Ricci Curvature for Graphs

Debdas Paul
Institute for Systems Theory and Automatic Control
University of Stuttgart

July 24, 2015

Abstract

The purpose of this informal report is to explain Ollivier-Ricci curvature for graphs with examples.

1 Ollivier-Ricci curvature for graphs

The concept of Ollivier-Ricci curvature [7] (a coarse version of usual Ricci curvature on manifolds) for graphs can be described in terms of Optimal Transportation problem due to Monge and Kantorovich (Monge-Kantorovich transportation problem) [9]. Before going into the details of transportation problem and subsequent derivation of Ollivier-Ricci curvature from it, we should revise some basic mathematical concepts/terminologies which will be used frequently in this report.

1.1 Algebra, measure, probability measure, metric and graphs

First, we begin with algebra of sets.

Definition 1. *(Algebra) Let Ω denotes a universal set. A collection $\mathcal{A}$ of subsets of Ω is called an algebra or field if:*

1. $\Omega \in \mathcal{A}$
2. *if $A \in \mathcal{A}$ then complement of A or A^c is also $\in \mathcal{A}$*
3. *if $A \in \mathcal{A}$ and if $B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$*

Example 1. 1. $\mathcal{A} = \{\phi, \Omega\}$ *is an algebra*

2. *If Ω is a finite set, the the power set of Ω is an algebra*

Definition 2 (σ-algebra). *A collection $\mathcal{A}$ of subsets of Ω is called a σ-algebra if :*

1. $\Omega \in \mathcal{A}$
2. *if $A \in \mathcal{A}$ then complement of A or A^c is also $\in A$*
3. *if $A_n \in \mathcal{A}$ for each n in a countable collection $(A_n)_{n=1}^{\infty} then \cup_{n=1}^{\infty} A_n \in \mathcal{A}$*

Example 2. 1. *Let $\Omega = \{T, H\}$ in a coin-tossing. Then*

$$\mathcal{F} = \{\{H\}, \{T\}, \{H, T\}, \phi\}$$

is a $\sigma - algebra$ and contains also all possible subsets of Ω

2. *Let $\Omega = \{1, 2, 3\}$. Then*

$$\mathcal{F} = \{\{1\}, \{2, 3\}, \{1, 2, 3\}, \phi\}$$

is a $\sigma - algebra$.

3. *Let $\Omega = \{1, 2, 3\}$. Then*

$$\mathcal{F} = \{\{1\}, \{2\}, \{3\}, \{1, 2, 3\}, \phi\}$$

is NOT a $\sigma - algebra$.

Now,we are going to define general notion of measure. Intuitively one can think of measures are the length on the real line, area in two dimension, volume in three dimensions, when properly defined.

Definition 3. *(Measure) Let Ω be a set. A measure (mainly positive measure) over a σ-algebra ($\mathcal{F}$ defined over Ω)is a function $\mu : \mathcal{F} \to \mathcal{R}_+$ satisfying*

1. *$\mu(A) \geq 0$ for all $A \in \mathcal{F}$*
2. $\mu(\phi) = 0$
3. *if $A_i \in \mathcal{F}$ for all A_i in the collection $(A_i)_{i=1}^{\infty}$ of pairwise disjoint sets, then*

$$\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

The pair $(\Omega, \mathcal{F})$ is called a **measurable space**. Now if $\mu(\Omega) = 1$ then μ is called a *probability measure* and denoted by P. The triplet $(\Omega, \mathcal{F}, P)$ is called a *probability space*, where Ω is a set of outcomes, $\mathcal{F}$ is a set of events, and $P : \mathcal{F} \to [0, 1]$ is a function assigning probabilities to events. $\mathcal{F}$ is taken to be a $\sigma-$ algebra .

Definition 4. *(Metric space) A metric space is a non-empty set X together with a function d (called a metric or "distance function") which assigns a real number $d(x, y)$ to every pair x, y belongs X satisfying the properties (or axioms):*

1. $d(x,y) \geq 0$ *and* $d(x,y) = 0$ *iff* $x = y$*,*

2. $d(x,y) = d(y,x)$

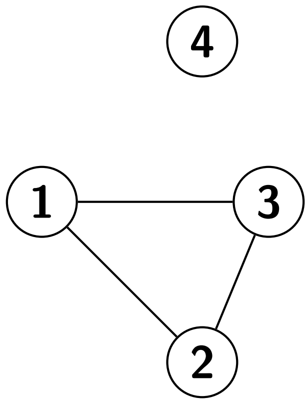
3. $d(x,y) + d(y,z) \geq d(x,y)$*, for any* $z \in X$*.*

Now let us consider a simple (without self-loops) undirected weighted finite graph g over a set of vertices X, then we can think of g as a function $g : X \times X \to [0,\infty)$, i.e, g takes two vertices and maps to a non-negative real number such that:

1. $g(x,x) = 0$ (As the graph is simple)
2. $g(x,y) = g(y,x)$ (As the graph is undirected)
3. the set $\{y \in X | g(x,y) > 0\}$ is finite for every $x \in X$ (local finiteness)

Now, two vertices $x, y \in X$ are connected by an edge of weight $g(x,y)$ whenever $g(x,y) > 0$. In this case, we can write $x \sim y$ and say x is adjacent to y or x and y are neighbors.

Example 3. *Let us consider the following graph (Disconnected): The set*



of vertices is $X = 1, 2, 3, 4$ *and*

$$g(1,2) = g(2,1) = g(1,3) = g(3,1) = g(2,3) = g(3,2) = 1$$

and $g(i,4) = g(4,i) = 0$ *for* $i = 1,2,3$ *Now the degree of a vertex* $a \in X$ *is defined as*

$$w_a = \sum_{x \in X} g(x,a)$$

In case $g(x,y)$ *takes* 0 *or* 1 *for* $x, y \in X$*, the degree* w_a *is the number of neighbors* a*. In the above graph* $w_1 = w_2 = w_3 = 2$*,* $w_4 = 0$*.*

Now, for a connected graph (simple and undirected and finite) g, if $d(x,y)$ is the minimal n such that vertices $x, y \in X$ can be connected by a path length of n then d is a metric i.e, it satisfies all three properties of a metric described above. Below is the proof.

***proof* 1.** *Since g is connected the is a path connecting every two vertices of the graph. This $d(x,y) < \infty$ for all $x, y \in X$. Now, the path from x to x has length 0, therefore $d(x,x) = 0$. Clearly, if $x \neq y$, the $d(x,y) > 0$ and thus $d(x,y) \neq 0$.*
Let $(x_0, \ldots, x_n)$ be the shortest path from $x := x_0$ to $y := y_0$ by remembering $i \mapsto n - i, i = 0, \ldots, n$, we get $(x_n, \ldots, x_0)$ from $ytox$. Since the length of the path did not change we have $d(x,y) = d(y,x)$.
Now, for the last axiom, let us consider another vertex z which not in the minimal path from x to y, then definitely the path becomes longer. Hence $d(x,y) \leq d(x,z) + d(z,y)$

Now, we are ready to proceed towards defining the Kantorovich formalism of Monge's transportation problem (Monge-Kantorovich transportation problem).

1.2 The Monge-Kantorovich transportation problem

In simple words, the problem is to *transport a pile of sand into a hole.* Obviously, the pile and hole must have the same volume (or **measures!**). Now consider the following notations:

1. measure μ_{sand} := amount of sand defined on measurable space $(\Omega_{sand}, \mathcal{F}_{sand})$
2. measure μ_{hole} := size of hole defined on measurable space $(\Omega_{hole}, \mathcal{F}_{hole})$
3. Let A_{sand} and A_{hole} are measurable subsets of Ω_{sand} and Ω_{hole} respectively, then we can interpret the following:
    - $\mu_{sand}(A_{sand})$ gives a measure **how much sand is located inside** A_{sand}
    - $\mu_{hole}(A_{hole})$ gives a measure **how much sand can be piled in** A_{hole}

Now, if we normalize the measure (for both the amount of sand and volume of hole) to 1, i.e, $\mu_{sand}(\Omega_{sand})$ and $\mu_{hole}(\Omega_{hole})$ to 1, then μ_{sand} and μ_{hole} are probability measures!. The sand moving from a **position** $A_{sand} \subseteq \Omega_{sand}$ to a **position** $A_{hole} \subseteq \Omega_{hole}$ has to be transported over a **distance** d. So, now we need a **Transference plan** (T) which indicates **how the transport of the mass is managed (amount − measure!)**. Therefore formally we can think of a **measure** $T(A_{sand} \times A_{hole})$ which tells us **how much sand of position** A_{sand} **is filled in at position** A_{hole}. Moreover as we normalized the amount of sand and the volume of the hole to 1, the transference plan is nothing but a probability measure!. Now, there might be several transference plans but we need the optimal one which minimize the cost of transporting a unit mass from a position in the pile to a position in the hole. Let us formally define a transference plan.

Definition 5. *(Transference plan) The* ***measure*** $T \in P(\Omega_{sand} \times \Omega_{hole})$ *(as it is a probability measure) is a transference plan if it satisfies for* $A_{sand} \in \Omega_{sand}$

$$T(A_{sand} \times \Omega_{hole}) = \int_{A_{sand} \times \Omega_{hole}} dT(x, y) = \mu_{sand}(A_{sand})$$

and for $A_{hole} \in \Omega_{hole}$

$$T(\Omega_{sand} \times A_{hole}) = \int_{\Omega_{sand} \times A_{hole}} dT(x, y) = \mu_{hole}(A_{hole})$$

Now, we can denote the set of all transference plans in the following way:

$$\begin{aligned} T_{all}(\mu_{sand}, \mu_{hole}) =& \{T \in P(\Omega_{sand} \times \Omega_{hole}) | T(A_{sand} \times \Omega_{hole}) = \mu_{sand}(A_{sand}), \\ & T(\Omega_{sand} \times A_{hole}) = \mu_{hole}(A_{hole}), A_{sand} \in \mathcal{F}_{sand}, A_{hole} \in \mathcal{F}_{hole}\} \end{aligned}$$

Now, a measure of **how good a transference plan** works is the **transportation distance** which can be defined the following way:

Definition 6. *(Transportation distance : L. V. Kantorovich) For two* ***probability measures*** α *and* β *on a* ***metric space*** (X, d) *the transportation distance between* α *and* β *is defined as*

$$W_1(\alpha, \beta) = \inf_{T \in T_{all}(\alpha, \beta)} \int_{X \times X} d(x, y) dT(x, y)$$

Now, the question is how we can reformulate the above concept for graphs!. We have already seen graph as a metric space. What left is to associate probability measure to graph (actually to each vertices of graph). As a first step, for each vertex $a \in X$ (set of vertices), we attach a non-negative function is the following way:

$$m_a : X \to [0, 1], x \mapsto \frac{g(x, a)}{w_a}$$

Now this is a function but we need a measure (esp. probability measure) ! let us symbolize our measure as $\hat{m}_a$ and if we can prove somehow $\hat{m}_a(X) = 1$, we are done! To prove this we need to somehow relate the non-negative function m_a and our measure $\hat{m}_a$. Fortunately, there is a theorem to do that!.

Theorem 1. *Let* X *be an at most countable set. The mapping*

$$\{m : X \to [0, \infty]\} \to M(X), m \mapsto \hat{m}$$

with

$$\hat{m}_(A) = \sum_{x \in A} m(x), A \subseteq X$$

is a bijection Where, $M(X)$ *is the set of all measures on* $\{X, 2^X\}$

Basically, the theorem tell us for every measure (say, ν) on $\{X, 2^X\}$ there is $m : X \to [0, \infty]$ such that $\nu = \hat{m} = \sum_{x \in A} m(x), A \subseteq X$. Now with this theorem back to our graph problem. So every m_a can induce a measure $\hat{m}_a$ such that:

$$\hat{m}_a(A) = \sum_{x \in A} m_a(x)$$

for every subset $A \subseteq X$ of vertices. For $A = X$, we have

$$\hat{m}_a(X) = \sum_{x \in X} m_a(x) = \sum_{x \in X} \frac{g(x,a)}{w_a} = \frac{1}{w_a} \sum_{x \in X} g(x,a) = \frac{w_a}{w_a} = 1 (\text{Probability measure!})$$

Now, we are in a position to define transportation distance for graphs. Before the actual definition let us consider some notations:

1. For a vertex $a \in X$ in a graph g, we have
$$m_a(x) = \frac{g(x,a)}{w_a} = \frac{g(a,x)}{w_a}, x \in X$$
2. The associated measure with vertices a and b are $\hat{m}_a$ and $\hat{m}_b$
3. Let the set of all transference plans (for graphs) between $a, b \in X$ be P_{ab}, then
$$P_{ab} = \Bigg\{ \zeta : X \times X \to [0, \infty) \mid \sum_{x \in X} \zeta(x,z) = m_b(z), \sum_{y \in X} \zeta(z,y) = m_a(z) \forall z \in X,$$
$$\sum_{x,y \in X} \zeta(x,y) = 1 \Bigg\}$$

Now consider the following theorems regarding the transportation distance for graphs

Theorem 2. *Let X be an at most countable set and g a graph over X. For all $a, b \in X$ the transportation distance $W_1(\hat{m}_a, \hat{m}_b)$ satisfies*

$$W_1(\hat{m}_a, \hat{m}_b) = \inf_{\zeta \in P_{ab}} \sum_{x, x \sim a} \sum_{y, y \sim b} d(x,y) \zeta(x,y)$$

$d(x,y)$ is the minimal path length between x and y. The sign $j \sim k$ means that j and k are neighbors. Now, in case of finite vertex set X of graph g, we have the next theorem.

Theorem 3. *Let X be a finite set and g a graph over X. For all $a, b \in X$ the transportation distance $W_1(\hat{m}_a, \hat{m}_b)$ satisfies*

$$W_1(\hat{m}_a, \hat{m}_b) = \min_{\zeta \in P_{ab}} \sum_{x, x \sim a} \sum_{y, y \sim b} d(x,y) \zeta(x,y)$$

Next, we will define the Ollivier-Ricci curvature for graphs.

Definition 7. *(**Ollivier - Ricci curvature**) Let g be a graph over X. For any two distinct points $x, y \in X$, the **Ollivier-Ricci curvature** along the path from x to y is defined as*

$$\kappa(x,y) = 1 - \frac{W_1(\hat{m}_x, \hat{m}_y)}{d(x,y)} \tag{1}$$

In general, so far what we achieved is an upper bound on the Wasserstein metric (W_1) and subsequently a lower bound on the Ricci curvature. There is another theorem called **Kantorovich-Rubenstein theorem** which provides an upper bound to the Ricci curvature based on 1-Lipschitz function.

1.2.1 The Kantorovich-Rubenstein theorem

Let X be an at most countable set. $Lip_M(X)$ or M-Lipschitz ($M \in \mathcal{R}$) is defined as the set of all functions $f : X \to \mathcal{R}$ on a graph g that satisfy

$$|f(x) - f(y)| \leq Md(x,y) \; \forall x, y \in X$$

Theorem 4. *(Kanorovich-Rubenstein theorem) Let X be a finite set and $a, b \in X$. Then, the transportation distance W_1 is equal to:*

$$W_1(\hat{m}_x, \hat{m}_y) = \sup \left\{ \sum_{z \in X, z \sim a} f(z) m_a(z) - \sum_{z \in X, z \sim b} f(z) m_b(z) \mid f \in Lip_1(X) \right\}$$

and we have an associated lemma of inequality:

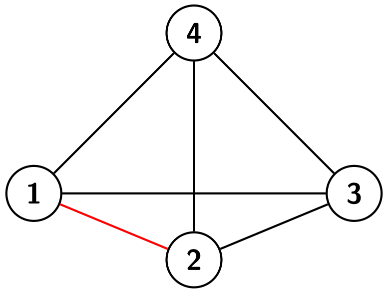
Lemma 1. *Let X be an at most countable set and $a, b \in X$. Then, the transportation distance W_1 is equal to:*

$$W_1(\hat{m}_x, \hat{m}_y) \geq \sup \left\{ \sum_{z \in X, z \sim a} f(z) m_a(z) - \sum_{z \in X, z \sim b} f(z) m_b(z) \mid f \in Lip_1(X) \right\}$$

We are going to use of this theorem in the next section to derive an upper bound of the Ricci curvature for our example graph.

2 Examples

In this section, we will calculate the Ollivier-Ricci curvature (coarse) for some graphs. First take a simple example of a regular tetrahedron (unweighted). This is a simple undirected graph with vertex set $X = \{1, 2, 3, 4\}$. Now consider $g(x, y)$ is either 0 ($x \nsim y$) or 1 ($x \sim y$ or neighbors). Let



us take vertices $a = 1$ and $b = 2$ and we will calculate W_1 between these two vertices. Now, the degree of vertex 1 is 3 denoted by w_a and degree of vertex 2 is also 3 denoted by w_b. Now consider N_a be the set of neighbors of a and N_b be the set of neighbors of b. Therefore, in our case $N_a = \{2, 3, 4\}$ and $N_b = \{1, 3, 4\}$. Because Now consider the following notation:

$$m_i(j) = \frac{g(i,j)}{\sum_{x \in X} g(i,x)}$$

Now using the above formula let us calculate $m_1(j)$ and $m_2(j)$. Here we will see two calculations to understand the difference that when $i \nsim j$, $m_i(j) = 0$.

$$m_1(1) = \frac{g(1,1)}{\sum_{x \in X} g(1,x)} = \frac{g(1,1)}{g(1,1) + g(1,2) + g(1,3) + g(1,4)} = \frac{0}{0+1+1+1} = 0$$
$$m_1(2) = \frac{g(1,2)}{\sum_{x \in X} g(1,x)} = \frac{g(1,2)}{g(1,1) + g(1,2) + g(1,3) + g(1,4)} = \frac{1}{0+1+1+1} = 1/3$$

In similar fashion, we can obtain the M matrix whose elements are $m_i(j)$ or $m(i,j)$ as following (for vertices 1 and 2 only):

$$M = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 1/3 \end{bmatrix}$$

As a physical interpretation of $m_i(j)$, we can say that it represents the distribution of mass around the neighbors of location i. Now we will calculate $\zeta(x, y)$. But first rewrite the structure of the set contains ζ.

$$P_{ab} = \Bigg\{ \zeta : X \times X \to [0, \infty) \mid \sum_{x \in X} \zeta(x,z) = m_b(z), \sum_{y \in X} \zeta(z,y) = m_a(z) \forall z \in X,$$
$$\sum_{x,y \in X} \zeta(x,y) = 1 \Bigg\}$$

Following the above structure, we can have some initial results for the entries in the matrix ζ.

$$\sum_{x \in X} \zeta(x,2) = m_2(2) = 0 \implies \zeta(x,2) = 0 \forall x \in X$$

This is because for $x, y \in X, \zeta(x,y) \geq 0$ Similarly have,

$$\sum_{y \in X} \zeta(1,y) = m_1(1) = 0 \implies \zeta(1,y) = 0 \forall y \in X$$

So the initial construction of ζ matrix is as follows:

$$\zeta = \begin{bmatrix} 0 & 0 & 0 & 0 \\ * & 0 & * & * \\ * & 0 & * & * \\ * & 0 & * & * \end{bmatrix}$$

Now, let us form rest of the equations for ζ. Varying the rows, we have

$$\begin{aligned}
\zeta(1,1) + \zeta(2,1) + \zeta(3,1) + \zeta(4,1) &= m_2(1) = 1/3 \\
\zeta(1,3) + \zeta(2,3) + \zeta(3,3) + \zeta(4,3) &= m_2(3) = 1/3 \\
\zeta(1,4) + \zeta(2,4) + \zeta(3,4) + \zeta(4,4) &= m_2(4) = 1/3
\end{aligned}$$

Now varying the columns, we have

$$\begin{aligned}
\zeta(2,1) + \zeta(2,2) + \zeta(2,3) + \zeta(2,4) &= m_1(2) = 1/3 \\
\zeta(3,1) + \zeta(3,2) + \zeta(3,3) + \zeta(3,4) &= m_1(3) = 1/3 \\
\zeta(4,1) + \zeta(4,2) + \zeta(4,3) + \zeta(4,4) &= m_1(4) = 1/3
\end{aligned}$$

We also have another constraint that the sum of all the elements in ζ is equal to 1. From the above equations we can construct a linear programming problem with constraint in the form of $AY = B$. Now let us see what these A, Y and B are. Y and B are column vectors with the following entries:

$$Y = \left[\begin{array}{c} \zeta(1,1) \\ \zeta(1,2) \\ \zeta(1,3) \\ \zeta(1,4) \\ \hline \zeta(2,1) \\ \zeta(2,2) \\ \zeta(2,3) \\ \zeta(2,4) \\ \hline \zeta(3,1) \\ \zeta(3,2) \\ \zeta(3,3) \\ \zeta(3,4) \\ \hline \zeta(4,1) \\ \zeta(4,2) \\ \zeta(4,3) \\ \zeta(4,4) \end{array}\right] \qquad B = \left[\begin{array}{c} m_1(1) = 0 \\ m_1(2) = 1/3 \\ m_1(3) = 1/3 \\ m_1(4) = 1/3 \\ \hline m_2(1) = 1/3 \\ m_2(2) = 0 \\ m_2(3) = 1/3 \\ m_2(4) = 1/3 \end{array}\right]$$

Therefore the matrix A will be the following:

$$
A = \left[\begin{array}{ccccccc}
1 & \dots & 1 & & & & \\
 & & & \dots & & & \\
 & & & & 1 & \dots & 1 \\
\hline
1 & & & & 1 & & \\
 & \ddots & & \dots & & \ddots & \\
 & & 1 & & & & 1
\end{array}\right]
$$

A is *totally unimodular* and has entries 0 and 1. Moreover, A has no more than 2 non-zero entries on each column. A very simple choice for ζ would be

$$
\zeta = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1/9 & 0 & 1/9 & 1/9 \\ 1/9 & 0 & 1/9 & 1/9 \\ 1/9 & 0 & 1/9 & 1/9 \end{bmatrix} = \frac{1}{3 \times 3} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}
$$

So, the number of non-zero entries in ζ is at most the number of neighbors of a multiplied by the number of neighbors of b (in this particular type of uniform measure). Now the question is that if this choice is optimal or not. Solving the linear programming problem (in MATLAB or else) we can easily get the optimal choice which is:

$$
\zeta_{opt} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 \end{bmatrix}
$$

Now the Wasserstein distance or $W_1(\hat{m}_1, \hat{m}_2)$ for the optimal choice is:

$$
\begin{aligned}
W_1(\hat{m}_1, \hat{m}_2) &= \sum_{x, x\sim 1} \sum_{y, y \sim 2} d(x,y)\zeta_{opt}(x,y) \\
&= \sum_{x, x\sim 1} d(x,1)\zeta_{opt}(x,1) + d(x,3)\zeta_{opt}(x,3) + d(x,4)\zeta_{opt}(x,4) \\
&= d(2,1)\zeta_{opt}(2,1) + d(2,3)\zeta_{opt}(2,3) + d(2,4)\zeta_{opt}(2,4) \\
&+ d(3,1)\zeta_{opt}(3,1) + d(3,3)\zeta_{opt}(3,3) + d(3,4)\zeta_{opt}(3,4) \\
&+ d(4,1)\zeta_{opt}(4,1) + d(4,3)\zeta_{opt}(4,3) + d(4,4)\zeta_{opt}(4,4) \\
&= d(2,1)\zeta_{opt}(2,1) + d(3,3)\zeta_{opt}(3,3) + d(4,4)\zeta_{opt}(4,4) \\
&= 1 \times 1/3 + 0 \times 1/3 + 0 \times 1/3 \\
&= 1/3
\end{aligned}
$$

Now let us see what could be the value of $W_1(\hat{m}_1, \hat{m}_2)$ for our previous simple choice i.e. each non-zero entry is 1/9. It will be 7/9 (as we have two zero entries $d(3,3)$ and $d(4,4)$. Definitely this is not the minimum!. Now, we are going to use the **Kantorovich-Rubenstein theorem** to derive a lower bound on the W_1 metric. For any $f \in Lip_1(X)$ we have the following

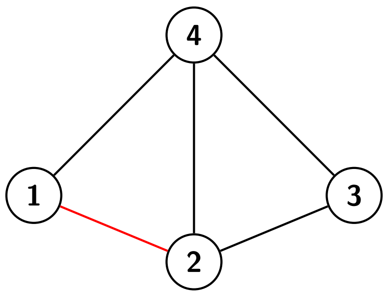
$$
\begin{aligned}
W_1(\hat{m}_x, \hat{m}_y) \geq & \sum_{z \in X, z \sim 1} f(z) m_1(z) - \sum_{z \in X, z \sim 2} f(z) m_2(z) \\
& f(2)m_1(2) + f(3)m_1(3) + f(4)m_1(4) \\
& - f(1)m_2(1) - f(3)m_2(3) - f(4)m_2(4) \\
& = 1/3(f(2) - f(1))
\end{aligned}
$$

Now, $f : X \to \mathcal{R}$ be such that $f(2) = 1$ and $f(x) = 0$ for $x \neq 2$. Therefore, $W_1(\hat{m}_x, \hat{m}_y) \geq 1/3$ Therefore, the Ricci curvature of the edge $1 \sim 2$ for our optimal choice is:

$$\kappa(1,2) = 1 - \frac{1/3}{1} = 2/3$$

For regular tetrahedron, this is same for all the edges due to symmetry. Note that the upper and lower bound for W_1 may not be same for all kinds of graphs.

Now, for example if we remove one of the edges from the tetrahedron, say for example edge $1 \sim 3$, we have the resulting figure:



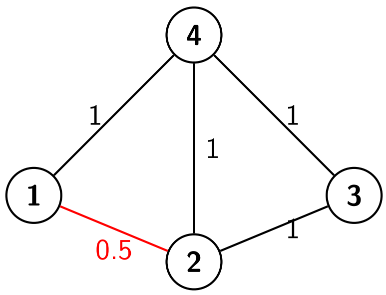
Now, if we calculate the Ricci curvature for the edge $1 \sim 2$, it will be lesser. Let us see what will happen. Our M matrix will be :

$$M = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 1/3 \end{bmatrix}$$

and in the same way (using Linear programming), we obtain the ζ matrix

$$\zeta_{opt} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \\ 0.0833 & 0 & 0.0833 & 0.33 \end{bmatrix}$$

The Wasserstein distance is 0.667 (obtained in the same way as above). Finally the Ricci curvature for $1 \sim 2$ is 0.33 which is lesser than that of the regular tetrahedron!. Next, we consider a weighted graph as follows:



Here, we redefine our measure function ($m_i(j)$) as follows:

$$m_i(j) = \begin{cases} \frac{\rho_{ij}}{\sum_{x \in N_i} \rho_{ix}} \text{ if } ij \in E \\ 0 \text{ otherwise} \end{cases}$$

Where ρ_{ij} is the weight of the edge $i \sim j$, N_i is the set of neighbors of i, E is the set of edges of the graph. In this set up, we have the M matrix as follows:

$$M = \begin{bmatrix} 0 & 0.33 & 0 & 0.667 \\ 0.2 & 0 & 0.4 & 0.4 \end{bmatrix}$$

and subsequently,

$$\zeta_{opt} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.2 & 0 & 0.1333 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2667 & 0.4 \end{bmatrix}$$

The corresponding the Wasserstein distance is 0.6 and the Ollivier-Ricci curvature is 0.4– slightly improvement from the previous one!.

3 Discussion

In this report we have discussed the coarse Ollivier-Ricci curvature for the graphs. There are quite a handful of research papers [4, 3, 1, 5] which discussed and modified the definition of Ollivier-Ricci in a slightly different way. For example, [4] considers the measure $m_x^\alpha(y)$ as a *lazy random walk* in the following way:

$$m_x^\alpha(y) = \begin{cases} \alpha \text{ if } y = x \\ \frac{1-\alpha}{d_x} \text{ if } y \in N_x \\ 0 \end{cases}$$

Where, N_x denotes the set of neighbor of vertex x and d_x denotes the degree of x and $\alpha \in [0, 1]$. The corresponding α-Ricci curvature is:

$$\kappa^{\alpha}(x,y) = 1 - \frac{W_1(\hat{m}_x^{\alpha}, \hat{m}_y^{\alpha})}{d(x,y)} \tag{2}$$

We can observe that, if $\alpha \to 1$, $\kappa^{\alpha}(x,y) \to 0$ and the asymptotic Ollivier-Ricci curvature is:

$$Ric(x,y) = \lim_{\alpha \to 1} \inf \frac{\kappa^{\alpha}(x,y)}{1-\alpha} \tag{3}$$

This definition is more appropriate for a Markov chain defined on the graph. It corresponds to a continuous Markov process. The authors show that with *Ric* the curvature for cube is 2/3 where as with our previous coarse form it is 0. Moreover, for dodecahedron *Ric* is 0 or at-least non-negative but with our usual coarse form it is negative $-1/3$ which should not be as it is an abstraction from sphere and should have non-negative Ricci curvature.

Some practical graph theoretical applications of Ollivier-Ricci curvature can be found in [6, 8, 10]. Papers [6, 10] used the definition of $\kappa^{\alpha}(x,y)$. [6] interprets that in a complex network, negative Ricci curvature edges act as bridges. Hence, removal of these edges makes the whole graph disconnected very fast! Hence it is not desirable to have more number of negative Ricci curvature edges in a graph/network. It affects the robustness of the graph/network. The same conclusion is drawn by the paper [8]. Here the authors relate the Ricci curvature to the entropy of the network and applied it to measure the robustness in cancer networks (interactions between genes). In an example they show that the well studied E-coli transcription network (which claimed to be a robust network structure) has less number of negative curvature edges than its random counterpart. Paper [6] also infer that the positive curvature edges have more number of triangles associated with them which is also natural according to the findings of the paper [3] which shows that the inclusion of triangles increases the Ricci curvature. The authors in [3] provide a sharp lower bound on Ricci curvature based on the number of triangles in the graph and subsequently relate the curvature to the clustering coefficient of the graph. Sharp inequality (lower bound) on Ricci curvature for a undirected, weighted, connected, finite (multi) graph is provided by [2] which is claimed to be inappropriate for example in case of Euclidean square lattice [10]. Hence the usual way of linear programming is necessary to estimate the exact value of the curvature.

From the above discussion, it is evident that the Ricci-curvature based approach especially on discrete structures like graphs is promising. Therefore it will be interesting also to investigate further on how to apply the concept on graphical abstraction of mechanical structures and infer the qualitative properties of the structure most importantly robustness.

References

[1] Frank Bauer, Bobo Hua, Jürgen Jost, Shiping Liu, et al. Generalized ricci curvature and the geometry of graphs. 2015.

[2] Frank Bauer, Jürgen Jost, and Shiping Liu. Ollivier-ricci curvature and the spectrum of the normalized graph laplace operator. *arXiv preprint arXiv:1105.3803*, 2011.

[3] Jürgen Jost and Shiping Liu. Ollivier?s ricci curvature, local clustering and curvature-dimension inequalities on graphs. *Discrete & Computational Geometry*, 51(2):300–322, 2014.

[4] Yong Lin, Linyuan Lu, Shing-Tung Yau, et al. Ricci curvature of graphs. *Tohoku Mathematical Journal*, 63(4):605–627, 2011.

[5] Benoît Loisel and Pascal Romon. Ricci curvature on polyhedral surfaces via optimal transportation. *Axioms*, 3(1):119–139, 2014.

[6] Chien-Chun Ni, Yu-Yao Lin, Jie Gao, Xianfeng David Gu, and Emil Saucan. Ricci curvature of the internet topology. *arXiv preprint arXiv:1501.04138*, 2015.

[7] Yann Ollivier. Ricci curvature of markov chains on metric spaces. *Journal of Functional Analysis*, 256(3):810–864, 2009.

[8] Allen Tannenbaum, Chris Sander, Liangjia Zhu, Romeil Sandhu, Ivan Kolesov, Eduard Reznik, Yasin Senbabaoglu, and Tryphon Georgiou. Ricci curvature and robustness of cancer networks. *arXiv preprint arXiv:1502.04512*, 2015.

[9] Cédric Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.

[10] Chi Wang, Edmond Jonckheere, and Reza Banirazi. Wireless network capacity versus ollivier-ricci curvature under heat-diffusion (hd) protocol. In *American Control Conference (ACC), 2014*, pages 3536–3541. IEEE, 2014.