ESCAPE RATE OF MARKOV CHAINS ON INFINITE GRAPHS

XUEPING HUANG

Abstract. We study the escape rate of continuous time symmetric Markov chains associated with weighted graphs. The upper rate functions are given in terms of volume growth of the weighted graph. For a class of symmetric birth and death processes, we obtain sharp upper rate functions.

Introduction

0.1. **History.** An increasing function $R(t)$ is called an upper rate function for a symmetric (sub)-Markov process $\left((X_t)_{t\geq 0}, \{\mathbb{P}_x\}_{x\in V}\right)$ on a metric space (V,d) if for some reference point $\bar{x}$,

$$\mathbb{P}_{\bar{x}}\{d(X_t,\bar{x}) \leq R(t) \text{ for all sufficiently large t } \} = 1.$$

It explicitly describes how far the process can escape after a long time. The boundary of the ball $B(\bar{x}, R(t))$ is intuitively the "forefront" of the process. Note that the existence of an upper rate function already implies the stochastic completeness (or non-explosion) of the process, namely the process almost surely cannot run out of the space V in finite time. The upper rate function can be viewed as quantitative information for stochastic completeness. There is also a matching concept of lower rate function (cf. [24]), the existence of which implies transience of the process.

The problem of escape rate for a Markov process has a long history and was studied extensively in many different settings. The celebrated Khinchin's theorem, i.e. the law of iterated logarithm, gives $R(t) = C\sqrt{t\ln\ln t}$ as an upper rate function for the (discrete time) simple random walk on $\mathbb{Z}$ for some constant $C > 0$. The same upper function works for the Brownian motion on $\mathbb{R}$. Both results lead to fertile developments that we will list some in the following.

A large body of literature is devoted to the escape rate for discrete time symmetric Markov chains on graphs or groups. Dvoretzky and Erdös [16] studied the lower rate function for the simple random walk on $\mathbb{Z}^N$ $(N \geq 3)$ which was later generalized by Hebisch and Saloff-Coste [29]. Barlow and Perkins [3] studied the escape rate for symmetric Markov chains on subgraphs of $\mathbb{Z}^N$. Varopoulos [44] studied the linear type escape rate for random walks on groups of exponential growth. See the survey of Vershik [45] for further references.

Date: February 8, 2012.

2010 *Mathematics Subject Classification.* Primary 60J27, Secondary 05C81.

Key words and phrases. escape rate, upper rate function, Markov chains, weighted graphs, birth and death process.

Research supported by Project CRC701.

For the Brownian motion on geodesically complete Riemannian manifolds with polynomial type volume growth, Grigor'yan [24] first proved that $R(t) = C\sqrt{t\ln t}$ is an upper rate function. Grigor'yan and Hsu [25] then obtained upper rate functions for Cartan-Hardamad manifolds with volume growth $\mathrm{vol}\,(B(r))$ satisfying

$$\int^{\infty} \frac{rdr}{\ln \mathrm{vol}\,(B(r))} = \infty. \tag{0.1}$$

One important significance is that their upper rate functions naturally relate to the stochastic completeness criterion (0.1) of Grigor'yan [22]. Recently Hsu and Qin proved the following theorem.

Theorem 0.1 (Hsu and Qin [30])**.** *Let M be a complete Riemannian manifold and let $z \in M$. Let $B(R)$ be the geodesic ball on M of radius R and centered at z. Define*

$$\psi(R) = \int_6^R \frac{rdr}{\ln \mathrm{vol}\,(B(r)) + \ln\ln r}.$$

Then there is a constant $C > 0$ such that $C\psi^{-1}(Ct)$ is an upper rate function of Brownian motion $(X_t)_{t\geq 0}$ on M.

This sharpness of Theorem 0.1 can be examined through examples of model manifolds. Let M be a model manifold with a pole z, that is, the Riemannian metric on M is spherically symmetric with respect to z. Consider the function $m(r) = \frac{V''(r)}{V'(r)}$ where V is the volume growth function $V(r) = \mu(B(z,r))$. In [25], Grigor'yan and Hsu proved that if

$$m(r) > 0, m'(r) \geq 0 \text{ for all large enough } r, \text{ and } \int_0^\infty \frac{dr}{m(r)} = \infty,$$

then the function $\phi(ct)$ is an upper rate function for the Brownian motion on M for any $c > 1$, where ϕ is defined by

$$\int_0^{\phi(t)} \frac{ds}{m(s)} = t.$$

0.2. **Main results.** In this paper, we are interested in the escape rate of *continuous time* symmetric Markov chains associated with weighted graphs. A weighted graph is an undirected graph with two weight functions, one on the vertices which serves as a measure and one on the edges controlling the jump probabilities. We prove upper bounds on the escape rate in terms of volume growth of the weighted graph. Explicit definitions of weighted graphs and constructions of the corresponding Markov chains will be given in Section 2. Here for simplicity, we focus on a family of simple but important examples: symmetric birth and death processes.

Example 0.2. Consider $V = \mathbb{N} = \{0, 1, 2, \cdots\}$ with the undirected graph structure that n and $n+1$ are neighbors for all $n \in \mathbb{N}$. Let μ be a positive function on V that can be viewed as a measure on V. Let w be a nonnegative function on $V \times V$ such that $w(x,y) = w(y,x)$ and

$$w(x,y) > 0 \Leftrightarrow |x - y| = 1.$$

For simplicity, we impose two further technical conditions:

(1) the function μ is nondecreasing;
(2) $\mu(n) \le 2w(n, n+1)$ for all $n \in V$.

The two functions μ and w naturally gives a discrete analogue of Laplace operator on the space of real valued functions on V:

$$\Delta f(x) := \frac{1}{\mu(x)} \sum_{y \in \mathbb{N}, |x-y|=1} w(x,y)\left(f(x) - f(y)\right).$$

The minimal Markov process $\left((X_t)_{t\ge 0}, \{\mathbb{P}_x\}_{x\in V}\right)$ (see Section 2 for more details) generated by Δ is a symmetric birth and death process.

The stochastic completeness of the process $\left((X_t)_{t\ge 0}\right)$ (or equivalently, of the weighted graph (V, w, μ)) in Example 0.2 is known (see [36], Theorem 5) to be equivalent to the following condition:

$$\sum_{n=0}^{\infty} \frac{\mu(0) + \cdots + \mu(n)}{w(n, n+1)} = \infty. \tag{0.2}$$

By (0.2), we cannot expect a general criterion of stochastic completeness in terms of volume growth in the obvious metric ρ on V given by $\rho(x, y) = |x - y|$, since it does not reflect the information of weights. In [26], [32], a notion of adapted metrics is applied to the stochastic completeness problem of weighted graphs. In the case of Example 0.2, we can explicitly write down an adapted metric on V. Define

$$\sigma_n = \sqrt{\frac{\mu(n)}{2w(n, n+1)}} \text{ and } d(m, n) = \sum_{k=n}^{m-1} \sigma_k, \tag{0.3}$$

where $m > n, m, n \in \mathbb{N}$. The key feature of the metric d is the following adaptedness condition:

$$\frac{1}{\mu(x)} \sum_{y \in V} w(x, y) d^2(x, y) \le 1, \tag{0.4}$$

which is an analogue of the property

$$|\nabla d(x, \cdot)|^2 \le 1$$

for the geodesic distance d on a Riemannian manifold.

To give a flavor of our main result, Theorem 4.2, which holds for more general weighted graphs, we first state it for the case of Example 0.2.

Theorem 0.3. *Let (V, w, μ) be the weighted graph defined as in Example 0.2. Let d be the adapted metric on V defined by (0.3). We assume that for all $r \ge 2$, the volume of closed balls centered at 0 satisfies*

$$\mu\left(B_d(0, r)\right) = \sum_{d(x,0) \le r} \mu(x) \le \exp f(r), \tag{0.5}$$

where f is a positive, increasing continuous function on $[0, \infty)$. For technical reasons, we consider two special classes of functions f.

(1) *There is some constant $M > 0$ such that*

$$\frac{f(r)}{r} \leq M \tag{0.6}$$

for all $r \geq 2$. In this case, there exists some constant $C > 0$ such that the inverse function $\psi^{-1}(t)$ of

$$\psi(R) = C \int_8^R \frac{rdr}{f(r) + \ln\ln(r)} \tag{0.7}$$

is an upper rate function for $(X_t)_{t\geq 0}$ with respect to d.

(2) *For all $r \geq 2$ the function $\frac{f(r)}{r}$ is increasing and*

$$f(r) < cr \ln r \tag{0.8}$$

for some constant $0 < c < \frac{1}{2}$. Then for any constant $2 < C_0 \leq \frac{1}{c}$, there is some constant $C > 0$, such that the inverse function $\psi^{-1}(t)$ of

$$\psi(R) = C \int_1^R \frac{rdr}{f(r) \exp\left(C_0 \frac{f(r)}{r}\right)} \tag{0.9}$$

is an upper rate function for $(X_t)_{t\geq 0}$ with respect to d.

Remark 0.4. We expect the sharpness the upper rate function given by (0.7) for at most exponential type volume growth. It is likely to be a sharp upper rate function for a much larger range of volume growth. The upper rate function given by (0.9) seems not sharp and the restriction on volume growth is only technical. We will show this later through examples.

Since in Example 0.2 we assume that μ is nondecreasing and positive, the following condition is satisfied:

$$C_\mu := \inf_{x\in V} \mu(x) > 0. \tag{0.10}$$

When the measures of balls are finite, this condition ensures the compactness (finiteness) of the balls $B_d(0, r)$ in the adapted metric, which is the analogue to geodesic completeness of Riemannian manifolds. As proven in [26] and [32], under assumption (0.10), the weighted graph is stochastically complete if $f(r) < cr \ln r$ where $0 < c < \frac{1}{2}$. See [38] for further improvements. The locally finiteness of V in d will also be used in Section 5 when we apply tools from martingale theory.

However, since symmetric birth and death processes are very special Markov chains, there is an alternative approach to determine their upper rate functions. The idea is to view them as analogues to Brownian motions on model manifolds. Consider $r(x) := d(x, 0)$ as a function on V. Then the following process is a martingale under $\mathbb{P}_0$:

$$M_t := r(X_t) + \int_0^t (\Delta r)(X_s)ds.$$

Furthermore, the martingale M satisfies that for all $t > 0$,

$$\mathbb{E}_0(M_t^2) \leq t,$$

which is guaranteed by (0.4), the adaptedness of the metric d. By Doob's maximal martingale inequality, we see that $|M_t|$ is "small" compared to t. So when the function $-\Delta r$ is bounded from below by some positive constant $C > 0$, we have that $\mathbb{P}_0$ almost surely,

$$r(X_t) \asymp -\int_0^t (\Delta r)(X_s)ds.$$

(Throughout the paper, we use the notation $f(x) \asymp g(x)$ when there is a constant $C > 1$ such that $g(x)/C \leq f(x) \leq Cg(x)$.) This eventually leads to sharp upper rate functions for a family of symmetric birth and death processes.

Theorem 0.5. *Let (V, w, μ) be the weighted graph defined in Example 0.2. Let d be the adapted metric on V defined by (0.3) and $r(x) := d(x, 0)$. Suppose that for all $x \in V$ large enough, and for some constant $c > 0$,*

$$-\Delta r(x) > c.$$

Let m be a positive, continuous, non-decreasing function such that

$$m(r(x)) \geq -\Delta r(x), \tag{0.11}$$

for all $x \in V$ large enough. Assume that

$$\sum_{n=0}^{\infty} \frac{\sigma_n}{m(r(n))} = \infty \tag{0.12}$$

Define the function $\phi(t)$ by the identity

$$t = \int_0^{\phi(t)} \frac{ds}{m(s)}. \tag{0.13}$$

Then the function $\phi(ct)$ is an upper rate function for the corresponding birth and death process for any $c > 1$. If instead of (0.11) we assume $m(r(x)) = -\Delta r(x)$ for all large enough x, then $\phi(ct)$ is not an upper rate function for any $0 < c < 1$.

Remark 0.6. By (0.12), we have

$$\sum_{n=0}^{\infty} \sigma_n \geq m(r(0)) \sum_{k=0}^{\infty} \frac{\sigma_k}{m(r(k))} = \infty.$$

So the space (V, d) is locally finite and the function m is well defined on $[0, \infty)$. By the monotonicity of m, it is easy to show (see Remark 5.2) that (0.12) is equivalent to either one of following:

$$\int_0^{\infty} \frac{ds}{m(s)} = \infty \text{ or } \sum_{n=0}^{\infty} \frac{\sigma_n}{m(r(n+1))} = \infty.$$

Remark 0.7. If we view the quantity $-\Delta r$ as kind of "mean curvature" for the weighted graph (V, w, μ), then the above theorem is very similar to that for Brownian motions on model manifolds (cf. [25], Theorem 5).

0.3. **Examples.** We can apply Theorem 0.3 to calculate some upper rate functions related to volume growth.

Example 0.8. In the following $c > 0$.

(1) Take
$$f(r) = c \ln r.$$
This corresponds to that the volume is bounded from above by some power function of distance. In this case, the $\ln \ln r$ term will be small compared to $f(r)$. So we have
$$R(t) = C\sqrt{t \ln t}$$
is an upper rate function for some constant $C > 0$.

(2) Take
$$f(r) = cr^{\alpha}$$
for some $0 < \alpha \leq 1$. Then
$$R(t) = Ct^{\frac{1}{2-\alpha}}$$
is an upper rate function for some constant $C > 0$.

(3) Take
$$f(r) = cr \ln r.$$
Here the constant $c > 0$ becomes important. If $0 < c < \frac{1}{2}$, then for any $1/c > C_0 > 2$, there exists some constant $C > 0$ (depending on C_0) such that
$$R(t) = C\,(t \ln t)^{\frac{1}{1-cC_0}}$$
is an upper rate function. If $c \geq 1/2$, we are not able to determine the existence of an upper rate function.

Remark 0.9. Our upper rate functions in Theorem 0.3 and Theorem 0.5 are given in terms of inverse functions. For the calculations of inverse functions, it is very convenient to adopt the notion of de Bruijn conjugate for slowly varying functions. We refer to Appendix 5 of [5] for details.

To compare the upper rate function $\psi^{-1}(t)$ given by Theorem 0.3 with the function $\phi(t)$ by Theorem 0.5, it suffices to compare
$$\frac{1}{m(r)} \text{ and } \frac{r}{\ln \mu(B_d(0,r))}.$$
For simplicity, we assume $m(r(x)) = -\Delta r(x)$ for all $x \in \mathbb{N}$. The heuristic argument in [25] becomes subtler in the discrete setting. For "nice" functions μ and σ, we have for all $n \geq 1$,
$$\frac{1}{m(r(n))} = \frac{\mu(n)}{\frac{\mu(n)}{\sigma_n} - \frac{\mu(n-1)}{\sigma_{n-1}}} \asymp \frac{\mu(n) + \cdots + \mu(0)}{\frac{\mu(n)}{\sigma_n}} = \frac{\sigma_n}{\frac{\mu(n)}{\mu(n)+\cdots+\mu(0)}}.$$
For example, the above holds if, $\frac{\mu(n)}{\mu(n-1)} \geq c > 1$, or if for all n large enough, $b_n \leq Cb_{[\frac{n}{2}]}$ where $C > 1$ and $b_n = \frac{\mu(n)}{\sigma_n} - \frac{\mu(n-1)}{\sigma_{n-1}}$.

However, in general

$$\sum_{k=0}^{n} \frac{\mu(k)}{\mu(k)+\cdots+\mu(0)} \asymp \ln\left(\mu(n)+\cdots+\mu(0)\right),$$

does *not* hold if

$$\frac{\mu(n)}{\mu(n)+\cdots+\mu(0)} \not\to 0,$$

as $n \to \infty$. So even when μ and σ are "nice" functions, we cannot expect the upper rate function given by (0.7) to be sharp. We can see this phenomenon more concretely through the following two examples.

Example 0.10. Let (V, w, μ) be the weighted graph defined in Example 0.2. Let $\mu(n) = (n+1)^\alpha$ for some $\alpha \geq 0$. Let

$$w(n, n+1) = \frac{1}{2}\mu(n)(n+1)^2 \ln(n+3)^\beta \ln\ln(n+30)^\gamma,$$

where $\beta \geq 0$ and $\gamma > 0$. Then

$$\sigma_n = \frac{1}{(n+1)\ln(n+3)^{\beta/2}\ln\ln(n+30)^{\gamma/2}}.$$

We have for all $n \geq 1$,

$$\begin{aligned}
m(r(n)) = -\Delta(r(n)) &= \frac{w(n,n+1)\sigma_n - w(n-1,n)\sigma_{n-1}}{\mu(n)} \\
&= \frac{w(n,n+1)\sigma_n}{\mu(n)} - \frac{w(n-1,n)\sigma_{n-1}}{\mu(n-1)} + \left(\frac{1}{\mu(n-1)} - \frac{1}{\mu(n)}\right) w_{n-1}\sigma_{n-1} \\
&= \frac{1}{2}\left(\frac{1}{\sigma_n} - \frac{1}{\sigma_{n-1}}\right) + \frac{1}{2}\left(1 - \frac{\mu(n-1)}{\mu(n)}\right)\frac{1}{\sigma_{n-1}} \\
&= \frac{1}{2}\left(\left(\frac{\sigma_{n-1}}{\sigma_n} - 1\right) + \left(1 - \frac{\mu(n-1)}{\mu(n)}\right)\right)\frac{1}{\sigma_{n-1}}.
\end{aligned}$$

We can easily see that

$$\frac{\sigma_{n-1}}{\sigma_n} - 1 \asymp \frac{1}{n} \asymp 1 - \frac{\mu(n-1)}{\mu(n)}, \tag{0.14}$$

whence

$$m(r(n)) \asymp \left(\frac{\sigma_{n-1}}{\sigma_n} - 1\right)\frac{1}{\sigma_{n-1}} = \frac{1}{\sigma_n} - \frac{1}{\sigma_{n-1}},$$

and

$$\frac{\sigma_n}{m(r(n))} \asymp \frac{1}{(n+1)\ln(n+3)^\beta \ln\ln(n+30)^\gamma}.$$

The condition (0.12) boils down to

$$0 \leq \beta < 1 \text{ or } \beta = 1, \gamma \leq 1.$$

Since

$$d(n,0) = \sigma_0 + \cdots \sigma_n \asymp \frac{\ln(n+3)^{1-\beta/2}}{\ln\ln(n+30)^{\gamma/2}},$$

we have that for all r large enough,

$$\ln\left(\mu\left(B_d(0,r)\right)\right) \asymp r^{\frac{2}{2-\beta}}(\ln r)^{\frac{\gamma}{2-\beta}},$$

and

$$\int_0^r \frac{dr}{m(r)} \asymp \begin{cases} \dfrac{r^{\frac{2-2\beta}{2-\beta}}}{(\ln r)^{\frac{\gamma}{2-\beta}}}, & \text{if } 1 > \beta \geq 0, \gamma > 0, \\ (\ln r)^{1-\gamma}, & \text{if } \beta = 1, 0 < \gamma < 1, \\ \ln \ln r, & \text{if } \beta = \gamma = 1. \end{cases}$$

Then the upper rate function $\phi(t)$ given by Theorem 0.5 satisfies

$$\phi(t) \asymp \begin{cases} t^{\frac{2-\beta}{2-2\beta}} (\ln t)^{\frac{\gamma}{2-2\beta}}, & \text{if } 1 > \beta \geq 0, \gamma > 0, \\ \exp\left(C t^{\frac{1}{1-\gamma}}\right), & \text{if } \beta = 1, 0 < \gamma < 1, \\ \exp \exp Ct, & \text{if } \beta = \gamma = 1, \end{cases}$$

for some constant $C > 0$. This coincides with the upper rate function given by (0.7) although if $\beta > 0$ the volume growth here is out of the range of Theorem 0.3.

For the case $\beta = 0$ and $0 < \gamma < 1$, the volume growth function has the form

$$\ln\left(\mu\left(B_d(0,r)\right)\right) \asymp r(\ln r)^{\frac{\gamma}{2}},$$

and the upper rate function given by (0.9) satisfies

$$\ln \psi^{-1}(t) = \ln t + C(\ln t)^{\gamma/2} + o(1),$$

for some $C > 0$ (by truncated Lagrange inversion, see [5]), while $\phi(t) \asymp t(\ln t)^{\gamma/2}$. This comparison indicates that the upper rate function given by (0.9) in Theorem 0.3 is not sharp.

Example 0.11. Let (V, w, μ) be the weighted graph defined in Example 0.2. Let $w(n, n+1) = \frac{1}{2}\mu(n)(n+1)^\beta \ln(n+3)^\gamma$, where $\beta > 0, \gamma > 0$. Then

$$\sigma_n = \frac{1}{(n+1)^{\beta/2} \ln(n+3)^{\gamma/2}}.$$

We assume that for some constant $C > 1$, $\mu(n+1) \geq C\mu(n)$ for all n. This in particular implies that

$$\frac{\sigma_{n-1}}{\sigma_n} - 1 \asymp \frac{1}{n} = o\left(1 - \frac{\mu(n-1)}{\mu(n)}\right), \tag{0.17}$$

in contrast to (0.14). It follows that for $n \geq 1$

$$m(r(n)) \asymp \left(1 - \frac{\mu(n-1)}{\mu(n)}\right) \frac{1}{\sigma_{n-1}} \asymp \frac{1}{\sigma_{n-1}},$$

and

$$\frac{\sigma_n}{m(r(n))} \asymp \sigma_{n-1}\sigma_n \asymp \sigma_n^2 = \frac{1}{(n+1)^\beta \ln(n+3)^\gamma}.$$

By Theorem 0.5, we obtain the upper rate function $\phi(t)$ through similar calculations as in Example 0.10,

$$\phi(t) \asymp \begin{cases} t^{\frac{2-\beta}{2-2\beta}} (\ln t)^{\frac{\gamma}{2-2\beta}}, & \text{if } 1 > \beta > 0, \gamma > 0, \\ \exp\left(C t^{\frac{1}{1-\gamma}}\right), & \text{if } \beta = 1, 0 < \gamma < 1, \\ \exp \exp Ct, & \text{if } \beta = \gamma = 1, \end{cases}$$

for some constant $C > 0$. If we choose $\mu(n) = e^n$, then the volume growth satisfies

$$\ln\left(\mu\left(B_d(0,r)\right)\right) \asymp r^{\frac{2}{2-\beta}} (\ln r)^{\frac{\gamma}{2-\beta}},$$

for all r large enough. The upper rate functions given by Theorem 0.5 and (0.7) still coincide.

However, in this case there is no upper bound on the growth of $\mu(n)$. We can choose $\mu(n)$ growing arbitrarily fast such that the upper rate function given by (0.7) is no longer sharp or may be even meaningless. For instance, if we choose $\mu(n) = \exp\left(n\left(\ln(n+3)\right)^\alpha\right)$ for some $\alpha > 0$, and assume $\beta = 1$ and $1 > \gamma > 0$, then

$$\ln\left(\mu\left(B_d(0,r)\right)\right) \asymp r^2 (\ln r)^{\alpha+\gamma},$$

for all r large enough. It follows that the integral $\int^\infty \frac{rdr}{\ln(\mu(B_d(0,r)))}$ is no longer divergent for $\alpha > 1-\gamma$. And if $1-\gamma > \alpha > 0$, the upper rate function given by (0.7) has the form

$$\psi^{-1}(t) \asymp \exp C\left(t^{\frac{1}{1-\alpha-\gamma}}\right)$$

which is not sharp. Note that the volume growth in this case depends only on $\alpha+\gamma$, while the sharp upper rate $\phi(t)$ depends only on γ. So this also gives an example that there are two weighted graphs with different type upper rate functions while having the same type volume growth.

Remark 0.12. From the above two examples, we are inclined to believe that the upper rate function given by (0.9) in Theorem 0.3 is not sharp in general while for certain range of volume growth, the one given by (0.7) should be nearly sharp. Above certain growth, there is no hope for a sharp upper rate function purely in terms of volume function with respect to the adapted metric.

Note that when $\sigma_n \equiv 1$, i.e. when the graph metric is adapted, the relation (0.17) is always true, independent of the choice of μ. It is not surprising that in this case the upper rate function can not be expressed in terms of volume growth.

0.4. **Organization.** Before proceeding to the main part, we briefly explain the main strategy to obtain upper rate functions in the general case, so as to motivate the preliminary materials we presented later. We will adopt the Borel-Cantelli lemma type argument to transfer the problem of escape rate to that of tail estimates of exit time from a ball. This is classical and is used by Grigor'yan and Hsu [25] (see also Hsu and Qin [30]) to study the escape rate for Brownian motions on manifolds. Using the construction of the Markov chain, we can view the tail probability of exit time as a solution to certain heat equation with zero initial condition. A discrete integrated maximum principle is developed to give good estimates for such type solutions. This is inspired by the work of Grigor'yan in the manifold setting [23]. The concept of adapted metrics then emerges naturally when constructing auxiliary functions in the estimate of solutions to the heat equation. In our opinion, it is a good analogue of the geodesic metric on Riemannian manifolds for weighted graphs.

Our paper is organized as follows. Section 1 is devoted to the basic settings of weighted graphs and Markov chains. Using the weights on a graph, we briefly summarize the construction of the associated minimal, continuous time, càdlàg, symmetric

Markov chains explicitly. The discrete integrated maximum principle is developed in Section 2. In Section 3 adapted metrics will be introduced as an analogue of the geodesic metric on Riemannian manifolds. After these preparations, in Section 4, we state and prove our main result that gives upper rate functions in terms of volume growth with respect to adapted metrics. In the last section, we apply probabilistic tools to escape rate of Markov chains associated with weakly spherically symmetric weighted graphs, which can be viewed as a generalization of symmetric birth and death processes. An appendix is added in order to collect some results from martingale theory required in Section 5.

1. Generalities of weighted graphs and Markov chains

The most general framework of weighted graphs, through the theory of Dirichlet forms (see [21], [7]), is due to Keller and Lenz [35]. For the classical theory of continuous time Markov chains, we refer to Feller [17], Chung [9], Freedman [20], and Norris [39]. For the general theory of Markov processes, we refer to Blumenthal and Getoor [6]. We will present neither the complete theory nor the most general setting here. Only some definitions and basic facts about locally finite, connected weighted graphs and the associated continuous time minimal Markov chain will be summarized.

Let (V, E) be a simple undirected graph with a countably infinite vertex set V and an edge set $E \subseteq V \times V$. Here, "simple" means "loopless and without multiedges". A pair of vertices $x, y \in V$ are called neighbors if $(x, y) \in E$, which we write $x \sim y$ for short. The graph is called connected if for any pair of distinct vertices $x, y \in V$, there exists a path of some length $n \in \mathbb{N}_+$ connecting them, that is, a sequence of vertices $x_0, \cdots, x_n$ in V such that

$$x_0 = x, x_n = y, x_k \sim x_{k+1} \text{ for all } 0 \le k \le n-1.$$

For $x \in V$, we define the degree of x to be $\deg(x) = \#\{y \in V : y \sim x\}$, i.e. the number of neighbors of x. We call a graph locally finite if $\deg(x) < \infty$ for any $x \in V$.

Throughout this paper, we *only consider* locally finite, countably infinite, connected, simple, undirected graphs without specification.

There is a natural graph metric ρ on the graph (V, E), defined through

$$\rho(x, y) := \inf\{n : \exists x_0, \cdots, x_n \text{ s.t. } x_0 = x, x_n = y, x_k \sim x_{k+1} \text{ for all } 0 \le k \le n-1\},$$

for any pair of $x \neq y$. The metric ρ induces the discrete topology on V.

We simultaneously consider two types of weights on (V, E): a positive function μ on V as a measure and a nonnegative function w on $V \times V$ such that

(1) $w(x, y) = w(y, x)$ for all $x, y \in V$;
(2) $(x, y) \in E \Leftrightarrow w(x, y) > 0$.

Remark 1.1. Note that $w(x, x) = 0$ for all $x \in V$. We can also start from the data (V, w) and determine the edge set E through the condition (2) above.

The triple (V, w, μ) is called a weighted graph. An important quantity is the weighted degree:

$$\mathrm{Deg}(x) = \frac{1}{\mu(x)} \sum_{y \in V} w(x, y).$$

Keller and Lenz [35] defined the formal Laplacian on a weighted graph (V, w, μ) as follows:

$$\Delta f(x) = \frac{1}{\mu(x)} \sum_{y \in V} w(x, y)(f(x) - f(y))$$

where f is an arbitrary function on V. The operator Δ plays a similar role on weighted graphs as the Laplace-Beltrami operator on Riemannian manifolds.

Remark 1.2. This kind of Laplacians on graphs with general weights, treated as analogue to elliptic differential operators, appears first in Colin de Verdière [12] to the knowledge of the author.

Example 1.3. Let (V, E) be a graph. Let the weight function $w = \mathbf{1}_E$, that is, the characteristic function of the edge set. Choose the vertex weight to be $\mu = \deg$. The formal Laplacian is given by

$$\Delta f(x) = \frac{1}{\deg(x)} \sum_{y, y \sim x} (f(x) - f(y)).$$

It is called the normalized Laplacian because the weighted degree function is constant 1. We call the corresponding minimal Markov chain the normalized random walk. The reason for the name "normalized" can also be seen from the fact that the expected holding time of the Markov chain at each vertex is constant 1. It is easy to see that the minimal Markov chain is always stochastically complete. This is the most common setting in the analytical study of random walks. See Chung [8], Saloff-Coste [41] for example.

Example 1.4. Let (V, E) be a graph. Let the weight function $w = \mathbf{1}_E$. Choose the vertex weight to be $\mu \equiv 1$. In this case the formal Laplacian is

$$\Delta f(x) = \sum_{y, y \sim x} (f(x) - f(y)). \tag{1.1}$$

This is the so-called physical Laplacian studied first by Dodziuk [14], Dodziuk and Matthai [15]. Its stochastic completeness and essential self-adjointness was then studied by Weber [46] and Wojciechowski [48], [49] independently. See also the work of Keller, Lenz and Wojciechowski [36] for recent developments. The weighted degree function in this case is $\mathrm{Deg} = \deg$ that is not necessarily bounded. The physical Laplacian case offers a large family of weighted graphs whose stochastic completeness problem is interesting.

There are various approaches to rigorously construct the minimal càdlàg (right continuous with left limits) Markov chain corresponding to a weighted graph. For instance, we can canonically define a minimal regular Dirichlet form on the weighted graph (see [35]) and the theory of Dirichlet forms guarantees the existence and uniqueness of the corresponding Hunt process (see [21]). There are also direct constructions of the sample space, filtration and probability measures, see for example section 12, chapter 1 of Blumenthal and Getoor [6]. We will not repeat these well-known constructions here. Instead, we describe an intuitive construction following Norris [39] and then construct the canonical process (cf. Bass [4]).

In the probabilistic language, the formal Laplacian amounts to the concept of Q-matrices. We follow the notations of Norris [39] in the following.

Let (V, w, μ) be a weighted graph. Define

$$q_{xy} = \frac{w(x,y)}{\mu(x)}$$

for $x \neq y$ and

$$q_{xx} = -\mathrm{Deg}(x).$$

The matrix $Q = (q_{xy})_{V\times V}$ satisfies that

(1) $0 \leq -q_{xx} < \infty$ for all x;
(2) $q_{xy} \geq 0$ for all $x \neq y$;
(3) $\sum_{y\in V} q_{xy} = 0$ for all x.

This kind of matrices is called Q-matrices in [39]. Viewing a function f on V also as a column vector, we have

$$(Qf)_x = -\Delta f(x).$$

There is a natural jump matrix $\Pi = (\pi_{xy})_{V\times V}$ associated with Q as:

$$\pi_{xy} = \frac{q_{xy}}{|q_{xx}|} \text{ if } x \neq y; \pi_{xx} = 0.$$

Note that $q_{xx} \neq 0$ by our basic assumption that the weighted graph is infinite and connected.

Following [39], we can construct a minimal càdlàg Markov chain $(X_t)_{t\geq 0}$ corresponding to the Q-matrix. First we adjoin a cemetery point ∞ to V and denote the set $V \cup \infty$ by V_∞. From the jump matrix Π we can construct a discrete time Markov chain $(Y_n)_{n\in\mathbb{N}}$ on V. Let $T_1, T_2, \cdots$ be a sequence of independent exponential random variables of parameter 1 that are independent of $\{Y_n\}_{n\in\mathbb{N}}$. Set $S_n = T_n/\mathrm{Deg}(Y_{n-1})$ and define the jump times $J_n = S_1 + \cdots + S_n$ with the convention $J_0 = 0$. Define the explosion time ζ by

$$\zeta = \sup_n J_n,$$

which is the first time that X_t jumps out of V. Then the minimal continuous time Markov chain $(X_t)_{t\geq 0}$ on V_∞ is defined as

$$X_t = \begin{cases} Y_n, & \text{if } J_n \leq t < J_{n+1} \text{ for some } n \in \mathbb{N}, \\ \infty, & \text{if } t \geq \zeta. \end{cases}$$

For all $n \in \mathbb{N}_+$, conditioning on $Y_0 = x_0, \cdots, Y_{n-1} = x_{n-1}$, $S_1, \cdots, S_n$ are independent exponential random variables with parameters $\mathrm{Deg}(x_0), \cdots, \mathrm{Deg}(x_{n-1})$ respectively. The process $(X_t)_{t\geq 0}$ has the (time homogenous) Markov property in the sense that for all $n \in \mathbb{N}$, all sequences of time $0 \leq t_0 \leq \cdots \leq t_{n+1}$, and all sequences of points $x_0, \cdots, x_{n+1}$ in V_∞,

$$\mathbb{P}(X_{t_{n+1}} = x_{n+1} | X_{t_0} = x_0, \cdots, X_{t_n} = x_n) = \mathbb{P}(X_{t_{n+1}-t_n} = x_{n+1} | X_0 = x_n).$$

Remark 1.5. (1) From the above construction, we see that $\frac{1}{\mathrm{Deg}(x)}$ gives the expected holding time of the process at $x \in V$. This is a probabilistic interpretation of the weighted degree function.

(2) It is direct to see that at the time when the process $(X_t)_{t\geq 0}$ leaves a point $x \in V$, it can only jump to the neighbors of x.

Definition 1.6. *We call a function $\omega : [0,\infty) \to V_\infty$ a step function if for some $\zeta(\omega) \in (0,\infty]$, the following holds:*

- *On $[0, \zeta(\omega))$ the function ω takes values in V and is càdlàg (in particular, piecewise constant);*
- *If $\zeta(\omega) < \infty$, then $\omega \equiv \infty$ on $[\zeta(\omega), \infty)$.*

Let Ω be the space of step functions with an additional point ω_∞ satisfying $\omega_\infty(t) \equiv \infty$ and define $\tilde{X}_t : \Omega \to V_\infty$ by $\tilde{X}_t(\omega) = \omega(t)$. There is a family of probability measures $(\mathbb{P}_x)_{x\in V}$ on Ω uniquely determined by (cf. Bass [4] p. 158)

$$\mathbb{P}_x(\tilde{X}_{t_1} = y_1, \tilde{X}_{t_2} = y_2, \cdots, \tilde{X}_{t_n} = y_n) = \mathbb{P}(X_{t_1} = y_1, X_{t_2} = y_2, \cdots X_{t_n} = y_n | X_0 = x),$$

for all $n \geq 1$, $0 < t_1 < t_2 < \cdots < t_n$, $y_1, \cdots, y_n \in V$ (here $\mathbb{P}$ is the probability measure in the previous intuitive construction). The construction of filtration is a bit subtle. First we define a filtration $\mathcal{F}^0 = (\mathcal{F}^0_t)_{t\geq 0}$ on Ω by $\mathcal{F}^0_t := \sigma\left(\tilde{X}_s : 0 \leq s \leq t\right)$, the σ-algebra generated by the random variables $\left(\tilde{X}_t\right)_{0\leq s\leq t}$.

Let $\mathcal{F}^0_\infty = \bigvee_{t\geq 0} \mathcal{F}^0_t$, the σ-algebra generated by $\bigcup_{t\geq 0} \mathcal{F}^0_t$. In this paper, we usually fix some reference point $\bar{x} \in V$ and consider the completion $\mathcal{F}^{\bar{x}}$ of $\mathcal{F}^0$ with respect to $\mathbb{P}_{\bar{x}}$. Namely, let $\mathcal{N}_{\bar{x}}$ be the set of $\mathbb{P}_{\bar{x}}$-null sets in $\mathcal{F}^0_\infty$ and define $\mathcal{F}^{\bar{x}} = (\mathcal{F}^{\bar{x}}_t)_{t\geq 0}$ by $\mathcal{F}^{\bar{x}}_t = \sigma\left(\mathcal{F}^0_t, \mathcal{N}_{\bar{x}}\right)$. The filtration $(\mathcal{F}^{\bar{x}}_t)_{t\geq 0}$ is right continuous and the process

$$\left(\left(\tilde{X}_t\right)_{t\geq 0}, \mathcal{F}^{\bar{x}}, \Omega, (\mathbb{P}_x)_{x\in V_\infty}\right)$$

(here $\mathbb{P}_\infty$ is given by $\mathbb{P}_\infty(\omega_\infty) = 1$) satisfies the strong Markov property in the sense that (cf. Theorem A.1.18 in [7]) for any $\mathcal{F}^{\bar{x}}$-stopping time T, $s \geq 0$, and $U \subseteq V_\infty$,

$$\mathbb{E}_{\bar{x}}\left(\tilde{X}_{T+s} \in U | \mathcal{F}^{\bar{x}}_T\right) = \mathbb{P}_{X_T}\left(\tilde{X}_s \in U\right), \ \mathbb{P}_{\bar{x}}\text{-a.s.}$$

In the following, we will only consider the canonical process corresponding to a weighted graph and denote it by $(X_t)_{t\geq 0}$ for simplicity from time to time. Note that in the setting of canonical process, the jump times can be defined inductively by

$$\tau_0 = 0, \tau_{n+1} = \inf\{t > \tau_n : X_t \neq X_{t-}\},$$

where we make the convention that $\inf \emptyset = \infty$. For each n, τ_n has the same distribution as J_n in the intuitive construction. As can be seen from the construction of $(X_t)_{t\geq 0}$, $\mathbb{P}_x(\tau_n = \infty) = 0$ for all n and $x \in V$.

The following definition of explosion of a Markov chain is taken from [39].

Definition 1.7. *The Markov chain $(X_t)_{t\geq 0}$ is called explosive (or stochastically incomplete) if for some $x \in V$,*

$$\mathbb{P}_x(\zeta < \infty) > 0.$$

Otherwise, $(X_t)_{t\geq 0}$ is called nonexplosive (or stochastically complete).

The quantity

$$\tilde{p}(t,x,y) = \mathbb{P}_x(X_t = y)$$

is called the transition probability of $(X_t)_{t\geq 0}$. For any $t \geq 0$, define the matrix $\mathcal{P}_t$ by $\mathcal{P}_t = (\tilde{p}(t,x,y))_{V\times V}$. The family $\{\mathcal{P}_t\}_{t\geq 0}$ is in fact a matrix semigroup as

$$\mathcal{P}_{t+s} = \mathcal{P}_t\mathcal{P}_s, \mathcal{P}_0 = \mathrm{Id},$$

for all $s \geq 0, t \geq 0$. From the construction of $(X_t)_{t\geq 0}$, we see that $(X_t)_{t\geq 0}$ is stochastically complete if and only if for any $x \in V$ and any $t > 0$,

$$(\mathcal{P}_t\mathbf{1}_V)(x) = \sum_{y\in V}\tilde{p}(t,x,y) = \mathbb{P}_x(X_t \in V) = 1,$$

where $\mathbf{1}_A$ denotes the characteristic function for a set A.

Now, we consider the restriction of the process $(X_t)_{t\geq 0}$ to a finite subset U of V. Let τ_U be the first exit time of U, that is

$$\tau_U = \inf\{t \geq 0 : X_t \in U^c\}.$$

We can define a semigroup $\mathcal{P}_t^U$ on the spaces of functions on U as (see (4.1.2) in [21] for example)

$$\mathcal{P}_t^U f(x) = \mathbb{E}_x\left(f(X_t)\mathbf{1}_{\{t<\tau_U\}}\right).$$

Since U is finite, $\{\mathcal{P}_t^U\}_{t\geq 0}$ can be viewed as a semigroup of finite dimensional matrices. The corresponding Markov chain $\{X_t^U\}_{t\geq 0}$ can be viewed as constructed from $(X_t)_{t\geq 0}$ in the way that at the first time when X_t runs out of U, we send it to the cemetery point ∞ and it never gets back. It is a standard calculation from the construction of $(X_t)_{t\geq 0}$ and the definition of τ_U to obtain that $\mathcal{P}_t^U$ satisfies

$$\frac{\partial}{\partial t}\mathcal{P}_t^U = Q^U\mathcal{P}_t^U, \mathcal{P}_0^U = \mathrm{Id}_{U\times U},$$

where the matrix $Q^U = (q_{xy}^U)_{U\times U}$ is

$$q_{xy}^U = \begin{cases} \dfrac{w(x,y)}{\mu(x)}, & \text{if } x \neq y,\ x \in U, y \in U \\ -\dfrac{1}{\mu(x)}\displaystyle\sum_{z\in V} w(x,z), & \text{if } x = y \in U. \end{cases}$$

Remark 1.8. Note that as U is finite, $\mathcal{P}_t^U$ is simply $\exp\left(tQ^U\right)$.

Definition 1.9. *Let (V, w, μ) be a weighted graph and U be a subset of V. We define the (outer) boundary ∂U of U as*

$$\partial U = \{x \in U^c : \exists y \in U, x \sim y\}.$$

The closure $\bar{U}$ of U is defined to be

$$\bar{U} = U \cup \partial U.$$

It is easy to see that the closure of a finite subset of V is again finite because of locally finiteness of (V, E). The following proposition relates the tail probability of exit time of a finite subset U of V to the heat equation on U.

Proposition 1.10. *Let (V, w, μ) be a weighted graph with the formal Laplacian Δ. Let W and U be two finite subsets of V such that $\bar{U} \subseteq W$. Define a function u on $W \times [0, \infty)$ to be*

$$u(x, t) = \mathbb{P}_x(\tau_W \le t).$$

Then $u(x, t)$ is differentiable in t on $[0, \infty)$ and satisfies

$$\frac{\partial}{\partial t} u(x, t) + \Delta u(x, t) = 0,$$

for all $x \in U$ and $t \ge 0$ with initial condition $u(x, 0) \equiv 0$.

Proof. By the definition of $\mathcal{P}_t^W$, we have that

$$u(x, t) = \mathbb{P}_x(\tau_W \le t) = 1 - \mathbb{E}_x\left(\mathbf{1}_{\{\tau_W > t\}}\right) = 1 - \mathcal{P}_t^W \mathbf{1}_W(x).$$

It is clear that $u(x, 0) \equiv 0$ on W. Moreover, viewing $\mathbf{1}_W$ as a column vector, for all $x \in W$,

$$\frac{\partial}{\partial t} u(x, t) = -\frac{\partial}{\partial t} \mathcal{P}_t^W \mathbf{1}_W(x) = -Q^W \mathcal{P}_t^W \mathbf{1}_W(x).$$

Since $\bar{U} \subseteq W$, for all $x \in U$, we have that $y \in W$ if $w(x, y) > 0$. Hence for all $x \in U$,

$$\begin{aligned} Q^W \mathcal{P}_t^W \mathbf{1}_W(x) &= \sum_{y \in W} \frac{w(x, y)}{\mu(x)} \mathcal{P}_t^W \mathbf{1}_W(y) - \sum_{y \in V} \frac{w(x, y)}{\mu(x)} \mathcal{P}_t^W \mathbf{1}_W(x) \\ &= \frac{1}{\mu(x)} \sum_{y \in V} w(x, y)(\mathcal{P}_t^W \mathbf{1}_W(y) - \mathcal{P}_t^W \mathbf{1}_W(x)) \\ &= \Delta u(x, t). \end{aligned}$$

The equation

$$\frac{\partial}{\partial t} u(x, t) + \Delta u(x, t) = 0$$

then holds for all $x \in U$ and $t \ge 0$. □

2. Discrete integrated maximum principle

As briefly explained before, to estimate the tail probability of exit time, we study the solutions to the heat equation with zero initial condition. A powerful tool for such estimates is the so-called "integrated maximum principle" which dates back to Aronson [1], [2]. The following lemma can be viewed as a discrete version of Grigor'yan's "integrated maximum principle" [23].

Lemma 2.1. *Let (V, w, μ) be a weighted graph. Let $\Omega \subseteq V$ be a finite subset of V. Fix some $T > 0$. Let $u(x,t)$ be a function on $\bar{\Omega} \times [0,T]$ that is differentiable in t on $[0,T]$ and $u(x,0) \equiv 0$. Assume further that $u(x,t)$ solves the heat equation*

$$\frac{\partial}{\partial t}u(x,t) + \Delta u(x,t) = 0, \tag{2.1}$$

on $\Omega \times [0,T]$. Take two auxiliary functions $\eta(x)$ on V and $\xi(x,t)$ on $V \times [0,T]$ such that

(1) *the function $\eta(x) \geq 0$ is finitely supported and* $\mathrm{supp}\eta \subseteq \Omega$*;*
(2) *$\xi(x,t)$ is continuously differentiable in t on $[0,T]$ for each $x \in V$;*
(3) *the inequality $(\eta^2(x) - \eta^2(y))(e^{\xi(x,t)} - e^{\xi(y,t)}) \geq 0$ holds for all $x \sim y$ and $t \in [0,T]$;*
(4) *the inequality $\mu(x)\frac{\partial}{\partial t}\xi(x,t) + \frac{1}{2}\sum_{y\in V} w(x,y)(1 - e^{\xi(y,t)-\xi(x,t)})^2 \leq 0$ holds for any $x \in V$ and $t \in [0,T]$.*

Then for any $\tau \in (0,T]$, we have the following estimate:

$$\sum_{x\in\Omega} u^2(x,\tau)\eta^2(x)e^{\xi(x,\tau)}\mu(x) \leq 2\int_0^\tau \sum_{x\in\bar{\Omega}}\sum_{y\in\bar{\Omega}} w(x,y)(\eta(x)-\eta(y))^2u^2(x,t)e^{\xi(y,t)}dt. \tag{2.2}$$

Remark 2.2. Folz [18] develops a different version of the discrete "integrated maximum principle" independently of us. Note that in [10], Coulhon, Grigor'yan and Zucca developed a discrete integrated maximum principle to study discrete time Markov chains.

Before giving a proof of Lemma 2.1, we present an elementary fact that we will use frequently.

Lemma 2.3. *Let $\{a_i\}_{i=1}^\infty$ and $\{b_i\}_{i=1}^\infty$ be two sequences of real numbers such that*

$$\sum_{i=1}^\infty a_i^2 < \infty, \sum_{i=1}^\infty b_i^2 < \infty.$$

The inequality

$$\sum_{i=1}^\infty a_ib_i \leq \frac{\delta}{2}\sum_{i=1}^\infty a_i^2 + \frac{1}{2\delta}\sum_{i=1}^\infty b_i^2 \tag{2.3}$$

holds for all $\delta > 0$.

Proof. By the Cauchy-Schwarz inequality we have that

$$\sum_{i=1}^\infty a_ib_i \leq \left(\delta\sum_{i=1}^\infty a_i^2\right)^{\frac{1}{2}}\left(\frac{1}{\delta}\sum_{i=1}^\infty b_i^2\right)^{\frac{1}{2}}.$$

Then the desired inequality follows by the AM-GM inequality. □

Proof of Lemma 2.1. We multiply the heat equation in (2.1) by $u(x,t)\eta^2(x)e^{\xi(x,t)}\mu(x)$ and sum over $x \in \Omega$:

$$\sum_{x\in\Omega} \frac{\partial}{\partial t}u(x,t) \cdot u(x,t)\eta^2(x)e^{\xi(x,t)}\mu(x) + \sum_{x\in\Omega}\sum_{y\in\bar{\Omega}} w(x,y)(u(x,t)-u(y,t)) \cdot u(x,t)\eta^2(x)e^{\xi(x,t)} = 0. \tag{2.4}$$

Note that since $\eta(x)$ is finitely supported, the sums in (2.4) are of finite type. Furthermore, since $\mathrm{supp}\eta \subseteq \Omega$, if we make a sum over $x \in \Omega$ of some multiple of $\eta^2(x)$, it is equivalent to do it over $\bar{\Omega}$. By symmetry of $w(x,y)$, we have

$$\sum_{x\in\bar{\Omega}} \frac{\partial}{\partial t}u^2(x,t) \cdot \eta^2(x)e^{\xi(x,t)}\mu(x) + \sum_{x\in\bar{\Omega}}\sum_{y\in\bar{\Omega}} w(x,y)(u(x,t)-u(y,t))(u(x,t)\eta^2(x)e^{\xi(x,t)} - u(y,t)\eta^2(y)e^{\xi(y,t)}) = 0. \tag{2.5}$$

In this proof, from now on, the sums without specification of range will be understood to be over $\bar{\Omega}$.

Using the fact that

$$\left(\frac{\partial}{\partial t}u^2(x,t)\right) \cdot e^{\xi(x,t)} = \frac{\partial}{\partial t}(u^2(x,t)e^{\xi(x,t)}) - u^2(x,t) \cdot e^{\xi(x,t)}\frac{\partial}{\partial t}\xi(x,t),$$

and the finiteness of the sums, we obtain

$$\frac{\partial}{\partial t}(\sum_x u^2(x,t)\eta^2(x)e^{\xi(x,t)}\mu(x)) = \sum_x u^2(x,t)\eta^2(x)e^{\xi(x,t)}\mu(x)\frac{\partial}{\partial t}\xi(x,t) - \sum_x\sum_y w(x,y)(u(x,t)-u(y,t))(u(x,t)\eta^2(x)e^{\xi(x,t)} - u(y,t)\eta^2(y)e^{\xi(y,t)}). \tag{2.6}$$

We split the sum in the last line:

$$-\sum_x\sum_y w(x,y)(u(x,t)-u(y,t))(u(x,t)\eta^2(x)e^{\xi(x,t)} - u(y,t)\eta^2(y)e^{\xi(y,t)}) \tag{2.7}$$

$$= -\sum_x\sum_y w(x,y)(u(x,t)-u(y,t))^2\eta^2(x)e^{\xi(x,t)} \tag{2.8}$$

$$-\sum_x\sum_y w(x,y)(u(x,t)-u(y,t))u(y,t)(\eta^2(x)-\eta^2(y))e^{\xi(x,t)} \tag{2.9}$$

$$-\sum_x\sum_y w(x,y)(u(x,t)-u(y,t))u(y,t)\eta^2(y)(e^{\xi(x,t)}-e^{\xi(y,t)}), \tag{2.10}$$

which is a discrete analogue to the Leibniz rule. Then we apply Lemma 2.3 to (2.9) and (2.10) to cancel the term in (2.8).

First, for any $\delta_1 > 0$,

$$\begin{aligned}
&-\sum_x \sum_y w(x,y)(u(x,t)-u(y,t))u(y,t)(\eta^2(x)-\eta^2(y))e^{\xi(x,t)} \\
&\le \frac{\delta_1}{2}\sum_x \sum_y w(x,y)(u(x,t)-u(y,t))^2(\eta(x)+\eta(y))^2 e^{\xi(x,t)} \\
&+ \frac{1}{2\delta_1}\sum_x \sum_y w(x,y)(\eta(x)-\eta(y))^2 u^2(y,t)e^{\xi(x,t)}.
\end{aligned}$$

Applying the elementary fact

$$(\eta(x)+\eta(y))^2 \le 2(\eta^2(x)+\eta^2(y)),$$

we have that by the symmetry of $w(x,y)$,

$$\begin{aligned}
&-\sum_x \sum_y w(x,y)(u(x,t)-u(y,t))u(y,t)(\eta^2(x)-\eta^2(y))e^{\xi(x,t)} \\
&\le \delta_1 \sum_x \sum_y w(x,y)(u(x,t)-u(y,t))^2(\eta^2(x)+\eta^2(y))e^{\xi(x,t)} \\
&+ \frac{1}{2\delta_1}\sum_x \sum_y w(x,y)(\eta(x)-\eta(y))^2 u^2(y,t)e^{\xi(x,t)} \\
&= \frac{\delta_1}{2}\sum_x \sum_y w(x,y)(u(x,t)-u(y,t))^2(\eta^2(x)+\eta^2(y))(e^{\xi(x,t)}+e^{\xi(y,t)}) \\
&+ \frac{1}{2\delta_1}\sum_x \sum_y w(x,y)(\eta(x)-\eta(y))^2 u^2(y,t)e^{\xi(x,t)}.
\end{aligned}$$

Using the condition (3) of Lemma 2.1, it follows that

$$\begin{aligned}
&-\sum_x \sum_y w(x,y)(u(x,t)-u(y,t))u(y,t)(\eta^2(x)-\eta^2(y))e^{\xi(x,t)} \\
&\le \delta_1 \sum_x \sum_y w(x,y)(u(x,t)-u(y,t))^2(\eta^2(x)e^{\xi(x,t)}+\eta^2(y)e^{\xi(y,t)}) \\
&+ \frac{1}{2\delta_1}\sum_x \sum_y w(x,y)(\eta(x)-\eta(y))^2 u^2(y,t)e^{\xi(x,t)} \\
&= 2\delta_1 \sum_x \sum_y w(x,y)(u(x,t)-u(y,t))^2\eta^2(x)e^{\xi(x,t)} \\
&+ \frac{1}{2\delta_1}\sum_x \sum_y w(x,y)(\eta(x)-\eta(y))^2 u^2(x,t)e^{\xi(y,t)}.
\end{aligned} \tag{2.11}$$

Similarly, for any $\delta_2 > 0$,

$$
\begin{aligned}
&-\sum_x \sum_y w(x,y)(u(x,t)-u(y,t))u(y,t)\eta^2(y)(e^{\xi(x,t)}-e^{\xi(y,t)})\\
&= -\sum_x \sum_y w(x,y)(u(x,t)-u(y,t))u(x,t)\eta^2(x)(e^{\xi(x,t)}-e^{\xi(y,t)})\\
&\le \frac{\delta_2}{2}\sum_x \sum_y w(x,y)(u(x,t)-u(y,t))^2\eta^2(x)e^{\xi(x,t)}\\
&+ \frac{1}{2\delta_2}\sum_x \sum_y w(x,y)(1-e^{\xi(y,t)-\xi(x,t)})^2 u^2(x,t)\eta^2(x)e^{\xi(x,t)}.
\end{aligned}
$$

Choose $\delta_1 = 1/4$ and $\delta_2 = 1$ and apply the above estimates of (2.9) and (2.10) to (2.7). It follows that

$$
\begin{aligned}
-\sum_x \sum_y &w(x,y)(u(x,t)-u(y,t))(u(x,t)\eta^2(x)e^{\xi(x,t)}-u(y,t)\eta^2(y)e^{\xi(y,t)})\\
&\le 2\sum_x \sum_y w(x,y)(\eta(x)-\eta(y))^2 u^2(x,t)e^{\xi(y,t)}\\
&+ \frac{1}{2}\sum_x \sum_y w(x,y)(1-e^{\xi(y,t)-\xi(x,t)})^2 u^2(x,t)\eta^2(x)e^{\xi(x,t)}.
\end{aligned}
$$

Hence, by (2.6),

$$
\begin{aligned}
&\frac{\partial}{\partial t}(\sum_x u^2(x,t)\eta^2(x)e^{\xi(x,t)}\mu(x))\\
&\le \sum_x u^2(x,t)\eta^2(x)e^{\xi(x,t)}\mu(x)\frac{\partial}{\partial t}\xi(x,t)\\
&+ 2\sum_x \sum_y w(x,y)(\eta(x)-\eta(y))^2 u^2(x,t)e^{\xi(y,t)}\\
&+ \frac{1}{2}\sum_x \sum_y w(x,y)(1-e^{\xi(y,t)-\xi(x,t)})^2 u^2(x,t)\eta^2(x)e^{\xi(x,t)}\\
&\le 2\sum_x \sum_y w(x,y)(\eta(x)-\eta(y))^2 u^2(x,t)e^{\xi(y,t)},
\end{aligned}
$$

where in the last inequality we used the condition (4) on the auxiliary function. Integrating the above inequality with respect to t on $[0,\tau]$, we get the desired inequality as $u(x,0) \equiv 0$. □

3. Adapted metrics

The first work introducing metrics other than the graph metric for analysis on weighted graphs seems due to Davies [11]. The notion of adapted metrics for graphs is inspired by the intrinsic metrics for Dirichlet forms. The concept of intrinsic metrics for general regular Dirichlet forms is first introduced by Frank, Lenz and Wingert [19], based on earlier work of Sturm [43] in the strongly local case. Inspired by the integrability conditions for a Lévy measure, the work of Masamune and Uemura [37] implicitly contains a similar notion. Grigor'yan, Huang and Masamune applied the adapted metrics to the stochastic completeness problem of jump type Dirichlet forms and in particular to weighted graphs. Folz [18] also came up with similar ideas with the goal to obtain heat kernel estimates on weighted graphs. A different notion of weighted distance has also been introduced by Colin de Verdière, Torki-Hamza and Truc [13] in the context of essential self-adjointness for weighted graphs.

In the previous section, we reduced the problem of estimating solutions to the heat equation to that of constructing suitable auxiliary functions η and ξ. We observe that the key quantities take the form of

$$\frac{1}{\mu(x)}\sum_{y\in V} w(x,y)\left(\eta(x)-\eta(y)\right)^2, \frac{1}{\mu(x)}\sum_{y\in V} w(x,y)(1-e^{\xi(y,t)-\xi(x,t)})^2.$$

They are discrete analogues of $|\nabla\eta|^2(x)$ and $|\nabla\xi|^2(x,t)$ on manifolds. It is natural to construct these auxiliary functions in some simple form from a metric on V. For example, η as a cut off function takes the simplest form of a tent-like function

$$\eta(x) = C(R-d(x,\bar{x}))_+$$

where d is a metric on V. To get estimates in good shape, we would like to have that

$$\frac{1}{\mu(x)}\sum_{y\in V} w(x,y)\left(\eta(x)-\eta(y)\right)^2 \le C \tag{3.1}$$

for some $C>0$. Note that the inequality (3.1) is sensible to the weights μ and w while the graph metric ρ does not distinguish different weighted graphs with the same underlying graph structure. So in general, it is not suitable to use the graph metric to construct auxiliary functions.

Definition 3.1. *We call a metric d on a weighted graph (V,w,μ) adapted if*

(1)
$$\frac{1}{\mu(x)}\sum_{y\in V} w(x,y)d^2(x,y) \le 1 \tag{3.2}$$
for every $x\in V$;

(2) *$d(x,y)\le 1$ whenever $w(x,y)>0$.*

Remark 3.2. The constant 1 is not essential in the assumption that $d(x,y)\le 1$ for $w(x,y)>0$. For our purpose, it suffices to have some upper bound on the distances between neighboring points in the adapted metric.

Remark 3.3. Note that the quantity

$$\frac{1}{\mu(x)}\sum_{y\in V} w(x,y)\left(\eta(x)-\eta(y)\right)^2$$

can be viewed as a discrete analogue of $|\nabla\eta|^2(x)$. So (3.2) is an analogue of the fact that

$$|\nabla d|^2 \le 1$$

where d is the geodesic distance on a Riemannian manifold. Tent functions with respect to an adapted metric automatically satisfy (3.1) by the triangle inequality.

Example 3.4. Let (V, w, μ) be the weighted graph in Example 1.3, the normalized Laplacian case. In this case the graph metric is adapted as $\rho(x, y) = 1$ for $x \sim y$ and

$$\frac{1}{\mu(x)} \sum_{y \in V} w(x, y)\rho^2(x, y) = \frac{1}{\deg(x)} \sum_{y, y \sim x} 1 = 1.$$

Example 3.5. Let (V, w, μ) be the weighted graph in Example 1.4, the physical Laplacian case. In this case, we have

$$\frac{1}{\mu(x)} \sum_{y \in V} w(x, y)\rho^2(x, y) = \sum_{y, y \sim x} 1 = \deg(x).$$

So the graph metric is in general not adapted.

The following construction shows that there always exists an adapted metric on a weighted graph.

Definition 3.6. *Define a function $\sigma(x, y)$ for all pairs of neighbors $x \sim y$ by*

$$\sigma(x, y) = \min \left\{ \frac{1}{\sqrt{\mathrm{Deg}(x)}}, \frac{1}{\sqrt{\mathrm{Deg}(y)}}, 1 \right\}. \tag{3.3}$$

It naturally induces a metric d on X as follows: for all pairs of distinct points x, y,

$$d(x, y) := \inf\{\sum_{i=0}^{n-1} \sigma(x_i, x_{i+1}) : x_0, x_1, \cdots, x_n \text{ is a path connecting } x \text{ and } y\}. \tag{3.4}$$

Remark 3.7. It is easy to see by definition that

$$d(x, y) \le \sigma(x, y) \le 1$$

if $x \sim y$. A direct consequence is that for any $x, y \in V$,

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

So the volume growth with respect to an adapted metric is generally larger than that with respect to the graph metric. When the weighted degree function $\mathrm{Deg}(x)$ is bounded from above by some constant $C > 1$, we have that

$$\frac{1}{\sqrt{C}}\rho(x, y) \le d(x, y) \le \rho(x, y).$$

In this case, the adapted metric and the graph metric have similar properties.

Since we only consider locally finite and connected weighted graphs, the closed balls in the graph metric are compact (i.e. finite). It is is a nice topological property so that the cut off functions are finitely supported. This property is not necessarily true for an adapted metric on (V, w, μ). However, as already mentioned in Remark 0.6, we can make a mild technical assumption to get rid of this problem.

Assumption 3.8. *The weight function on vertices of the weighted graph (V, w, μ) have a positive lower bound, namely*

$$C_\mu = \inf_{x \in V} \mu(x) > 0.$$

Lemma 3.9. *Let (V, w, μ) be a weighted graph such that Assumption 3.8 holds. Let d be a metric on V such that every closed ball has finite measure. Then the closed balls are finite.*

Proof. Let $B_d(x, r)$ be a closed ball with finite measure. Then

$$\# B_d(x, r) \leq \frac{\mu(B_d(x, r))}{C_\mu} < \infty.$$

$\square$

We end this section by concrete examples of adapted metrics. First we define the so-called weakly spherically symmetric weighted graphs following Keller, Lenz and Wojciechowski [36]. The symmetric birth and death process introduced in Example 0.2 can be viewed as a special case.

Definition 3.10. *Let (V, w, μ) be a locally finite, connected, infinite weighted graph with the underlying graph structure (V, E) and the graph metric ρ. Fix a point $\bar{x} \in V$ as a reference point. Write $\rho(x) = \rho(x, \bar{x})$ for short. We will denote the set $\{x \in V : \rho(x) = n\}$ by S_n. Note that $S_0 = \{\bar{x}\}$.*

For $x \in V$, define for all $x \in V$

$$\kappa_+(x) = \frac{1}{\mu(x)} \sum_{y \in S_{\rho(x)+1}} w(x, y),$$

and for all $x \neq \bar{x}$,

$$\kappa_-(x) = \frac{1}{\mu(x)} \sum_{y \in S_{\rho(x)-1}} w(x, y).$$

The weighted graph (V, w, μ) is called weakly spherically symmetric (with respect to $\bar{x}$) if the functions $\kappa_\pm$ depend only on $\rho(x)$. In this paper we make one more technical assumption that $w(x, y) = 0$ for any pair of x and y such that $\rho(x) = \rho(y)$. Equivalently, there is no edge connecting points on the same sphere S_n for all n.

Remark 3.11. The technical assumption is for the purpose of applying the random measure theory in a more convenient way. It may not be essential.

The underlying graph (V, E) need not to have any symmetry. It is proven in [36] that the heat kernel of the corresponding minimal Markov chain is spherically symmetric if and only if the weighted graph is weakly spherically symmetric. Thus it is not surprising that we will use weakly spherically symmetric weighted graphs as analogues of model manifolds.

We can construct an adapted metric on a weakly spherically symmetric weighted graph in a similar way as (0.3).

Example 3.12. Let (V, w, μ) be a weakly spherically symmetric weighted graph. Define

$$\mu_n = \sum_{x \in S_n} \mu(x), \text{ and } w_n = \sum_{x \in S_n} \sum_{y \in S_{n+1}} w(x, y).$$

For simplicity, we assume that $(\mu_n)_{n \in \mathbb{N}}$ is an nondecreasing sequence and that $\mu_n \leq 2w_n$ for all $n \in \mathbb{N}$. Let

$$\sigma_n = \sqrt{\frac{\mu_n}{2w_n}}. \tag{3.5}$$

We can define a function $\sigma(x, y)$ by

$$\sigma(x, y) = \sigma_{\min\{\rho(x), \rho(y)\}}$$

for all pairs of neighboring points x, y with $|\rho(x) - \rho(y)| = 1$. It naturally induces a metric on V by

$$\begin{aligned} d(x, y) := \inf\{\sum_{i=0}^{n-1} \sigma(x_i, x_{i+1}) : x_0, x_1, \cdots, x_n \in V \text{ s.t. } \forall 0 \leq i \leq n-1, \\ x_i \sim x_{i+1} \text{ and } |\rho(x_i) - \rho(x_{i+1})| = 1\}. \end{aligned} \tag{3.6}$$

for all pairs of distinct points x, y. The existence of a path (x_i) connecting x and y with $|\rho(x_i) - \rho(x_{i+1})| = 1$ can be seen from the special case that $x = \bar{x}$. It is direct to check that d is an adapted metric and $d(x, \bar{x}) = \sigma_0 + \cdots + \sigma_{n-1}$ for all $x \in S_n$ and $n \geq 1$.

The following two examples are special cases of Example 3.12.

Example 3.13 (Model trees)**.** Let (V, E) be a locally finite, connected, infinite tree with the graph metric ρ. We consider the physical Laplacian structure on (V, E), i.e., the weight functions satisfying $\mu \equiv 1$ on V and $w \equiv 1$ on E. Fix some reference point $\bar{x} \in V$. We see that $\kappa_-(x) \equiv 1$ for $x \neq \bar{x}$ and $\kappa_+(x) = \#\{y \in V : x \sim y, \rho(y) = \rho(x) + 1\}$. The weighted graph (V, w, μ) is weakly spherically symmetric if and only if κ_+ is constant on S_n for each n. In this case the underlying graph (V, E) is in fact spherically symmetric.

Consider the adapted metric d constructed in Example 3.12. Let $\kappa_+(x) = [(n+1)^s]$ for $x \in S_n$ where $s > 0$ and $[c]$ is the integer part of c. Then for $x \sim y$ such that $x \in S_n, y \in S_{n+1}$, we have

$$\sigma(x, y) = \sqrt{\frac{\mu_n}{2w_n}} = \sqrt{\frac{1}{2k_+(x)}} \asymp \frac{1}{(n+1)^{s/2}}.$$

So for $x \in V$ such that $x \in S_n, n > 1$, we have

$$d(x, \bar{x}) \asymp \begin{cases} (n+1)^{1-s/2}, & \text{if } 0 < s < 2, \quad (3.7) \\ \ln(n+2), & \text{if } s = 2. \quad (3.7') \end{cases}$$

If $s > 2$, we see that (V, d) is bounded. The balls in the metric space (V, d) are all finite only when $0 < s \leq 2$.

Example 3.14 (Anti-trees (Wojciechowski [50])). Let $(S_n)_{n\in\mathbb{N}}$ be a sequence of disjoint, finite, nonempty sets with $S_0 = \{\bar{x}\}$. Let

$$V = \bigcup_{n\in\mathbb{N}} S_n.$$

Define $\rho(x) = n$ if $x \in S_n$. Then define

$$E = \{(x, y) \in V \times V : |\rho(x) - \rho(y)| = 1\}.$$

In other words, we connect every vertex in S_r to every vertex in S_{r+1} to get a graph (V, E) that is symmetric with respect to $\bar{x}$. Note that the graph metric ρ on (V, E) naturally gives $\rho(x) = \rho(x, \bar{x})$.

Consider the physical Laplacian structure on (V, E). Then the resulting weighted graph is weakly spherically symmetric with $\kappa_+(x) = \#S_{n+1}$ and $\kappa_-(x) = \#S_{n-1}$ for $x \in S_n$. Furthermore, we have $\mu_n = \#S_n$ and $w_n = \#S_n \times \#S_{n+1}$. Take for example $\#S_n = [(n+1)^s (\ln(n+e))^{s'}]$ where $s, s' > 0$. We have

$$\sigma_n = \sqrt{\frac{\mu_n}{2w_n}} = \sqrt{\frac{1}{2\#S_{n+1}}} \asymp \frac{1}{(n+1)^{s/2} (\ln(n+e))^{s'/2}}.$$

Hence, for $x \in V$ such that $x \in S_n, n > 1$, we have

$$d(x, \bar{x}) \asymp \begin{cases} \dfrac{(n+1)^{1-s/2}}{(\ln(n+e))^{s'/2}}, & \text{if } 0 < s < 2, \quad (3.8) \\ (\ln(n+e))^{1-s'/2}, & \text{if } s = 2, 0 < s' < 2, \quad (3.8') \\ \ln\ln(n+30), & \text{if } s = s' = 2. \quad (3.8'') \end{cases}$$

4. Upper rate function

In this section, we assume that (V, w, μ) is a weighted graph such that Assumption 3.8 holds. Let $(X_t)_{t\geq 0}$ be the corresponding minimal càdlàg Markov chain as constructed in Section 1. We first give an explicit definition of the upper rate function.

Definition 4.1. *Let (V, w, μ) be a weighted graph such that Assumption 3.8 holds. Let d be an adapted metric on (V, w, μ). Fix a reference point $\bar{x} \in V$. A function $R(t)$ is called an upper rate function (with respect to d) for the minimal process $(X_t)_{t\geq 0}$ (or equivalently, for (V, w, μ)) if*

$$\mathbb{P}_{\bar{x}}\{d(X_t, \bar{x}) \leq R(t) \text{ for all sufficiently large } t\} = 1.$$

The main result in its full generality of our paper is given as follows:

Theorem 4.2. *Under the setting in Definition 4.1, we assume that for all $r \geq 2$,*

$$\ln \mu(B_d(\bar{x}, r)) \leq f(r), \tag{4.1}$$

where f is a positive, increasing continuous function on $[0, \infty)$. For technical reasons, we consider two special classes of functions f.

(1) *There is some constant* $M > 0$ *such that*

$$\frac{f(r)}{r} \leq M \tag{4.2}$$

for all $r \geq 2$. *In this case, there exists some constant* $C > 0$ *such that the inverse function* $\psi^{-1}(t)$ *of*

$$\psi(R) = C \int_8^R \frac{rdr}{f(r) + \ln\ln(r)} \tag{4.3}$$

is an upper rate function for $(X_t)_{t\geq 0}$.

(2) *The function* $\frac{f(r)}{r}$ *is increasing for* $r \geq 2$ *and*

$$\int_1^\infty \frac{rdr}{f(r)\exp\left(C_0 \frac{f(r)}{r}\right)} = \infty \tag{4.4}$$

for some constant $C_0 > 2$. *Then there is some constant* $C > 0$, *such that the inverse function* $\psi^{-1}(t)$ *of*

$$\psi(R) = C \int_1^R \frac{rdr}{f(r)\exp\left(C_0 \frac{f(r)}{r}\right)} \tag{4.5}$$

is an upper rate function for $(X_t)_{t\geq 0}$.

Remark 4.3. (1) Note that by Lemma 3.9, the closed balls of (V, w, μ) in the metric d are finite since we assume Assumption 3.8 and (4.1).

(2) The upper rate function as the inverse function of ψ in (4.3) seems to be sharp since it coincides with results in the manifold case [25, 30]. It would be interesting to see whether it is true even without the restriction (4.2).

The proof of Theorem 4.2 consists of two parts. In the first part, using probabilistic arguments, one reduces the question to estimates of certain solutions to the heat equation. The probabilistic argument based on the Borel-Cantelli lemma is standard (see, for example, [25], [30]). The required estimates are obtained through the discrete integrated maximum principle Lemma 2.1.

4.1. **Borel-Cantelli lemma.** We generally follow the strategy in [24] and [25] for upper rate functions of the Brownian motion on Riemannian manifolds. Recall that $B_d(x, r)$ denotes a closed ball centered at x with radius r in d. Let $\{R_n\}_{n=0}^\infty$ be a strictly increasing sequence of positive numbers to be chosen later such that $\lim_{n\to\infty} R_n = \infty$. Denote the balls $B_d(\bar{x}, R_n)$ by B_n and define a sequence of stopping times τ_n by

$$\tau_n = \tau_{B_n}.$$

Suppose that for a sequence of positive numbers $\{c_n\}_{n=1}^\infty$ we have that

$$\sum_1^\infty \mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n) < \infty.$$

Then, by the Borel-Cantelli lemma, it follows that $\mathbb{P}_{\bar{x}}$ almost surely

$$\tau_n - \tau_{n-1} > c_n$$

for all n large enough. Let $T_n = \sum_1^n c_k$. With $\mathbb{P}_{\bar{x}}$ probability 1, we have that $\tau_n > T_n - T_0$ for all n large enough, where T_0 is some random number. Suppose that we can find a strictly increasing homeomorphism $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$T_{n-1} - \psi(R_n) \to +\infty \tag{4.6}$$

as $n \to \infty$. Then with $\mathbb{P}_{\bar{x}}$ probability 1, for n large enough and for t such that

$$\psi(R_{n-1}) < t \leq \psi(R_n) < T_{n-1} - T_0,$$

we have that

$$d(X_t) \leq R_{n-1} \leq \psi^{-1}(t). \tag{4.7}$$

Notice that $\lim_{n\to\infty} \psi(R_n) = \infty$, so (4.7) holds for all large enough t, $\mathbb{P}_{\bar{x}}$ almost surely. In other words, $\psi^{-1}(t)$ is an upper rate function for $(X_t)_{t\geq 0}$.

4.2. **Exit time estimate.** With the main strategy in hand, the key technical problem is to estimate the quantity

$$\mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n).$$

By the strong Markov property of the minimal Markov chain $(X_t)_{t\geq 0}$, we have

$$\mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n) = \mathbb{E}_{\bar{x}}(\mathbb{P}_{X_{\tau_{n-1}}}\{\tau_n \leq c_n\}). \tag{4.8}$$

By the construction of $(X_t)_{t\geq 0}$, since (V, w, μ) is locally finite, we know that

$$X_{\tau_{n-1}} \in \partial B_{n-1}, X_{\tau_n} \in \partial B_n.$$

Define

$$r_n = R_n - R_{n-1} - 1$$

and assume that $r_n > 2$ for $n \geq 1$. Hence, X_t must run out of a ball $B_d(X_{\tau_{n-1}}, r_n)$ before it leaves B_n. So it follows that

$$\mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n) \leq \sup_{z\in\partial B_{n-1}} \mathbb{P}_z\{\tau_{B_d(z,r_n)} \leq c_n\}.$$

For a fixed $z \in \partial B_{n-1}$, define

$$u(x,t) = \mathbb{P}_x(\tau_{B_d(z,r_n)} \leq t)$$

as a function on $B_d(z, r_n) \times [0, \infty)$. It is direct to see that $0 \leq u(x,t) \leq 1$. Note that $\bar{B}_d(z, r_n - 1) \subseteq B_d(z, r_n)$. By Proposition 1.10, the function

$$u(x,t) = 1 - \mathcal{P}_t^{B_d(z,r_n)} 1,$$

is a solution to the heat equation

$$\frac{\partial}{\partial t} u(x,t) + \Delta u(x,t) = 0,$$

on $B_d(z, r_n - 1) \times [0, \infty)$ with $u(x, 0) \equiv 0$.

Fixing some n, we can then apply Lemma 2.1 to $u(x,t)$. Choose the auxiliary functions to be

$$\xi(x,t) = -2\alpha_n^2 e^{4\alpha_n} t - 2\alpha_n d(x,z) \text{ and } \eta(x) = \frac{(e^{\alpha_n(r_n-1)} - e^{\alpha_n d(x,z)})_+}{e^{\alpha_n(r_n-1)} - 1}$$

with $\alpha_n > 0$ to be determined later. Concerning $\eta(x)$, we have the estimate

$$
\begin{aligned}
&\sum_{y\in V} w(x,y)(\eta(x)-\eta(y))^2\\
\le&\frac{1}{(e^{\alpha_n(r_n-1)}-1)^2}\sum_{y\in V} w(x,y)(e^{\alpha_n d(x,z)}-e^{\alpha_n d(y,z)})^2\\
\le&\frac{\alpha_n^2 e^{2\alpha_n d(x,z)+2\alpha_n}}{(e^{\alpha_n(r_n-1)}-1)^2}\sum_{y\in V} w(x,y)d^2(x,y)\\
\le&\frac{\alpha_n^2 e^{2\alpha_n d(x,z)+2\alpha_n}}{(e^{\alpha_n(r_n-1)}-1)^2}\mu(x).
\end{aligned}
\tag{4.9}
$$

It is direct to check that they fulfill the conditions in Lemma 2.1, so we have that for any $\tau > 0$

$$
\begin{aligned}
&\sum_{x\in B_d(z,r_n-1)} u^2(x,\tau)\eta^2(x)e^{\xi(x,\tau)}\mu(x)\\
\le&2\int_0^\tau \sum_{x\in \bar{B}_d(z,r_n-1)}\sum_{y\in \bar{B}_d(z,r_n-1)} w(x,y)(\eta(x)-\eta(y))^2u^2(y,t)e^{\xi(x,t)}dt\\
\le&2\int_0^\tau \sum_{x\in \bar{B}_d(z,r_n-1)}\sum_{y\in \bar{B}_d(z,r_n-1)} w(x,y)(\eta(x)-\eta(y))^2e^{\xi(x,t)}dt,
\end{aligned}
\tag{4.10}
$$

where we used the symmetry of $w(x,y)$ and the fact that $0 \le u \le 1$. Plugging in the explicit forms of the auxiliary functions, we have

$$
u^2(z,\tau)\mu(z)e^{-2\alpha_n^2 e^{4\alpha_n}\tau} \tag{4.11}
$$

$$
\le 2\int_0^\tau \sum_{x\in B_d(z,r_n)}\sum_{y\in B_d(z,r_n)} w(x,y)(\eta(x)-\eta(y))^2e^{\xi(x,t)}dt \tag{4.12}
$$

$$
\le 2\int_0^\tau \sum_{x\in B_d(z,r_n)} \frac{\alpha_n^2 e^{2\alpha_n d(x,z)+2\alpha_n}}{(e^{\alpha_n(r_n-1)}-1)^2}\times e^{-2\alpha_n d(x,z)-2\alpha_n^2e^{4\alpha_n}t}\mu(x)dt \tag{4.13}
$$

$$
=\frac{e^{-2\alpha_n}(1-e^{-2\alpha_n^2e^{4\alpha_n}\tau})}{(e^{\alpha_n(r_n-1)}-1)^2}\mu(B_d(z,r_n)). \tag{4.14}
$$

Recall that we assume that for any $x \in V$,

$$
\mu(x) \ge C_\mu > 0.
$$

It follows that

$$
u^2(z,\tau) \le \frac{1}{C_\mu}\frac{e^{-2\alpha_n}(e^{2\alpha_n^2e^{4\alpha_n}\tau}-1)}{(e^{\alpha_n(r_n-1)}-1)^2}\mu(B_d(z,r_n)) \tag{4.15}
$$

Let $\tau = c_n$, we have

$$
\mathbb{P}_{\bar{x}}(\tau_n-\tau_{n-1}\le c_n) \le \sup_{z\in\partial B_{n-1}} u(z,c_n) \le \sqrt{\frac{\mu(B_n)}{C_\mu}}\frac{e^{\alpha_n^2e^{4\alpha_n}c_n}}{e^{\alpha_n r_n}-e^{\alpha_n}}.
$$

Note the elementary fact that if $r_n > 2$, then

$$\frac{1}{2}\alpha_n r_n < \alpha_n(r_n - 1)$$

and hence

$$\frac{1}{e^{\alpha_n r_n} - e^{\alpha_n}} \leq \frac{\alpha_n(r_n-1)+1}{\alpha_n(r_n-1)} e^{-\alpha_n r_n} \leq (1 + \frac{2}{\alpha_n r_n}) e^{-\alpha_n r_n}.$$

Together with (4.1), we have

$$\mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n) \leq \frac{1}{\sqrt{C_\mu}}(1 + \frac{2}{\alpha_n r_n}) \exp\{\alpha_n^2 e^{4\alpha_n} c_n + \frac{1}{2} f(R_n) - \alpha_n r_n\}. \tag{4.16}$$

Now we prove Theorem 4.2 by specifying R_n, c_n and α_n. We give separate proofs for the two classes of functions f.

4.3. **Upper rate function for case (1).**

Proof of Theorem 4.2, case (1). We choose $R_n = 2^{n+3}$ for $n \geq 0$ and then

$$r_n = R_n - R_{n-1} - 1 = 2^{n+2} - 1 \geq \frac{1}{4} R_n \geq 4$$

or $n \geq 1$. Let

$$\alpha_n = \frac{2f(R_n) + 2\ln\ln R_n}{r_n}.$$

Then we have that by condition (4.2)

$$\alpha_n \leq \frac{8f(R_n)}{R_n} + \frac{2\ln(n+3)}{2^{n+2}-1} \leq 8M + 1,$$

and

$$1 + \frac{2}{\alpha_n r_n} = 1 + \frac{2}{2f(R_n) + 2\ln\ln R_n} \leq 1 + \frac{1}{f(0)}.$$

Write

$$C_1 = \frac{1}{8e^{32M+4}}$$

for short. Choose

$$c_n = C_1 \frac{r_n^2}{f(R_n) + \ln\ln R_n}.$$

The estimate (4.16) gives

$$\begin{aligned} \mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n) &\leq \frac{1 + \frac{1}{f(0)}}{\sqrt{C_\mu}} \exp\{\frac{1}{2}(f(R_n) + \ln\ln R_n) + \frac{1}{2} f(R_n) - 2(f(R_n) + \ln\ln R_n)\} \\ &\leq \frac{1 + \frac{1}{f(0)}}{\sqrt{C_\mu}} \exp\{-\frac{3}{2}\ln\ln R_n\}. \end{aligned} \tag{4.17}$$

It easily follows that

$$\sum_{n=1}^{\infty} \mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n) < \infty.$$

We can now determine a function ψ such that (4.6) holds. Consider

$$\begin{aligned} T_n &= \sum_{m=1}^{n} c_m \\ &\geq \frac{C_1}{16}\sum_{m=1}^{n} \frac{R_m^2}{f(R_m)+\ln\ln R_m} \\ &= \frac{C_1}{32}\sum_{m=1}^{n} \frac{R_{m+1}(R_{m+1}-R_m)}{f(R_m)+\ln\ln R_m} \\ &\geq \frac{C_1}{32}\int_{R_1}^{R_{n+1}} \frac{rdr}{f(r)+\ln\ln r}. \end{aligned}$$

Set

$$\psi(R) = \frac{C_1}{64}\int_{8}^{R} \frac{rdr}{f(r)+\ln\ln r}.$$

By condition (4.2) we can see that

$$T_n - \psi(R_{n+1}) \to \infty \text{ as } n \to \infty.$$

Thus $\psi^{-1}(t)$ is the desired upper rate function. □

4.4. **Upper rate function for case (2).**

Proof of Theorem 4.2, case (2). Let

$$\varepsilon = (\frac{C_0}{2} - 1) \wedge 1 > 0$$

and

$$\frac{8}{\varepsilon} + 1 \geq k_0 = [\frac{8}{\varepsilon}] + 1 \geq \frac{8}{\varepsilon} \geq 8.$$

We choose $R_n = k_0^n$ such that for all $n \geq 1$,

$$r_n = k_0^n - k_0^{n-1} - 1 \geq (1 - \frac{2}{k_0})R_n \geq (1 - \frac{\varepsilon}{4})R_n > 2.$$

Choose

$$\alpha_n = \frac{(\frac{1}{2}+\frac{\varepsilon}{4})f(R_n)}{r_n}.$$

We have that

$$1 + \frac{2}{\alpha_n r_n} = 1 + \frac{2}{(\frac{1}{2}+\frac{\varepsilon}{4})f(R_n)} < 1 + \frac{4}{f(0)}.$$

Then choose

$$c_n = \frac{\varepsilon f(R_n)}{8\alpha_n^2 e^{4\alpha_n}}.$$

The estimate (4.16) gives

$$\mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n) \leq \frac{1+\frac{4}{f(0)}}{\sqrt{C_\mu}} \exp\{-\frac{\varepsilon}{8} f(R_n)\}. \tag{4.18}$$

By the assumed monotonicity of $\frac{f(r)}{r}$ for $r \geq 1$, we obtain that

$$\begin{aligned}\sum_1^\infty \mathbb{P}_{\bar{x}}(\tau_n - \tau_{n-1} \leq c_n) &\leq \frac{1+\frac{4}{f(0)}}{\sqrt{C_\mu}} \sum_1^\infty \exp\{-\frac{\varepsilon}{8} f(R_n)\} \\ &\leq \frac{1+\frac{4}{f(0)}}{\sqrt{C_\mu}} \sum_1^\infty \exp\{-\frac{\varepsilon}{8} f(1) R_n\} < \infty.\end{aligned}$$

We can now determine a function ψ such that (4.6) holds. Recall that

$$r_n = (1 - \frac{1}{k_0} - \frac{1}{k_0^n}) R_n \geq (1 - \frac{\varepsilon}{4}) R_n.$$

Note the elementary fact that

$$\frac{1+\frac{\varepsilon}{2}}{1-\frac{\varepsilon}{4}} \leq 1 + \varepsilon,$$

where $0 < \varepsilon \leq 1$. Then we can estimate c_n as

$$\begin{aligned}c_n = \frac{\varepsilon f(R_n)}{8\alpha_n^2 e^{4\alpha_n}} &= \frac{\varepsilon}{8(\frac{1}{2}+\frac{\varepsilon}{4})^2} \frac{r_n^2}{f(R_n)\exp\{2(1+\frac{\varepsilon}{2})\frac{f(R_n)}{r_n}\}} \\ &\geq \frac{\varepsilon(1-\frac{\varepsilon}{4})^2}{8(\frac{1}{2}+\frac{\varepsilon}{4})^2} \frac{R_n^2}{f(R_n)\exp\{2(1+\varepsilon)\frac{f(R_n)}{R_n}\}}.\end{aligned}$$

Write

$$C = \frac{(1-\frac{\varepsilon}{4})^2 \varepsilon}{16(\frac{1}{2}+\frac{\varepsilon}{4})^2 \frac{8}{\varepsilon}(\frac{8}{\varepsilon}+1)}$$

for short. Hence

$$\begin{aligned}T_n &= \sum_{m=1}^n c_m \\ &\geq \frac{\varepsilon(1-\frac{\varepsilon}{4})^2}{8(\frac{1}{2}+\frac{\varepsilon}{4})^2} \sum_{m=1}^n \frac{R_m^2}{f(R_m)\exp\{2(1+\varepsilon)\frac{f(R_m)}{R_m}\}} \\ &= \frac{(1-\frac{\varepsilon}{4})^2 \varepsilon}{8(\frac{1}{2}+\frac{\varepsilon}{4})^2 k_0(k_0-1)} \sum_{m=1}^n \frac{R_{m+1}(R_{m+1}-R_m)}{f(R_m)\exp\{2(1+\varepsilon)\frac{f(R_m)}{R_m}\}} \\ &\geq 2C \int_{R_1}^{R_{n+1}} \frac{r\,dr}{f(r)\exp\{C_0 \frac{f(r)}{r}\}},\end{aligned}$$

where in the lase inequality we have applied the assumption that $\frac{f(r)}{r}$ is increasing for $r \geq 1$. Set

$$\psi(R) = C \int_1^R \frac{r\,dr}{f(r)\exp\{C_0 \frac{f(r)}{r}\}}.$$

By condition (4.4) we can see that

$$T_n - \psi(R_{n+1}) \to \infty \text{ as } n \to \infty.$$

Thus $\psi^{-1}(t)$ is the desired upper rate function. □

5. Weakly spherically symmetric weighted graphs

In this section (V, w, μ) will be the weakly spherically symmetric weighted graph considered in Example 3.12. Let $\left((X_t)_{t\geq 0}, \{\mathbb{P}_x\}_{x\in V}\right)$ be the corresponding minimal symmetric càdlàg Markov chain as constructed in Section 1. Without further specification, the filtration we use here is the augmented one $(\mathcal{F}_t^{\bar{x}})_{t\geq 0}$ for some fixed reference point $\bar{x}$. We essentially follow the strategy in Grigor'yan and Hsu [25]: split the radical process into a dominating part and a relatively small part. However, the technical details are more involved due to the discreteness of the space. We first state a more general form of Theorem 0.5.

Theorem 5.1. *Let (V, w, μ) be the weakly spherically symmetric weighted graph in Example 3.12 with reference point $\bar{x}$. Let d be the adapted metric on V defined by (3.5) and (3.6), and assume that V is unbounded with respect to d. Denote $r(x) := d(x, \bar{x})$. Suppose that there are some constant $c > 0$ and a positive, continuous, non-decreasing function m on $[0, \infty)$, such that for all $x \in V$ outside a finite set,*

$$m(r(x)) \geq -\Delta r(x) > c. \tag{5.1}$$

Assume that

$$\int_0^\infty \frac{ds}{m(s)} = \infty. \tag{5.2}$$

Define the function $\phi(t)$ by the identity

$$t = \int_0^{\phi(t)} \frac{ds}{m(s)}. \tag{5.3}$$

Then the function $\phi(ct)$ is an upper rate function for the corresponding birth and death process for any $c > 1$. If instead of (5.1) we assume $m(r(x)) = -\Delta r(x) > c$ for all but finite many $x \in V$, then $\phi(ct)$ is not an upper rate function for any $0 < c < 1$.

Remark 5.2. By monotonicity of m, (5.2) implies the following

$$\sum_{n=0}^{\infty} \frac{\sigma_n}{m(r_n)} = \infty, \tag{5.4}$$

where $r_n := d(x, \bar{x})$ for $x \in S_n$ ($d(x, \bar{x})$ depends only on $\rho(x)$).

If in the theorem we assume (5.4) instead, then the space (V, d) is necessarily unbounded and the function m is well defined on $[0, \infty)$. By the monotonicity of m, it is easy to show that (5.2) is equivalent to either one of following:

$$\int_0^\infty \frac{ds}{m(s)} = \infty \text{ or } \sum_{n=0}^{\infty} \frac{\sigma_n}{m(r_{n+1})} = \infty.$$

Indeed, since m is nondecreasing, there are two possibilities: either $\lim_{n\to\infty} \frac{1}{m(r_n)} > 0$ or $\lim_{n\to\infty} \frac{1}{m(r_n)} = 0$. For the first case, there are two constants $0 < c_1 < c_2$ such that $c_1 < \frac{1}{m(r_n)} < c_2$ for any $n \in \mathbb{N}$. The assertion is obvious. For the second case, we have

$$0 \leq \sum_{n=0}^{\infty} \left(\frac{\sigma_n}{m(r_n)} - \frac{\sigma_n}{m(r_{n+1})} \right) \leq \sum_{n=0}^{\infty} \left(\frac{1}{m(r_n)} - \frac{1}{m(r_{n+1})} \right) < \infty.$$

Before proving Theorem 5.1, we start with several auxiliary results.

Proposition 5.3. *Let (V, w, μ) be a weakly spherically symmetric weighted graph. Recall the notations*

$$\mu_n = \sum_{x \in S_n} \mu(x), \text{ and } w_n = \sum_{x \in S_n} \sum_{y \in S_{n+1}} w(x, y).$$

The corresponding process $\left((X_t)_{t \geq 0}, \{\mathbb{P}_x\}_{x \in V}\right)$ is transient if and only if

$$\sum_{n=0}^{\infty} \frac{1}{w_n} < \infty. \tag{5.5}$$

Proof. It is known that transience of a symmetric Markov process is equivalent to the following condition on capacity of point $\bar{x}$ (see [47], Theorem 4.5.1):

$$\operatorname{cap}(\bar{x}) := \inf \frac{1}{2} \sum_{x \in V} \sum_{y \in V} w(x, y) \left(f(x) - f(y)\right)^2 > 0, \tag{5.6}$$

where the infimum is taken over all finitely supported functions f on V with $f(\bar{x}) = 1$.

We first prove that (5.5) implies transience. Consider the finite sum

$$\begin{aligned} &\sum_{x \in S_n} \sum_{y \in S_{n+1}} w(x, y) \left(f(x) - f(y)\right)^2 \\ \geq &\frac{1}{w_n} \left(\sum_{x \in S_n} \sum_{y \in S_{n+1}} w(x, y) f(x) - \sum_{x \in S_n} \sum_{y \in S_{n+1}} w(x, y) f(y) \right)^2 \\ = &w_n \times \left(\sum_{x \in S_n} f(x) \cdot \frac{\sum_{y \in S_{n+1}} w(x, y)}{w_n} - \sum_{y \in S_{n+1}} f(y) \cdot \frac{\sum_{x \in S_n} w(x, y)}{w_n} \right)^2, \end{aligned}$$

where we applied the Cauchy-Schwarz inequality. From Definition 3.10 of weakly spherically symmetry, we see that for $x \in S_n$

$$\frac{\sum_{y \in S_{n+1}} w(x, y)}{w_n} = \frac{\mu(x)\kappa_+(x)}{\sum_{z \in S_n} \mu(z)\kappa_+(z)} = \frac{\mu(x)}{\mu_n},$$

and similarly for $y \in S_{n+1}$

$$\frac{\sum_{x \in S_n} w(x, y)}{w_n} = \frac{\mu(y)}{\mu_{n+1}}.$$

We obtain for all finitely supported f with $f(\bar{x}) = 1$,

$$\frac{1}{2}\sum_{x\in V}\sum_{y\in V} w(x,y)\left(f(x)-f(y)\right)^2$$

$$\geq \sum_{n=0}^{\infty}\sum_{x\in S_n}\sum_{y\in S_{n+1}} w(x,y)\left(f(x)-f(y)\right)^2$$

$$\geq \sum_{n=0}^{\infty} w_n \times \left(\sum_{x\in S_n} f(x)\cdot\frac{\mu(x)}{\mu_n} - \sum_{y\in S_{n+1}} f(y)\cdot\frac{\mu(y)}{\mu_{n+1}}\right)^2 \tag{5.7}$$

$$\geq \frac{1}{\sum_{n=0}^{\infty}\frac{1}{w_n}}. \tag{5.8}$$

For the last inequality we applied Cauchy-Schwarz inequality and the fact that f is finitely supported with $f(\bar{x}) = 1$. So the "if" part of the assertion follows directly.

For the other direction, assume that

$$\sum_{n=0}^{\infty}\frac{1}{w_n} = \infty.$$

Define a sequence of functions $(f_N)_{N\in\mathbb{N}}$ by

$$f_N(x) = \begin{cases} \dfrac{\frac{1}{w_n} + \cdots \frac{1}{w_N}}{\frac{1}{w_0} + \cdots \frac{1}{w_N}}, & \text{if } x \in S_n, 0 \leq n \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

By direct calculation,

$$\frac{1}{2}\sum_{x\in V}\sum_{y\in V} w(x,y)\left(f_N(x)-f_N(y)\right)^2 = \frac{1}{\frac{1}{w_0} + \cdots \frac{1}{w_N}} \to 0,$$

as $N \to \infty$, whence

$$0 \leq \operatorname{cap}(\bar{x}) \leq \inf_{N\in\mathbb{N}} \frac{1}{2}\sum_{x\in V}\sum_{y\in V} w(x,y)\left(f_N(x)-f_N(y)\right)^2 = 0.$$

□

Remark 5.4. In fact, Theorem 5.9 in Woess [47] only treated the transience in discrete time case. However, the transience property of a continuous time Markov chain only depends on the associated discrete time Markov chain, see for example [39], Theorem 3.4.1 or [42], Theorem 4.4.5.

Proposition 5.5. *Let (V, w, μ) be a weakly spherically symmetric weighted weighted graph as in Example 3.12. Let d be the adapted metric on V defined by (3.5) and $r(x) = d(x, \bar{x})$. Suppose that for all $x \in V$ large enough, and for some constant $c > 0$,*

$$-\Delta r(x) > c.$$

Then the corresponding process $\left((X_t)_{t\geq 0}, \{\mathbb{P}_x\}_{x\in V}\right)$ is transient.

Proof. Fix some $N \geq 1$ such that for all $n \geq N$ and all $x \in S_n$,

$$\begin{aligned}-(\Delta r)(x) =& \frac{1}{\mu(x)} \sum_{y \in S_{n+1}} w(x,y)\sigma_n - \frac{1}{\mu(x)} \sum_{y \in S_{n-1}} w(x,y)\sigma_{n-1} \\ =& \kappa_+(n)\sigma_n - \kappa_-(n)\sigma_{n-1} = \frac{w_n\sigma_n - w_{n-1}\sigma_{n-1}}{\mu_n} > c > 0.\end{aligned}$$

Then for all $n > N$

$$w_n\sigma_n \geq c\left(\mu_N + \cdots + \mu_n\right).$$

By definition of the metric d, $2w_n\sigma_n^2 = \mu_n$ and hence

$$\frac{1}{w_n} \leq \frac{w_n\sigma_n^2}{c^2\left(\mu_N + \cdots + \mu_n\right)^2} = \frac{\mu_n}{2c^2\left(\mu_N + \cdots + \mu_n\right)^2}.$$

Note that $\sum_{n=N}^{\infty} \mu_n = \infty$ by monotonicity of μ, we have

$$\begin{aligned}&\sum_{n=N}^{\infty} \frac{\mu_n}{\left(\mu_N + \cdots + \mu_n\right)^2} \\ \leq& \frac{1}{\mu_N} + \sum_{n=N}^{\infty}\left(\frac{1}{\mu_N + \cdots + \mu_n} - \frac{1}{\mu_N + \cdots + \mu_{n+1}}\right) < \infty.\end{aligned}$$

Condition (5.5) follows directly. □

Remark 5.6. The monotonicity of μ is in fact not used in the proof. We only need that $\sum_{n=0}^{\infty} \mu_n = \infty$.

Proposition 5.7. *Under the assumptions of Theorem 5.1, the process* $\left((X_t)_{t \geq 0}, \{\mathbb{P}_x\}_{x \in V}\right)$ *is stochastically complete.*

Proof. Keller, Lenz and Wojciechowski proved the following equivalent condition for stochastic completeness of a weakly spherically symmetric weighted graph.

$$\sum_{n=0}^{\infty} \frac{\mu_0 + \cdots + \mu_n}{w_n} = \infty. \tag{5.10}$$

Fix some $N \geq 1$ such that for all $n \geq N$ and all $x \in S_n$,

$$m(r(x)) \geq -(\Delta r)(x) = \frac{w_n\sigma_n - w_{n-1}\sigma_{n-1}}{\mu_n} > 0.$$

By monotonicity of m, we have that for all $K > N$,

$$\begin{aligned}\mu_0 + \cdots + \mu_K \geq& \sum_{n=N}^{K} \frac{w_n\sigma_n - w_{n-1,n}\sigma_{n-1}}{m(r_n)} \\ \geq& \frac{1}{m(r_K)} \sum_{n=N}^{K} \left(w_n\sigma_n - w_{n-1}\sigma_{n-1}\right) \\ =& \frac{w_K\sigma_K - w_{N-1}\sigma_{N-1}}{m(r_K)}.\end{aligned}$$

Hence

$$\sum_{n=0}^{\infty} \frac{\mu_0 + \cdots + \mu_n}{w_n} \geq \sum_{n=N+1}^{\infty} \left(\frac{\sigma_n}{m(r_n)} - \frac{w_{N-1}\sigma_{N-1}}{m(r_n)w_n} \right).$$

In Proposition 5.5, we proved that

$$\sum_{n=0}^{\infty} \frac{1}{w_n} < \infty.$$

It follows that

$$\sum_{n=N+1}^{\infty} \frac{w_{N-1}\sigma_{N-1}}{m(r_n)w_n} \leq \frac{w_{N-1}\sigma_{N-1}}{m(r_N)} \sum_{n=N+1}^{\infty} \frac{1}{w_n} < \infty.$$

Hence for some constant $c' > 0$,

$$\sum_{n=0}^{\infty} \frac{\mu_0 + \cdots + \mu_n}{w_n} \geq \sum_{n=N+1}^{\infty} \frac{\sigma_n}{m(r_n)} - c' = \infty.$$

The assertion follows. □

The following result is proved by tools from martingale theory. It is a direct consequence of Proposition A.2 in the Appendix.

Proposition 5.8. *Let $M_t = r(X_t) + \int_0^t (\Delta r)(X_s)ds$. Under the assumptions of Theorem 5.1, the process $(M, \mathbb{P}_{\bar{x}})$ is a martingale. Furthermore, for all $t > 0$, $\mathbb{E}_{\bar{x}} \left(M_t^2 \right) \leq t$.*

Proof. We only need to check that the function $r(x)$ satisfies the conditions in Proposition A.1 and Proposition A.2. The assumption (3) on f in Proposition A.1 is fulfilled by our technical assumption that $w(x, y) = 0$ for all x, y with $\rho(x) = \rho(y)$ in Definition 3.10. Other conditions are clear. □

Now we can prove our Theorem 5.1.

Proof of Theorem 5.1. We divide the proof into several steps.

Step 1: For any $\varepsilon > 0$, by Doob's maximal martingale inequality, we have

$$\mathbb{P}_{\bar{x}} \left(\sup_{0 \leq s \leq 2^n} |M_s| \geq \varepsilon 2^n \right) \leq \frac{1}{\varepsilon^2 2^{2n}} \sup_{0 \leq s \leq 2^n} \mathbb{E}_{\bar{x}} \left(M_s^2 \right) \leq \frac{1}{\varepsilon^2 2^n}.$$

It follows that

$$\sum_{n=1}^{\infty} \mathbb{P}_{\bar{x}} \left(\sup_{0 \leq s \leq 2^n} |M_s| \geq \varepsilon 2^n \right) < \infty.$$

By Borel-Cantelli lemma, we have that

$$\mathbb{P}_{\bar{x}} \{ \sup_{0 \leq s \leq 2^n} |M_s| < \varepsilon 2^n \text{ for all sufficiently large n } \} = 1.$$

For all large enough $t > 0$, we can find $n \in \mathbb{N}$ such that $t \leq 2^n < 2t$. And hence,

$$\mathbb{P}_{\bar{x}} \{ |M_t| < 2\varepsilon t \text{ for all sufficiently large t } \} = 1. \tag{5.11}$$

Step 2: Since the process $(X_t)_{t\geq 0}$ is transient (Proposition 5.5) and the space (V, d) has infinite diameter, we have with probability 1, $r(X_t) \to \infty$ as $t \to \infty$. It follows that, with probability 1,

$$\liminf_{t\to\infty} \frac{\int_0^t m(r(X_s))ds}{-\int_0^t (\Delta r)(X_s)ds} \geq 1,$$

where we used the facts that m is a positive nondecreasing function and that $m(r(x)) \geq -(\Delta r)(x)$ for all $x \in S_n$ with large enough n.

By definition of M_t, we have

$$r(X_t) = M_t - \int_0^t (\Delta r)(X_s)ds,$$

which implies that, for any $\varepsilon > 0$, the following inequality holds,

$$r(X_t) \leq (1+\varepsilon)\int_0^t m(r(X_s))ds, \tag{5.12}$$

for all sufficiently large t, with probability 1. All further inequalities in this argument are also understood for all large enough t, with probability 1 in the sense similar to (5.11).

Step 3: Consider the following functions:

$$v(t) := \int_0^t m(r(X_s))ds,$$
$$u_1(t) := \int_0^t m\left((1+\varepsilon)v(s)\right)ds,$$

Note that with probability 1, v is continuous and u_1 are continuously differentiable. By (5.12), we have

$$r(X_t) \leq (1+\varepsilon)v(t), \tag{5.13}$$

whence by monotonicity of m,

$$m\left(r(X_t)\right) \leq m\left((1+\varepsilon)v(t)\right).$$

Integrating the above inequality with respect to t, we obtain

$$v(t) \leq (1+\varepsilon)u_1(t). \tag{5.14}$$

Applying the monotonicity of m to (5.14), since $\frac{du_1(t)}{dt} = m\left((1+\varepsilon)v(t)\right)$, we obtain a differential inequality for $u_1(t)$:

$$\frac{du_1(t)}{dt} \leq m\left((1+\varepsilon)^2 u_1(t)\right).$$

Rename $C_1 = (1+\varepsilon)^2$. We have

$$\int_0^{C_1 u_1(t)} \frac{dr}{m(r)} \leq C_1 t + C_0 \tag{5.15}$$

where C_0 is a large enough random constant. Note that by (5.13) and (5.14),

$$r(X_t) \leq C_1 u_1(t).$$

Combining (5.15) with the above inequality, we obtain by definition of $\phi(t)$ in (0.13),

$$r(X_t) \leq C_1 u_1(t) \leq \phi(C_1 t + C_0) \leq \phi(C_1^2 t).$$

Thus $\phi(ct)$ is an upper rate function for any $c > 1$.

Step 4: Suppose now $m(r(x)) = -\Delta r(x)$ for all large enough x. Then with probability 1,

$$\lim_{t\to\infty} \frac{\int_0^t m(r(X_s))ds}{-\int_0^t (\Delta r)(X_s)ds} = 1.$$

And hence for any $\varepsilon > 0$,

$$(1-\varepsilon)\int_0^t m(r(X_s))ds \leq r(X_t) = M_t - \int_0^t (\Delta r)(X_s)ds. \tag{5.16}$$

Consider another function:

$$u_2(t) := \int_0^t m\left((1-\varepsilon)v(s)\right)ds.$$

By (5.16),

$$(1-\varepsilon)v(t) \leq r(X_t). \tag{5.17}$$

whence by monotonicity of m,

$$m\left((1-\varepsilon)v(t)\right) \leq m\left(r(X_t)\right).$$

Integrating the above inequality in t, we obtain

$$(1-\varepsilon)u_2(t) \leq v(t). \tag{5.18}$$

Applying the monotonicity of m once more to (5.18), since $\frac{du_2(t)}{dt} = m\left((1-\varepsilon)v(t)\right)$, we obtain a differential inequality for $u_2(t)$:

$$\frac{du_2(t)}{dt} \geq m\left((1-\varepsilon)^2 u_2(t)\right).$$

Rename $C_2 = (1-\varepsilon)^2$. We have

$$\int_0^{C_2 u_2(t)} \frac{dr}{m(r)} \geq C_2 t - C_0, \tag{5.19}$$

where C_0 is a large enough random constant. Note that by (5.17) and (5.18),

$$C_2 u_2(t) \leq r(X_t).$$

Combining (5.19) with the above inequality, we obtain by definition of $\phi(t)$ in (0.13),

$$\phi(C_2^2 t) \leq \phi(C_2 t - C_0) \leq C_2 u_2(t) \leq r(X_t),$$

for all large enough t with probability 1. Thus $\phi(ct)$ is not an upper rate function for any $0 < c < 1$. In fact, as we can see from the proof, $\phi(ct)$ is a lower rate function for any $0 < c < 1$.

□

For illustration, let us show how to apply Theorem 5.1 to Example 3.13 and Example 3.14. We omit the details of calculations as they are essentially the same as in Example 0.10 and Example 0.11.

In Example 3.13, if we choose $\kappa_+(x) = [(n+1)^s]$ for $x \in S_n$, $0 < s \le 1$, then the upper rate function $\phi(t)$ has the form

$$\phi(t) \asymp t^{\frac{2-s}{2-2s}}, \text{ if } 0 < s < 1, \text{ and } \phi(t) \asymp e^{Ct}, \text{ if } s = 1, \tag{5.20}$$

where $C > 0$ is some constant. The volume growth function satisfies

$$\ln\left(B_d(\bar{x}, r)\right) \asymp r^{\frac{2}{2-s}} \ln r,$$

for all r large enough.

In Example 3.14, if we choose $\kappa_+(x) = [(n+1)^2 \ln(n+e)^s]$ for $x \in S_n$, then the upper rate function $\phi(t)$ again is given by (5.20), whereas the volume growth function satisfies that

$$\ln\left(B_d(\bar{x}, r)\right) \asymp r^{\frac{2}{2-s}},$$

for all r large enough.

Although the upper rate functions for the two examples have the same form, the volume growth functions behave differently. Hence, even for the physical Laplacian case, obtaining a sharp upper rate function purely in terms of the volume growth is impossible.

Appendix A

In this appendix, we collect some basic facts for the stochastic analysis of pure jump martingales and provide a proof for Proposition 5.8. Although results here are not new, it is hard to find explicit references. We include proofs here for the sake of completeness.

The following proposition is a generalization of Theorem 2 in Hamza and Klebaner [27].

Proposition A.1. *Let (V, w, μ) be a locally finite, connected weighted graph with a fixed reference point $\bar{x} \in V$ and $\left((X_t)_{t\ge 0}, \mathcal{F}^{\bar{x}}, (\mathbb{P}_x)_{x\in V_\infty}\right)$ be the corresponding minimal càdlàg Markov chain. Assume that the process is stochastically complete. Let f be an unbounded function on V that satisfies the following conditions:*

1. *$f(\bar{x}) = 0$;*
2. *the sets $\{x \in V : |f(x)| \le n\}$ are finite for all $n \in \mathbb{N}$;*
3. *$f(x) \ne f(y)$ for any pair $x \sim y$.*

Then under $\mathbb{P}_{\bar{x}}$, the process M defined by

$$M_t := f(X_t) + \int_0^t (\Delta f)(X_s)\, ds$$

is a local martingale with respect to the filtration $\mathcal{F}^{\bar{x}}$.

Proof. We will follow Jacod and Shiryaev [34] for the theory of random measures and compensators (or dual predictable projections). Every process in the proof will be considered under the probability $\mathbb{P}_{\bar{x}}$ and the filtration $(\mathcal{F}_t^{\bar{x}})_{t\geq 0}$. We make the convention that $\inf\emptyset = \infty$. Let in the following

$$X_{t-} = \lim_{s\to t, s<t} X_s,$$

the existence of which is guaranteed by the càdlàg property of the process.

Consider the process $Y_t := f(X_t)$ which is a càdlàg pure jump type process on $\mathbb{R}$. Let μ^f be the random measure counting the jumps of Y:

$$\mu^f(\omega; dt, dx) := \sum_s \mathbf{1}_{\{Y_s(\omega)-Y_{s-}(\omega)\neq 0\}}\delta_{(s, Y_s(\omega)-Y_{s-}(\omega))}(dt, dx),$$

where $\delta_{(s,x)}$ is the Dirac measure on $\mathbb{R}\times\mathbb{R}$ at point (s,x). Note that by the assumption (3) on f, $f(X_s) - f(X_{s-}) \neq 0$ if and only if $X_s \neq X_{s-}$ (since we have $X_s \sim X_{s-}$ when $X_s \neq X_{s-}$). Define a sequence of stopping times $(\tau_n)_{n\geq 0}$ by

$$\tau_0 = 0, \tau_{n+1} = \inf\{t > \tau_n : X_t \neq X_{t-}\},$$

which are the jumping times of the process X. We can rewrite μ^f as

$$\mu^f(dt, dx) = \sum_{n\geq 1} \delta_{(\tau_n, Y_{\tau_n} - Y_{\tau_n -})}(dt, dx),$$

where we omit the dependence on the sample ω for simplicity.

By Theorem 1.33 in [34] and the strong Markov property of the process X, we have the compensator of μ^f is given by (see for example [27], [28], [33])

$$\nu^f(dt, \{x\}) = \sum_{n\geq 0} \mathbf{1}_{\{\tau_n < t \leq \tau_{n+1}\}} \mathrm{Deg}\,(X_{\tau_n})\,\mathbb{P}_{X_{\tau_n}}[Y_{\tau_{n+1}} - Y_{\tau_n} = x]dt.$$

Let $S_n = \inf\{t > 0 : |Y_t| > n\}$. Define another sequence of stopping times $(T_n)_{n\geq 1}$ by $T_n = \min\{n, S_n\}$. Since the process X is stochastically complete, on each finite time interval $[0,t]$, there are only finitely many jumps of X, $\mathbb{P}_{\bar{x}}$ almost surely. It follows that $\mathbb{P}_{\bar{x}}$ almost surely, for all $t > 0$, there is some n such that $S_n > t$, as f is unbounded. Then we see that $\mathbb{P}_{\bar{x}}$ almost surely $\lim_{n\to\infty} T_n = \infty$.

Consider the functions $F_n(t,x) = x \cdot \mathbf{1}_{\{t\leq T_n\}}$. By the property of compensator of a random measure (see for example Theorem 1.8 in [34]), we have for any n,

$$\begin{aligned}
\mathbb{E}_{\bar{x}}\left(\int |F_n(t,x)|\mu^f(dt,dx)\right) &= \mathbb{E}_{\bar{x}}\left(\int |F_n(t,x)|\nu^f(dt,dx)\right) \\
&= \mathbb{E}_{\bar{x}}\left(\int_0^{T_n} \frac{1}{\mu(X_{s-})}\sum_{y\in V} w(X_{s-},y)|f(X_{s-}) - f(y)|ds\right) &\text{(A.1)} \\
&= \mathbb{E}_{\bar{x}}\left(\int_0^{T_n} \frac{1}{\mu(X_s)}\sum_{y\in V} w(X_s,y)|f(X_s) - f(y)|ds\right). &\text{(A.2)}
\end{aligned}$$

By locally finiteness of (V, w, μ) and finiteness of the sets $\{x \in V : |f(x)| \leq n\}$,

$$A_n := \sup_{\{(t,\omega):0<t\leq T_n(\omega)\}} \frac{1}{\mu(X_s)}\sum_{y\in V} w(X_s,y)|f(X_s) - f(y)| < \infty.$$

By (A.2), we have

$$\mathbb{E}_{\bar{x}}\left(\int |F_n(t,x)|\mu^f(dt,dx)\right) \leq A_n T_n \leq nA_n < \infty,$$

whence the process

$$|Y|_t := \int \mathbf{1}_{\{s\leq t\}}|x|\mu^f(dt,dx)$$

is locally integrable.

Then we have by direct calculation

$$Y_t = f(X_t) = f(X_t) - f(X_0) = \int \mathbf{1}_{\{s\leq t\}} x\mu^f(dt,dx),$$

$$\int_0^t (\Delta f)(X_s)\, ds = \int_0^t \frac{1}{\mu(X_s)} \sum_{y\in V} w(X_s,y)\left(f(X_s)-f(y)\right) ds = -\int \mathbf{1}_{\{s\leq t\}} x\nu^f(dt,dx)$$

$\mathbb{P}_{\bar{x}}$ almost surely, since the right hand side of each equality is well defined. Moreover, by Theorem 1.8 in [34], we have that the process M is a local martingale since

$$M_t = Y_t + \int_0^t (\Delta f)(X_s)\, ds = \int \mathbf{1}_{\{s\leq t\}} x\left(\mu^f(dt,dx) - \nu^f(dt,dx)\right).$$

□

Proposition A.2. *Under the same assumptions of Proposition A.1, if we assume furthermore that*

$$\frac{1}{\mu(x)} \sum_{y\in V} w(x,y)(f(x)-f(y))^2 \leq 1 \tag{A.3}$$

for all $x \in V$, then the process M is a martingale, and $\mathbb{E}_{\bar{x}}(M_t^2) \leq t$ for all $t > 0$.

Proof. We adopt the same notations as in the proof of Proposition A.1. The quadratic variation process $[M,M]$ of the local martingale M is given by

$$[M,M]_t = \sum_{0<s\leq t} (f(X_s) - f(X_{s-}))^2 = \int \mathbf{1}_{\{s\leq t\}} x^2\mu^f(dt,dx).$$

Note that the quadratic predictable covariation process $< M,M >$ of M satisfies

$$< M,M >_t = \int \mathbf{1}_{\{s\leq t\}} x^2\nu^f(dt,dx) = \int_0^t \frac{1}{\mu(X_s)} \sum_{y\in V} w(X_s,y)\left(f(X_s)-f(y)\right)^2 ds.$$

By (A.3), we have that for all $t > 0$,

$$\mathbb{E}_{\bar{x}}\left([M,M]_t\right) = \mathbb{E}_{\bar{x}}\left(< M,M >_t\right) \leq t.$$

Then it follows from standard results in martingale theory that M is a martingale with

$$\mathbb{E}_{\bar{x}}(M_t^2) = \mathbb{E}_{\bar{x}}\left([M,M]_t\right) \leq t$$

for all $t > 0$ (see for example Protter [40], Corollary 3, p. 72). □

Remark A.3. Condition (A.3) corresponds to the adaptedness condition when f is given by the distance from a reference point. From this martingale theory point of view, we also see that the adapted metric is a natural notion.

Acknowledgement

This paper is based on part of the author's Ph.D. thesis [31] at Bielefeld University. The author is grateful to his supervisor, Prof. Grigor'yan for his continuous encouragement. The author also would like to thank Naotaka Kajino and Yutao Ma for inspiring discussions.

References

1. D. G. Aronson, *Bounds for the fundamental solution of a parabolic equation*, Bull. Amer. Math. Soc. **73** (1967), 890–896.
2. D.G. Aronson, *Non-negative solutions of linear parabolic equations*, Ann. Scuola Norm. Sup. Pisa. Cl. Sci. **22** (1968. Addendum **25**:221–228, 1971), no. 4, 607–694.
3. M. Barlow and E.A. Perkins, *Symmetric Markov chains in $\mathbf{Z}^d$: how fast can they move?*, Probab. Theory Related Fields **82** (1989), no. 1, 95–108.
4. R. Bass, *Stochastic processes*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 33, Cambridge University Press, Cambridge, 2011.
5. N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1987.
6. R.M. Blumenthal and R.K. Getoor, *Markov processes and potential theory*, Academic Press, New York-London, 1968.
7. Z.-Q. Chen and M. Fukushima, *Symmetric Markov processes, time change, and boundary theory*, London Mathematical Society Monographs Series, vol. 35, Princeton University Press, Princeton, NJ, 2012.
8. F. R. K. Chung, *Spectral graph theory*, CBMS Regional Conference Series in Mathematics, vol. 92, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1997.
9. K.L. Chung, *Markov chains with stationary transition probabilities*, Second edition. Die Grundlehren der mathematischen Wissenschaften, Band 104, Springer-Verlag New York, Inc., New York, 1967.
10. Thierry Coulhon, Alexander Grigor'yan, and Fabio Zucca, *The discrete integral maximum principle and its applications*, Tohoku Math. J. (2) **57** (2005), no. 4, 559–587.
11. E. B. Davies, *Analysis on graphs and noncommutative geometry*, J. Funct. Anal. **111** (1993), no. 2, 398–430.
12. Y. Colin de Verdière, *Spectres de graphes*, Cours Spécialisés [Specialized Courses], vol. 4, Société Mathématique de France, Paris, 1998.
13. Y. Colin de Verdière, N. Torki-Hamza, and F. Truc, *Essential self-adjointness for combinatorial Schrödinger operators II*, Math. Phys. Anal. Geom. **14** (2011), no. 1, 21–38.
14. J. Dodziuk, *Elliptic operators on infinite graphs*, Analysis, geometry and topology of elliptic operators, World Sci. Publ., Hackensack, NJ, 2006, pp. 353–368.
15. J. Dodziuk and V. Mathai, *Kato's inequality and asymptotic spectral properties for discrete magnetic Laplacians*, The ubiquitous heat kernel, Contemp. Math., vol. 398, Amer. Math. Soc., Providence, RI, 2006, pp. 69–81.
16. A. Dvoretzky and P. Erdös, *Some problems on random walk in space*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950 (Berkeley and Los Angeles), University of California Press, 1951, pp. 353–367.
17. W. Feller, *An introduction to probability theory and its applications*, John Wiley & Sons, Ltd., New York-London-Sydney, 1966.

18. M. Folz, *Gaussian upper bounds for heat kernels of continuous time simple random walks*, Elec. J. Prob. **16** (2011), 1693–1722.
19. R. Frank, D. Lenz, and D. Wingert, *Intrinsic metrics for (non-local) symmetric Dirichlet forms and applications to spectral theory*, preprint, arXiv:1012.5050v1.
20. D. Freedman, *Markov chains*, Holden-Day, San Francisco, Calif., 1971.
21. M. Fukushima, Oshima Y., and M. Takeda, *Dirichlet forms and symmetric Markov processes*, Walter de Gruyter, Berlin, 1994.
22. A. Grigor'yan, *On stochastically complete manifolds*, DAN SSSR **290** (1986), 534–537, in Russian. Engl. transl.: Soviet Math. Dokl., 34 (1987) no.2, 310-313.
23. ______, *Integral maximum principle and its applications*, Proc. Roy. Soc. Edinburgh Sect. A **124** (1994), no. 2, 353–362.
24. ______, *Escape rate of Brownian motion on Riemannian manifolds*, Appl. Anal. **71** (1999), no. 1-4, 63–89.
25. A. Grigor'yan and E. P. Hsu, *Volume growth and escape rate of Brownian motion on a Cartan-Hadamard manifold*, Sobolev spaces in mathematics. II, Int. Math. Ser. (N. Y.), vol. 9, Springer, New York, 2009, pp. 209–225.
26. A. Grigor'yan, X. Huang, and J. Masamune, *On stochastic completeness for nonlocal Dirichlet forms*, Math. Z. (to appear), DOI: 10.1007/s00209-011-0911-x.
27. K. Hamza and F. C. Klebaner, *Conditions for integrability of Markov chains*, J. Appl. Probab. **32** (1995), no. 2, 541–547.
28. S.-W. He, J.-G. Wang, and J.-A. Yan, *Semimartingale theory and stochastic calculus*, Kexue Chubanshe (Science Press), Beijing, 1992, in Chinese.
29. W. Hebisch and L. Saloff-Coste, *Gaussian estimates for Markov chains and random walks on groups*, Ann. Probab. **21** (1993), no. 2, 673–709.
30. E. P. Hsu and G. Qin, *Volume growth and escape rate of Brownian motion on a complete Riemannian manifold*, Ann. Probab. **38** (2010), no. 4, 1570–1582.
31. X. Huang, *On stochastic completeness of weighted graphs*, Ph.D. thesis, Universität Bielefeld.
32. ______, *On uniqueness class for a heat equation on graphs*, (2011), preprint.
33. J. Jacod, *Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **31** (1974/75), 235–253.
34. J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003.
35. M. Keller and D. Lenz, *Dirichlet forms and stochastic completeness of graphs and subgraphs*, J. Reine Angew. Math. (accepted), DOI: 10.1515/CRELLE.2011.122.
36. M. Keller, D. Lenz, and R. K. Wojciechowski, *Volume growth, spectrum and stochastic completeness of infinite graphs*, (2011), preprint, arXiv:1105.0395v1 [math.SP].
37. J. Masamune and T. Uemura, *Conservation property of symmetric jump processes*, Ann. Inst. Henri. Poincare Probab. Statist. **47** (2011), no. 3, 650–662.
38. J. Masamune, T. Uemura, and J. Wang, *On the conservativeness and recurrence of symmetric jump-diffusions*, (2011), preprint.
39. J. R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998, Reprint of 1997 original.
40. P. E. Protter, *Stochastic integration and differential equations*, second ed., Applications of Mathematics (New York), vol. 21, Springer-Verlag, Berlin, 2004, Stochastic Modelling and Applied Probability.
41. L. Saloff-Coste, *Lectures on finite Markov chains*, Lectures on probability theory and statistics (Saint-Flour, 1996), Lecture Notes in Math., vol. 1665, Springer, Berlin, 1997, pp. 301–413. MR 1490046 (99b:60119)
42. D. Stroock, *An introduction to Markov processes*, Graduate Texts in Mathematics, vol. 230, Springer-Verlag, Berlin, 2005.
43. K. T. Sturm, *Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and L^p-Liouville properties*, J. Reine Angew. Math. **456** (1994), 173–196.

44. N. Th. Varopoulos, *Long range estimates for Markov chains*, Bull. Sci. Math. (2) **109** (1985), no. 3, 225–252.
45. A. M. Vershik, *Dynamic theory of growth in groups: entropy, boundaries, examples*, Uspekhi Mat. Nauk **55** (2000), no. 4, 59–128.
46. A. Weber, *Analysis of the laplacian and the heat flow on a locally finite graph*, J. of Math. Anal. and App. **370** (2010), 146–158.
47. W. Woess, *Denumerable Markov chains*, EMS Textbooks in Mathematics, European Mathematical Society (EMS), Zürich, 2009, Generating functions, boundary theory, random walks on trees.
48. R. K. Wojciechowski, *Stochastic completeness of graphs*, ProQuest LLC, Ann Arbor, MI, 2008, Thesis (Ph.D.)–City University of New York.
49. ———, *Heat kernel and essential spectrum of infinite graphs*, Indiana Univ. Math. J. **58** (2009), no. 3, 1419–1441.
50. ———, *Stochastically incomplete manifolds and graphs*, Progress in Probability **64** (2011), 163–179.

Department of Mathematics, University of Bielefeld, 33501 Bielefeld , Germany

E-mail address: `xhuang1@math.uni-bielefeld.de`