OLLIVIER–RICCI CURVATURE AND THE SPECTRUM OF THE NORMALIZED GRAPH LAPLACE OPERATOR

Frank Bauer, Jürgen Jost and Shiping Liu

ABSTRACT. We prove the following estimate for the spectrum of the normalized Laplace operator Δ on a finite graph G,

$$1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}}, \ \forall \text{ integers } t \geq 1.$$

Here $k[t]$ is a lower bound for the Ollivier–Ricci curvature on the neighborhood graph $G[t]$, which was introduced by Bauer–Jost. In particular, when $t = 1$ this is Ollivier's estimates $k \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 2 - k$. For sufficiently large t, we show that, unless G is bipartite, our estimates for λ_1 and λ_{N-1} are always nontrivial and improve Ollivier's estimates for all graphs with $k \leq 0$. By definition neighborhood graphs are weighted graphs which may have loops. To understand the Ollivier–Ricci curvature on neighborhood graphs, we generalize a sharp estimate of the Ricci curvature given by Jost–Liu to weighted graphs with loops and relate it to the relative local frequency of triangles and loops.

1. Introduction

In this paper, we utilize techniques inspired by Riemannian geometry and the theory of stochastic processes in order to control eigenvalues of graphs. In particular, we shall quantify the deviation of a (connected, undirected, weighted, finite) graph G from being bipartite (a bipartite graph is one without cycles of odd lengths; equivalently, its vertex set can be split into two classes such that edges can be present only between vertices from different classes) in terms of a spectral gap. The operator whose spectrum we shall consider here is the normalized graph Laplacian Δ. This is the operator underlying random walks on graphs, and so, this leads to a natural connection with the theory of stochastic processes. We observe that on a bipartite graph, a random walker, starting at a vertex x at time 0 and at each step hopping to one of the neighbors of the vertex where it currently sits, can revisit x only at even times. This connection then will be explored via the eigenvalues of Δ. More precisely, the largest eigenvalue λ_{N-1} of Δ is 2 iff G is bipartite and is < 2 else. Therefore, $2 - \lambda_{N-1}$ quantifies the deviation of G from being bipartite, and we want to understand this aspect in more detail. In more general terms, we are asking for a quantitative connection between the geometry (of the graph G) and the analysis (of the operator Δ, or the random walk encoded by it). Now, such connections have been explored systematically in Riemannian geometry, and many eigenvalue estimates are known there that connect the corresponding Laplace operator with the geometry of the underlying space M; see, e.g., Li–Yau [17], Chavel [5]. The crucial role here is played by the Ricci curvature of M. In recent years, a kind of axiomatic approach to curvature has been developed. This approach encodes the abstract formal properties

of curvature and thereby makes the notion extendible to spaces more general than Riemannian manifolds. By now, there exist many notions of generalized curvature, and several of them have found important applications, see Sturm [24], Lott–Villani [19], Ollivier [21], Ohta [20], Bonciocat–Sturm [3], Joulin–Ollivier [14] and the references therein. The curvature notion that turns out to be most useful for our purposes is the one introduced by Ollivier [21]. In his paper, Olliver actually showed that the eigenvalues of the normalized Laplace operator satisfy

$$(1.1) \qquad k \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 2 - k.$$

In fact, one of the main points of the present paper is to relate lower bounds for λ_1 and upper bounds for λ_{N-1} via random walks. As in Bauer–Jost [2], we translate this relationship into the geometric concept of a neighborhood graph. The idea here is that in the tth neighborhood graph $G[t]$ of G, vertices x and y are connected by an edge with a weight given by the probability that a random walker starting at x reaches y after t steps times the degree of x. We note that even though the original graph may have been unweighted, the neighborhood graphs $G[t]$ are necessarily weighted. In addition, they will in general possess self-loops, because the random walker starting at x may return to x after t steps. Therefore, we need to develop our theory on weighted graphs with self-loops even though the original G might have been unweighted and without such loops. Since Ollivier's curvature is defined in terms of transportation distances (Wasserstein metrics), we can then use our neighborhood graphs in order to geometrically control the transportation costs and thereby to estimate the curvature of the neighborhood graphs in terms of the curvature of the original graph. As it turns out that lower bounds for the smallest eigenvalue of $G[t]$, t even, are related to upper bounds for the largest eigenvalue of G, we obtain the following more general estimate

$$(1.2) \qquad 1 - (1 - k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}}, \quad \forall \text{ integers } t \geq 1.$$

Whereas (1.1) is only useful under the restrictive assumption that k be positive, our estimate (1.2) is nontrivial for any graph that is not bipartite. In fact, for a non-bipartite graph, we obtain an exponential decay of $1 - k[t]$ with a rate that can be controlled by the geometry of the graph.

For controlling the smallest eigenvalue, besides Ollivier [21], we also refer to Lin–Yau [18] and Jost–Liu [13]. In particular, in the last paper, we could relate λ_1 to the local clustering coefficient introduced in Watts–Strogatz [27]. The local clustering coefficients measures the relative local frequency of triangles, that is, cycles of length 3. Since bipartite graphs cannot possess any triangles, this then is obviously related to our question about quantifying the deviation of the given graph G from being bipartite. In fact, in Jost–Liu [13], this local clustering has been controlled in terms of Ollivier's Ricci curvature. Thus, in the present paper we are closing the loop between the geometric properties of a graph G, the spectrum of its graph Laplacian, random walks on G, and the generalized curvature of G, drawing upon deep ideas and concepts originally developed in Riemannian geometry and the theory of stochastic processes.

2. The normalized Laplace operator, neighborhood graphs, and Ollivier–Ricci curvature

In this paper, $G = (V, E)$ will denote an undirected, weighted, connected, finite graph of N vertices. We do not exclude loops, i.e., we permit the existence of an edge between a vertex and itself. V denotes the set of vertices and E denotes the set of edges. If two vertices $x, y \in V$ are connected by an edge, we say x and y are neighbors, in symbols $x \sim y$. The associated weight function $w \colon V \times V \to \mathbb{R}$ satisfies $w_{xy} = w_{yx}$ (because the graph is undirected) and we assume $w_{xy} > 0$ whenever $x \sim y$ and $w_{xy} = 0$ iff $x \nsim y$. For a vertex $x \in V$, its degree d_x is defined as $d_x := \sum_{y \in V} w_{xy}$. If $w_{xy} = 1$ whenever $x \sim y$, we shall call the graph an unweighted one. We will also consider a locally finite graph $\tilde{G} = (\tilde{V}, \tilde{E})$, which is an undirected, weighted, connected graph with a possible infinite number of vertices that satisfies the property that for every $x \in \tilde{V}$, the number of edges connected to x is finite.

2.1. The normalized graph Laplace operator and its eigenvalues. In this subsection, we recall the definition of the normalized graph Laplace operator and state some of its basic properties. In particular, we will emphasize the relations between eigenvalues of the Laplace operator and random walks on graphs.

Let $C(V)$ denote the space of all real-valued functions on the set V and let $m_x(\cdot)$ be a probability measure attached to a vertex $x \in V$.

Definition 1. The Laplace operator $\Delta : C(V) \to C(V)$ is pointwise defined by

$$(2.1) \qquad \Delta f(x) = \sum_{y \in V} f(y) m_x(y) - f(x), \quad \forall x \in V.$$

The measure $m_x(\cdot)$ can also be considered as the distribution of a 1-step random walk starting from x. We will choose

$$(2.2) \qquad m_x(y) = \begin{cases} \dfrac{w_{xy}}{d_x}, & \text{if } y \sim x, \\ 0, & \text{otherwise,} \end{cases}$$

in the following. Note that $x \sim x$ is possible when x has a loop. On a graph G without loops, we can also consider a lazy random walk. A lazy random walk is a random walk that does not move with a certain probability, i.e., for some x we might have $m_x(x) \neq 0$. In this case, the lazy random walk on G is equivalent to the usual random walk on the graph G^{lazy} that is obtained from G by adding for every vertex x a loop with the weight $(d_x + m_x(x)) m_x(x)$, where d_x is the degree of x in G.

With the family (2.2) of probability measures $\{m_x(\cdot)\}$, Δ is just the normalized graph Laplace operator studied for instance in [2, 11] and is unitarily equivalent to the Laplace operator studied in [9].

We also have a natural measure μ on the whole set V, $\mu(x) := d_x$, which gives an inner product structure on $C(V)$.

Definition 2. The inner product of two functions $f, g \in C(V)$ is defined as

$$(2.3) \qquad (f, g)_\mu = \sum_{x \in V} f(x) g(x) \mu(x).$$

With this inner product $C(V)$ becomes a Hilbert space, and we can write $C(V) = l^2(V, \mu)$. By the definition of the degree and the symmetry of the weight function, we can check that

- μ is invariant w.r.t. $\{m_x(\cdot)\}$, i.e., $\sum_{x \in V} m_x(y)\mu(x) = \mu(y), \forall y \in V$;
- μ is reversible w.r.t. $\{m_x(\cdot)\}$, i.e., $m_x(y)\mu(x) = m_y(x)\mu(y), \forall x, y \in V$.

These two facts imply immediately that the operator Δ is nonpositive and self-adjoint on the space $l^2(V, \mu)$. We call λ an eigenvalue of Δ if there exists some $f \not\equiv 0$ such that $\Delta f = -\lambda f$. Using this convention it follows from the observation that Δ is self-adjoint and nonpositive that all its eigenvalues are real and nonnegative. In fact, it's well known that (see, e.g., Chung [9]) $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 2$. Since our graph is connected we actually have $0 < \lambda_1$. In Chung [9] it is shown, by proving a discrete version of the Cheeger inequality, that λ_1 is a measure for how easy/difficult it is to cut the graph into two large pieces. Furthermore, it is well known that $\lambda_{N-1} = 2$ if and only if G is bipartite. In Bauer–Jost [2] a Cheeger type estimate for the largest eigenvalue λ_{N-1} was obtained. The results in Bauer–Jost [2] show that λ_{N-1} is a measure for how close (the meaning of close is made precise in [2]) a graph is to a bipartite one. In the following, we will call λ_1 the first eigenvalue and λ_{N-1} the largest eigenvalue of the operator Δ.

2.2. Neighborhood graphs. In this section, we discuss the deep relationship between eigenvalues estimates for the Laplace operator Δ and random walks on the graph G. In particular, we recall the neighborhood graph method developed by Bauer–Jost [2].

We first introduce the following notation. For a probability measure μ, we denote

$$\mu P(\cdot) := \sum_x \mu(x) m_x(\cdot).$$

Let δ_x be the Dirac measure at x, then we can write $\delta_x P^1(\cdot) := \delta_x P(\cdot) = m_x(\cdot)$. Therefore the distribution of a t-step random walk starting from x with a transition probability m_x is

$$(2.4) \qquad \delta_x P^t(\cdot) = \sum_{x_1, \ldots, x_{t-1}} m_x(x_1) m_{x_1}(x_2) \cdots m_{x_{t-1}}(\cdot)$$

for $t > 1$. The idea is now to define a family of graphs $G[t]$, $t \geq 1$ that encodes the transition probabilities of the t-step random walks on the graph G.

Definition 3. The neighborhood graph $G[t] = (V, E[t])$ of the graph $G = (V, E)$ of order $t \geq 1$ has the same vertex set as G and the weights of the edges of $G[t]$ are defined in terms of the transitions probabilities of the t-step random walk,

$$(2.5) \qquad w_{xy}[t] := \delta_x P^t(y) d_x.$$

In particular, $G = G[1]$ and $x \sim y$ in $G[t]$ if and only if there exists a path of length t between x and y in G. It is easy to see that the neighborhood graph $G[t]$ is in general a weighted graph with loops, even if the original graph G is an unweighted, simple graph. Moreover, we note here that the neighborhood graph method is related to the

discrete heat kernel $p_t(x, y)$ (see, e.g., [1] and the references therein) on graphs by

$$p_t(x, y) = \frac{w_{xy}[t]}{d_x d_y}.$$

Lemma 1 (**Bauer–Jost [2]**)**.** *The neighborhood graph $G[t]$ has the following properties:*

(i) *If t is even, then $G[t]$ is connected if and only if G is not bipartite. Furthermore, if t is even, $G[t]$ is not bipartite.*

(ii) *If t is odd, then $G[t]$ is always connected and $G[t]$ is bipartite iff G is bipartite.*

(iii) $d_x[t] = d_x$ *for all $x \in V$.*

We mention the following crucial observation which can for instance be found in [2]:

Observation 1. *The Laplace operator Δ on G and the Laplace operator $\Delta[t]$ on $G[t]$ are related to each other by the following identity:*

$$\Delta[t] = -\mathrm{id} + (\mathrm{id} + \Delta)^t.$$

An easy consequence of this observation is that the eigenvalues of $\Delta[t]$ satisfy

$$(2.6) \qquad 0 = \lambda_0[t] \leq \lambda_1[t] \leq \cdots \leq \lambda_{N-1}[t] \leq 1$$

if t is even. The reason why all eigenvalues of $G[t]$ (for t even) are less or equal to one is that every vertex in the neighborhood graph $G[t]$ has a loop. Because of equation (2.6) we can assume in the following that bounds for the eigenvalues of $\Delta[t]$, t even, are less or equal to one.

The importance of Observation 1 comes from the following corollary that establishes a connection between estimates for the smallest and the largest eigenvalue on G and $G[t]$, respectively.

Corollary 1 (**Bauer–Jost [2]**)**.** (i) *Let $\mathcal{A}[t]$ be a lower bound for the eigenvalue $\lambda_1[t]$ of $\Delta[t]$, i.e., $\lambda_1[t] \geq \mathcal{A}[t]$. Then*

$$(2.7) \qquad 1 - (1 - \mathcal{A}[t])^{\frac{1}{t}} \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 1 + (1 - \mathcal{A}[t])^{\frac{1}{t}}$$

if t is even and

$$(2.8) \qquad 1 - (1 - \mathcal{A}[t])^{\frac{1}{t}} \leq \lambda_1$$

if t is odd.

(ii) *Let $\mathcal{B}[t]$ be an upper bound for the largest eigenvalue $\lambda_{N-1}[t]$ of $\Delta[t]$, i.e., $\lambda_{N-1}[t] \leq \mathcal{B}[t]$. Then all eigenvalues of Δ are contained in the union of the intervals*

$$\left[0, 1 - (1 - \mathcal{B}[t])^{\frac{1}{t}}\right] \bigcup \left[1 + (1 - \mathcal{B}[t])^{\frac{1}{t}}, 2\right]$$

if t is even and

$$\lambda_{N-1} \leq 1 - (1 - \mathcal{B}[t])^{\frac{1}{t}}$$

if t is odd.

These results indicate the deep connection between random walks on graphs and eigenvalue estimate of the Laplace operator. In the rest of this paper we will use these insights to derive lower bounds for λ_1 and upper bounds for λ_{N-1} in terms of the Ollivier–Ricci curvature of a graph.

2.3. Ollivier–Ricci curvature from a probabilistic view.
We consider the usual graph metric $d : V \times V \to \mathbb{R}^+$ on the set of vertices V, i.e., for two distinct points $x, y \in V$, $d(x, y)$ is the number of edges in the shortest path connecting x and y. Then, including the family of probability measures $m := \{m_x(\cdot)\}$, we have a structure (V, d, m), on which the definition of Ricci curvature proposed by Ollivier [21] can be stated.

Definition 4 (Ollivier [21]). For any two distinct points $x, y \in V$, the (Ollivier-) Ricci curvature of (V, d, m) along (xy) is defined as

$$(2.9) \qquad \kappa(x, y) := 1 - \frac{W_1(m_x, m_y)}{d(x, y)}.$$

Here, $W_1(m_x, m_y)$ is the transportation distance between the two probability measures m_x and m_y, in a formula,

$$(2.10) \qquad W_1(m_x, m_y) = \inf_{\xi^{x,y} \in \prod(m_x, m_y)} \sum_{(x', y') \in V \times V} d(x', y') \xi^{x,y}(x', y'),$$

where $\prod(m_x, m_y)$ is the set of probability measures $\xi^{x,y}$ that satisfy

$$(2.11) \qquad \sum_{y' \in V} \xi^{x,y}(x', y') = m_x(x'), \quad \sum_{x' \in V} \xi^{x,y}(x', y') = m_y(y').$$

Conditions (2.11) simply ensure that we start with the measure m_x and end up with m_y. Intuitively, $W_1(m_x, m_y)$ is the minimal cost to transport the mass of m_x to that of m_y with the distance as the cost function. We also call such a $\xi^{x,y}$ a transfer plan between m_x and m_y, or a coupling of two random walks governed by m_x and m_y, respectively. Those $\xi^{x,y}$ ($\xi^{x,y}$ might not be unique) which attain the infimum value in (2.10), are called optimal couplings. The optimal coupling exists in a very general setting. For locally finite graphs the existence follows from a simple and interesting argument in Remark 14.2 in [15].

By the Kantorovich duality formula for transportation distances $W_1(m_x, m_y)$ is also given in the form,

$$(2.12) \qquad W_1(m_x, m_y) = \sup_{f : \mathrm{Lip}(f) \leq 1} \left[\sum_{x' \in V} f(x') m_x(x') - \sum_{y' \in V} f(y') m_y(y') \right],$$

where $\mathrm{Lip}(f) := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}$. For more details about this concept, we refer to Villani [25, 26], and Evans [10].

For the rest of this paper, let k be a lower bound for the Ollivier–Ricci curvature, i.e.,

$$(2.13) \qquad \kappa(x, y) \geq k, \ \forall x \sim y.$$

The Ricci curvature satisfies the following properties (see [21]):

Lemma 2.   (i) *If $\kappa(x,y) \geq k$ for all neighbors $x \sim y$, then this is true for all pairs of vertices $x, y \in V$.*

(ii) *We have $-2 \leq \kappa(x,y) \leq 1$.*

We will derive more precise lower and upper bounds for κ on a locally finite graph with loops in Theorems 2 and 3, respectively (see also Lin–Yau [18] and Jost–Liu [13] for related results).

Combining (2.9) and (2.13), we obtain

$$(2.14) \qquad W_1(m_x, m_y) \leq (1-k)d(x,y) = 1 - k, \quad \forall x \sim y,$$

which is essentially equivalent to the well known path coupling criterion on the state space of Markov chains used to study the mixing time of them (see [4, 15, 23]). We will utilize this idea to interpret the lower bound of the Ollivier–Ricci curvature as a control on the expectation value of the distance between two coupled random walks.

We reformulate Bubley–Dyer's theorem (see [4, 15, 23]) in our language.

Theorem 1 (Bubley–Dyer). *On (V, d, m), if for each pair of neighbors $x, y \in V$, we have the contraction*

$$W_1(m_x, m_y) \leq (1-k)d(x,y) = 1 - k,$$

then for any two probability measures μ and ν on V, we have

$$W_1(\mu P, \nu P) \leq (1-k)W_1(\mu, \nu).$$

With this at hand, it is easy to see that if for any pair of neighbors x, y, $\kappa(x,y) \geq k$, then for any time t and any two $\bar{x}, \bar{y}$, which are not necessarily neighbors, the following is true,

$$(2.15) \qquad W_1(\delta_{\bar{x}} P^t, \delta_{\bar{y}} P^t) \leq (1-k)^t d(\bar{x}, \bar{y}).$$

We consider two coupled discrete time random walks $(\bar{X}_t, \bar{Y}_t)$, whose distributions are $\delta_{\bar{x}} P^t$, $\delta_{\bar{y}} P^t$ respectively. They are coupled in a way that the probability

$$p(\bar{X}_t = \bar{x}', \bar{Y}_t = \bar{y}') = \xi_t^{\bar{x}, \bar{y}}(\bar{x}', \bar{y}'),$$

where $\xi_t^{\bar{x}, \bar{y}}(\cdot, \cdot)$ is the optimal coupling of $\delta_{\bar{x}} P^t$ and $\delta_{\bar{y}} P^t$. In this language, we can interpret the term $W_1(\delta_{\bar{x}} P^t, \delta_{\bar{y}} P^t)$ as the expectation value of the distance $\mathbf{E}^{\bar{x}, \bar{y}} d(\bar{X}_t, \bar{Y}_t)$ between the coupled random walks $\bar{X}_t$ and $\bar{Y}_t$.

Corollary 2. *On (V, d, m), if $\kappa(x,y) \geq k$, $\forall x \sim y$, then we have for any two $\bar{x}, \bar{y} \in V$,*

$$(2.16) \qquad \mathbf{E}^{\bar{x}, \bar{y}} d(\bar{X}_t, \bar{Y}_t) = W_1(\delta_{\bar{x}} P^t, \delta_{\bar{y}} P^t) \leq (1-k)^t d(\bar{x}, \bar{y}).$$

3. Estimates for Ollivier–Ricci curvature on locally finite graphs with loops

In [13] Jost–Liu obtained a sharp estimate for Ollivier–Ricci curvature on locally finite graphs without loops. As mentioned above, neighborhood graphs are in general weighted graphs with loops. Therefore, for our purposes, we need to understand

the curvature of graphs with loops. In this section, we generalize the estimates in Jost–Liu [13] for locally finite graphs $\tilde{G} = (\tilde{V}, \tilde{E})$ that may have loops. This is done by considering a novel optimal transportation plan.

We first fix some notations. For any two real numbers a, b,

$$a_+ := \max\{a, 0\}, \quad a \wedge b := \min\{a, b\}, \text{ and } a \vee b := \max\{a, b\}.$$

We denote $\tilde{N}_x := \{z \in \tilde{V} | z \sim x\}$ as the neighborhood of x and $N_x := \tilde{N}_x \cup \{x\}$. Then $N_x = \tilde{N}_x$ if x has a loop. For every pair of neighbors x, y, we divide N_x, N_y into disjoint parts as follows.

$$(3.1) \qquad N_x = \{x\} \cup \{y\} \cup N_x^1 \cup N_{xy}, \quad N_y = \{y\} \cup \{x\} \cup N_y^1 \cup N_{xy},$$
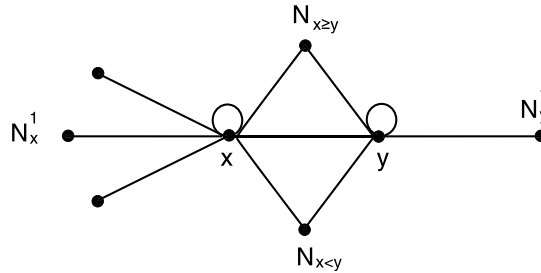
where

$$N_{xy} = N_{x \geq y} \cup N_{x < y}$$

and

$$N_x^1 := \{z | z \sim x, z \nsim y, z \neq y\},$$

$$N_{x \geq y} := \left\{ z | z \sim x, z \sim y, z \neq x, z \neq y, \frac{w_{xz}}{d_x} \geq \frac{w_{zy}}{d_y} \right\},$$

$$N_{x < y} := \left\{ z | z \sim x, z \sim y, z \neq x, z \neq y, \frac{w_{xz}}{d_x} < \frac{w_{zy}}{d_y} \right\}.$$

In the next figure, we illustrate this partition of the vertex set.

Theorem 2. *On $\tilde{G} = (\tilde{V}, \tilde{E})$, we have for any pair of neighbors $x, y \in \tilde{V}$,*

$$\kappa(x, y) \geq k(x, y) := -\left(1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} \vee \frac{w_{x_1 y}}{d_y} \right)_+$$

$$-\left(1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y} \right)_+$$

$$+ \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y} + \frac{w_{xx}}{d_x} + \frac{w_{yy}}{d_y}.$$

Moreover, this inequality is sharp.

Remark 1. On an unweighted graph, the lower bound for the Ricci curvature $k(x, y)$ for $x \sim y$ becomes

$$k(x, y) = -\left(1 - \frac{1}{d_x} - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_x \wedge d_y}\right)_+ - \left(1 - \frac{1}{d_x} - \frac{1}{d_y} - \frac{\sharp(x, y)}{d_x \vee d_y}\right)_+$$
$$+ \frac{\sharp(x, y)}{d_x \vee d_y} + \frac{c(x)}{d_x} + \frac{c(y)}{d_y},$$

where $\sharp(x, y) := \sum_{x_1 \in N_{xy}} 1$ is the number of triangles containing x, y, $c(x) = 0$ or 1 is the number of loops at x.

Proof. Since the total mass of m_x is equal to one, we obtain from (3.1) the following identity for neighboring vertices x and y:

$$(3.2) \qquad 1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} = \frac{w_{xx}}{d_x} + \sum_{x_1 \in N_x^1} \frac{w_{x_1 x}}{d_x}$$

A similar identity holds for y.

We denote

$$A_{x,y} := 1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} \vee \frac{w_{x_1 y}}{d_y},$$

$$B_{x,y} := 1 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y}.$$

Obviously, $A_{x,y} \leq B_{x,y}$. We firstly try to understand these two quantities.

If $A_{x,y} \geq 0$, we have

$$(3.3) \qquad 1 - \frac{w_{xy}}{d_y} - \sum_{x_1 \in N_{xy}} \frac{w_{x_1 y}}{d_y} \geq \frac{w_{xy}}{d_x} + \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1 y}}{d_y}\right),$$

i.e., using (3.2) we observe that the mass of m_y at y and N_y^1 is no smaller than that of m_x at y and the excess mass at $N_{x \geq y}$. Rewriting (3.3) in the form

$$\frac{w_{xy}}{d_y} + \sum_{x_1 \in N_{xy}} \frac{w_{x_1 y}}{d_y} \leq 1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1 y}}{d_y}\right),$$

and subtracting the term $\sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1 y}}{d_y}$ on both sides we obtain

$$(3.4) \qquad \frac{w_{xy}}{d_y} + \sum_{x_1 \in N_{x < y}} \left(\frac{w_{x_1 y}}{d_y} - \frac{w_{xx_1}}{d_x}\right) \leq 1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x},$$

i.e., the mass of m_x at x and N_x^1 is larger than that of m_y at x and the excess mass at $N_{x < y}$.

If $B_{x,y} \geq 0$, we have

$$(3.5) \qquad 1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} + \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1 y}}{d_y}\right) \geq \frac{w_{xy}}{d_y},$$

i.e., the mass of m_x at x and N_x^1 and the excess mass at $N_{x \geq y}$ is no smaller than that of m_y at x.

In Jost–Liu [13] it is explicitly described how much mass has to be moved from a vertex in $\tilde{N}_x$ to which point in $\tilde{N}_y$, i.e., the exact value of $\xi^{x,y}(x',y')$, for any $x' \in \tilde{N}_x$, $y' \in \tilde{N}_y$. However, in the case with loops it would be too complicated if we try to do the same thing. Instead, we adopt here a dynamic strategy. That is, we think of a discrete-time flow of mass. After one unit time, the mass flows forward for distance 1 or stays there. We only need to determine the direction of the flow according to different cases.

As in Jost–Liu [13], we divide the discussion into three cases.

- $0 \leq A_{x,y} \leq B_{x,y}$. In this case we use the following transport plan: suppose the initial time is $t = 0$.

$t = 1$: Move all the mass at N_x^1 to x and the excess mass at $N_{x \geq y}$ to y. We denote the distribution of the mass after the first time step by m^1. We have

$$W_1(m_x, m^1) \leq \left(1 - \frac{w_{xx}}{d_x} - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x}\right) \times 1 + \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1y}}{d_y}\right) \times 1$$

$t = 2$: Move one part of the excess mass at x now to fill the gap at $N_{x<y}$ and the other part to y. By (3.4) the mass at x after $t = 1$ is enough to do so. The distribution of the mass is now denoted by m^2. We have

$$W_1(m^1, m^2) \leq \sum_{x_1 \in N_{x<y}} \left(\frac{w_{x_1y}}{d_y} - \frac{w_{xx_1}}{d_x}\right) \times 1 + \left[\left(1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x}\right)\right.$$
$$\left. - \sum_{x_1 \in N_{x<y}} \left(\frac{w_{x_1y}}{d_y} - \frac{w_{xx_1}}{d_x}\right) - \frac{w_{xy}}{d_y}\right] \times 1$$

$t = 3$: Move the excess mass at y now to N_y^1. We denote the mass after the third time step by $m^3 = m_y$. We have

$$W_1(m^2, m_y) \leq \left[\left(1 - \frac{w_{xy}}{d_x} - \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x}\right) - \sum_{x_1 \in N_{x<y}} \left(\frac{w_{x_1y}}{d_y} - \frac{w_{xx_1}}{d_x}\right) - \frac{w_{xy}}{d_y}\right.$$
$$\left. + \frac{w_{xy}}{d_x} + \sum_{x_1 \in N_{x \geq y}} \left(\frac{w_{xx_1}}{d_x} - \frac{w_{x_1y}}{d_y}\right) - \frac{w_{yy}}{d_y}\right] \times 1$$

By triangle inequality and (2.10), we obtain

$$W_1(m_x, m_y) \leq W_1(m_x, m^1) + W_1(m^1, m^2) + W_1(m^2, m_y)$$
$$= 3 - 2\frac{w_{xy}}{d_x} - 2\frac{w_{xy}}{d_y} - 2 \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1y}}{d_y}$$
$$- \sum_{x_1 \in N_{xy}} \frac{w_{xx_1}}{d_x} \vee \frac{w_{x_1y}}{d_y} - \frac{w_{xx}}{d_x} - \frac{w_{yy}}{d_y}.$$

Moreover, if the following function can be extended as a function on the graph such that $\mathrm{Lip}(f) \leq 1$, (i.e., if there are no paths of length 1 between N_x^1 and

$N_{x<y}$, nor paths of length 1 between N_y^1 and $N_{x\geq y}$, nor paths of length 1 or 2 between N_x^1 and N_y^1,)

$$f(z) = \begin{cases} 0, & \text{if } z \in N_y^1; \\ 1, & \text{if } z \in \{y\} \cup N_{x<y}; \\ 2, & \text{if } z \in \{x\} \cup N_{x\geq y}; \\ 3, & \text{if } z \in N_x^1, \end{cases}$$

then by Kantorovich duality (2.12), we can show that the inequality above is actually an equality. Recalling the definition of $\kappa(x,y)$, we have proved the theorem in this case.

- $A_{x,y} < 0 \leq B_{x,y}$. We use the following transfer plan:
$t = 1$: We divide the excess mass of m_x at $N_{x\geq y}$ into two parts. One part together with the mass of m_x at y is enough to fill gaps at y and N_y^1. Since (3.3) does not hold in this case, this is possible. We move this part of mass to y and the other part to x. We also move all the mass of m_x at N_x^1 to x.
$t = 2$: We move the excess mass at x now to $N_{x<y}$ and the excess mass at y to N_y^1.

Applying this transfer plan, we can prove (we omit the calculation here)

$$W_1(m_x, m_y) \leq 2 - \frac{w_{xy}}{d_x} - \frac{w_{xy}}{d_y} - 2\sum_{x_1 \in N_{xy}} \left(\frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1 y}}{d_y} \right) - \frac{w_{xx}}{d_x} - \frac{w_{yy}}{d_y}.$$

Moreover, if the following function can be extended as a function on the graph such that $\mathrm{Lip}(f) \leq 1$, (i.e., if there are no paths of length 1 between $N_x^1 \cup N_{x\geq y}$ and $N_y^1 \cup N_{x<y}$,)

$$f(z) = \begin{cases} 0, & \text{if } z \in N_y^1 \cup N_{x<y}; \\ 1, & \text{if } z = x \text{ or } z = y; \\ 2, & \text{if } z \in N_x^1 \cup N_{x\geq y}, \end{cases}$$

then by Kantorovich duality (2.12), we can check that the inequality above is actually an equality.

- $A_{x,y} \leq B_{x,y} < 0$. We use the following transport plan:
$t = 1$: Move the mass of m_x at N_x^1 and $N_{x\geq y}$ to x. Since now (3.5) does not hold, we need to move one part of the mass $m_y(y)$ to x and the other part to N_y^1 and $N_{x<y}$.
Applying this transfer plan, we can calculate

$$W_1(m_x, m_y) \leq 1 - \sum_{x_1 \in N_{xy}} \left(\frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1 y}}{d_y} \right) - \frac{w_{xx}}{d_x} - \frac{w_{yy}}{d_y}.$$

Since the following function can be extended as a function on the graph such that $\mathrm{Lip}(f) \leq 1$,

$$f(z) = \begin{cases} 0, & \text{if } z \in \{x\} \cup N_{x<y} \cup N_y^1; \\ 1, & \text{if } z \in \{y\} \cup N_{x\geq y} \cup N_x^1, \end{cases}$$

we can check the inequality above is in fact an equality by Kantorovich duality. That is, in this case for any $x \sim y$,

$$\kappa(x,y) = \sum_{x_1 \in N_{xy}} \left(\frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1 y}}{d_y} \right) + \frac{w_{xx}}{d_x} + \frac{w_{yy}}{d_y}.$$

$\square$

We also have a generalization of the upper bound in Jost–Liu [13] on $\tilde{G}$.

Theorem 3. *On $\tilde{G} = (\tilde{V}, \tilde{E})$, we have for every pair of neighbors x, y,*

$$\kappa(x,y) \leq \sum_{x_1 \in \{x\} \cup \{y\} \cup N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y}.$$

Proof. $I := \sum_{x_1 \in \{x\} \cup \{y\} \cup N_{xy}} \frac{w_{x_1 x}}{d_x} \wedge \frac{w_{x_1 y}}{d_y}$ is exactly the mass of m_x which we need not move. The other mass need to be moved for at least distance 1. So we have $W_1(m_x, m_y) \geq 1 - I$, which implies $\kappa(x,y) \leq I$, for $x \sim y$. $\square$

Example 1. We consider a lazy random walk on an unweighted complete graph $\mathcal{K}_N$ with N vertices governed by $m_x(y) = 1/N, \forall x, y$. Or equivalently, we consider the graph $\mathcal{K}_N^{\text{lazy}}$. Using Theorem 2 and Theorem 3, we get for any x, y

$$1 = \frac{N-2}{N} + \frac{1}{N} + \frac{1}{N} \leq \kappa(x,y) \leq \frac{1}{N} \cdot N = 1.$$

That is, in this case, both the lower and the upper bound are sharp.

An immediate consequence of Theorem 3 is the following important observation.

Corollary 3. *If there exist two vertices $x \sim y$ in G such that $\sharp(x,y) = c(x) = c(y) = 0$ then $\kappa(x,y) \leq 0$ and hence $k \leq 0$.*

This corollary shows that positive Ricci curvature is a quite strong requirement. For instance, in a loopless graph, already the existence of a single edge that is not contained in a triangle prevents the graph from having a positive Ricci curvature lower bound. We will show in the following that the neighborhood graph technique can be used to reduce the influence of such edges. This observation is particularly important in the next section when we study eigenvalue estimates in terms of the Ricci curvature.

Neighborhood graphs are nothing but coarse representations of the original graph. More precisely, the neighborhood graphs $G[t]$ encode the larger scale structure of the original graph G, where larger values of t stand for larger scales, in the sense that an edge between two nodes in $G[t]$ is equivalent to the existence of a path of length t in the original graph G between these two nodes. In order to see how neighborhood graphs can reduce the influence of single edges, we state the following simple observations that follow immediately from the definition of the neighborhood graphs.

Observation 2.

- *Triangles and loops are preserved when we go to higher order neighborhood graphs, i.e., if (xyz) form a triangle in $G[s]$ (x has a loop in $G[s]$) then they from a triangle in $G[t]$ (x has a loop in $G[t]$) for all $t > s$.*
- *If t is even, every vertex has a loop in $G[t]$.*

- *If t is odd, the edge set of G is a subset of the edge set of $G[t]$, i.e., $E \subseteq E[t]$.*
- *If in G a vertex x is not contained in a triangle but contained in a cycle of length $3t$ then x is contained in a triangle in $G[t]$.*
- *If in G a vertex x is contained in a cycle of odd length $2l+1$, $l \geq 1$, then x is contained in a triangle (in fact in a complete graph $\mathcal{K}_{2l+1}$) in $G[t]$ if $t \geq 2l-1$.*
- *If in G a vertex x is not contained in a triangle but $x \sim y$ where y is contained in a triangle, then x is also contained in a triangle in $G[t]$ for all $t \geq 2$.*

These observations show, unless G is bipartite, that the number of triangles and loops will monotonically increase when we go from G to $G[t]$. Hence even though the Ricci curvature of the original graph is negative, Corollary 3 does not exclude that the Ricci curvature of the neighborhood graph $G[t]$ is positive. Indeed, we will show in Theorem 5 that for all graphs that are not bipartite there exists a $t' \in \mathbb{N}$ such that the Ricci curvature of the neighborhood graph $G[t]$ satisfies $k[t] := \min_{x,y} \kappa[t](x,y) > 0$ for all $t > t'$.

4. Estimates of the spectrum in terms of Ollivier–Ricci curvature

In this section, we obtain nontrivial estimates for the extremal eigenvalues of the normalized Laplace operator in terms of the Olliver–Ricci curvature of the neighborhood graphs. In particular, our new estimates improve the eigenvalue estimates obtained by Olliver in [21].

In Proposition 30 of [21], Ollivier proved a spectral radius estimate which works on a general metric space with random walks. In particular, on finite graphs, it can be stated as follows.

Theorem 4 (Ollivier). *On (V, d, m), if $\kappa(x, y) \geq k$, $\forall x \sim y$, then the eigenvalues of the normalized graph Laplace operator Δ satisfy*

$$k \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 2 - k.$$

The lower bound for λ_1 is a discrete analogue of the estimate for the smallest nonzero eigenvalue of the Laplace–Beltrami operator on a Riemannian manifold by Lichnerowicz [16]. As pointed out by Ollivier [21], this result is also related to the coupling method for estimates of the first eigenvalue in the Riemannian setting developed by Chen–Wang [8] (which leads to a refinement of the eigenvalue estimate of Li–Yau [17]); see also the surveys Chen [6,7]. The corresponding result of Corollary 2 in the smooth case, i.e., controlling the expectation distance of two coupled Markov chains in terms of the lower bound of Ricci curvature on a Riemannian manifold, is a key step in Chen–Wang's method.

A direct proof of Theorem 4 can be found in [21]. Here for readers' convenience, we present an analogue of Chen–Wang's method in the discrete setting, which motivated us to combine the Ollivier–Ricci curvature and the neighborhood graph method via random walks. It reflects the deep connection between eigenvalue estimates and random walks or heat equations.

Proof. We consider the transition probability operator $P : l^2(V, \mu) \to l^2(V, \mu)$ defined by $Pf(x) := \sum_y f(y)m_x(y) = \sum_y f(y)\delta_x P(y)$. Then we have

$P^t f(x) = \sum_y f(y)\delta_x P^t(y)$. We construct a discrete time heat equation,

(4.1)
$$\begin{cases} f(x,0) = f_1(x), \\ f(x,1) - f(x,0) = \Delta f(x,0), \\ f(x,2) - f(x,1) = \Delta f(x,1), \\ \cdots \\ f(x,t+1) - f(x,t) = \Delta f(x,t), \end{cases}$$

where $f_1(x)$ satisfies $\Delta f_1(x) = -\lambda f_1(x) = P f_1(x) - f_1(x)$ for $\lambda \neq 0$. Iteratively, one can find the solution of the above system of equations as

(4.2)
$$f(x,t) = P^t f_1(x) = (1-\lambda)^t f_1(x).$$

We remark here that the solution of the heat equation on a Riemannian manifold with the eigenfunction as the initial value is $f(x,t) = f_1(x)e^{-\lambda t}$, which also involves information about both the eigenvalue λ and the eigenfunction $f_1(x)$.

Then we have for any $\bar{x}, \bar{y} \in V$

$$|1-\lambda|^t |f_1(\bar{x}) - f_1(\bar{y})| = |f(\bar{x},t) - f(\bar{y},t)| = |P^t f_1(\bar{x}) - P^t f_1(\bar{y})|$$

$$\leq \sum_{\bar{x}',\bar{y}'} |f(\bar{x}') - f(\bar{y}')| \xi_t^{\bar{x},\bar{y}}(\bar{x}',\bar{y}')$$

$$\leq \mathrm{Lip}(f_1)\mathbf{E}^{\bar{x},\bar{y}} d(\bar{X}_t, \bar{Y}_t) \leq \mathrm{Lip}(f_1)(1-k)^t d(\bar{x},\bar{y}).$$

Here, $\mathrm{Lip}(f)$ is always finite since the underlying space V is a finite set. In the last inequality, we used Corollary 2. From an analytic point of view, the above calculation can be seen as a gradient estimate for the solution of the heat equation.

Since the eigenfunction f_1 for the eigenvalue λ is orthogonal to the constant function, i.e., $(f_1, \mathbf{1})_\mu = 0$, we can always find $x_0, y_0 \in V$ such that $|f_1(x_0) - f_1(y_0)| > 0$. It follows that

$$0 < |f_1(x_0) - f_1(y_0)| \leq \left(\frac{1-k}{|1-\lambda|} \right)^t \mathrm{Lip}(f_1) d(x_0, y_0) \text{ for all } t.$$

To prevent a contradiction when $t \to \infty$, we need $|1-\lambda| \leq 1-k$, which completes the proof. □

As an immediate consequence of Theorem 4 and Theorem 2 we obtain an estimate for the largest eigenvalue in terms of the number of triangles and loops in the graph.

Corollary 4. *On $G = (V,E)$, the largest eigenvalue satisfies*

$$\lambda_{N-1} \leq 2 - \min_{x \sim y} k(x,y),$$

where $k(x,y)$ is defined in Theorem 2.

By considering the graph $\mathcal{K}_N^{\mathrm{lazy}}$ in Example 1 it is easy to see that Ollivier's estimates in Theorem 4 can be sharp for certain graphs. However, from Corollary 3 we know that a positive lower curvature bound is a strong restriction on a graph. In the open Problem G in [22], Olliver asks for the possibility to relax this strong assumption. We will show in the following how to obtain nontrivial estimates for all graphs that are not bipartite by using the neighborhood graph technique. This gives an answer to Ollivier's problem in the finite graph setting.

Before we show how one can improve Olliver's result by using the neighborhood graph technique, we show how this technique can be used to obtain upper bounds for λ_{N-1} from lower bounds for λ_1. We do this by carefully comparing the Olliver–Ricci curvature on a graph G and its neighborhood graphs $G[t]$.

If we interpret the graph $G = (V, E)$ as a structure $(V, d, m = \{\delta_x P\})$, then by (2.5) its neighborhood graph $G[t] = (V, E[t])$ can be considered as a structure $(V, d[t], \{\delta_x P^t\})$. So the first step should be to estimate the graph metric on G and the graph metric on $G[t]$ by each other.

Lemma 3. *For any $x, y \in V$, we have*

(4.3) $$\frac{1}{t}d(x,y) \leq d[t](x,y),$$

where we use the convention that $d[t](x,y) = \infty$ if $G[t]$ is not connected. By Lemma 1 this happens iff G is bipartite and t is even.

Proof. If $G[t]$ is not connected, then (4.3) is trivially satisfied. Otherwise, we can find a shortest path $x_0 = x, x_1, \ldots, x_l = y$, between x and y in $G[t]$, i.e., $l = d[t](x, y)$. For x_i, x_{i+1}, $i = 0, \ldots, l-1$, by definition of neighborhood graph, we have $d(x_i, x_{i+1}) \leq t$ in G. Equivalently,

$$\frac{1}{t}d(x_i, x_{i+1}) \leq 1 = d[t](x_i, x_{i+1}).$$

Summing over all i, we get

$$\frac{1}{t}\sum_{i=0}^{l-1} d(x_i, x_{i+1}) \leq d[t](x, y).$$

Then the triangle inequality of d on G gives (4.3). $\qquad\square$

Remark 2. In fact, when t is larger than the diameter D of the graph G, we have a better estimate

(4.4) $$\frac{1}{t}d(x,y) < \frac{1}{D}d(x,y) \leq 1 \leq d[t](x,y).$$

Lemma 4. *If $E \subseteq E[t]$, then $d[t](x,y) \leq d(x,y)$.*

Proof. The proof is obvious. $\qquad\square$

The importance of Lemma 4 comes from the observation that when the Ollivier–Ricci curvature of the graph G is positive, $E \subseteq E[t]$ is satisfied for all t and hence Lemma 4 is applicable. This can be seen as follows. Corollary 3 implies that if $k > 0$, then for all $(x, y) \in E$ we have $\sharp(x, y) \neq 0$ or $c(x) \neq 0$ or $c(y) \neq 0$ which in turn implies (see Observation 2) that $(x, y) \in E[t]$ for all t.

Lemma 5. *Let k be a lower bound for κ on G. If $E \subseteq E[t]$, then the curvature $\kappa[t]$ of the neighborhood graph $G[t]$ satisfies*

$$(4.5) \qquad \kappa[t](x,y) \geq 1 - t(1-k)^t, \quad \forall x,y \in V.$$

Proof. Let $W_1^{d[t]}$, W_1^d indicate the different cost functions used in these two quantities. By Lemma 4, Corollary 2 and Lemma 3, we get

$$W_1^{d[t]}(\delta_x P^t, \delta_y P^t) \leq W_1^d(\delta_x P^t, \delta_y P^t) \leq (1-k)^t d(x,y) \leq t(1-k)^t d[t](x,y),$$

where we used in the first inequality that the transportation distance (2.10) is linear in the graph distance $d(\cdot,\cdot)$. Recalling the definition of the curvature, we have proved (4.5). $\qquad\qquad\square$

Remark 3. Now we have reached the point where we can give a short geometric proof of the upper bound of the largest eigenvalue in Theorem 4. First assume that $k > 0$. In this case, $E \subset E[t]$ and thus we can use Lemma 5. From Lemma 5 and $\lambda_1 \geq k$, we know on $G[t]$,

$$\lambda_1[t] \geq 1 - t(1-k)^t.$$

Then by using Corollary 1 (i), we get for any even number t,

$$\lambda_{N-1} \leq 1 + t^{\frac{1}{t}}(1-k).$$

Letting $t \to +\infty$, we get $\lambda_{N-1} \leq 2 - k$. If we assume that $k \leq 0$, then $\lambda_{N-1} \leq 2 - k$ is trivially satisfied.

Using the neighborhood graph technique, we further obtain the following generalization of Theorem 4:

Theorem 5. *Let $k[t]$ be a lower bound of the Ollivier–Ricci curvature of the neighborhood graph $G[t]$. Then for all $t \geq 1$ the eigenvalues of Δ on G satisfy*

$$(4.6) \qquad 1 - (1-k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 1 + (1-k[t])^{\frac{1}{t}}.$$

Moreover, if G is not bipartite, then there exists a $t' \geq 1$ such that for all $t \geq t'$ the eigenvalues of Δ on G satisfy

$$0 < 1 - (1-k[t])^{\frac{1}{t}} \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 1 + (1-k[t])^{\frac{1}{t}} < 2.$$

Remark 4. Olliver–Ricci curvature is not well defined for two vertices, which belong to two different connected components. However, by Lemma 1, $G[t]$ is disconnected iff G is bipartite and t is even. In this case, we use the convention in (4.6) that $k[t] = -\infty$.

Proof. Combining Theorem 4, and Corollary 1 immediately yields (4.6).

The second part of this Theorem is proved in two steps. In the first step, we will show that if G is not bipartite then there exists a t' such that for all $t \geq t'$ the neighborhood graph $G[t]$ of G satisfies $w_{xy}[t] \neq 0$ for all $x,y \in V$, i.e., $G[t]$ is a complete graph and each vertex has a loop. In the second step, we show that any graph that satisfies $w_{xy} \neq 0$ for all $x,y \in V$ has a positive lower curvature bound, i.e., $k > 0$. This then completes the proof.

Step 1: By the definition of the neighborhood graph it is sufficient to show that for all $t \geq t'$ there exists a path of length t between any pair of vertices. Since G is

not bipartite it follows from the definition of bipartiteness that there exists a path of even and a path of odd length between any pair of vertices in the graph. Given a path of length L between x and y then we can find a path of length $L+2$ between x and y as follows: We go in L steps from x to y and then from y to one of its neighbors and then back to y. This is a path of length $L+2$ between x and y. Since G is finite, it follows that there exists a t' such that for every pair of vertices there exists paths of length t for all $t \geq t'$.

Step 2: Given a graph that satisfies $w_{xy} \neq 0$ for all $x, y \in V$.

Since each vertex in the graph is a neighbor of all other vertices, it is clear that we can move the excess mass of m_x for distance 1 to anywhere. Therefore

$$W_1(m_x, m_y) \leq 1 - \sum_{x_1 \in V} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1 y}}{d_y},$$

which implies

$$\kappa(x, y) \geq \sum_{x_1 \in V} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1 y}}{d_y}.$$

By Theorem 3, it follows that the above inequality is in fact an equality. Hence for all $x, y \in V$, we have

$$\kappa(x, y) = \sum_{x_1 \in V} \frac{w_{xx_1}}{d_x} \wedge \frac{w_{x_1 y}}{d_y} \geq N \frac{\min_{x,y} w_{xy}}{\max_x d_x} \geq \frac{\min_{x,y} w_{xy}}{\max_{x,y} w_{xy}} > 0,$$

since the weights w_{xy} are positive for every pair (x, y) of vertices.

This completes the proof. $\qquad\square$

Example 2. We consider the unweighted cycle $\mathcal{C}_5$ consisting of 5 vertices. The first and largest eigenvalue of Δ on $\mathcal{C}_5$ are $\lambda_1 = 1 - \cos\frac{2\pi}{5} \doteq 0.6910$ and $\lambda_4 = 1 - \cos\frac{4\pi}{5} \doteq 1.8090$. It is easy to check that the optimal lower bound k for the curvature is 0. So in this case Ollivier's estimates in Theorem 4 only yield trivial estimates.

For the neighborhood graphs $\mathcal{C}_5[2], \mathcal{C}_5[3]$, and $\mathcal{C}_5[4]$ we find that the optimal lower bound for the Olliver–Ricci curvature is $k[2] = 1/4, k[3] = 3/8$, and $k[4] = 1/2$, respectively. Hence Theorem 5 yields nontrivial estimates, even if the lower bound for the Ricci curvature of the original graph is zero.

From the proof of Theorem 5, we see that for all graphs G, $k[t]$ eventually becomes positive for sufficiently large t. The next two theorems are concerned with the behavior of $k[t]$ as $t \to \infty$.

Theorem 6. *If G is not bipartite, the limit*

$$\lim_{t \to \infty} \frac{\log(1 - k[t])}{t} := -a$$

exists with $a \in (0, +\infty]$. That means, $k[t]$ behaves like $1 - P(t)e^{-at}$ as $t \to \infty$ where $P(t)$ is a polynomial in t.

Proof. Let t' be as in the the proof of Theorem 5 and let $s, t \geq t'$. This immediately implies that

$$(4.7) \qquad d[t + s](x, y) = d[s](x, y) = d[t](x, y) = 1, \quad \forall x \neq y.$$

Let x, y be the pair of vertices in $G[t+s]$ which attains $\min_{x,y} \kappa[t+s](x, y)$, we have

$$
\begin{aligned}
1 - k[t+s] = W_1^{d[t+s]}(\delta_x P^{t+s}, \delta_y P^{t+s}) &= W_1^{d[s]}(\delta_x P^{t+s}, \delta_y P^{t+s}) \\
&\leq (1 - k[s]) W_1^{d[s]}(\delta_x P^t, \delta_y P^t) = (1 - k[s]) W_1^{d[t]}(\delta_x P^t, \delta_y P^t) \\
&\leq (1 - k[s])(1 - k[t]),
\end{aligned}
$$

where we used several times (4.7), in the first inequality Theorem 1 and in the last inequality (2.14) and (4.7). It follows that $\log(1 - k[t])$ is subadditive, i.e.,

$$
\log(1 - k[t+s]) \leq \log(1 - k[t]) + \log(1 - k[s]) \qquad \forall s, t \geq t'.
$$

We can suppose that $\log(1 - k[t])$ is finite for every t. Otherwise there exists t_0 such that $k[t_0] = 1$, which implies $k[t] = 1$, $\forall t \geq t_0$ and then the limit exists and is equal to $-\infty$. Therefore we can use an extension of Fekete's subadditivity Lemma by Hammersley [12] to conclude that the limit $-a$ exists and $-\infty \leq -a \leq 0$. Furthermore, let $\tilde{t} := t't$, $t = 1, 2, \ldots$, be a subsequence. Since $E[\tilde{t}] = E[t']$, we can use Lemma 5 and obtain

$$
(1 - k[\tilde{t}])^{\frac{1}{\tilde{t}}} \leq t^{\frac{1}{\tilde{t}}} (1 - k[t'])^{\frac{t}{\tilde{t}}} = \left(\frac{\tilde{t}}{t'}\right)^{\frac{1}{\tilde{t}}} (1 - k[t'])^{\frac{1}{t'}}.
$$

Therefore $\lim_{\tilde{t} \to \infty} \log(1 - k[\tilde{t}])^{\frac{1}{\tilde{t}}} \leq \log(1 - k[t'])^{\frac{1}{t'}} < 0$, which implies $a > 0$. $\qquad \square$

Remark 5. If G is unweighted, non-bipartite and has no self-loop, we have $\frac{N}{N-1} \leq \lambda_{N-1} \leq 1 + (1 - k[t])^{\frac{1}{t}}$. Therefore, in this case $a \leq \log(N - 1)$.

Theorem 7. *If G is not bipartite, we have for $t \geq t'$*

$$
k[t] \geq 1 - 2N e^{-t(1 - \overline{\lambda})} \frac{\max_x d_x}{\min_x d_x},
$$

where $\overline{\lambda} = \max_{i \neq 0} |1 - \lambda_i| > 0$ and again t' is as in the proof of Theorem 5.

Proof. Let π $(\pi(x) = \frac{d_x}{\operatorname{vol}(G)})$ be the stationary distribution of the random walk. Let $t \geq t'$ and x, y be the pair of vertices in $G[t]$ which attains $\min_{x,y} \kappa[t](x, y)$, noting (4.7) we have

$$
\begin{aligned}
k[t](x, y) = 1 - W_1^{d[t]}(m_x[t], m_y[t]) &\geq 1 - N \max_{x,y,z} |m_x[t](z) - m_y[t](z)| \\
&\geq 1 - 2N \max_{x,z} |\delta_x P^t(z) - \pi(z)|,
\end{aligned}
$$

where we used in the first inequality that one has to move at most N times the maximal difference between any two m_x and m_y. Chung [9] showed that

$$
\frac{\max_{x,z} |\delta_x P^t(z) - \pi(z)|}{\max_x \pi(x)} \leq \max_{x,z} \frac{|\delta_x P^t(z) - \pi(z)|}{\pi(z)} \leq e^{-t(1 - \overline{\lambda})} \frac{\operatorname{vol}(G)}{\min_x d_x}.
$$

Thus, we have

$$
\max_{x,z} |\delta_x P^t(z) - \pi(z)| \leq e^{-t(1 - \overline{\lambda})} \frac{\max_x d_x}{\min_x d_x},
$$

which completes the proof. $\qquad \square$

5. Estimates for the largest eigenvalue in terms of the number of joint neighbors

In Bauer–Jost [2], it is shown that the next lemma is a simple consequence of Observation 1.

Lemma 6. *Let u be an eigenfunction of Δ for the eigenvalue λ. Then,*

$$(5.1) \qquad 2 - \lambda = \frac{(u, \Delta[2]u)_\mu}{(u, \Delta u)_\mu} = \frac{\sum_{x,y} w_{xy}[2](u(x) - u(y))^2}{\sum_{x,y} w_{xy}(u(x) - u(y))^2}.$$

Lemma 6 can be used to derive further estimates for the largest eigenvalue λ_{N-1} from above and below. We introduce the following notations:

Definition 5. Let $\tilde{N}_x$ be the neighborhood of vertex x as in Section 3. The minimal and the maximal numbers of joint neighbors of any two neighboring vertices is defined as $\tilde{\sharp}_1 := \min_{x \sim y}(\sharp(x,y) + c(x) + c(y))$ and $\tilde{\sharp}_2 := \max_{x \sim y}(\sharp(x,y) + c(x) + c(y))$, respectively. Furthermore, we define $W := \max_{x,y} w_{xy}$ and $w := \min_{x,y;x \sim y} w_{xy}$.

Theorem 8. *We have the following estimates for λ_{N-1}:*

  (i)

$$\lambda_{N-1} \leq 2 - \frac{w^2}{W}\frac{\tilde{\sharp}_1}{\max_x d_x}.$$

  (ii) If $E(G[2]) \subseteq E(G)$ then

$$2 - \frac{W^2}{w}\frac{\tilde{\sharp}_2}{\min_x d_x} \leq \lambda_{N-1}.$$

Proof. We observe that (i) is non-trivial only when $\tilde{\sharp}_1 > 0$, in which case $E(G) \subseteq E(G[2])$ holds (see also the discussion after Lemma 4). On the one hand, we observe that if $E(G) \subseteq E(G[2])$, then for every pair of neighboring vertices $x \sim y$ in G

$$(5.2) \qquad \frac{w_{xy}[2]}{w_{xy}} = \frac{\sum_z \frac{1}{d_z} w_{xz} w_{zy}}{w_{xy}} \geq \frac{w^2}{W}\frac{\tilde{\sharp}_1}{\max_x d_x}.$$

On the other hand, if $E(G[2]) \subseteq E(G)$ then for every pair of neighboring vertices $x \sim y$ in $G(2)$, we have

$$(5.3) \qquad \frac{w_{xy}[2]}{w_{xy}} = \frac{\sum_z \frac{1}{d_z} w_{xz} w_{zy}}{w_{xy}} \leq \frac{W^2}{w}\frac{\tilde{\sharp}_2}{\min_x d_x}.$$

Substituting the inequalities (5.2) and (5.3) into equation (5.1) completes the proof. $\square$

For unweighted regular graphs, Theorem 8 (i) improves the estimate $\lambda_{N-1} \leq 2 - k$. From Theorem 3 it follows that for an unweighted graph

$$\kappa(x,y) \leq \frac{\sharp(x,y)}{d_x \vee d_y} + \frac{c(x)}{d_x} + \frac{c(y)}{d_y}$$

for all pairs of neighboring vertices x, y. In the case of a d-regular graph G this implies that a lower bound k for the Ollivier–Ricci curvature must satisfy

$$k \leq \frac{\tilde{\sharp}_1}{d}.$$

Hence for an unweighted d-regular graph Theorem 8 implies

$$\lambda_{N-1} \le 2 - \frac{\tilde{\natural}_1}{d} \le 2 - k.$$

Acknowledgments

References

[1] P. Auscher, T. Coulhon and A. Grigor'yan, eds, *Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces*, Contemporary Mathematics, **338**, American Mathematical Society, Providence, RI, 2003.

[2] F. Bauer and J. Jost, *Bipartite and neighborhood graphs and the spectrum of the normalized graph Laplacian*, `http://arxiv.org/abs/0910.3118v3`, to appear in Comm. Anal. Geom.

[3] A.-I. Bonciocat and K.-T. Sturm, *Mass transportation and rough curvature bounds for discrete spaces*, J. Funct. Anal. **256**(9) (2009), 2944–2966.

[4] R. Bubley and M.E. Dyer, *Path coupling: a technique for proving rapid mixing in Markov chains*, 38th Annual Symposium on Foundations of Computer Science (FOCS '97), (1997), 223–231.

[5] I. Chavel, *Eigenvalues in Riemannian Geometry*, Academic Press, 1984.

[6] M. Chen, *Trilogy of couplings and general formulas for lower bound of spectral gap*, Probability towards 2000 (New York, 1995), 123–136, Lecture Notes in Statistics, **128**, Springer, New York, 1998.

[7] M. Chen, *Coupling, spectral gap and related topics I*, Chinese Sci. Bull. **42**(16) (1997), 1321–1327.

[8] M. Chen and F. Wang, *Application of coupling method to the first eigenvalue on manifold*, Sci. China Ser. A **37**(1) (1994), 1–14.

[9] F.R.K. Chung, *Spectral graph theory*, CBMS Regional Conference Series in Mathematics, **92**, 1997.

[10] L.C. Evans, *Partial differential equations and Monge–Kantorovich mass transfer*, Current developments in mathematics, 1997 (Cambridge, MA), 65–126, Int. Press, Boston, MA, 1999.

[11] A. Grigor'yan, *Analysis on graphs*, Lecture notes University of Bielefeld, 2009.

[12] J.M. Hammersley, *Generalization of the Fundamental theorem on subadditive functions*, Mathematical Proceedings of the Cambridge Philosophical Society **58** (1962), 235–238.

[13] J. Jost and S. Liu, *Ollivier's Ricci curvature, local clustering and curvature dimension inequalities on graphs*, `http://arxiv.org/pdf/1103.4037v2`.

[14] A. Joulin and Y. Ollivier, *Curvature, concentration and error estimates for Markov chain Monte Carlo*, Ann. Probab. **38**(6) (2010), 2418–2442.

[15] D.A. Levin, Y. Peres and E.L. Wilmer, *Markov chains and mixing times*, With a chapter by James G. Propp and David B. Wilson, American Mathematical Society, Providence, RI, 2009.

[16] A. Lichnerowicz, *Géométrie des groupes de transformations*, Travaux et Recherches Mathématiques, III, Dunod, Paris, 1958.

[17] P. Li and S.T. Yau, *Estimates of eigenvalues of a compact Riemannian manifold*, AMS Proc. Symp. Pure Math. **36** (1980), 205–240.

[18] Y. Lin and S.T. Yau, *Ricci curvature and eigenvalue estimate on locally finite graphs*, Math. Res. Lett. **17**(2) (2010), 343–356.

[19] J. Lott and C. Villani, *Ricci curvature for metric measure spaces via optimal transport*, Ann. Math. **169**(3) (2009), 903–991.

[20] S.-I. Ohta, *On the measure contraction property of metric measure spaces*, Comment. Math. Helv. **82**(4) (2007), 805–828.

[21] Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. **256**(3) (2009), 810–864.

[22] Y. Ollivier, *A survey of Ricci curvature for metric spaces and Markov chains*, in 'Probabilistic approach to geometry' (M. Kotani, M. Hino and T. Kumagai, eds.), Adv. Stud. Pure Math. 57, Math. Soc. Japan, Tokyo, 2010, pp. 343–381.

[23] Y. Peres, *Mixing for Markov chains and spin systems*, Lecture notes, `http://www.stat.berkeley.edu/~peres/ubc.pdf`, 2005.

[24] K.-T. Sturm, *On the geometry of metric measure spaces, I and II*, Acta Math. **196**(1) (2006), 65–131, 133–177.

[25] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, **58**, American Mathematical Society, Providence, RI, 2003.

[26] C. Villani, *Optimal transport, Old and new*, Grundlehren der Mathematischen Wissenschaften, **338**, Springer-Verlag, Berlin, 2009.

[27] D.J. Watts and S.H. Strogatz, *Collective dynamics of "small-world" networks*, Nature **393** (1998), 440–442.

Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany
E-mail address: `Frank.Bauer@mis.mpg.de`

Current Address: Harvard University, Department of Mathematics, One Oxford Street, Cambridge, MA 02138, USA
E-mail address: `fbauer@math.harvard.edu`

Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany; Department of Mathematics and Computer Science, University of Leipzig, 04109 Leipzig, Germany and Santa Fe Institute for the Sciences of Complexity, Santa Fe, NM 87501, USA
E-mail address: `jost@mis.mpg.de`

Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany and Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, People's Republic of China
E-mail address: `shiping@mis.mpg.de`