Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 124 (2014) 3965–3985

stochastic processes and their applications

www.elsevier.com/locate/spa

Resolvent metric and the heat kernel estimate for random walks

András Telcs [a,b,*], Vincenzo Vespri [c]

[a] Department of Quantitative Methods, Faculty of Economics, University of Pannonia, Veszprém, Hungary
[b] Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Magyar tudósok kőrútja 2, H-1117, Budapest, Hungary
[c] Dipartimento di Matematica ed Informatica Ulisse Dini, Universitá degli studi di Firenze Viale Morgagni, 67/a I-50134 Firenze, Italy

Received 17 September 2012; received in revised form 17 June 2014; accepted 12 July 2014
Available online 19 July 2014

Abstract

In this paper we introduce the resolvent metric, the generalization of the resistance metric used for strongly recurrent walks. By using the properties of the resolvent metric we show heat kernel estimates for recurrent and transient random walks.

1. Introduction

This paper is motivated by Kigami's results [10,11] on resistance forms, resistance metric, and several further studies in the same spirit, among others the paper by Barlow et al. [3] and Grigor'yan et al. [7]. All these works present heat kernel estimates under mild conditions thanks to the resistance metric which links the potential theoretic properties of the space to the properties determining the heat propagation. Unfortunately in transient spaces the resistance metric does not carry enough information to describe heat propagation. In the present paper we introduce a new metric, the resolvent metric and apply it to analyze random walks on weighted graphs. The new

* Corresponding author at: Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Magyar tudósok kőrútja 2, H-1117 Budapest, Hungary. Tel.: +36 30 3753896.
E-mail addresses: telcs.szit.bme@gmail.com, telcsrw@gmail.com (A. Telcs), vespri@math.unifi.it (V. Vespri).

http://dx.doi.org/10.1016/j.spa.2014.07.012
0304-4149/

metric has similar properties as the resistance metric and heat kernel estimates can be achieved under similar conditions used in the recurrent case. In that respect the paper is a generalization of [3,7]. It is particularly interesting that in the strongly recurrent case Barlow, Coulhon and Kumagai [3] obtained the elliptic Harnack inequality basically from volume doubling and a resistance estimate:

$$R(x,y) \asymp \frac{d^{\beta}(x,y)}{V(x,r)}.$$

This relation connects the effective resistance $R(x,y)$ between points x, y with distance $r = d(x,y)$, the volume of the ball $V(x,r)$ about x with radius r and the gauge function $d^{\beta}(x,y)$ for the mean exit time $E(x,r)$. The mean exit time, $E(x,r) = \mathbb{E}\left(T_{B(x,r)} | X_0 = x\right)$, is the time needed for the walk to escape from a ball $B(x,r)$ starting the walk at x. Here β plays the role of the walk exponent which is in many cases responsible for the growth of the expected time. Kigami obtains several equivalent conditions for the two-sided heat kernel estimates (c.f. [11] Theorems 15.10, 15.11), among others volume doubling and the Einstein relation. The latter one can be considered as a resistance estimate too:

$$R\left(x, B^{c}(x,r)\right) \asymp \frac{E(x,r)}{V(x,r)}.$$

In what follows we need a similar but weaker assumption, a lower bound for the newly introduced iterated resistance (see below (2) and Definition 3.4), inspired by the condition used in [7] for the standard effective resistance. It might be interesting to observe that condition (2) is necessary and sufficient for the parabolic Harnack inequality and implies the elliptic one (though in an implicit way).

Our main results can be summarized as follows. We consider a weighted graph (Γ, μ) and a random walk on it. We assume that there is a $p_0 > 0$ such that for all $\mu_{x,y} > 0$ we have

$$P(x,y) \geq p_0 \tag{p_0}$$

uniformly, this condition will be assumed in the whole paper and will be referred by (p_0). We construct the resolvent metric ρ based on the iterated resistance $R_m(A,B)$ (for $m > 0$ and A, B disjoint sets), (see Definition 3.4 and Corollary 3.1) and consider $B_\rho(x,r)$, balls in ρ (more precisely, connected subset of them containing x), their volume $V_\rho(x,r)$ and define the scaling function $F(x,r) = \left(r^2 V_\rho(x,r)\right)^{1/m}$ for a well chosen m. Denote $f(x,\cdot) = F^{-1}(x,\cdot)$ and $\widetilde{p}_n(x,y) = p_n(x,y) + p_{n+1}(x,y)$ the sum of the transition kernel.

Definition 1.1. We define a set W_0 of strictly increasing, doubling functions, $F \in W_0$ if $F : \Gamma \times [0,\infty] \to \mathbb{R}^+$ is strictly increasing and there is a $C > 0$ such that for all $x \in \Gamma, r > 0$

$$\frac{F(x,2r)}{F(x,r)} \leq C.$$

Volume doubling (VD), is a particular case of that, means that $V_\rho(x,r) = \mu\left(B_\rho(x,r)\right) \in W_0$.

Definition 1.2. In the whole paper we will use the standing assumption (M) that there is an $m \in \mathbb{Z}^+$ such that for all $x \in \Gamma$

$$\sum_{n=0}^{\infty} n^m p_n(x,x) = \infty. \tag{1}$$

The key notions of the paper are the iterated resistance $R_m(A, B)$ and the resolvent metric $\rho(x, y)$. The proper definition of them needs some preparation hence deferred to Sections 2 and 3. The next definition is adopted from [7].

Definition 1.3. We say that (R_2) holds if there is a $c > 0$ such that for all $x \in \Gamma, r > 0$

$$R_m\left(\{x\}, B_\rho^c(x, r)\right) > cr^2. \tag{2}$$

Here m is the same as in (1) and in the whole sequel.

Theorem 1.1. *If (p_0), volume doubling and (M) hold then there is a $C > 1$ and an $F \in W_0$ such that for all $x \in \Gamma$ and $n > 0$ (DUE), the diagonal upper estimate holds:*

$$p_n(x, x) \leq \frac{C}{V_\rho(x, f(x, n))}. \tag{3}$$

Theorem 1.2. *Assume (p_0) and (M). Volume doubling and (R_2) for μ with respect to ρ hold if and only if there are $C > c > 0, \beta > 1, \delta > 0$ and an $F \in W_0$ such that for all $x, y \in \Gamma$ and $n > 0$ (UE):*

$$p_n(x, y) \leq \frac{C}{V_\rho(x, f(x, n))} \exp\left[-\left(\frac{F(x, \rho(x, y))}{n}\right)^{\frac{m}{\beta-1}}\right] \tag{4}$$

holds and the particular lower estimate (PLE) holds: there are $c, \delta > 0$ such that for all $x \in \Gamma, r > 0, B = B(x, r), n < \delta F(x, r)$

$$\widetilde{p}_n^B(x, y) \geq \frac{C}{V_\rho(x, f(x, n))}. \tag{5}$$

Here p^B is the heat kernel of the random walk killed when exits from B.

We will see that the scaling function is $F(x, r) = \left(r^2 V_\rho(x, r)\right)^{1/m}$.

It is well-known that (5) is the key to obtain heat kernel bounds. Surprisingly enough the upper estimate follows from it as well as the parabolic Harnack inequality (c.f. [7,9]).

To obtain our results we assume (p_0), volume doubling and (M), that is the existence of an m which makes the $\lambda = 0$ resolvent divergent. It is clear that the latter one is a very weak assumption, but essential in the construction of the resolvent metric. The former one is needed only in the very last step in the proof of the diagonal upper estimate. The two-sided estimate needs a bit more. In general two-sided estimates are equivalent with much stronger conditions, like the elliptic Harnack inequality (c.f. [8]) or the cut-off Sobolev inequality (c.f. [2]). Here we use (2), analogous used in several works among others in [7].

The main results are given in Theorems 4.1, 6.1 and 6.2. The paper ends with examples that we hope will be useful to the readers.

2. Basic definitions

We consider (Γ, μ), weighted graph, Γ is a countable infinite set of vertices and $\mu_{x,y} = \mu_{y,x} \geq 0$ a symmetric weight. Edges are formed by the pairs for which $\mu_{x,y} > 0$. We assume that the graph is connected. These weights define a measure on vertices:

$$\mu(x) = \sum_{y \in \Gamma} \mu_{x,y}$$

and on sets $A \subset \Gamma$

$$\mu(A) = \sum_{z \in A} \mu(z).$$

Due to the connectedness $\mu(x) > 0$ for all x. It is natural to define the random walk on weighted graphs, which is a reversible Markov chain given by the one-step transition probabilities:

$$P(x, y) = \frac{\mu_{x,y}}{\mu(x)}.$$

For any $A \subset \Gamma$, P^A stands for the transition probability of the above random walk killed at exiting from A. Similar superscript notation will refer to the corresponding killed walk objects. In what follows we always assume the condition (p_0): there is a constant $p_0 > 0$ such that for all x, y with $\mu_{x,y} > 0$

$$P(x, y) \geq p_0$$

holds. One can define the transition operator P on $c_0(\Gamma)$ functions $Pf(x) = \sum P(x, y) f(y)$. The inner product for $l_2(\Gamma, \mu)$ is defined by $(f, g) = (f, g)_{l_2(\Gamma),\mu} = \sum_x f(x) g(x) \mu(x)$.

If ρ is a metric, balls are defined with respect to it by

$$\widehat{B}_\rho(x, r) = \{y : \rho(x, y) < r\}.$$

Consider the induced subgraph $\Gamma_\rho(x, r)$ of $\widehat{B}_\rho(x, r)$, it contains only the points of $\widehat{B}_\rho(x, r)$ and edges between them. Two points are belonging to the same component of $\Gamma_\rho(x, r)$ if they can be connected with a path running in $\widehat{B}_\rho(x, r)$. Denote by $B = B_\rho(x, r)$ the connected component of $\widehat{B}_\rho(x, r)$ containing x. The volume of the connected part B is denoted by $V_\rho(x, r) = \mu\left(B_\rho(x, r)\right)$.

3. The resolvent metric

Denote by P the transition operator on $l_1(\Gamma)$. $Pf(x) = \sum_{y \sim x} P(x, y) f(y)$.

For a finite $A \subset \Gamma$ let P^A be the transition operator restricted to A, corresponding to the random walk killed on exiting A.

Definition 3.1. The λ-Laplace operator is defined as $\Delta = P - I$. The Dirichlet form corresponding to the Laplace operator is given by

$$\mathcal{E}(f, g) = (-\Delta f, g) = ((I - P) f, g)$$

for $f, g \in l_2(\Gamma, \mu)$.

Definition 3.2. For $A, B \subset \Gamma$, $A \cap B = \emptyset$ we define the resistance

$$R(A, B) = \inf_{f \in l_2(\Gamma,\mu)} [\mathcal{E}(f, f) : f|_A = 1, f|_B = 0]^{-1}.$$

In case of recurrent spaces Kigami's observation (c.f. [10,11]) is that the effective resistance is a metric. The existence of the resistance metric has a particular consequence that, for any f in the domain of the Dirichlet form $\mathcal{E}$

$$|f(x) - f(y)|^2 \leq R(x, y) \mathcal{E}(f, f). \tag{6}$$

Unfortunately in the transient case, the resistance metric is bounded, in particular $R\big(x, B_R^c(x,r)\big) \to R_0 > 0$ as $r \to \infty$ and the rate of convergence can be understood from the decay of $R\left(B_R(x,r), B_R^c(x,2r)\right)$. It is not suitable for the derivation of heat kernel bounds.

In several previous works resolvents are used with success to analyze transient walks and diffusions. The simplest resolvent is the following:

$$\sum_{n=0}^{\infty} n^m (1+\lambda)^{-n} P_n(x,y).$$

For technical reasons we prefer instead the resolvent which is based on $Q^A = \left(P^A\right)^2$ for any $A \subset \Gamma$. We drop the index A for $A = \Gamma$.

$$\sum_{n=0}^{\infty} n^m (1+\lambda)^{-n} Q_n^A(x,y)$$

and our final choice is

$$G_{\lambda,m}^A(x,y) = \sum_{n=0}^{\infty} (1+\lambda)^{-(m+n)} \Omega_m(n) Q_n^A(x,y),$$

where $\Omega_m(n) = \binom{n+m-1}{m-1}$ and it is worth noting that $\Omega_m(n) \simeq n^{m-1}$ and $g_{\lambda,m}^A(x,y) = \frac{1}{\mu(y)} G_{\lambda,m}^A(x,y)$ its kernel.

It is clear that this resolvent is always convergent, and monotonically increasing in m and λ. We fix an $m \in \mathbb{N}$ which will be specified later and reserved as the parameter of the resolvent.

In [8] we started the utilization of polyharmonic functions, Green function as well as Green operators (or resolvents). Now we follow this direction with slight modification and find a new metric for non strongly recurrent graphs (weakly recurrent and transient) which possess nice features.

Definition 3.3. Let $q_n^A(x,y)$ be the transition density corresponding to Q^A. Let us recall that the stationary distribution of Q is μ.

Let $A \subset \Gamma$, $\Delta_\lambda^A = Q^A - I(1+\lambda)$, $D_\lambda^A = -\Delta_\lambda^A$ and

$$D_{\lambda,m}^A = \left(D_\lambda^A\right)^m$$

for $m \geq 1$ integer and $\lambda \geq 0$. If $A = \Gamma$ or $\lambda = 0$ we drop it from the notation.

Consider the bilinear form $\mathcal{D}_{\lambda,m}^A$ on $A \subset \Gamma$

$$\mathcal{D}_{\lambda,m}^A(f,g) = \left(D_{\lambda,m}^A f, g\right)_{l_2(A,\mu)}$$

and define the domain of the λ, m-Dirichlet form on A by $\mathcal{F}_{\lambda,m}^A = \mathcal{F}^A\left(\mathcal{D}_{\lambda,m}^A\right) = \left\{f \in l_2(\Gamma,\mu), \mathcal{D}_{\lambda,m}^A(f,f) < \infty, f|_{A^c} = 0\right\}$. It is clear that $G_{\lambda,m}^A = \left(D_{\lambda,m}^A\right)^{-1}$.

Now we define the iterated resistance setting $\lambda = 0$.

Definition 3.4. For $A, B \subset D \subset \Gamma$, $A \cap B = \varnothing$

$$R_m^D(A,B) = \sup_f \left\{ \frac{|f(x) - f(y)|^2}{\mathcal{D}_m^A(f,f)} : f|_A = 1, f|_B = 0, f \in \mathcal{F}_m^D \right\}.$$

The quasi resolvent metric on A is defined as

$$R_m^A(x,y)=\sup_f\left\{\frac{|f(x)-f(y)|^2}{\mathcal{D}_m^A(f,f)}:f(x)\neq f(y),f\in\mathcal{F}_m^A\right\}.$$

Note that R_m^A is decreasing in A since $\mathcal{D}_m^A$ is increasing by definition, consequently

$$R_m(x,y)=\sup_f\left\{\frac{|f(x)-f(y)|^2}{\mathcal{D}_m(f,f)}:f(x)\neq f(y),f\in\mathcal{F}_m\right\},\tag{7}$$

exists.

Lemma 3.1. 1. *For any $f\in\mathcal{F}(\mathcal{D}_m)$*

$$|f(x)-f(y)|^2\leq R_m(x,y)\,\mathcal{D}_m(f,f).\tag{8}$$

2. *Let $A,B\subset\Gamma$ and $A\cap B=\emptyset$ and belong to the same connected component then,*

$$0<R_m(A,B)<\infty.$$

3. *For any $A\subset\Gamma$, $\lambda\geq 0$ $f,g\in\mathcal{F}_{\lambda,m}^A$*

$$\mathcal{D}_{\lambda,m}^A(f,g)=\mathcal{D}_{\lambda,m}^A(g,f).\tag{9}$$

Proof. The statements follow from the definition. ■

Remark 3.1. It is clear that $g(x)=g_{\lambda,m}^A(x,\cdot)$ possesses the reproducing property: for any $u\in\mathcal{F}_{\lambda,m}^A$ we have that $\mathcal{D}_m(g,u)=u(x)$, in particular $\mathcal{D}_m(g,g)=g_{\lambda,m}^A(x,x)$.

Lemma 3.2. *The minimal value in the definition of $R_{\lambda,m}(x,A^c)$ of $\mathcal{D}_{\lambda,m}(f,f)$ is attained at $g(y)=\frac{1}{g_{\lambda,m}^A(x,x)}g_{\lambda,m}^A(x,y)$ and*

$$R_{\lambda,m}\left(x,A^c\right)=g_{\lambda,m}^A(x,x).\tag{10}$$

Proof. Let h be another function with $h(x)=1$ $h|_{A^c}=0$ and let $d=h-g$ then $h=d+g$,

$$\mathcal{D}_{\lambda,m}(h,h)=\mathcal{D}_{\lambda,m}(g,g)+\mathcal{D}_{\lambda,m}(d,d)+2\mathcal{D}_{\lambda,m}(d,g)$$

but $\mathcal{D}_{\lambda,m}(d,d)\geq 0$ while $\mathcal{D}_{\lambda,m}(d,g)=cd(x)=0$. ■

Lemma 3.3. *If $A\subset B\subset D\subset\Gamma$ then*

$$R_{\lambda,m}\left(A,B^c\right)\leq R_{\lambda,m}\left(A,D^c\right).\tag{11}$$

Proof. 1. By the definition of $D^c\subset B^c$

$$\begin{aligned}R_m^{-1}\left(A,B^c\right)&=\inf\{\mathcal{D}_m(f,f):f|_A(x)=1,\ f|_{B^c}=0\}\\&\geq\inf\{\mathcal{D}_m(f,f):f|_A(x)=1,\ f|_{D^c}=0\}\\&=R_m^{-1}\left(A,D^c\right).\end{aligned}$$ ■

Lemma 3.4. *$R_m(x,y)$ is a quasi metric: for any $x, y \in \Gamma$,*

$$R_m(x,y) = R_m(y,x), \tag{12}$$

$$R_m(x,y) = 0 \quad \text{if and only if } x = y \tag{13}$$

$$R_m(x,y) \le 2\left(R_m(x,z) + R_m(z,y)\right). \tag{14}$$

Proof. The first statement is ensured by the definition. For the second see the end of the proof of [12, Proposition 3.1]. The weak triangular inequality can be seen as follows:

$$\begin{aligned} R_m(x,y) &= \sup_g \left\{ |g(x) - g(y)|^2 : 0 < \mathcal{D}_m(g,g) \right\} \\ &\le \sup_g \left\{ 2|g(x) - g(z)|^2 + 2|g(z) - g(y)|^2 : 0 < \mathcal{D}_m(g,g) \right\} \\ &\le \sup_g \left\{ 2|g(x) - g(z)|^2 : 0 < \mathcal{D}_m(g,g) \right\} \\ &\quad + \sup_g \left\{ 2|g(z) - g(y)|^2 : 0 < \mathcal{D}_m(g,g) \right\} \\ &= 2\left(R_m(x,z) + R_m(z,y)\right). \quad \blacksquare \end{aligned}$$

The next result of Mac´ıas and Segovia [13] is essential in our work.

Theorem 3.1. *If X is a non-empty set and d is quasi-symmetric on it with constant K:*

$$d(x,y) \le K\left(d(x,z) + d(z,y)\right)$$

then, there are a metric ρ, and a $C > 1$ such that for all $x, y \in X$

$$\frac{1}{C}\rho(x,y) \le d^p(x,y) \le C\rho(x,y)$$

with $p = \frac{1}{1+\log_2 K}$.

Corollary 3.1. *There are a metric ρ and $C > 1 > c > 0$ such that for all $x, y \in \Gamma$*

$$c\rho^2(x,y) \le R_m(x,y) \le C\rho^2(x,y). \tag{15}$$

Definition 3.5. In what follows ρ will play a crucial role and is called the resolvent metric.

Based on this theorem we define balls with respect to ρ: $\widetilde{B}_\rho(x,r) = \{y : \rho(x,y) < r\}$. As in the case of the resistance metric it can be that the balls are not connected. Let $B = B(x,r) = B_\rho(x,r) \subset \widetilde{B}_\rho(x,r)$ be the connected subset of $\widetilde{B}_\rho(x,r)$ containing x. With slight abuse of notation we shall use B_R for the sets (balls) with respect to the quasi-metric R_m.

Remark 3.2. We need nesting of balls in the quasi-metric R_m and in ρ, which can be ensured thanks to (15), to the comparability of R_m and ρ.

Lemma 3.5. *Let $r > 0$ and $x \in \Gamma$. There is a $c > 0$ such that for all $x \in \Gamma, r > 0$ if $s = cr^2$ then,*

$$B_R(x,s) \subset B_\rho(x,r).$$

Proof. For a given r we set $s = cr^2$, where c is the constant on the l.h.s. of (15). Let us pick up a vertex $z \in B_R(x,s)$, that means that $R_m(x,z) \leq s$ and z belongs to the connected component of the subgraph of $B_R(x,s)$ which contains x. From the choice of s we have that $c\rho^2(x,z) \leq R_m(x,z) \leq s = cr^2$, i.e. $\rho(x,z) \leq r$. If $z \in B_R(x,s)$, then on one hand $R_m(x,z) \leq s$, and on the other hand z belongs to the connected component of x within $B_R(x,s)$. Let us consider a path, $x_0 = x, x_1, \ldots, x_N = z$ running fully in $B_R(x,s)$ which connects x and z. (Note that this path is not necessarily the shortest path in any metric we use.) It is clear that for all x_i, $R_m(x,x_i) < s$ and x_i connected to x for all $i = 1, \ldots, N$, consequently $\rho(x,x_i) < r$ and the path provide connection from x along itself up to z, which means that z is in the connected component of $\Gamma_\rho(x,r)$ containing x, that is $z \in B_\rho(x,r)$. From that the statement follows. ■

4. The diagonal upper estimate

The following proofs are adaptation of the ones developed in the works [3,4,12,11], all on resistance forms in case of recurrent (or strongly recurrent) spaces, graphs.

A fairly simple but powerful method is developed in the mentioned works (see in particular [4]). The key observation is the following (see for the simplified proof [11, Theorem 10.4]). Without any further assumption for any finite set $A \subset \Gamma$

$$p_n(x,x) \leq \frac{2\overline{R}(x,A)}{n} + \frac{\sqrt{2}}{\mu(A)},$$

where with slight abuse of the notation $\overline{R}(x,A) = \sup_{y\in A} R(x,y)$. Let us introduce analogously $\overline{R}_m(x,A) = \sup_{y\in A} \overline{R}_m(x,y)$. We have the following version of the statement.

Proposition 4.1. *There is a $C > 0$ such that, for any finite set $A \subset \Gamma$ and $x \in A$*

$$p_{2n}(x,x) \leq C\left(\frac{\overline{R}_m(x,A)}{n^m} + \frac{1}{\mu(A)}\right) \tag{16}$$

holds for all $n > 0$.

Before we prove the statement we show how one can obtain the diagonal upper bound from it.

Theorem 4.1. *Assume (p_0), $(VD)_\rho$ and (M) then there is a $C > 1$ such that for all $n \geq 1$, $x \in \Gamma$*

$$p_n(x,x) \leq \frac{C}{V_\rho(x, f(x,n))}, \tag{17}$$

where $f(x,n)$ is the inverse of $F(x,r) = \left[r^2 V_\rho(x,r)\right]^{1/m}$ in the second variable, furthermore

$$p_n(x,y) \leq \frac{C}{\sqrt{V_\rho(x, f(x,n))\, V_\rho(y, f(y,n))}}. \tag{18}$$

Proof. Let $A = B = B_\rho(x,r)$ and choose in (16) r to have $\frac{r^2}{n^m} = \frac{C'}{V_\rho(x,r)}$ and $n = CF(x,r)$ then

$$p_{2\left[C\overline{R}_m(x,B)V_\rho(x,r)\right]^{1/m}}(x,x) \leq \frac{C}{\mu\left(B_\rho(x,r)\right)}$$

and by $(VD)_\rho$

$$p_{2n}(x,x) \le \frac{C}{V_\rho(x, f(x,n))}.$$

This shows the statement for even n, for odd n it can be seen together with (18) as in [8]. In that way Theorem 1.1 is shown. ■

Definition 4.1. We introduce the discrete difference for $f : \mathbb{N} \times \Gamma \to \mathbb{R}, k \ge 1, n \ge 0$

$$f_n^{(1)} = f_n - f_{n+1}$$
$$f_n^{(k)} = f_n^{(k-1)} - f_{n+1}^{(k-1)},$$

where $f_n(\cdot) = f(n,\cdot)$.

Let us recall that $D^A = \left(I - \left(P^A\right)^2\right)$ and for $f_n(y) = q_n(x,y)\ \ x, y \in \Gamma k \ge 1, n \ge 0$

$$D^A f_n = \left(D^A f\right)_n = f_n - f_{n+1}$$
$$\left(D^A\right)^k f_n = f_n^{(k)},$$

and the corresponding bilinear form $\mathcal{D}_k(f_n, f_n) = \left(\left(D^A\right)^k f_n, f_n\right)$. Here and in the next lemmas the case $A = \Gamma$ is also included, the observations will be used for Γ and for finite subsets as well.

Lemma 4.1. *If* $f_n(y) = q_n^A(x,y)$, *then*

$$\mathcal{D}_m(f_n, f_n) = f_{2n}^{(m)}(x).$$

Proof. First observe that $\left(P^A\right)^2 f_n = f_{n+2}$. We show the statement by induction

$$\mathcal{D}_m(f_n, f_n) = f_{2n}^{(m)}(x) \tag{19}$$
$$\mathcal{D}_m(f_{n+1}, f_n) = f_{2n+1}^{(m)}(x). \tag{20}$$

For $m = 1$ consider

$$\begin{aligned}\mathcal{D}_1(f_n, f_n) &= \left(\left[I - Q^A\right] f_n, f_n\right)\\ &= (f_n, f_n) - (f_{n+1}, f_n)\\ &= f_{2n} - f_{2n+1} = D^A f_{2n} = f_{2n}^{(1)},\end{aligned}$$

and with the same steps it is easy to see that (20) holds as well for $m = 1$. We assume that both (19), (20) hold for all $n \ge 1$ and for $k \le m$ and show for $m + 1$.

$$\begin{aligned}\mathcal{D}_{m+1}\left(f_{n,} f_n\right) &= \left(\left(I - Q^A\right)^{m+1} f_n, f_n\right) = \left(\left(I - Q^A\right)^m \left(I - P^2\right) f_n, f_n\right)\\ &= \left(\left(I - Q^A\right)^m f_n, f_n\right) - \left(\left(I - Q^A\right)^m Q^A f_n, f_n\right)\\ &= f_{2n}^{(m)} - f_{2n+1}^{(m)}(x)\\ &= f_{2n}^{(m+1)}.\end{aligned}$$

A similar trick leads to the proof of (20) for $m + 1$. ■

Lemma 4.2. *For all $k \geq 0$ there is a $C > 0$ such that for all $A \subset \Gamma, k \geq 0, n > 0, x \in \Gamma$, if $f_n(y) = q_n^A(x, y)$ then*

$$f_{2n}^{(k)}(x) \leq C \frac{1}{n^k} f_n(x) . \tag{21}$$

Proof. From the spectral decomposition of $p_{2n}^A(x, x)$ for any finite $A' \subset \Gamma$ we know that $\left[p_{2n}^{A'}(x, x)\right]^{(k)} \geq 0$; consequently it holds in the limit $A' \to A$; $f_n^{(k)}(x) \geq 0$ for all $k \geq 0$ and the same implies that the map $n \to \left(f_n^{(k-1)} - f_{n+1}^{(k-1)}\right)(x)$ is non-increasing. We show the statement by induction using a slightly stronger statement. Assume it holds for all $0 \leq i < k$,

$$f_{2n}^{(i)} \leq \frac{1}{(\lfloor s_i n \rfloor)^i} f_{n-\lfloor s_i n \rfloor}(x),$$

where $c_j = 2^{-(j+1)}$, $s_k = s_{k-1} + c_k$, and note that for $i = 0$ the assumption holds.

$$\begin{aligned} f_{2n}^{(k)} &= \left[f_{2n}^{(k-1)} - f_{2n+1}^{(k-1)}\right] \\ &\leq \frac{1}{2\lfloor c_k n \rfloor} \sum_{i=0}^{\lfloor c_k n \rfloor} \left[f_{2n-i}^{(k-1)} - f_{2n+1-i}^{(k-1)}\right] \\ &\leq \frac{1}{2\lfloor c_k n \rfloor} \left[f_{2n-\lfloor c_k n \rfloor}^{(k-1)} - f_{2n+1}^{(k-1)}\right] \\ &\leq \frac{1}{2\lfloor c_k n \rfloor} f_{2n-\lfloor c_k n \rfloor}^{(k-1)}. \end{aligned}$$

Now by induction, if $l = 2n - \lfloor c_k n \rfloor$

$$f_{2n}^{(k)} \leq \frac{1}{\lfloor c_k n \rfloor} f_l^{(k-1)} \leq \cdots \leq \frac{1}{\lfloor c_k n \rfloor} \frac{1}{(\lfloor s_{k-1} n \rfloor)^{k-1}} f_{2n-\lfloor s_{k-1} n \rfloor}(x) .$$

Let us recall that $f_n(x)$ is non-increasing in n and find that

$$2n - \lfloor s_{k-1} n \rfloor \geq 2n - \lfloor s_k n \rfloor ,$$

which leads to the needed inequality.

$$f_{2n}^{(k)} \leq \frac{1}{(2\lfloor s_k n \rfloor)^k} f_{2n-2\lfloor s_k n \rfloor}(x) .$$

Finally observing that $s_k = \sum_{i=0}^k 2^{-(i+2)}$ we have that $s_k < \frac{1}{2}$ and we obtain (21). ■

Proof of Proposition 4.1. Let $x \in A \subset \Gamma$ be a finite, connected set. Let $A_2 = \left\{z \in A : p_{2n}(x, z) > 0\right\}$ and choose y^* so that

$$\begin{aligned} p_{2n}\left(x, y^*\right) &:= \min_{y \in A_2} p_{2n}(x, y) \\ p_{2n}\left(x, y^*\right) \sum_{z \in A_2} \mu(z) &\leq \sum_{z \in A_2} p_{2n}(x, z) \mu(z) \\ &\leq \sum_{z \in \Gamma} P_{2n}(x, z) \leq 1, \end{aligned}$$

and

$$p_{2n}\left(x, y^*\right) \le \frac{1}{\mu\left(A_2\right)}.$$

Now we use the consequence (see [14, Proposition 2.1]) of the condition (p_0) that each vertex has a bounded neighborhood and for $x \sim y, \mu(x) \asymp \mu(y)$ from which we have that $\mu(A) \le C\mu(A_2)$ and

$$p_{2n}\left(x, y^*\right) \le \frac{C}{\mu(A)}.$$

For $f_n(y) = q_n(x, y)$ using the above lemmas we have that

$$\begin{aligned}
\frac{1}{2} f_n^2(x) &\le f_n^2\left(y^*\right) + \left|f_n(x) - f_n\left(y^*\right)\right|^2 \\
&\le \frac{C}{\mu^2(A)} + \overline{R}_m(x, A)\,\mathcal{D}_m\left(f_n, f_n\right) \\
&\le \frac{1}{\mu^2(A)} + \overline{R}_m(x, A)\,\frac{C}{n^m} f_n(x, x),
\end{aligned}$$

where in the last step Lemma 4.2 is used. Solving this for $f_n(x, x)$ we obtain

$$f_n(x) \le C_1 \frac{\overline{R}_m(x, A)}{t^m} + \left(\frac{C}{\mu^2(A)} + C_2 \frac{\overline{R}_m(x, A)}{t^{2m}}\right)^{1/2} \tag{22}$$

$$\le C\left(\frac{\overline{R}_m(x, A)}{n^m} + \frac{C}{\mu(A)}\right). \quad \blacksquare \tag{23}$$

5. The tail distribution of the exit time

This section contains two key results. One establishes an estimate similar to the Einstein relation, the other presents the estimate of the tail distribution of the exit time. In the sequel of the paper, the condition (R_2) (see (2)) will be used in several places. For the rest of the paper we need the additional assumption that (Γ, ρ) is an unbounded metric space, more precisely the following holds.

Condition 5.1. *In the whole sequel we assume that (Γ, μ) is unbounded and for all x and $r > 0$ the balls $B_\rho(x, r)$ are finite, and contain finite number of vertices.*

For brevity we will use the following notations:

$$\begin{aligned}
&\Omega_m(n) = \binom{n+m-1}{m-1}, \\
&E_m(A|x) = \mathbb{E}\left(\Omega_{m+1}\left(T_A\right)|X_0 = x\right), \\
&\overline{E}_m(A) = \max_{x\in A} \mathbb{E}\left(\Omega_{m+1}\left(T_A\right)|X_0 = x\right), \\
&E_m(x, r) = \mathbb{E}\left(\Omega_{m+1}\left(T_{B_\rho(x,r)}\right)|X_0 = x\right).
\end{aligned}$$

We will use the particular notation for $m = 1$,

$$E_\rho(x, r) = \mathbb{E}\left(T_{B_\rho(x,r)}|X_0 = x\right)$$

is the usual mean exit time, where ρ emphasizes the metric in use.

The scaling functions are $H(x,r) = r^2 V_\rho(x,r)$ and $F(x,r) = [H(x,r)]^{1/m}$.

Please note that the definition of E_m is based via Ω_{m+1} on T^m and not on an $(m-1)$-th power.

5.1. Preliminary estimates

First we need a little technical adjustment. We defined the exit time with respect to P while the resolvent is based on $Q = P^2$. The next Lemma shows that if we estimate the mean exit time, that does not make significant difference. Denote by $\mathbb{E}^P, \mathbb{E}^Q$ the expected value and the exit times $T^P_{B(x,r)}, T^Q_{B(x,r)}$ with respect to P and Q.

Lemma 5.1. *Let $A \subset \Gamma$ be a finite set, $x \in A$ then*

$$T^P_A \leq 2T^Q_A \tag{24}$$

and

$$\mathbb{E}^P\left(T^P_A | X_0 = x\right) \geq \mathbb{E}^Q\left(T^Q_A | X_0 = x\right). \tag{25}$$

Proof. If $X^Q_k = X_{2k}$ then X^Q_k has the transition probability Q. Let $T^Q_A = k$, then it is clear that $T^P_A \leq 2k$. For (25) consider the decomposition of the mean exit time:

$$\begin{aligned}\mathbb{E}^P\left(T^P_A | X^P_0 = x\right) &= \sum_{n=0}^{\infty}\sum_{y\in A} P^A_{2n}(x,y) + P^A_{2n+1}(x,y)\\ &\geq \sum_{n=0}^{\infty}\sum_{y\in A} P^A_{2n}(x,y) = \mathbb{E}^Q\left(T^P_A | X^Q_0 = x\right). \quad \blacksquare\end{aligned}$$

Remark 5.1. Having Lemma 5.1 we shall use interchangeably $\mathbb{E}(T_A | X_0 = x)$ for P and Q if the arguments are not sensitive to the constant multiplier.

Theorem 5.1. *If (Γ, μ) satisfies (p_0), (VD), (M) and (R_2) then, $(ER)_m$:*

$$E_m(x, 2r) \asymp R_m(x, 2r)\, V_\rho(x, 2r) \tag{26}$$

holds, where $B = B_\rho(x,r)$, $E_m(x,r) = \mathbb{E}_m(B|x)$ and $R_m(x,r) = R_m\left(x, B^c_\rho(x,r)\right)$.

From now on the notation $R_m(x,r) = R_m(x, B^c) = R_m\left(x, B^c_\rho(x,r)\right)$ will be used depending on the context.

Let us recall that $\mathbb{E}_m(B|x) = \mathbb{E}\left(\Omega_{m+1}(T_B) | X_0 = x\right)$.

Also let us note that we need that m is an integer for which the sum (1) is infinite. That condition, (1), is needed to ensure, that the resolvent provides enough information on the asymptotic of the heat kernel.

The first lemma is elementary.

Lemma 5.2. *Let $B = B_\rho(x,r)$, $T = T_B$, $m \in \mathbb{Z}^+$ then*

$$\mathbb{E}_x\left(\Omega_{m+1}(T)\right) \asymp \mathbb{E}_x\left(T^m\right).$$

Proof. Let $T = T_{B_\rho(x,r)}$. Assume that r is large enough to ensure $T > 2m$ i.e. $B_d(x, 2m+1) \subset B_\rho(x, r)$ and obtain

$$\frac{(2T)^m}{m!} \geq \frac{(T+m)^m}{m!} \geq \Omega_{m+1}(T) \geq \frac{(T-m)^m}{m!} \geq c\left(\frac{T-m}{m}\right)^m \geq \left(\frac{T}{2m}\right)^m.$$

For small r values the inequality follows by adjusting the constants. ■

Of course the statement holds for arbitrary finite set as well.

Lemma 5.3 (*Feynman–Kac Formula, c.f. [8] or [14, p. 102. Section 8.5.3]*)**.** *Let f be a function on Γ, $A \subset \Gamma$, $\lambda > 0$, satisfying*

$$\Delta f - \lambda f = 0 \quad \textit{in } B.$$

Then for any $x \in A$, $\omega = (1+\lambda)^{-1}$, $T = T_A$

$$f(x) = \mathbb{E}_x\left[\omega^T f(X_T)\right]$$

and for any $m \geq 0$

$$G^A_{\lambda,m} f(x) = \left(\left[\sum \omega_m^{n+m} \Omega_{m+1}(n) P^A_n\right] f\right)(x) = \mathbb{E}_x\left(\Omega_{m+1}(T)\,\omega^{T+m} f(X_T)\right). \tag{27}$$

Corollary 5.1. *If we choose $f \equiv 1$ we have from 5.3 that*

$$E_m(A|x) = \mathbb{E}_x(\Omega_{m+1}(T_A)) = \sum_{y \in B} G^A_m(x, y). \tag{28}$$

Proof of Theorem 5.1. Denote $B = B_\rho(x, 2r)$. We start with (8): If $f \in \mathcal{F}(\mathcal{D}_m)$

$$|f(x) - f(y)|^2 \leq R_m(x, y)\,\mathcal{D}_m(f, f)$$

in particular let $g(z) = g^B_m(x, z)$ and $z \in B_\rho(x, r)$ then

$$|g(x) - g(z)|^2 \leq R_m(x, z)\,\mathcal{D}_m(g, g). \tag{29}$$

From the reproducing property of $g^B_m(x, z)$ and (2) we have that $\mathcal{D}_m(g, g) = g^B_m(x, x) = R_m(x, B^c)$

$$|g(x) - g(z)|^2 \leq R_m(x, z)\,g(x) \leq Cr^2 g(x) \leq Cg^2(x), \tag{30}$$

where in the last step (2) has been used. If $g(y^*) = \max_{z \in B} g(z)$, (30) yields

$$g\left(y^*\right) \leq Cg(x) \tag{31}$$

and $E_m(x, 2r) = \sum_{y \in B} g^B_m(x, y)\,\mu(y) \leq g(y^*)\,V_\rho(x, r) \leq CR_m(x, B^c)\,V_\rho(x, r)$.

The proof of the lower estimate is similar. One will find that $(VD)_\rho$ is used once, only at the very end of the proof and condition (2) is essential in the next argument. We proceed from (29). From the reproducing property of $g^B_m(x, z)$ and from (R_2) we have that $\mathcal{E}_m(g, g) = g^B_m(x, x)$

$= R_m(x, B^c) \geq cr^2$

$$|g(x) - g(z)|^2 \leq C\delta^2 r^2 g(x) \leq C\delta^2 g(x)^2 .$$

We can choose δ such that $C\delta^2 \leq 2$, and we obtain from $g(z) \leq g(x)$, that for $z \in B_\rho(x, \delta r)$

$$g_m^B(x, z) \geq \frac{1}{2} g_m^B(x, x) .$$

Now we finish immediately using the definition and $(VD)_\rho$.

$$\begin{aligned} E_m(x, 2r) &= \sum_{y \in B} g_m^B(x, y)\, \mu(y) \geq \sum_{z \in B_\rho(x, \delta r)} g_m^B(x, z)\, \mu(z) \\ &\geq \frac{1}{2} g_m^B(x, x)\, V_\rho(x, \delta r) \\ &\geq c R_m\left(x, B_\rho^c(x, 2r)\right) V_\rho(x, 2r) , \end{aligned}$$

where the last step follows from (2), $cr^2 \leq R_m\left(x, B_\rho^c(x, 2r)\right) \leq 4r^2$ and $(VD)_\rho$. ■

Lemma 5.4. *For a set $A \subset \Gamma$, $x \in A$, there is a $C_0 > 1$ depending on $m \in \mathbb{Z}^+$, such that*

$$\mathbb{P}_x(T_A < n) \leq 1 - \frac{E_m(A)}{C_0 \overline{E}_m(A)} + \frac{C_0 n^m}{\overline{E}_m(A)} . \tag{32}$$

Proof.

$$\begin{aligned} &T_A \leq 2n + I(T_A > n)\, T_A \circ \Theta_n , \\ &T_A^m \leq 2^m \left((2n)^m + I(T_A > n)\, T_A^m \circ \Theta_n\right) , \end{aligned}$$

where Θ_n is the time shift operator. From the strong Markov property one obtains with $C = 2^{\lceil m \rceil}$

$$\begin{aligned} E_m(A) &\leq C^2 n^m + C\mathbb{E}_x\left(I(T_A > n)\, \mathbb{E}_{X_n}\left(T_A^m\right)\right) \\ &\leq C^2 n^m + C\mathbb{P}_x(T_A > n)\, \overline{E}_m(A) . \\ \frac{E_m(A)}{C\overline{E}_m(A)} &\leq \frac{Cn^m}{\overline{E}_m(A)} + \mathbb{P}_x(T_A > n) \end{aligned}$$

and the statement follows. ■

Let us recall here that under $(VD)_\rho$ the scaling function $H(x, r) = r^2 V_\rho(x, r)$ has nice regularity properties.

Corollary 5.2. *If (p_0), $(VD)_\rho$, (M) and (R_2) hold then there are $c_0 > 0$ and $C_0 > 1$ such that if $n = \left(\frac{1}{2} C_0^{-2} E_m(x, r)\right)^{1/m}$*

$$\mathbb{P}_x\left(T_{B_\rho(x,r)} \geq n\right) \geq c_0 . \tag{33}$$

Here C_0 is given by Lemma 5.4.

Proof. From (32) follows the statement if we have that

$$\overline{E}_m\left(B_\rho(x, r)\right) \leq C E_m(x, r) . \tag{34}$$

Let $x_0 \in B(x,r)$ such that $\overline{E}_m\left(B_\rho(x,r)\right) = E_m(x_0,r)$ then it is clear that

$$\overline{E}_m\left(B_\rho(x,r)\right) = E_m\left(B_\rho(x,r)|X_0 = x_0\right) \le E_m\left(B(x_0,2r)|X_0 = x_0\right)$$

but from Theorem 5.1 we have that

$$E_m\left(B(x,2r)|X_0 = x\right) \asymp R_m(x,2r)\, V_\rho(x,2r) \asymp R_m(x,r)\, V(x,r) \asymp E_m(x,r)$$

and we obtain (34). Let us observe that the conditions $(VD)_\rho$, (M) and (R_2) are needed in the application of Theorem 5.1. ■

Theorem 5.2. *If (Γ,ρ) satisfies (p_0), $(VD)_\rho$, (M) and (R_2) then, for $B = B_\rho(x,r)$*

$$\mathbb{P}_x(T_B < n) \le C\exp\left(-ck_m(x,n,r)\right) \tag{35}$$

where $k = k_m(x,n,r) > 1$ is the maximal integer for which

$$\frac{n^m}{k} \le c \min_{y\in B_\rho(x,r)} E_m\left(B_\rho\left(y,\frac{r}{k}\right)\right), \tag{36}$$

where c is a small fixed constant (c.f. [14, p. 72. Definition 6.1]).

Definition 5.1. Let us define $\beta = \beta_m$ as the smallest possible exponent for which

$$\frac{R^2 V_\rho(x,R)}{r^2 V_\rho(x,r)} \le C\left(\frac{R}{r}\right)^{\beta_m}, \tag{37}$$

and observe that (37) is equivalent to $(VD)_\rho$.

From the definition of β and $(VD)_\rho$ it follows that $\beta > 2$.

Remark 5.2. There are several further equivalent forms of (35). In the simplest case if $\left[r^2 V_\rho(x,r)\right] \asymp r^\beta$, $B = B_\rho(x,r)$ one has

$$\mathbb{P}_x(T_B < n) \le C\exp\left(-c\left(\frac{r^\beta}{n^m}\right)^{\frac{1}{\beta-1}}\right). \tag{38}$$

Remark 5.3. From (38) one can see that the estimate is weaker as m increases. However it should be recognized that the increase of m not only increases the upper bound but the probability on the left hand side of (38).

Proof of Theorem 5.2. The proof follows the old, nice idea of [1] (see also [5, Lemma 3.14]). The only modification is that we use the very rough estimate:

$$T^m_{B_\rho(x,r)} \ge \sum_{i=1}^{k} \tau_i^m$$

where τ_i is the exit time of $B_\rho\left(\xi_i,\frac{r}{k}\right)$, $\xi_i = X_{\tau_{i-1}}$ and $k \ge 1$ will be chosen later. From Lemma 5.4 we have that with $t = \frac{n}{k}$

$$P(\tau < t) \le p + at^m \tag{39}$$

where $p \in \left[\frac{1}{2}, 1-\varepsilon\right]$ and $a = \frac{2^m}{\overline{E}_m(x,\frac{r}{k})}$. Let η be such that $P(\tau < t) = (p + at^m) \wedge 1$. The relation (39) can be rewritten as

$$P\left(\tau^m < s\right) \le p + as$$
$$\mathbb{E}\left(\exp\left(-\lambda\tau^m\right)\right) \le \mathbb{E}\left(\exp\left(-\lambda\eta^m\right)\right) \le p + a\lambda^{-1}.$$

From that point the proof can be finished as in [5]. ∎

Remark 5.4. From the definitions, from Theorem 5.2, $(VD)_\rho$ and $(ER)_\rho$ it is immediate that

$$\frac{n^m}{k+1} \ge q \min_{y\in B_\rho(x,r)} E_m\left(B_\rho\left(y,\frac{r}{k}\right)\right) \tag{40}$$

$$\ge cq \min_{y\in B_\rho(x,r)} E_m\left(B_\rho(y,r)\right)k^{-\beta_m}, \tag{41}$$

$$(k+1)^{\beta_m - 1} \ge cq\frac{E_m\left(B_\rho\left(\underline{y},r\right)\right)}{n^m}, \tag{42}$$

$$k+1 \ge c\left(\frac{E_m(B)}{n^m}\right)^{\frac{1}{\beta_m-1}}, \tag{43}$$

where $B = B(x,r)$, which yields

$$\mathbb{P}_x(T_B < n) \le C\exp\left(-c\left(\frac{E_m(B)}{n^m}\right)^{\frac{1}{\beta_m-1}}\right), \tag{44}$$

$$\mathbb{P}_x(T_B < n) \le C\exp\left(-c\left(\frac{H(x,r)}{n^m}\right)^{\frac{1}{\beta_m-1}}\right). \tag{45}$$

5.2. *The Einstein relation*

The relation between the mean exit time of a ball, its volume and resistance is regarded as a key tool to obtain heat kernel estimates. In this section we obtain the corresponding relation with respect to the distance ρ.

Theorem 5.3. *If (Γ,μ) satisfies (p_0), $(VD)_\rho$, (M) and (R_2) then, it satisfies the Einstein relation, $(ER)_\rho$:*

$$E_\rho(x,2r) \asymp \left[R_m\left(x,B^c\right)V_\rho(x,2r)\right]^{1/m} \tag{46}$$

with $B = B_\rho(x,2r)$, $E_\rho(x,r) = \mathbb{E}_m\left(B_\rho(x,r)\,|X_0 = x\right)$.

Corollary 5.3. *Under the same conditions*

$$E_\rho(x,2r) \asymp \left[r^2 V_\rho(x,2r)\right]^{1/m}. \tag{47}$$

Theorem 5.3 will follow from Theorem 5.1, the next statement and from the tail estimate (38) of the exit time.

Lemma 5.5. *Under the same conditions of Theorem 5.3*

$$E_\rho(x,r) \asymp E_m(x,r)^{1/m}.$$

Proof. Let $B = B_\rho(x,r)$, $T = T_B$ From the Jensen inequality we obtain that for $m \geq 1$

$$E_m(x,r) = \mathbb{E}_x\left(\Omega_{m+1}\left(T_{B_R(x,r)}\right)\right) \asymp \mathbb{E}_x\left(T^m\right) \geq \left[E_\rho(x,r)\right]^m.$$

For the opposite estimate denote $E = E_m(x,r)$ and

$$\begin{aligned} E_\rho(x,r) &= \sum_n P(T>n) \geq \sum_{n=c_0(E)^{1/m}}^{2c_0(E)^{1/m}} P(T>n) \\ &\geq c_0 E^{1/m} P\left(T_B > 2c_0 E^{1/m}\right). \end{aligned}$$

Now we use Theorem 5.2, in particular (45)

$$\mathbb{P}_x(T<n) \leq C\exp\left(-c\left(\frac{H(x,r)}{n^m}\right)^{\frac{1}{\beta-1}}\right). \tag{48}$$

Given that we assume that $B(x,r)$ is the connected component containing x it and finite as assumed (c.f. Condition 5.1), it follows that $\mathbb{P}_x(T<\infty) = 1$, that means that

$$\begin{aligned} \mathbb{P}_x(T \geq n) &= 1 - \mathbb{P}_x(T<n) \\ &\geq 1 - C\exp\left(-c\left(\frac{H(x,r)}{n^m}\right)^{\frac{1}{\beta-1}}\right) \geq 1/2 \end{aligned}$$

if we chose $n^m \geq H(x,r)$ and c_0 such that $\log C - c\left(\frac{1}{2c_0}\right)^{\frac{1}{\beta-1}} = 1/2$ i.e. $c_0 = \frac{1}{2}\left(\frac{c}{\log C - 1/2}\right)^{\beta-1}$, the proof is complete. ■

Proof of Theorem 5.3. The statement is immediate from Lemmas 5.2, 5.5 and Theorem 5.1. ■

6. Two-sided estimates

In this section we show that $(VD)_\rho$, (R_2) implies the off-diagonal upper and near diagonal lower estimates.

6.1. The off-diagonal upper estimate

The off-diagonal estimate can be easily obtained from the diagonal one.

Theorem 6.1. *Assume* (p_0), $(VD)_\rho$, (M) *and* (R_2) *and* $(DUE)_F$ *then*

$$\begin{aligned} p_n(x,y) &\leq \frac{C}{V(x,f(x,n))}\exp(-ck(x,n,r)) \\ &\leq \frac{C}{V_\rho(x,f(x,n))}\exp\left(-c\left(\frac{F_\rho(x,d(x,y))}{n}\right)^{\frac{m}{\beta m-1}}\right). \end{aligned}$$

The proof is word by word the same as for Theorem 8.5 *in* [14, *p. 110*] *or an alternative proof is a combination of Theorems* 8.6 *and* 8.10 *in* [14, *pp. 113, 129*].

6.2. Lower estimates

It is standard to deduce a diagonal lower estimate and its stronger form from (33) (see also in [14, p. 73 Theorem 6.2 and p. 76 Remark 6.1]).

Proposition 6.1. *If (33) holds there is a $c > 0$ such that for all $n : \varepsilon F \geq n > 0, R \geq f(x, 2n), B = B_\rho(x, r)$*

$$p_{2n}^{B}(x, x) \geq \frac{c}{V_\rho(x, f(x, n))} \tag{LDLE}$$

$$\widetilde{p}_n(x, x) \geq \frac{c}{V_\rho(x, f(x, n))}, \tag{DLE}$$

where $f(x, \cdot)$ is the inverse of $F(x, \cdot)$ in the second variable.

Corollary 6.1. *We have that (p_0), (VD_ρ), (M) and (R_2) imply both inequalities (DLE) and (LDLE).*

The next task is to show the Near Diagonal Lower Estimate $(NDLE)$: There are δ and $c > 0$ such that, for all $x \in \Gamma, r > 0, y \in B(x, r), n > 0$ if $\rho(x, y) \leq \delta f(x, n)$ then,

$$\widetilde{p}_n(x, y) \geq \frac{c}{V_\rho(x, f(x, n))}$$

and its stronger form the Particular Lower Estimate (PLE): There are $\varepsilon, \delta, c > 0$ constants such that for all $x, y \in \Gamma, n \geq r > 0$

$$\widetilde{p}_n^{B_\rho(x,R)}(x, y) \geq \frac{c}{V(x, f(x, n))} \tag{49}$$

provided that $d(x, y) \leq \delta f(x, n), n \leq \varepsilon F(x, r)$.

Theorem 6.2. *If (Γ, μ) satisfies $(VD)_\rho$, (DLE) and (DUE) then, $(NDLE)$ and (PLE) hold.*

Proof. First we prove

$$\widetilde{p}_{2n}^{B_\rho(x,r)}(x, y) \geq \frac{c}{V_\rho(x, f(x, n))}$$

for $x, y \in \Gamma$ satisfying $d(x, y) \equiv 0 \mod 2$. Let us choose r such that $n = F_\rho(x, r) = \left[r^2 V_\rho(x, r)\right]^{1/m}$ and denote $B = B_\rho(x, r)$ $f_n(y) = p_n^B(x, y) + p_{n+1}^B(x, y)$, then

$$|f_n(x) - f_n(y)|^2 \leq R_m(x, y)\mathcal{D}_m(f_n, f_n).$$

By Lemma 4.1 we have that

$$|f_n(x) - f_n(y)|^2 \leq R_m(x, y)\mathcal{D}_m(f_n, f_n) = R_m(x, y) f_{2n}^{(m)}.$$

Now we can apply Lemma 4.2 to f_n and the diagonal upper and lower estimates

$$\begin{aligned} |f_n(x) - f_n(y)|^2 &\leq R_m(x, y)\frac{1}{n^m} f_n(x) \\ &\leq \frac{\overline{R}_m(x, \delta r)}{n^m} f_n(x) \\ &\leq C\frac{\delta^2 r^2}{r^2 V_\rho(x, r)} f_n(x) \leq \frac{1}{4} f_n^2(x), \end{aligned}$$

if δ is small enough. The above inequality and Proposition 6.1 mean that

$$\widetilde{p}_n(x,y) \geq \widetilde{p}_n^B(x,y) \geq \frac{1}{2}\widetilde{p}_n^B(x,x) \geq \frac{c}{V_\rho(x,r)}. \quad \blacksquare$$

Proof of Theorem 1.2. The diagonal upper estimate follows from Theorem 4.1, The off-diagonal upper estimate is stated in Theorem 6.1. The combination of Proposition 6.1 and Theorem 6.2 shows (5).

The reverse direction can be seen easily. The implication $(UE)+(PLE) \Longrightarrow (DUE)+(DLE) \Longrightarrow (VD)_\rho$ is immediate. The inequality (R_2) follows from the connection of $g_m^B(x,x)$ and $R_m\left(x, B(x,r)^c\right)$ and direct calculation using (DLE): Let $F = F(x,r) = \left[r^2 V_\rho(x,r)\right]^{1/m}$

$$R_m(x,r) = g_m^B(x,x) \geq c\sum_{n=F}^{2F} n^m p_n^B(x,x)$$

then we have

$$R_m(x,r) \geq c\sum_{n=F}^{2F} n^m p_n^B(x,x) \geq c\frac{r^2 V_\rho(x,r)}{V_\rho(x,r)}$$

which yields (R_2). $\blacksquare$

Remark 6.1. Kigami in [11] constructed a metric which is quasi-symmetric to $d(x,y)V_R(x,d(x,y)) + d(x,y)V_R(y,d(y,x))$. This procedure can be applied to $\rho^2(x,y)V_\rho(x,\rho(x,y)) + \rho^2(x,y)V_\rho(y,\rho(x,y))$ as well. All the conditions are satisfied to obtain a new metric σ which is quasi symmetric to ρ. We know that $(VD)_\rho$ implies $(VD)_\sigma$ and our heat kernel estimates can be obtained in a new form:

$$p_n(x,y) \leq \frac{C}{V_\sigma\left(x, g^{-1}(n)\right)}$$
$$p_n(x,y) \leq \frac{C}{V_\sigma\left(x, g^{-1}(n)\right)} \exp\left(-c\left(\frac{\sigma(x,y)}{n}\right)^{\frac{m}{\beta-1}}\right)$$
$$\widetilde{p}_n(x,y) \geq \frac{c}{V_\sigma\left(x, g^{-1}(n)\right)},$$

under the same conditions as in Theorems 1.1 and 1.2, where $g^{-1}(n)$ is the inverse of $g(r) = r^a V_\sigma(x,r)$ and a is the exponent determined by construction of σ from ρ. The parabolic Harnack inequality follows for σ as well. It should be noted here, that with the introduction of the second new metric the dependence from x in the exponential term is eliminated and $F(x,r)$ replaced by $g(r)$.

7. Examples

Example 7.1. Let us consider first the simplest possible example. Collapsing a slowly opening spherically symmetric tree with fractal dimension $\alpha = 3$ and walk dimension $\beta = 2$ (c.f. [6]) to $\mathbb{Z}^+ = \{0, 1, \ldots\}$ one obtains a transient weighted graph. The edge weights are $\mu_{0,1} = 1$ and $\mu_{j-1,j} = 2^{2(i-1)}$ if $j \in \left[2^i, 2^{i+1}\right)$ for $i \in \mathbb{Z}, i \geq 1$. Denote $D = \mathbb{Z}^+ \setminus \left\{2^i\right\}, k = 2^i$. It is clear that $\alpha = 3$, $\beta = 2$ and with $m = 2$ we have

$$p_n(0,0) \asymp n^{-3/2}$$

$$\begin{aligned} G_m^D(0,0) &\asymp \sum_{n=0}^{\infty} n^{m-1} p_n^D(0,0) \asymp \sum_{n=0}^{k^\beta} n p_n^D(0,0) \\ &\asymp \sum_{n=0}^{k^2} n p_n(0,0) \asymp \sum_{n=0}^{k^2} n^{-1/2} \asymp \left(k^2\right)^{1/2} = 2^i \end{aligned}$$

and in general for $N > 0$

$$G_m^{\mathbb{Z}^+ \setminus \{N\}}(0,0) \asymp N.$$

This means that

$$R_m(0,N) \asymp N$$

and

$$\rho(0,N) \asymp N^{1/2}.$$

At the same time (R_2) is satisfied.

Example 7.2. Now we consider the simple symmetric random walk on $\mathbb{Z}^d$, and first the case $d = 2k$. The vertex set of the graph is $\mathbb{Z}^d$, vertices $x, y \in \mathbb{Z}^d$ form an edge if $\|x - y\| = 1$. Edge weights are uniform, $\mu_{x,y} = 1/$. From the diagonal estimates we have that

$$\alpha_\rho / \beta_\rho = d/2. \tag{50}$$

On the other hand we know from (47) that the proper scaling:

$$\begin{aligned} r^{\beta_\rho} &= \left[r^2 V_\rho(0,r)\right]^{1/m} \qquad (51) \\ \beta_\rho &= \frac{2 + \alpha_\rho}{m}. \end{aligned}$$

Due to the spherical symmetry we can assume that $d = \omega\alpha_\rho$ for some $\omega > 0$. Consequently we have $\beta_\rho = 2\omega$ as well. The solution of these identities is:

$$\begin{aligned} \alpha_\rho &= \frac{2d}{2m - d} \\ \beta_\rho &= \frac{4}{2m - d} \end{aligned}$$

and

$$\omega = \frac{2m - d}{2}.$$

We have the restriction $\omega > 0$ and $m \geq \left\lceil \frac{d-2}{2} \right\rceil$ an integer, that yields for $d = 2k$, $m = k+1$, $\omega = 1$. As a result we have

$$\rho(x,y) \asymp d(x,y).$$

For $d = 2k + 1$ we have $\omega = 1/2$, $d(x,y) \asymp \rho(x,y)^{1/2}$. The correction would be the use of $m = \frac{d+1}{2}$ but at present we cannot handle non-integer derivation in Lemma 4.2. If we use the quasi-metric R_m instead of ρ we obtain for $\mathbb{Z}^d, d = 2k + 1$ that $R_m(x, y) \asymp d(x, y)$. In fact the use of the metric ρ is not essential in the whole work. One can see that the use of the weak triangular inequality would cause controllable cumulation of constants and the heat kernel estimates hold with respect to R_m.

Acknowledgments

Both authors are supported by the program "Hungarian–Italian intergovernmental S&T cooperation program for 2011–2013". Sincere thanks are due to the referees; their suggestions helped a lot to improve the paper.

References

[1] M.T. Barlow, F.R. Bass, Brownian motion and harmonic analysis on Sierpinski carpets, Canad. J. Math. 51 (1999) 673–744.
[2] M.T. Barlow, F.R. Bass, Stability of parabolic Harnack inequalities, Trans. Amer. Math. Soc. 356 (2004) 1501–1533.
[3] M.T. Barlow, T. Coulhon, T. Kumagai, Characterization of sub-Gaussian heat kernel estimates on strongly recurrent graphs, Comm. Pure Appl. Math. LVIII (2005) 1642–1677.
[4] M.T. Barlow, A.A. Jarai, T. Kumagai, G. Slade, Random walk on the incipient infinite cluster for oriented percolation in high dimensions, Comm. Math. Phys. 278 (2008) 385–431.
[5] M.T. Barlow, D. Nualart, Diffusion on fractals, in: P. Bernard. (Ed.), Lectures on Probability Theory and Statistics, in: Lecture Notes in Mathematics, vol. 1690, Springer-Verlag, Berlin, 1998, Lectures from the 25th Saint-Flour Summer School held July 10–26, 1995.
[6] P.G. Doyle, J.L. Snell, Random Walks and Electric Networks, in: Carus Mathematical Monographs, vol. 22, Mathematical Association of America, Washington, DC, 1984.
[7] A. Grigor'yan, J. Hu, K.-S. Lau, Obtaining upper bounds of heat kernels from lower bounds, Comm. Pure Appl. Math. 61 (2008) 639–660.
[8] A. Grigor'yan, A. Telcs, Harnack inequalities and sub-Gaussian estimates for random walks, Math. Ann. 324 (3) (2002) 521–556.
[9] W. Hebish, L. Saloff-Coste, On the relation between elliptic and parabolic Harnack inequalities, Ann. Inst. Fourier (Grenoble) 51 (5) (2001) 1437–1481.
[10] J. Kigami, Harmonic calculus on limits of networks and its application to dendrites, J. Funct. Anal. 128 (1995) 48–86.
[11] J. Kigami, Resistance forms, quasi-symmetric maps and heat kernel estimates, Mem. Amer. Math. Soc. 216 (2012) 132.
[12] T. Kumagai, Heat kernel estimates and parabolic Harnack inequalities on graphs and resistance forms, Publ. Res. Inst. Math. Sci. 40 (2004) 793–818.
[13] R. Mac´ıas, C. Segovia, Lipschitz functions on spaces of homogeneous type, Adv. Math. 33 (1979) 257–270.
[14] A. Telcs, The Art of Random Walks, in: Lecture Notes in Mathematics, vol. 1885, Springer, 2006.