathbb{Z}^d} G dM := \mathbb{E}[\sum_{x\in\mathbb{Z}^d} Q_{0,x}G(\cdot,x)], \quad \text{for } G : \Omega\times\mathbb{Z}^d \to \mathbb{R}.$$

We say $G : \Omega \times \mathbb{Z}^d \to \mathbb{R}$ has the *cocycle property* (or *shift-covariance property*) if the following holds,

$$G(T_x\omega, y-x) = G(\omega,y) - G(\omega,x), \qquad \mathbb{P}-\text{a.s.}$$

Let $\overline{L}^2 = \{G \in \mathbb{L}^2(\Omega\times\mathbb{Z}^d, M) : G \text{ has the cocycle property.}\}$ and for $G \in \mathbb{L}^2(\Omega\times\mathbb{Z}^d, M)$, write $\|G\|_{\overline{L}^2}^2 := \mathbb{E}[\sum_{x\in\mathbb{Z}^d} Q_{0,x}G(\cdot,x)^2]$. It is easy to see that $\overline{L}^2$ is a Hilbert space. Also, for $G \in \overline{L}^2$, it holds that $G(\omega,0) = 0$ and $G(T_x\omega,-x) = -G(\omega,x)$.

Define $\nabla : \mathbb{L}^2 \to \overline{L}^2$ by

$$\nabla F(\omega,x) := F(T_x\omega) - F(\omega), \qquad \text{for } F \in \mathbb{L}^2.$$

Since $\|\nabla F\|_{\overline{L}^2}^2 = \hat{\mathcal{E}}(F,F) < \infty$ (due to Lemma 7.17 (ii)) and $\nabla F(T_x\omega, y-x) = F(T_y\omega) - F(T_x\omega) = \nabla F(\omega,y) - \nabla F(\omega,x)$, we can see that ∇F is indeed in $\overline{L}^2$.

Now we introduce an orthogonal decomposition of $\overline{L}^2$ as follows

$$\overline{L}^2 = L^2_{\text{pot}} \oplus L^2_{\text{sol}}.$$

Here $L^2_{\text{pot}} = Cl\{\nabla F : F \in \mathbb{L}^2\}$ where the closure is taken in $\overline{L}^2$, and L^2_{sol} is the orthogonal complement of L^2_{pot}. Note that 'pot' stands for 'potential', and 'sol' stands for 'solenoidal'.

Before giving some properties of L^2_{pot} and L^2_{sol}, we now give definition of the corrector. Let $\Pi : \mathbb{R}^d \to \mathbb{R}^d$ be the identity and denote byΠ_j the j-th coordinate ofΠ. Then,$\Pi_j \in \overline{L}^2$. Indeed, $\Pi_j(y-x) = \Pi_j(y) - \Pi_j(x)$ so it has the cocycle property, and by (7.11), $\|\Pi_j\|^2_{\overline{L}^2} < \infty$. Now define $\chi_j \in L^2_{\text{pot}}$ and$\Phi_j \in L^2_{\text{sol}}$ by

$$\Pi_j = \chi_j \oplus \Phi_j \in L^2_{\text{pot}} \oplus L^2_{\text{sol}}.$$

This gives the definition of the corrector $\chi = (\chi_1, \cdots, \chi_d) : \Omega \times \mathbb{Z}^d \to \mathbb{R}^d$.

Remark 7.18 *The corrector can be defined by using spectral theory as in Kipnis and Varadhan [80]. This 'projection' definition is given in Giacomin, Olla and Spohn [58], and Mathieu and Piatnitski [91] (also in [16, 32, 33] etc. mentioned above).*

Lemma 7.19 *For $G \in L^2_{\text{sol}}$, it holds that*

$$\sum_{x \in \mathbb{Z}^d} Q_{0,x} G(\omega, x) = 0, \quad \mathbb{P} - a.s. \tag{7.28}$$

Hence $M_n := G(\omega, Y_n)$ is a P^0_ω-martingale for $\mathbb{P}$-a.e. ω.

Proof. Recall that $G(T_x\omega, -x) = -G(\omega, x)$ for $G \in \overline{L}^2$. Using this, we have for each $F \in \mathbb{L}^2$

$$\begin{aligned}\sum_x \mathbb{E}[Q_{0,x} G(\cdot, x) F_x(\cdot)] &= \sum_x \mathbb{E}[T_{-x} Q_{0,x} G(T_{-x}\cdot, x) F_x(T_{-x}\cdot)] \\ &= \sum_x \mathbb{E}[Q_{0,-x}(-G(\cdot, -x)) F(\cdot)] = -\sum_x \mathbb{E}[Q_{0,x} G(\cdot, x) F(\cdot)].\end{aligned}$$

So $\sum_x \mathbb{E}[Q_{0,x} G(\cdot, x)(F + F_x)(\cdot)] = 0$. If $G \in L^2_{\text{sol}}$, then since $\nabla F \in L^2_{\text{pot}}$, we have

$$0 = \int_{\Omega \times \mathbb{Z}^d} G \nabla F dM = \sum_x \mathbb{E}[Q_{0,x} G(\cdot, x)(F_x - F)].$$

So we have $\mathbb{E}[\sum_x Q_{0,x} G F] = 0$. Since this holds for all $F \in \mathbb{L}^2$, we obtain (7.28). Since

$$\begin{aligned}E^0_\omega[G(\omega, Y_{n+1}) - G(\omega, Y_n) | Y_n = x] &= \sum_y Q_{xy}(\omega)(G(\omega, y) - G(\omega, x)) \\ &= \sum_y Q_{0,y-x}(T_x\omega) G(T_x\omega, y - x) = 0,\end{aligned}$$

where the cocycle property is used in the second equality and (7.28) is used in the last inequality, we see that $M_n := G(\omega, Y_n)$ is a P^0_ω-martingale. □

Verification of (1)-(3) in Theorem 7.13: (1) and (3) in Theorem 7.13 is clear by definition of χ and the definition of the inner product on $\mathbb{L}^2(\Omega \times \mathbb{Z}^d, M)$. By Lemma 7.19, we see that (2) in Theorem 7.13 hold. – Here, note that by the cocycle property of φ,

$$(\mathcal{L}_Q \varphi^j_\omega)(x) = \sum_{y \in \mathbb{Z}^d} Q_{x,y}(\omega)(\varphi^j_\omega(y) - \varphi^j_\omega(x)) = \sum_{y \in \mathbb{Z}^d} Q_{0,y-x}(T_x\omega)\varphi^j_{T_x\omega}(y-x) = 0,$$

for $1 \le j \le d$ where the last equality is due to (7.28). □

In [33, Section 5], there is a nice survey of the potential theory behind the notion of the corrector. There it is shown that$\Pi\ - \chi$ (here Π is the identity map) spans L^2_{pot} (see [33, Corollary 5.6]).

7.6 Proof of Theorem 7.15

The basic idea of the proof of Theorem 7.15 is that sublinearity on average plus heat kernel upper bounds imply pointwise sublinearity, which is (according to [32]) due to Y. Peres. In Theorem 7.15, the inputs that should come from the heat kernel upper bounds are reduced to the assumptions (7.19), (7.20). In order to prove Theorem 7.15, the next lemma plays a key role.

Lemma 7.20 *Let*

$$\bar{R}_n := \max_{\substack{x \in \mathcal{C}_{\infty,\alpha} \\ |x| \le n}} \big|\psi_\omega(x)\big|.$$

Under the conditions (1,2,4,5) of Theorem 7.15, for each $\epsilon > 0$ and $\delta > 0$, there exists $c_1 > 1$ and a random variable $n_0 = n_0(\omega, \epsilon, \delta)$ which is finite a.s. finite such that

$$\bar{R}_n \le \epsilon n + \delta \bar{R}_{2c_1 n}. \qquad n \ge n_0. \tag{7.29}$$

Before proving this lemma, let us prove Theorem 7.15 by showing how this lemma and (7.18) imply (7.21).

Proof of Theorem 7.15. Suppose that $\bar{R}_n/n \not\to 0$ and choose c with $0 < c < \limsup_{n\to\infty} \bar{R}_n/n$. Let θ be is as in (7.18) and choose $\epsilon := c/2$ and $\delta := (2c_1)^{-\theta-1}$. Note that then $c - \epsilon \ge (2c_1)^\theta \delta c$. If $\bar{R}_n \ge cn$ (this happens for infinitely many n's) and $n \ge n_0$, then (7.29) implies

$$\bar{R}_{2c_1 n} \ge \frac{c-\epsilon}{\delta} n \ge (2c_1)^\theta cn$$

and, inductively, $\bar{R}_{(2c_1)^k n} \ge (2c_1)^{k\theta} cn$. However, (7.18) says $\bar{R}_{(2c_1)^k n}/(2c_1)^{k\theta} \to 0$ as $k \to \infty$ for each fixed n, which is a contradiction. □

We now prove Lemma 7.20. The idea behind is the following. Let $\{\hat{Y}_t\}$ be the process corresponding to $\mathcal{L}_{\hat{Q}}$ started at the maximizer of $\bar{R}_n$, say z^*. Using the martingale property, we will estimate $E^{z^*}_\omega[\psi_\omega(\hat{Y}_t)]$ for time $t = o(n^2)$. The right-hand side of (7.29) expresses two situations that may occur at time t: (i) $|\psi_\omega(\hat{Y}_t)| \le \epsilon n$ (by 'sublinearity on average', this happens with high probability),

(ii) $\hat{Y}$ has not yet left the box $[-2c_1n, 2c_1n]^d$ and so $\psi_\omega(\hat{Y}_t) \le \bar{R}_{2c_1n}$. It turns out that these two are the dominating strategies.

Proof of Lemma 7.20. Fix $\epsilon, \delta > 0$ and let z_* be the site where the maximum $\bar{R}_n$ is achieved. Denote

$$\mathcal{O}_n := \{x \in \mathcal{C}_{\infty,\alpha} \colon |x| \le n,\, |\psi_\omega(x)| \ge \tfrac{1}{2}\epsilon n\}.$$

Recall that $\{\hat{Y}_t\}$ is the continuous-time random walk on $\mathcal{C}_{\infty,\alpha}$ corresponding to $\mathcal{L}_{\hat{Q}}\omega$ in (7.16). We denote its expectation for the walk started at z_* by $E_\omega^{z_*}$. Define

$$S_n := \inf\{t \ge 0 \colon |\hat{Y}_t - z_*| \ge 2n\}.$$

Note that, by Theorem 7.15 (5), there exists $n_1(\omega)$ which is a.s. finite such that we have $|\hat{Y}_{t\wedge S_n} - z_*| \le 2c_1n$ for all $t > 0$ and $n \ge n_1(\omega)$. Using the harmonicity of $x \mapsto x + \psi_\omega(x)$ and the optional stopping theorem, we have

$$\bar{R}_n = |\psi_\omega(z_*)| \le E_\omega^{z_*}\Big[\big|\psi_\omega(\hat{Y}_{t\wedge S_n}) + \hat{Y}_{t\wedge S_n} - z_*\big|\Big]. \tag{7.30}$$

We will consider t that satisfies

$$t \ge b_{3c_1n}, \tag{7.31}$$

where $b_n = o(n^2)$ is the sequence in Theorem 7.15 (4), and we will estimate the expectation separately on $\{S_n < t\}$ and $\{S_n \ge t\}$.

Case i $\{S_n < t\}$: On this event, the integrand of (7.30)$\le \bar{R}_{2c_1n} + 2c_1n$. To estimate $P_{\omega,z_*}(S_n < t)$, we consider the subcases $|\hat{Y}_{2t} - z_*| \ge n$ and $\le n$. For the former case, (7.31) and (7.19) imply

$$P_{\omega,z_*}\big(|\hat{Y}_{2t} - z_*| \ge n\big) \le \frac{E_\omega^{z_*}|\hat{Y}_{2t} - z_+|}{n} \le \frac{c_2\sqrt{t}}{n},$$

for some $c_2 = c_2(\omega)$. For the latter case, we have

$$\big\{|\hat{Y}_{2t} - z_*| \le n\big\} \cap \{S_n < t\} \subset \big\{|\hat{Y}_{2t} - \hat{Y}_{S_n}| \ge n\big\} \cap \{S_n < t\}.$$

Noting that $2t - S_n \in [t, 2t]$ and using (7.31), (7.19) again, we have

$$P_{\omega,x}\big(|\hat{Y}_s - x| \ge n\big) \le \frac{c_3\sqrt{t}}{n} \qquad \text{when } x := \hat{Y}_{S_n} \text{ and } s := 2t - S_n,$$

for some $c_3 = c_3(\omega)$. From the strong Markov property we have the same bound for $P_{\omega,z_*}(S_n < t, |\hat{Y}_{2t} - z_*| \ge n)$. Combining both cases, we have $P_{\omega,z_*}(S_n < t) \le (c_2 + c_3)\sqrt{t}/n$. So

$$E_\omega^{z_*}\Big[\big|\psi_\omega(\hat{Y}_{t\wedge S_n}) + \hat{Y}_{t\wedge S_n} - z_*\big|\mathbf{1}_{\{S_n<t\}}\Big] \le (\bar{R}_{2c_1n} + 2c_1n)\frac{(c_2 + c_3)\sqrt{t}}{n}. \tag{7.32}$$

Case ii $\{S_n \ge t\}$: On this event, the expectation in (7.30) is bounded by

$$E_\omega^{z_*}\big[|\psi_\omega(\hat{Y}_t)|\mathbf{1}_{\{S_n \ge t\}}\big] + E_\omega^{z_*}|\hat{Y}_t - z_*|.$$

By (7.19), the second term is less than $c_4\sqrt{t}$ for some $c_4 = c_4(\omega)$ provided $t \geq b_n$. The first term can be estimated depending on whether $\hat{Y}_t \notin \mathcal{O}_{2n}$ or not:

$$E_\omega^{z_*}\left[|\psi_\omega(\hat{Y}_t)|\mathbf{1}_{\{S_n\geq t\}}\right] \leq \frac{1}{2}\epsilon n + \bar{R}_{2c_1 n} P_{\omega,z_*}(\hat{Y}_t \in \mathcal{O}_{2n}).$$

For the probability of $\hat{Y}_t \in \mathcal{O}_{2n}$ we get

$$\begin{aligned} P_{\omega,z_*}(\hat{Y}_t \in \mathcal{O}_{2n}) &= \sum_{x\in\mathcal{O}_{2n}} P_{\omega,z_*}(\hat{Y}_t = x) \\ &\leq \sum_{x\in\mathcal{O}_{2n}} P_{\omega,z_*}(\hat{Y}_t = z_*)^{1/2} P_{\omega,x}(\hat{Y}_t = x)^{1/2} \leq c_5 \frac{|\mathcal{O}_{2n}|}{t^{d/2}}, \end{aligned}$$

where $c_5 = c_5(\omega)$ denote the suprema in (7.20). Here is the first inequality we used the Schwarz inequality and the second inequality is due to (7.20).

$$E_\omega^{z_*}\Big[\big|\psi_\omega(\hat{Y}_{t\wedge S_n}) + \hat{Y}_{t\wedge S_n} - z_*\big|\mathbf{1}_{\{S_n\geq t\}}\Big] \leq c_4\sqrt{t} + \frac{1}{2}\epsilon n + \bar{R}_{2c_1 n} c_5 \frac{|\mathcal{O}_{2n}|}{t^{d/2}}. \tag{7.33}$$

By (7.32) and (7.33), we conclude that

$$\bar{R}_n \leq (\bar{R}_{2c_1 n} + 2c_1 n)\frac{(c_2+c_3)\sqrt{t}}{n} + c_4\sqrt{t} + \frac{1}{2}\epsilon n + \bar{R}_{2c_1 n} c_5 \frac{|\mathcal{O}_{2n}|}{t^{d/2}}. \tag{7.34}$$

Since $|\mathcal{O}_{2n}| = o(n^d)$ as $n \to \infty$ due to (7.17), we can choose $t := \xi n^2$ with $\xi > 0$ so small that (7.31) applies and (7.29) holds for the given ϵ and δ once n is sufficiently large. □

7.7 Proof of Proposition 7.16

In this subsection, we prove Proposition 7.16 for Case 1 and Case 2 separately.

7.7.1 Case 1

For Case 1, we need to work on $\mathcal{C}_{\infty,\alpha}$, but we do not need to use Proposition 7.12 seriously yet (although we use it lightly in (7.39)). By definition, it is clear that χ satisfies Theorem 7.13 (3) for $\mathbb{P}_\alpha$ for all $\alpha \geq 0$ with $\mathbb{P}(0 \in \mathcal{C}_{\infty,\alpha}) > 0$. The next lemma proves Theorem 7.13(2) for $\mathbb{P}_\alpha$.

Lemma 7.21 *Let χ be the corrector on $\mathcal{C}_\infty$ and define $\varphi_\omega(x) = (\varphi_\omega^1(x), \cdots, \varphi_\omega^d(x)) := x + \chi(\omega, x)$. Then $\mathcal{L}_{\hat{Q}}\varphi_\omega^j(x) = 0$ for all $x \in \mathcal{C}_{\infty,\alpha}$ and $1 \leq j \leq d$.*

Proof. Note that $(\mathcal{L}_{\hat{Q}}\varphi_\omega)(x) = E_\omega^x\big[\varphi_\omega(Y_{\sigma_1})\big] - \varphi_\omega(x)$. Also, since $\{Y_t\}_{t\geq 0}$ moves in a finite component of $\mathcal{C}_\infty \setminus \mathcal{C}_{\infty,\alpha}$ for $t \in [0, \sigma_1]$, $\varphi_\omega(Y_t)$ is bounded. Since $\{\varphi_\omega(Y_t)\}_{t\geq 0}$ is a martingale and $\sigma_1 < \infty$ a.s., the optional stopping theorem gives $E_\omega^x\varphi_\omega(Y_{\sigma_1}) = \varphi_\omega(x)$. □

Proof of Proposition 7.16 (1). In this Case 1, the result holds for all $\theta > d$ as we shall prove. Let $\theta > d$ and write

$$R_n := \max_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\le n}} \big|\chi(\omega,x)\big|.$$

By Proposition 7.11 (ii),

$$\lambda(\omega) := \sup_{x\in\mathcal{C}_{\infty,\alpha}} \frac{d_\omega^{(\alpha)}(0,x)}{|x|} < \infty, \qquad \mathbb{P}_\alpha\text{-a.s.,} \tag{7.35}$$

where we let $d_\omega^{(\alpha)}(x,y)$ be the graph distance between x and y on $\mathcal{C}_{\infty,\alpha}$. So it is enough to prove $R_n/n^\theta \to 0$ on $\{\lambda(\omega)\le\lambda\}$ for all $\lambda<\infty$. Note that on $\{\lambda(\omega)\le\lambda\}$ every $x\in\mathcal{C}_{\infty,\alpha}\cap[-n,n]^d$ can be connected to the origin by a path that is inside $\mathcal{C}_{\infty,\alpha}\cap[-\lambda n,\lambda n]^d$. Using this fact and $\chi(\omega,0)=0$, we have, on $\{\lambda(\omega)\le\lambda\}$,

$$R_n \le \sum_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\le\lambda n}}\sum_{b\in B} 1_{\{\omega_{x,x+b}\in\mathcal{C}_{\infty,\alpha}\}}\big|\chi(\omega,x+b)-\chi(\omega,x)\big| \le \sum_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\le\lambda n}}\sum_{b\in B}\sqrt{\frac{\omega_{x,x+b}}{\alpha}}\big|\chi(\omega,x+b)-\chi(\omega,x)\big|, \tag{7.36}$$

where $B=\{\mathbf{e}_1,\cdots,\mathbf{e}_d,-\mathbf{e}_1,\cdots,-\mathbf{e}_d\}$ is the set of unit vectors. Using the Schwarz and (7.14), we get

$$\mathbb{E}_\alpha[(R_n\mathbf{1}_{\{\lambda(\omega)\le\lambda\}})^2] \le \frac{2d(\lambda n)^d}{\alpha}\mathbb{E}_\alpha\Big[\sum_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\|x|\le\lambda n}}\sum_{b\in B}\omega_{x,x+b}\big|\chi(\omega,x+b)-\chi(\omega,x)\big|^2\Big] \le Cn^{2d} \tag{7.37}$$

for some constant $C=C(\alpha,\lambda,d)<\infty$. Applying Chebyshev's inequality, for $\theta'\in(d,\theta)$,

$$\mathbb{P}_\alpha(R_n\mathbf{1}_{\{\lambda(\omega)\le\lambda\}}\ge n^{\theta'}) \le \frac{C}{n^{2(\theta'-d)}}.$$

Taking $n=2^k$ and using the Borel-Cantelli, we have $R_{2^k}/2^{k\theta'}\le C$ on $\{\lambda(\omega)\le\lambda\}$ for large k, so $R_{2^k}/2^{k\theta}\to 0$ a.e. Since $R_{2^k}/2^{k+1}\le R_n/n\le 2R_{2^{k+1}}/2^{k+1}$ for $2^k\le n\le 2^{K+1}$, we have $R_n/n^\theta\to 0$ a.s. on $\{\lambda(\omega)\le\lambda\}$. □

We will need the following lemma which is easy to prove using the Hölder inequality (see [30, Lemma 4.5] for the proof).

Lemma 7.22 *Let $p>1$ and $r\in[1,p)$. Let $\{X_n\}_{n\in\mathbb{N}}$ be random variables such that $\sup_{j\ge1}\|X_j\|_p<\infty$ and let N be a random variable taking values in positive integers such that $N\in L^s$ for some $s>r(1+\frac{1}{p})/(1-\frac{r}{p})$. Then $\sum_{j=1}^N X_j\in L^r$. Explicitly, there exists $C=C(p,r,s)>0$ such that*

$$\|\sum_{j=1}^N X_j\|_r \le C\big(\sup_{j\ge1}\|X_j\|_p\big)\big(\|N\|_s\big)^{s\left[\frac{1}{r}-\frac{1}{p}\right]},$$

where C is a finite constant depending only on p, r and s.

Proof of Proposition 7.16 (3-1). Let Z be a random variable satisfying the properties (a)–(c). Since $\chi \in L^2_{\rm pot}$, there exists a sequence $\psi_n \in \mathbb{L}^2$ such that

$$\chi_n(\cdot, x) := \psi_n \circ T_x - \psi_n \underset{n\to\infty}{\longrightarrow} \chi(\cdot, x) \qquad \text{in } \mathbb{L}^2(\Omega \times \mathbb{Z}^d).$$

Without loss of generality we may assume that $\chi_n(\cdot, x) \to \chi(\cdot, x)$ almost surely. Since Z is $\mathbb{P}_\alpha$-preserving, we have $\mathbb{E}_\alpha[\chi_n(\cdot, Z(\cdot))] = 0$ once we can show that $\chi_n(\cdot, Z(\cdot)) \in \mathbb{L}^1$. It thus suffices to prove that

$$\chi_n\big(\cdot, Z(\cdot)\big) \underset{n\to\infty}{\longrightarrow} \chi\big(\cdot, Z(\cdot)\big) \qquad \text{in } \mathbb{L}^1. \tag{7.38}$$

Abbreviate $K(\omega) := d^{(\alpha)}_\omega(0, Z(\omega))$ and note that, as in (7.36),

$$\big|\chi_n(\omega, Z(\omega))\big| \le \sum_{\substack{x\in\mathcal{C}_{\infty,\alpha}\\ |x|\le K(\omega)}} \sum_{b\in B} \sqrt{\frac{\omega_{x,x+b}}{\alpha}} \big|\chi_n(\omega, x+b) - \chi_n(\omega, x)\big|.$$

Note that $\sqrt{\omega_{x,x+b}}\,\big|\,\chi_n(\omega, x+b) - \chi_n(\omega, x)\big|\mathbf{1}_{\{x\in\mathcal{C}_{\infty,\alpha}\}}$ is bounded in $\mathbb{L}^2$, uniformly in x, b and n, and assumption (c) shows that $K \in \mathbb{L}^q$ for some $q > 3d$. Thus, by choosing $p = 2$, $s = q/d$ and $N = 2d(2K+1)^d$ in Lemma 7.22 (note that $N \in \mathbb{L}^s$), we obtain,

$$\sup_{n\ge1} \|\chi_n(\cdot, Z(\cdot))\|_r < \infty,$$

for some $r > 1$. Hence, the family $\{\chi_n(\cdot, Z(\cdot))\}$ is uniformly integrable, so (7.38) follows by the fact that $\chi_n(\cdot, Z(\cdot))$ converge almost surely. □

The proof of Proposition 7.16 (2) is quite involved since there are random holes in $\mathcal{C}_{\infty,\alpha}$. We follow the approach in Berger and Biskup [30].

Lemma 7.23 *For $\omega \in \{0 \in \mathcal{C}_{\infty,\alpha}\}$, let $\{x_n(\omega)\}_{n\in\mathbb{Z}}$ be the intersections of $\mathcal{C}_{\infty,\alpha}$ and one of the coordinate axis so that $x_0(\omega) = 0$. Then*

$$\lim_{n\to\infty} \frac{\chi(\omega, x_n(\omega))}{n} = 0, \qquad \mathbb{P}_\alpha\text{-a.s.}$$

Proof. Similarly to (7.25), let $\sigma(\omega) := T_{x_1(\omega)}(\omega)$ denote the "induced" shift. As before, it is standard to prove that σ is $\mathbb{P}_\alpha$-preserving and ergodic (cf. [30, Theorem 3.2]). Further, using Proposition 7.12,

$$\mathbb{E}_\alpha\big(d^{(\alpha)}_\omega(0, x_1(\omega))^p\big) < \infty, \qquad \forall p < \infty. \tag{7.39}$$

Now define $\Psi(\omega) := \chi(\omega, x_1(\omega))$. Using (7.39), we can use Proposition 7.16 (3-1) with $Z(\omega) = x_1(\omega)$ and obtain

$$\Psi \in \mathbb{L}^1(\Omega, \mathbb{P}_\alpha) \quad \text{and} \quad \mathbb{E}_\alpha\Psi(\omega) = 0.$$

Using the cocycle property of χ, we can write

$$\frac{\chi(\omega, x_n(\omega))}{n} = \frac{1}{n}\sum_{k=0}^{n-1} \Psi \circ \sigma^k(\omega)$$

and so the left-hand side tends to zero $\mathbb{P}_\alpha$-a.s. by Birkhoff's ergodic theorem. $\square$

Since the hole is random, we should 'build up' averages over higher dimensional boxes inductively, which is the next lemma. We first give some definition and notation.

Given $K > 0$ and $\epsilon > 0$, we say that a site $x \in \mathbb{Z}^d$ is (K,ϵ)-*good* if $x \in \mathcal{C}_\infty(\omega)$ and

$$\big|\chi(y,\omega) - \chi(x,\omega)\big| < K + \epsilon|x - y|$$

holds for every $y \in \mathcal{C}_{\infty,\alpha}(\omega)$ of the form $y = \ell\mathbf{e}$, where $\ell \in \mathbb{Z}$ and $\mathbf{e}$ is a unit vector. We denote by $\mathcal{G}_{K,\epsilon} = \mathcal{G}_{K,\epsilon}(\omega)$ the set of (K,ϵ)-good sites.

For each $\nu = 1, \dots, d$, let Λ_n^ν be the ν-dimensional box

$$\Lambda_n^\nu = \big\{k_1 e_1 + \cdots + k_\nu e_\nu \colon k_i \in \mathbb{Z},\ |k_i| \le n\ \forall i = 1, \dots, \nu\big\}.$$

For each $\omega \in \Omega$, define

$$\varrho_\nu(\omega) = \lim_{\epsilon \downarrow 0} \limsup_{n\to\infty} \inf_{y \in \mathcal{C}_{\infty,\alpha}(\omega)\cap\Lambda_n^1} \frac{1}{|\Lambda_n^\nu|} \sum_{x \in \mathcal{C}_{\infty,\alpha}(\omega)\cap\Lambda_n^\nu} \mathbf{1}_{\{|\chi(\omega,x)-\chi(\omega,y)|\ge\epsilon n\}}. \tag{7.40}$$

Note that the infimum is taken only over sites in one-dimensional boxΛ_n^1. The idea is to prove $\varrho_\nu = 0$ $\mathbb{P}_\alpha$-a.s. inductively for all $1 \le \nu \le d$.

Our goal is to show by induction that $\varrho_\nu = 0$ almost surely for all $\nu = 1, \dots, d$. The induction step is given the following lemma which is due to [30, Lemma 5.5].

Lemma 7.24 *Let $1 \le \nu < d$. If $\varrho_\nu = 0$, $\mathbb{P}_\alpha$-almost surely, then also $\varrho_{\nu+1} = 0$, $\mathbb{P}_\alpha$-almost surely.*

Proof. Let $p_\infty = \mathbb{P}_\alpha(0 \in \mathcal{C}_{\infty,\alpha})$. Let $\nu < d$ and suppose that $\varrho_\nu = 0$, $\mathbb{P}_\alpha$-a.s.. Fix δ with $0 < \delta < \frac{1}{2}p_\infty^2$. Consider the collection of ν-boxes

$$\Lambda_{n,j}^\nu = \tau_{j\mathbf{e}_{\nu+1}}(\Lambda_n^\nu), \qquad j = 1, \dots, L.$$

Here L is a deterministic number chosen so that

$$\Delta_0 = \big\{x \in \Lambda_n^\nu \colon \exists j \in \{0, \dots, L-1\},\ x + j\mathbf{e}_{\nu+1} \in \Lambda_{n,j}^\nu \cap \mathcal{C}_{\infty,\alpha}\big\}$$

is so large that $|\Delta_0| \ge (1-\delta)|\Lambda_n^\nu|$ once n is sufficiently large.

Choose $\epsilon > 0$ so that

$$L\epsilon + \delta < \frac{1}{2}p_\infty^2. \tag{7.41}$$

Pick $\epsilon > 0$ and a large value $K > 0$. Then for $\mathbb{P}_\alpha$-a.e. ω, all sufficiently large n, and for each $j = 1, \dots, L$, there exists a set of sites$\Delta_j \subset \Lambda_{n,j}^\nu \cap \mathcal{C}_{\infty,\alpha}$ such that

$$\big|(\Lambda_{n,j}^\nu \cap \mathcal{C}_{\infty,\alpha}) \setminus \Delta_j\big| \le \epsilon|\Lambda_{n,j}^\nu|$$

and

$$\big|\chi(x,\omega) - \chi(y,\omega)\big| \le \epsilon n, \qquad x, y \in \Delta_j. \tag{7.42}$$

Moreover, for n sufficiently large, Δ_j could be picked so that $\Delta_j \cap \Lambda_n^1 \neq \emptyset$ and, assuming K large, the non-(K,ϵ)-good sites could be pitched out with little loss of density to achieve even $\Delta_j \subset \mathcal{G}_{K,\epsilon}$. (These claims can be proved directly by the ergodic theorem and the fact that $\mathbb{P}_\alpha(0 \in \mathcal{G}_{K,\epsilon})$ converges to the density of $\mathcal{C}_{\infty,\alpha}$ as $K \to \infty$.) Given $\Delta_1, \dots, \Delta_L$, let Λ be the set of sites in $\Lambda_n^{\nu+1} \cap \mathcal{C}_{\infty,\alpha}$ whose projection onto the linear subspace $\mathbb{H} = \{k_1\mathbf{e}_1 + \cdots + k_\nu \mathbf{e}_\nu \colon k_i \in \mathbb{Z}\}$ belongs to the corresponding projection of $\Delta_1 \cup \cdots \cup \Delta_L$. Note that the Δ_j could be chosen so that $\Lambda \cap \Lambda_n^1 \neq \emptyset$.

By the construction, the projections of the Δ_j's, $j = 1, \dots, L$, onto $\mathbb{H}$ 'fail to cover' at most $L\epsilon|\Lambda_n^\nu|$ sites in Δ_0, and so at most $(\delta + L\epsilon)|\Lambda_n^\nu|$ sites in Λ_n^ν are not of the form $x + i\mathbf{e}_{\nu+1}$ for some $x \in \bigcup_j \Delta_j$. It follows that

$$\big|(\Lambda_n^{\nu+1} \cap \mathcal{C}_{\infty,\alpha}) \setminus \Lambda\big| \le (\delta + L\epsilon)|\Lambda_n^{\nu+1}|, \tag{7.43}$$

i.e., Λ contains all except at most $(L\epsilon + \delta)$-fraction of all sites in $\Lambda_n^{\nu+1}$. Next note that if K is sufficiently large, then for all $1 \le i < j \le L$, $\mathbb{H}$ contains at least $\frac{1}{2}p_\infty^2$-fraction of sites x such that

$$z_i := x + i\mathbf{e}_\nu \in \mathcal{G}_{K,\epsilon} \quad \text{and} \quad z_j := x + j\mathbf{e}_\nu \in \mathcal{G}_{K,\epsilon}.$$

By (7.41), if n is large enough, for each pair (i,j) with $1 \le i < j \le L$ such z_i and z_j can be found so that $z_i \in \Delta_i$ and $z_j \in \Delta_j$. But the Δ_j's were picked to make (7.42) true and so via these pairs of sites we have

$$\big|\chi(y,\omega) - \chi(x,\omega)\big| \le K + \epsilon L + 2\epsilon n \tag{7.44}$$

for every $x, y \in \Delta_1 \cup \cdots \cup \Delta_L$.

From (7.42) and (7.44), we can conclude that for all $r, s \in \Lambda$,

$$\big|\chi(r,\omega) - \chi(s,\omega)\big| \le 3K + \epsilon L + 4\epsilon n < 5\epsilon n, \tag{7.45}$$

provided that $\epsilon n > 3K + \epsilon L$. If $\varrho_{\nu,\epsilon}$ denotes the right-hand side of (7.40) before taking $\epsilon \downarrow 0$, the bounds (7.43) and (7.45) and $\Lambda \cap \Lambda_n^1 \neq \emptyset$ yield

$$\varrho_{\nu+1,5\epsilon}(\omega) \le \delta + L\epsilon,$$

for $\mathbb{P}_\alpha$-a.e. ω. But the left-hand side of this inequality increases as $\epsilon \downarrow 0$ while the right-hand side decreases. Thus, taking $\epsilon \downarrow 0$ and $\delta \downarrow 0$ proves that $\rho_{\nu+1} = 0$ holds for $\mathbb{P}_\alpha$-a.e. ω. □

Proof of Proposition 7.16 (2). First, by Lemma 7.23, we know that $\varrho_1(\omega) = 0$ for $\mathbb{P}_\alpha$-a.e. ω. Using induction, Lemma 7.24 then gives $\varrho_d(\omega) = 0$ for $\mathbb{P}_\alpha$-a.e. ω. Let $\omega \in \{0 \in \mathcal{C}_{\infty,\alpha}\}$. By Lemma 7.24, for each $\epsilon > 0$ there exists $n_0 = n_0(\omega)$ which is a.s. finite such that for all $n \ge n_0(\omega)$, we have $|\chi(x,\omega)| \le \epsilon n$ for all $x \in \Lambda_n^1 \cap \mathcal{C}_{\infty,\alpha}(\omega)$. Using this to estimate away the infimum in (7.40), $\varrho_d = 0$ now gives the desired result for all $\epsilon > 0$. □

7.7.2 Case 2

Recall that for Case 2, we only need to work when $\alpha = 0$, so $\mathcal{C}_{\infty,\alpha} = \mathcal{C}_\infty = \mathbb{Z}^d$ in this case. We follow [16, Section 5]. Let $Q^{(n)}_{x,y} = q^\omega_n(x,y)$. Since the based measure is $\nu_x \equiv 1$, we have $Q^{(n)}_{x,y} = \sum_z Q^{(n-1)}_{x,z} Q^{(1)}_{z,y}$. We first give a preliminary lemma.

Lemma 7.25 *For $G \in \overline{L}^2$, we have*

$$\mathbb{E}[\sum_y Q^{(n)}_{0,y} G(\cdot, y))^2] \le n\|G\|^2_{\overline{L}^2}. \tag{7.46}$$

Proof. Let a_n^2 be the left hand side of (7.46). Then, by the cocycle property, we have

$$\begin{aligned} a_n^2 &= \mathbb{E}[\sum_x \sum_y Q^{(n-1)}_{0,x} Q_{x,y} (G(T_x\cdot, y-x) + G(\cdot, x))^2] \\ &= \mathbb{E}[\sum_x \sum_y Q^{(n-1)}_{0,x} Q_{x,y} \{G(T_x\cdot, y-x)^2 + 2G(T_x\cdot, y-x)G(\cdot, x) + G(\cdot, x)^2\}] =: I + II + III. \end{aligned}$$

Then,

$$\begin{aligned} I &= \mathbb{E}[\sum_x \sum_y Q^{(n-1)}_{-x,0}(T_x\cdot) Q_{0,y-x}(T_x\cdot) G(T_x\cdot, y-x)^2] \\ &= \mathbb{E}[\sum_x \sum_z Q^{(n-1)}_{-x,0}(\cdot) Q_{0,z}(\cdot) G(\cdot, z)^2] = \mathbb{E}[\sum_z Q_{0,z}(\cdot) G(\cdot, z)^2] = \|G\|^2_{\overline{L}^2}, \\ III &= \mathbb{E}[\sum_x \sum_y Q^{(n-1)}_{0,x} Q_{x,y} G(\cdot, x)^2] = \mathbb{E}[\sum_x Q^{(n-1)}_{0,x} G(\cdot, x)^2] = a_{n-1}^2, \\ II &= 2\mathbb{E}[\sum_x \sum_z Q^{(n-1)}_{0,x}(\cdot) Q_{0,z}(T_x\cdot) G(T_x\cdot, z) G(\cdot, x)] \le 2 I^{1/2} II^{1/2} = 2a_{n-1}\|G\|_{\overline{L}^2}. \end{aligned}$$

Thus $a_n \le a_{n-1} + \|G\|_{\overline{L}^2} \le \cdots \le n\|G\|_{\overline{L}^2}$. □

The next lemma proves Proposition 7.16 (1). We use the heat kernel lower bound in the proof.

Lemma 7.26 *Let $G \in \overline{L}^2$.*
(i) For $1 \le p < 2$, there exists $c_p > 0$ such that

$$\mathbb{E}[|G(\cdot, x)|^p]^{1/p} \le c_p |x| \cdot \|G\|_{\overline{L}^2}, \qquad \forall x \in \mathbb{Z}^d. \tag{7.47}$$

(ii) It follows that $\lim_{n\to\infty} \max_{|x|\le n} n^{-d-4}|G(\omega, x)| = 0$, $\mathbb{P}$-a.e. ω.

Proof. Using the cocycle property and the triangle inequality iteratively, for $x = (x_1, \cdots, x_d)$, we have

$$\mathbb{E}[|G(\cdot, x)|^p]^{1/p} \le \sum_{i=1}^d |x_i| \mathbb{E}[|G(\cdot, \mathbf{e}_i)|^p]^{1/p},$$

so it is enough to prove $\mathbb{E}[|G(\cdot,\mathbf{e}_i)|^p] \le c_p\|G\|^p_{\overline{L}^2}$ for $1 \le i \le d$. By Theorem 7.3, there exists an integer valued random variable W_i with $W_i \ge 1$ such that $\mathbb{P}(W_i = n) \le c_1 \exp(-c_2 n^\eta)$ for some $\eta > 0$ and $q^\omega_t(0,\mathbf{e}_i) \ge c_3 t^{-d/2}$ for $t \ge W_i$. Set $\xi_{n,i} = q^\omega_n(0,\mathbf{e}_i)$. Then

$$\mathbb{E}[|G(\cdot,\mathbf{e}_i)|^p] = \sum_{n=1}^\infty \mathbb{E}[|G(\cdot,\mathbf{e}_i)|^p 1_{\{W_i=n\}}].$$

Let $\alpha = 2/p$ and $\alpha' = 2/(2-p)$ be its conjugate. Then, using the heat kernel bound and the Hölder,

$$\begin{aligned}
\mathbb{E}[|G(\cdot,\mathbf{e}_i)|^p 1_{\{W_i=n\}}] &= \mathbb{E}[\xi_n^{1/\alpha}|G(\cdot,\mathbf{e}_i)|^p \xi_n^{-1/\alpha} 1_{\{W_i=n\}}] \le \mathbb{E}[\xi_n G(\cdot,\mathbf{e}_i)^2]^{1/\alpha}\mathbb{E}[\xi_n^{-\alpha'/\alpha}1_{\{W_i=n\}}]^{1/\alpha'}\\
&\le \mathbb{E}[\sum_y Q^{(n)}_{0,y} G(\cdot,y)^2]^{1/\alpha}\Big((c_3 n^{-d/2})^{-\alpha'/\alpha} c_1\exp(-c_2 n^\eta)\Big)^{1/\alpha'}\\
&\le (n\|G\|^2_{\overline{L}^2})^{1/\alpha} c_4 n^{d/(2\alpha)}\exp(-c_5 n^\eta) = c_4 n^{(d+2)/(2\alpha)}\exp(-c_5 n^\eta)\|G\|^p_{\overline{L}^2},
\end{aligned}$$

where we used (7.46) in the last inequality. Summing over $n \ge 1$, we obtain $\mathbb{E}[|G(\cdot,\mathbf{e}_i)|^p] \le c_p\|G\|^p_{\overline{L}^2}$ for $1 \le i \le d$.
(ii) We have

$$\mathbb{P}(\max_{|x|\le n}|G(\cdot,x)| > \lambda_n) \le (2n+1)^d \max_{|x|\le n}\mathbb{P}(|G(\cdot,x)| > \lambda_n) \le \frac{cn^d}{\lambda_n}\max_{|x|\le n}\mathbb{E}[|G(\cdot,x)|] \le \frac{c' n^{d+1}}{\lambda_n}\|G\|_{\overline{L}^2},$$

where we used (7.47) in the last inequality. Taking $\lambda_n = n^{d+3}$ and using the Borel-Cantelli gives the desired result. □

We next prove Proposition 7.16 (3-2). One proof is to apply Proposition 7.16 (3-1) with $Z(\cdot) \equiv x$. Here is another proof.

Lemma 7.27 *For $G \in L^2_{\rm pot}$, it holds that $\mathbb{E}[G(\cdot,x)] = 0$ for all $x \in \mathbb{Z}^d$.*

Proof. If $G = \nabla F$ where $F \in \mathbb{L}^2$, then $\mathbb{E}[G(x,\cdot)] = \mathbb{E}[F_x - F] = \mathbb{E}F_x - \mathbb{E}F = 0$. In general, there exist $\{F_n\} \subset \mathbb{L}^2$ such that $G = \lim_n \nabla F_n$ in $\overline{L}^2$. Since $\mathbb{P}(Q_{0,x} > 0) = 1$ for all $x \in \mathbb{Z}^d$, it follows that $\nabla F_n(\cdot,x)$ converges to $G(\cdot,x)$ in $\mathbb{P}$-probability. By (7.47), for each $p \in [1,2)$, $\{\nabla F_n(\cdot,x)\}_n$ is bounded in $\mathbb{L}^p$ so that $\nabla F_n(\cdot,x)$ converges to $G(\cdot,x)$ in $\mathbb{L}^1$ for all $x \in \mathbb{Z}^d$. Thus $\mathbb{E}[G(\cdot,x)] = \lim_n \mathbb{E}[\nabla F_n(\cdot,x)] = 0$. □

By this lemma, we have $\mathbb{E}[\chi(\cdot,\mathbf{e}_i)] = 0$ where $\mathbf{e}_i$ is the unit vector for $1 \le i \le d$. Using the cocycle property inductively, we have

$$\chi(\omega,n\mathbf{e}_i) = \sum_{k=1}^n \chi(T_{(k-1)\mathbf{e}_i}\omega,\mathbf{e}_i) = \sum_{k=1}^n \Psi\circ T_{(k-1)\mathbf{e}_i},$$

where$\Psi(\omega) := \chi(\omega,\mathbf{e}_i)$. By the proof of Lemma 7.27, we see that$\Psi \in \mathbb{L}^1$. So Birkhoff's ergodic theorem implies

$$\lim_{n\to\infty}\frac{1}{n}\chi(\omega,n\mathbf{e}_i) = 0, \qquad \mathbb{P}\text{-a.e. } \omega. \tag{7.48}$$

Given this, which corresponds to Lemma 7.23 for Case 1, the proof of Proposition 7.16 (2) for Case 2 is the same (in fact slightly easier since $\mathcal{C}_\infty = \mathbb{Z}^d$ in this case) as that of Case 1.

7.8 End of the proof of quenched invariance principle

In this subsection, we will complete the proof of Theorem 7.2 and Theorem 7.4. Throughout this subsection, we assume that α satisfies (7.5) for Case 1 and $\alpha = 0$ for Case 2.

First, we verify some conditions in Theorem 7.15 to prove the sublinearity of the corrector. The heat kernel estimates and percolation estimates are seriously used here.

Lemma 7.28 *(i) The diffusive bounds (7.19), (7.20) hold for $\mathbb{P}_\alpha$-a.e. ω.*
(ii) Let $\tau_n = \{t \ge 0 : |\hat{Y}_t - \hat{Y}_0| \ge n\}$. Then for any $c_1 > 1$, there exists $N_\omega = N_\omega(c_1) > 0$ which is a.s. finite such that $|\hat{Y}_{\tau_n} - \hat{Y}_0| \le c_1 n$ for all $t > 0$ and $n \ge N_\omega$.

Proof. (i) For Case 2, this follows from Theorem 7.3. For Case 1, the proof is given in [32, Section 6]. (See also [90, Section 4]. In fact, in this case both sides Gaussian heat kernel bounds which are similar to the ones for simple random walk on the supercritical percolation clusters hold.)
(ii) For Case 1, this follows directly from Proposition 7.12 and the Borel-Cantelli. For Case 2, using Theorem 7.3, we have for any $z \in \mathbb{Z}^d$ with $|z| \le n$ and $c_1 > 1$,

$$\begin{aligned}\mathbb{P} \times P^z_\omega(|Y_1| \ge c_1 n) &\le \mathbb{P}(U_z \ge c_1 n) + \mathbb{E}[\sum_{z \in B(0,c_1 n)^c \cap \mathbb{Z}^d} q^{\cdot}_1(z,y) : U_z < c_1 n] \\ &\le c_2 \exp(-c_3 n^\eta) + c_4 n^{d-1} \exp(-c_5(c_1-1) n \log n),\end{aligned}$$

where $c_2, \cdots, c_5 > 0$ does not depend on z. So the result follows by the Borel-Cantelli. □

Proof of Theorem 7.2 and Theorem 7.4. By Proposition 7.16 and Lemmas 7.28, the corrector satisfies the conditions of Theorem 7.15. It follows that χ is sublinear on $\mathcal{C}_{\infty,\alpha}$ (for Case 2, since $\alpha = 0$, χ is sublinear on $\mathcal{C}_\infty$). However, for Case 1, by (7.6) the diameter of the largest component of $\mathcal{C}_\infty \setminus \mathcal{C}_{\infty,\alpha}$ in a box $[-2n, 2n]$ is less than $C(\omega) \log n$ for some $C(\omega) < \infty$. Using the harmonicity of φ_ω on $\mathcal{C}_\infty$, the optional stopping theorem gives

$$\max_{\substack{x \in \mathcal{C}_\infty \\ |x| \le n}} \big|\chi(\omega, x)\big| \le \max_{\substack{x \in \mathcal{C}_{\infty,\alpha} \\ |x| \le n}} \big|\chi(\omega, x)\big| + 2C(\omega) \log(2n),$$

hence χ is sublinear on $\mathcal{C}_\infty$ as well.

Having proved the sublinearity of χ on $\mathcal{C}_\infty$ for both Case 1 and Case 2, we proceed as in the $d = 2$ proof of [30]. Let $\{Y_t\}_{t \ge 0}$ be the VSRW and $X_n := Y_n$, $n \in \mathbb{N}$ be the discrete time process. Also, let $\varphi_\omega(x) := x + \chi(\omega, x)$ and define $M_n := \varphi_\omega(X_n)$. Fix $\hat{v} \in \mathbb{R}^d$. We will first show that (the piecewise linearization) of $t \mapsto \hat{v} \cdot M_{[tn]}$ scales to Brownian motion. For $K \ge 0$ and for $m \le n$, let

$$f_K(\omega) := E^0_\omega\big[(\hat{v} \cdot M_1)^2 \mathbf{1}_{\{|\hat{v} \cdot M_1| \ge K\}}\big], \quad \text{and} \quad V_{n,m}(K) := V^{(\omega)}_{n,m}(K) = \frac{1}{n} \sum_{k=0}^{m-1} f_K \circ T_{X_k}(\omega).$$

Since

$$V_{n,m}(\epsilon\sqrt{n}) = \frac{1}{n} \sum_{k=0}^{m-1} E^0_\omega\Big[(\hat{v} \cdot (M_{k+1} - M_k))^2 \mathbf{1}_{\{|\hat{v} \cdot (M_{k+1} - M_k)| \ge \epsilon\sqrt{n}\}} \Big| \mathcal{F}_k\Big],$$

for $\mathcal{F}_k = \sigma(X_0, X_1, \ldots, X_k)$, in order to apply the Lindeberg-Feller functional CLT for martingales, we need to verify the following for $\mathbb{P}_0$-almost every ω (recall $\mathbb{P}_0(\cdot) := \mathbb{P}(\cdot | 0 \in \mathcal{C}_\infty)$):

(i) $V_{n,[tn]}(0) \to Ct$ in P^0_ω-probability for all $t \in [0,1]$ and some $C \in (0,\infty)$,

(ii) $V_{n,n}(\epsilon\sqrt{n}) \to 0$ in P^0_ω-probability for all $\epsilon > 0$.

By Theorem 7.13 (3), $f_K \in \mathbb{L}^1$ for all K. Since $n \mapsto T_{X_n}(\omega)$ (c.f. (7.25)) is ergodic, we have

$$V_{n,n}(K) = \frac{1}{n}\sum_{k=0}^{n-1} f_K \circ T_{X_k}(\omega) \underset{n\to\infty}{\longrightarrow} \mathbb{E}_0 f_K, \tag{7.49}$$

for $\mathbb{P}_0$-a.e. ω and P^0_ω-a.e. path $\{X_k\}_k$ of the random walk. Taking $K = 0$ in (7.49), condition (i) above follows by scaling out the t-dependence first and working with tn instead of n. On the other hand, by (7.49) and the fact that $f_{\epsilon\sqrt{n}} \le f_K$ for sufficiently large n, we have, $\mathbb{P}_0$-almost surely,

$$\limsup_{n\to\infty} V_{n,n}(\epsilon\sqrt{n}) \le \mathbb{E}_0 f_K = \mathbb{E}_0\Big[E^0_\omega\big[(\hat{v}\cdot M_1)^2 \mathbf{1}_{\{|\hat{v}\cdot M_1| \ge K\}}\big]\Big] \underset{K\to\infty}{\longrightarrow} 0,$$

where we can use the dominated convergence theorem since $\hat{v}\cdot M_1 \in L^2$. Hence, conditions (i) and (ii) hold (in fact, even with P^0_ω-a.s. limits). We thus conclude that the following continuous random process

$$t \mapsto \frac{1}{\sqrt{n}}\big(\hat{v}\cdot M_{[nt]} + (nt-[nt])\,\hat{v}\cdot(M_{[nt]+1} - M_{[nt]})\big)$$

converges weakly to Brownian motion with mean zero and covariance $\mathbb{E}_0 f_0 = \mathbb{E}_0 E^0_\omega\big[(\hat{v}\cdot M_1)^2\big]$. This can be written as $\hat{v}\cdot D\hat{v}$ where D is the matrix defined by

$$D_{i,j} := \mathbb{E}_0 E^0_\omega\big((\mathbf{e}_i\cdot M_1)(\mathbf{e}_j\cdot M_1)\big), \qquad \forall i,j \in \{1,\cdots,d\}.$$

Thus, using the Cramér-Wold device and the continuity of the process, we obtain that the linear interpolation of $t \mapsto M_{[nt]}/\sqrt{n}$ converges to d-dimensional Brownian motion with covariance matrix D. Since $X_n - M_n = \chi(\omega, X_n)$, in order to prove the convergence of $t \mapsto X_{[nt]}/\sqrt{n}$ to Brownian motion $\mathbb{P}_0$-a.e. ω, it is enough to show that, for $\mathbb{P}_0$-a.e. ω,

$$\max_{1\le k\le n} \frac{|\chi(X_k,\omega)|}{\sqrt{n}} \underset{n\to\infty}{\longrightarrow} 0, \qquad \text{in } P^0_\omega\text{-probability}. \tag{7.50}$$

By the sublinearity of χ, we know that for each $\epsilon > 0$ there exists $K = K(\omega) < \infty$ such that

$$|\chi(x,\omega)| \le K + \epsilon|x|, \qquad \forall x \in \mathcal{C}_\infty(\omega).$$

Putting $x = X_k$ and substituting $X_k = M_k - \chi(X_k,\omega)$ in the right hand side, we have, if $\epsilon < \frac{1}{2}$,

$$|\chi(X_k,\omega)| \le \frac{K}{1-\epsilon} + \frac{\epsilon}{1-\epsilon}|M_k| \le 2K + 2\epsilon|M_k|.$$

But the above CLT for $\{M_k\}$ and the additivity of $\{M_k\}$ imply that $\max_{k\le n}|M_k|/\sqrt{n}$ converges in law to the maximum of Brownian motion $B(t)$ over $t \in [0,1]$. Hence, denoting the probability law of Brownian motion by P, we have

$$\limsup_{n\to\infty} P^0_\omega\big(\max_{k\le n}|\chi(X_k,\omega)| \ge \delta\sqrt{n}\big) \le P\Big(\max_{0\le t\le 1}|B(t)| \ge \frac{\delta}{2\epsilon}\Big).$$

The right-hand side tends to zero as $\epsilon \to 0$ for all $\delta > 0$. Hence we obtain (7.50). We thus conclude that $t \mapsto \hat{B}_n(t)$ converges to d-dimensional Brownian motion with covariance matrix D, where

$$\hat{B}_n(t) := \frac{1}{\sqrt{n}}\big(X_{[tn]} + (tn - [tn])(X_{[tn]+1} - X_{[tn]})\big), \qquad t \geq 0.$$

Noting that

$$\lim_{n\to\infty} P^0_\omega\Big(\sup_{0\leq s\leq T} \Big|\frac{1}{\sqrt{n}}Y_{sn} - \hat{B}_n(s)\Big| > u\Big) = 0, \qquad \forall u, T > 0,$$

which can be proved using the heat kernel estimates (using Theorem 7.3; see [16, Lemma 4.12]) for Case 2, and using the heat kernel estimates for $\hat{Y}$ and the percolation estimate Proposition 7.12 for Case 1 (see [90, (3.2)]; note that a VSRW version of [90, (3.2)] that we require can be obtained along the same line of the proof), we see that $t \mapsto Y_{tn}/\sqrt{t}$ converges to the same limit.

By the reflection symmetry of $\mathbb{P}_0$, we see that D is a diagonal matrix. Further, the rotation symmetry ensures that $D = (\sigma^2/d)\,\mathrm{I}$ where $\sigma^2 := \mathbb{E}_0 E^0_\omega |M_1|^2$. To see that the limiting process is not degenerate to zero, note that if $\sigma = 0$ then $\chi(\cdot, x) = -x$ holds a.s. for all $x \in \mathbb{Z}^d$. But that is impossible since, as we proved, $x \mapsto \chi(\cdot, x)$ is sublinear a.s.

Finally we consider the CSRW. For each $x \in \mathcal{C}_\infty$, let $\mu_x(\omega) := \sum_y \omega_{xy}$. Let $F(\omega) = \mu_0(\omega)$ and

$$A_t = \int_0^t \mu_{Y_s} ds = \int_0^t F(Z_s) ds.$$

Then $\tau_t = \inf\{s \geq 0 : A_s > t\}$ is the right continuous inverse of A and $W_t = Y_{\tau_t}$ is the CSRW. Since Z is ergodic, we have

$$\lim_{t\to\infty} \frac{1}{t} A_t = \mathbb{E}F = 2d\mathbb{E}\mu_e, \qquad \mathbb{P} \times P^0_\omega\text{-a.s.}$$

So if $\mathbb{E}\mu_e < \infty$, then $\tau_t/t \to (2d\mathbb{E}\mu_e)^{-1} =: M > 0$ a.s. Using the heat kernel estimates for Case 2 (i.e. using Theorem 7.3; see [16, Theorem 4.11]), and using the heat kernel estimates for $\hat{Y}$ and the percolation estimate Proposition 7.12 for Case 1 (i.e. using the VSRW version of [90, (3.2)]), we have

$$\lim_{n\to\infty} P^0_\omega\Big(\sup_{0\leq s\leq T} \Big|\frac{1}{\sqrt{n}}W_{sn} - \frac{1}{\sqrt{n}}Y_{sMn}\Big| > u\Big) = 0, \quad \forall T, u > 0, \quad \mathbb{P}\text{-a.e. } \omega.$$

Thus, $\frac{1}{\sqrt{n}}W_{sn} = \frac{1}{\sqrt{n}}Y_{sMn} + \frac{1}{\sqrt{n}}W_{sn} - \frac{1}{\sqrt{n}}Y_{sMn}$ converges to $\sigma_C B_t$ where $\{B_t\}$ is Brownian motion and $\sigma_C^2 = M\sigma_V^2 > 0$.

If $\mathbb{E}\mu_e = \infty$, then we have $\tau_t/t \to 0$, so $\frac{1}{\sqrt{n}}W_{sn}$ converges to a degenerate process. □

Remark 7.29 *Note that the approach of Mathieu and Piatnitski [91] (also [90]) is different from the above one. They prove sublinearity of the corrector in the L^2-sense and prove the quenched invariance principle using it. (Here L^2 is with respect to the counting measure on $\mathcal{C}_{\infty,\alpha}$ or on $\mathcal{C}_\infty$. In the above arguments, sublinearity of the corrector is proved uniformly on $\mathcal{C}_{\infty,\alpha}$, which is much stronger.) Since compactness is crucial in their arguments, they first prove tightness using the heat kernel estimates – recall that the above arguments do not require a separate proof of tightness. Then, using the Poincaré inequality and the so-called two-scale convergence, weak L^2 convergence of the corrector is proved.*

This together with the Poincaré inequality again, they prove strong L^2 convergence of the corrector (so the L^2-sublinearity of the corrector). In a sense, they deduce weaker estimates of the corrector (which is however enough for the quenched CLT) from weaker assumptions.

Acknowledgement. I would like to thank M.T. Barlow, M. Biskup, J-D. Deuschel, J. Kigami, G. Kozma, P. Mathieu, A. Nachmias and H. Osada for fruitful discussions and comments, A. Sapozhnikov for sending me his preprint [101]. I would also thank G. Ben Arous and F. den Hollander for their encouragements. For the preparations of St. Flour lectures, I had a series of pre-lectures in Kyoto. I would like to thank the audience of the lectures, especially H. Nawa, N. Yoshida, M. Shinoda, M. Takei, N. Kajino, M. Nakashima and D. Shiraishi for useful comments and for finding typos of the draft.

References

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. Preprint 2009. `arXiv:0903.4730`

[2] M. Aizenman and D.J. Barsky. Sharpness of the phase transition in percolation models. Comm. Math. Phys. **108** (1987), 489–526.

[3] M. Aizenman and C.M. Newman. Tree graph inequalities and critical behavior in percolation models. *J. Statist. Phys.* **36** (1984), 107–143.

[4] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** (1997), 812–854.

[5] D. Aldous. The continuum random tree. III. *Ann. Probab.* **21** (1993), 248-289.

[6] D. Aldous and J. Fill. *Reversible Markov chains and random walks on graphs.* Book in preparation. `http://www.stat.berkeley.edu/∼aldous/RWG/book.html`

[7] S. Alexander and R. Orbach. Density of states on fractals: "fractons". *J. Physique (Paris) Lett.*, **43**:L625–L631, (1982).

[8] O. Angel, J. Goodman, F. den Hollander, and G. Slade. Invasion percolation on regular trees. *Ann. Probab.* **36** (2008), 420–466.

[9] P. Antal and A. Pisztora. On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.* **24** (1996), 1036–1048.

[10] M. Barlow. *Random walks on graphs.* Lectures given at PIMS summer school: Lecture notes for a course at RIMS are available at `http://www.math.kyoto-u.ac.jp/~kumagai/rim10.pdf`

[11] M.T. Barlow. Random walks on supercritical percolation clusters. *Ann. Probab.* **32** (2004), 3024-3084.

[12] M.T. Barlow. Diffusions on fractals. Springer, New York, (1998). Lecture Notes in Mathematics **1690**, Ecole d'été de probabilités de Saint-Flour XXV–1995.

[13] M.T. Barlow and J. Černý. Convergence to fractional kinetics for random walks associated with unbounded conductances. *Probab. Theory Relat. Fields*, to appear.

[14] M.T. Barlow, T. Coulhon and A. Grigor'yan. Manifolds and graphs with slow heat kernel decay. *Invent. Math.* **144** (2001), 609–649.

[15] M.T. Barlow, T. Coulhon and T. Kumagai. Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs. *Comm. Pure Appl. Math.* **58** (2005), 1642–1677.

[16] M.T. Barlow and J.-D. Deuschel. Invariance principle for the random conductance model with unbounded conductances. *Ann. Probab.* **38** (2010), 234–276.

[17] M.T. Barlow and B.M. Hambly. Parabolic Harnack inequality and local limit theorem for random walks on percolation clusters. *Electron. J. Probab.* **14** (2009), 1–27.

[18] M.T. Barlow, A.A. Járai, T. Kumagai and G. Slade. Random walk on the incipient infinite cluster for oriented percolation in high dimensions. *Comm. Math. Phys.* **278** (2008), 385–431.

[19] M.T. Barlow and T. Kumagai. Random walk on the incipient infinite cluster on trees. *Illinois J. Math.* **50** (2006), 33–65 (electronic).

[20] M.T. Barlow and R. Masson. Spectral dimension and random walks on the two dimensional uniform spanning tree. Preprint 2009. `arXiv:0912.4765`

[21] M.T. Barlow, Y. Peres and P. Sousi. Collisions of random walks Preprint 2010. `arXiv:1003.3255`

[22] M.T. Barlow and X. Zheng. The random conductance model with Cauchy tails. *Ann. Appl. Probab.*, to appear.

[23] D.J. Barsky and M. Aizenman. Percolation critical exponents under the triangle condition. *Ann. Probab.* **19** (1991), 1520–1536.

[24] G. Ben Arous and J. Černý. Scaling limit for trap models on $\mathbb{Z}^d$. *Ann. Probab.* **35** (2007), 2356–2384.

[25] G. Ben Arous and J. Černý. Bouchaud's model exhibits two aging regimes in dimension one. *Ann. Appl. Probab.* **15** (2005), 1161–1192.

[26] G. Ben Arous, A. Fribergh, N. Gantert and A. Hammond. Biased random walks on Galton-Watson trees with leaves. Preprint 2007. `arXiv:0711.3686`

[27] D. Ben-Avraham and S. Havlin. *Diffusion and reactions in fractals and disordered systems.* Cambridge University Press, Cambridge, 2000.

[28] I. Benjamini, O. Gurel-Gurevich and R. Lyons. Recurrence of random walk traces. *Ann. Probab.* **35** (2007), 732-738.

[29] I. Benjamini and E. Mossel. On the mixing time of simple random walk on the super critical percolation cluster. *Probab. Theory Relat. Fields* **125** (2003), 408-420.

[30] N. Berger, M. Biskup. Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Relat. Fields* **137** (2007), 83–120.

[31] N. Berger, M. Biskup, C.E. Hoffman and G. Kozma. Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann. Inst. H. Poincaré Probab. Statist.* **44** (2008), 374–392.

[32] M. Biskup and T.M. Prescott. Functional CLT for random walk among bounded random conductances. *Electron. J. Probab.* **12** (2007), 1323–1348.

[33] M. Biskup and H. Spohn. Scaling limit for a class of gradient fields with non-convex potentials. *Ann. Probab.*, to appear.

[34] B. Bollobás. *Random graphs.* Cambridge University Press, second edition, 2001.

[35] E. Bolthausen and A-S. Sznitman. *Ten lectures on random media.* DMV Seminar, **32**, Birkhäuser Verlag, Basel, 2002.

[36] O. Boukhadra. Heat-kernel estimates for random walk among random conductances with heavy tail. *Stoch. Proc. Their Appl.* **120** (2010), 182–194.

[37] K. Burdzy and G.F. Lawler. Rigorous exponent inequalities for random walks. *J. Phys. A: Math. Gen.* **23** (1999), L23–L28.

[38] E.A. Carlen and S. Kusuoka and D.W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. Henri Poincaré-Probab. Statist.* **23** (1987), 245-287.

[39] D. Chen and X. Chen. Two random walks on the open cluster of $\mathbb{Z}^2$ meet infinitely often. Preprint 2009.

[40] T. Coulhon. Heat kernel and isoperimetry on non-compact Riemannian manifolds. Contemporary Mathematics **338**, pp. 65–99, Amer. Math. Soc., 2003.

[41] T. Coulhon. Ultracontractivity and Nash type inequalities. *J.* Funct. Anal., **141** (1996), 510–539.

[42] T. Coulhon and L. Saloff-Coste. Variétés riemanniennes isométriques à l'infini. *Rev. Mat. Iberoamericana,* **11** (1995), 687–726.

[43] D.A. Croydon. Scaling limit for the random walk on the giant component of the critical random graph. Preprint 2009.
`http://www2.warwick.ac.uk/fac/sci/statistics/staff/academic/croydon/research`

[44] D.A. Croydon. Random walk on the range of random walk. *J. Stat. Phys.* **136** (2009), 349–372.

[45] D.A. Croydon. Convergence of simple random walks on random discrete trees to Brownian motion on the continuum random tree. *Ann. Inst. H. Poincaré Probab. Statist.* **44** (2008), 987-1019.

[46] D.A. Croydon. Volume growth and heat kernel estimates for the continuum random tree. *Probab. Theory Relat. Fields* **140** (2008), 207–238.

[47] D.A. Croydon and B.M. Hambly. Local limit theorems for sequences of simple random walks on graphs. *Potential Anal.* **29** (2008), 351–389.

[48] D. Croydon and T. Kumagai. Random walks on Galton-Watson trees with infinite variance offspring distribution conditioned to survive. *Electron. J. Probab.* **13** (2008), 1419–1441.

[49] T. Delmotte. Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Math. Iberoamericana* **15** (1999), 181–232.

[50] 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. Statist. Phys.* **55** (1989), 787–855.

[51] P.G. Doyle and J.L. Snell. *Random walks and electric networks*. Mathematical Association of America, 1984. `http://xxx.lanl.gov/abs/math.PR/0001057`.

[52] L.R.G. Fontes, M. Isopi and C.M. Newman. Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension. *Ann. Probab.,* **30** (2002), 579–604.

[53] L.R.G. Fontes and P. Mathieu. On symmetric random walks with random conductances on $\mathbb{Z}^d$. *Probab. Theory Relat. Fields* **134** (2006), 565–602.

[54] I. Fujii and T. Kumagai. Heat kernel estimates on the incipient infinite cluster for critical branching processes. Proceedings of German-Japanese symposium in Kyoto 2006, RIMS Kôkyûroku Bessatsu **B6** (2008), 85–95.

[55] M. Fukushima, Y. Oshima and M. Takeda. *Dirichlet forms and symmetric Markov processes*. de Gruyter, Berlin, 1994.

[56] P. Gábor. A note on percolation on $\mathbb{Z}^d$: isoperimetric profile via exponential cluster repulsion. *Electron. Comm. Probab.* **13** (2008), 377–392.

[57] P.G. de Gennes. La percolation: un concept unificateur. *La Recherche*, **7** (1976), 919–927.

[58] G. Giacomin, S. Olla and H. Spohn. Equilibrium fluctuations for $\nabla\phi$ interface model. *Ann. Probab.* **29** (2001), 1138–1172.

[59] A. Grigor'yan. *Analysis on Graphs*. `http://www.math.uni-bielefeld.de/ grigor/aglect.pdf`

[60] A. Grigor'yan. Heat kernel upper bounds on a complete non-compact manifold, *Rev. Mat. Iberoamericana,* **10** (1994), 395–452.

[61] G. Grimmett. Percolation. 2nd ed., Springer, Berlin, 1999.

[62] G. Grimmett and P. Hiemer. Directed percolation and random walk. *In and out of equilibrium* (Mambucaba, 2000), pp. 273–297, Progr. Probab. **51**, Birkhäuser, Boston, 2002.

[63] G. Grimmett, H. Kesten and Y. Zhang. Random walk on the infinite cluster of the percolation model. *Probab. Theory Relat. Fields* **96** (1993), 33-44.

[64] B.M. Hambly and T. Kumagai. Diffusion on the scaling limit of the critical percolation cluster in the diamond hierarchical lattice. *Comm. Math. Phys.,* **295** (2010), 29–69.

[65] 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, Proc. of Symposia in Pure Math. **72**, Part 2, pp. 233–260, Amer. Math. Soc. 2004.

[66] T. Hara. Decay of correlations in nearest-neighbour self-avoiding walk, percolation, lattice trees and animals. *Ann. Prob.* **36** (2008), 530–593.

[67] T. Hara, R. van der Hofstad and G. Slade. Critical two-point functions and the lace expansion for spread-out high-dimensional percolation and related models. *Ann. Probab.*, **31** (2003), 349–408.

[68] R. van der Hofstad, F. den Hollander, and G. Slade. Construction of the incipient infinite cluster for spread-out oriented percolation above $4+1$ dimensions. *Comm. Math. Phys.*, **231** (2002) 435–461.

[69] R. van der Hofstad and A.A. Járai. The incipient infinite cluster for high-dimensional unoriented percolation. *J. Stat. Phys.* **114** (2004), 625–663.

[70] R. van der Hofstad and G. Slade. Convergence of critical oriented percolation to super-Brownian motion above $4+1$ dimensions. *Ann. Inst. H. Poincaré Probab. Statist.* **39** (2003), 415–485.

[71] B.D. Hughes. *Random walks and random environments*, volume 2: random environments. Oxford University Press, Oxford, 1996.

[72] O.D. Jones. Transition probabilities for the simple random walk on the Sierpinski graph. *Stoch. Proc. Their Appl.* **61** (1996), 45–69.

[73] M. Kanai. Analytic inequalities, and rough isometries between non-compact riemannian manifolds. *Lect. Notes Math.* **1201**, pp. 122–137, Springer, 1986.

[74] M. Kanai. Rough isometries, and combinatorial approximations of geometries of non-compact riemannian manifolds. *J. Math. Soc. Japan* **37** (1985), 391–413.

[75] A. Kasue. Convergence of metric graphs and energy forms. Preprint 2009.

[76] H. Kesten. The incipient infinite cluster in two-dimensional percolation. *Probab. Theory Relat. Fields* **73** (1986), 369–394.

[77] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.* **22** (1986), 425–487.

[78] J. Kigami. Dirichlet forms and associated heat kernels on the Cantor set induced by random walks on trees. *Adv. Math.*, to appear.

[79] J. Kigami. *Analysis on fractals.* Cambridge Univ. Press, Cambridge, 2001.

[80] C. Kipnis and S.R.S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** (1986), 1–19.

[81] S.M. Kozlov. The method of averaging and walks in inhomogeneous environments. *Russian Math. Surveys* **40** (1985), 73–145.

[82] G. Kozma. Percolation on a product of two trees. Preprint 2010. `arXiv:1003.5240`

[83] G. Kozma. The triangle and the open triangle. *Ann. Inst. Henri Poincaré-Probab. Statist.*, to appear.

[84] G. Kozma and A. Nachmias. Personal communications.

[85] G. Kozma and A. Nachmias. Arm exponents in high dimensional percolation. Preprint 2009. `arXiv:0911.0871`

[86] G. Kozma and A. Nachmias. The Alexander-Orbach conjecture holds in high dimensions. *Invent. Math.* **178** (2009), 635–654.

[87] T. Kumagai and J. Misumi. Heat kernel estimates for strongly recurrent random walk on random media. *J. Theoret. Probab.* **21** (2008), 910–935.

[88] R. Lyons and Y. Peres. *Probability on trees and networks.* Book in preparation. `http://mypage.iu.edu/~rdlyons/prbtree/prbtree.html`

[89] T. Lyons. Instability of the Liouville property for quasi-isometric Riemannian manifolds and reversible Markov chains. *J. Differential Geom.* **26** (1987), 33–66.

[90] P. Mathieu. Quenched invariance principles for random walks with random conductances. *J. Stat. Phys.* **130** (2008), 1025–1046.

[91] P. Mathieu and A. Piatnitski. Quenched invariance principles for random walks on percolation clusters. *Proc. Roy. Soc. A* **463** (2007), 2287-2307.

[92] P. Mathieu and E. Remy. Isoperimetry and heat kernel decay on percolation clusters. *Ann. Probab.* **32** (2004), 100–128.

[93] B. Morris and Y. Peres. Evolving sets, mixing and heat kernel bounds. *Probab. Theory Relat. Fields* **133** (2005), 245–266.

[94] J.C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. Henri Poincaré-Probab. Statist.*, to appear.

[95] A. Nachmias and Y. Peres, Critical random graphs: diameter and mixing time. *Ann. Probab.* **36** (2008), 1267–1286.

[96] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. Math. J.*, **80** (1958), 931–954.

[97] J. Norris. *Markov chains.* Cambridge University Press, Cambridge, 1998.

[98] F. Rassoul-Agha and T. Seppäläinen. An almost sure invariance principle for random walks in a space-time random environment. *Probab. Theory Relat. Fields* **133** (2005), 299–314.

[99] C. Rau. Sur le nombre de points visités par une marche aléatoire sur un amas infini de percolation. *Bull. Soc. Math. France* **135** (2007), 135–169.

[100] L. Saloff-Coste. A note on Poincaré, Sobolev, and Harnack inequalities. *Int. Math. Res. Not. IMRN* **2** (1992), 27–38.

[101] A. Sapozhnikov. Upper bound on the expected size of intrinsic ball. Preprint 2010. `arXiv:1006.1521v1`

[102] R. Schonmann. Multiplicity of phase transitions and mean-field criticality on highly non-amenable graphs. *Comm. Math. Phys.* **219** (2001), 271–322.

[103] D. Shiraishi. Exact value of the resistance exponent for four dimensional random walk trace. Preprint 2009.

[104] D. Shiraishi. Heat kernel for random walk trace on $\mathbb{Z}^3$ and $\mathbb{Z}^4$. *Ann. Inst. Henri Poincaré-Probab. Statist.*, to appear.

[105] V. Sidoravicius and A-S. Sznitman. Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Relat. Fields* **129** (2004), 219–244.

[106] G. Slade. *The lace expansion and its applications.* Springer, Berlin, (2006). Lecture Notes in Mathematics **1879**. Ecole d'Eté de Probabilités de Saint–Flour XXXIV–2004.

[107] P. M. Soardi. Potential theory on infinite networks. Lecture Note in Math. **1590**, Springer-Verlag, 1994.

[108] A-S. Sznitman. On the anisotropic walk on the supercritical percolation cluster. *Comm. Math. Phys.* **240** (2003), 123-148.

[109] W. Woess. Random walks on infinite graphs and groups. Cambridge Univ. Press, Cambridge, 2000.

[110] G.M. Zaslavsky. *Hamiltonian chaos and fractional dynamics.* Oxford University Press, Oxford, 2005.

[111] O. Zeitouni. *Random walks in random environment.* Springer, Berlin, (2004). Lecture Notes in Mathematics **1837**, pp. 189–312, Ecole d'Eté de Probabilités de Saint–Flour XXXI–2001.