mathematics

MDPI

Article

A Review of and Some Results for Ollivier–Ricci Network Curvature

Nazanin Azarhooshang †, Prithviraj Sengupta † and Bhaskar DasGupta *,†

Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA; nazarh2@uic.edu (N.A.); psengu4@uic.edu (P.S.)

* Correspondence: bdasgup@uic.edu; Tel.: +1-312-355-1319

† These authors contributed equally to this work.

Received: 31 July 2020; Accepted: 18 August 2020; Published: 24 August 2020

Abstract: Characterizing topological properties and anomalous behaviors of higher-dimensional topological spaces via notions of curvatures is by now quite common in mainstream physics and mathematics, and it is therefore natural to try to extend these notions from the non-network domains in a suitable way to the network science domain. In this article we discuss one such extension, namely Ollivier's discretization of Ricci curvature. We first motivate, define and illustrate the Ollivier–Ricci Curvature. In the next section we provide some "not-previously-published" bounds on the exact and approximate computation of the curvature measure. In the penultimate section we review a method based on the linear sketching technique for efficient approximate computation of the Ollivier–Ricci network curvature. Finally in the last section we provide concluding remarks with pointers for further reading.

Keywords: network science; network curvature; discrete Ricci curvature; earth-mover's distance

MSC: 68Q17; 68W40

1. Introduction

It is by now quite common in mainstream physics and mathematics [1,2] to characterize topological properties and anomalous behaviors of higher-dimensional topological spaces via notions of (local and global) curvatures of these spaces, e.g., in general relativity, extreme variations of four dimensional space-time curvatures via geodesic incompleteness lead to characterizations of black-holes [3]. It is therefore natural to try to extend these notions from the non-network domains e.g., from continuous metric spaces or from higher-dimensional geometric objects) in a suitable way to the network science domain so that non-trivial new topological characteristics of networks can be captured. There are several ways this can be achieved; we briefly mention two other approaches before proceeding with the approach that is the main topic of this paper. Note that such extensions need to overcome at least two key challenges, namely that (i) networks are discrete (non-continuous) objects, and that (ii) networks may not necessarily have an associated natural geometric embedding.

One notion of network curvature that has been well-studied in the network theory literature, first suggested by Gromov in a non-network group theoretic context [4], is the Gromov-hyperbolic curvature. First defined for infinite continuous metric space [2], the measure was later adopted for finite graphs. Usually the measure is defined via properties of geodesic triangles or via equivalent (in a sense that can be made precise) 4-node conditions, though Gromov originally defined the measure using Gromov-product nodes in [4]. Informally any infinite metric space has a finite Gromov-hyperbolicity measure if it behaves metrically in the large scale as a negatively curved Riemannian manifold, and thus the value of this measure can be correlated to the standard scalar curvature of a hyperbolic

manifold. Intuitively, for a finite network the measure is based on the properties of the set of exact and approximate geodesics of the network. There is a large body of research works dealing with theoretical and empirical aspects of this measure, e.g., see [5–10] for theoretical aspects, and see [11–13] for empirical aspects with applications to real-world networks.

A second notion of curvature is the applying Forman's discretization of Ricci curvature for (polyhedral or CW) complexes (the "Forman–Ricci curvature") [14] to networks. Informally, one applies the Forman-Ricci curvature to networks by topologically associating components (sub-graphs) of the given graphs with higher-dimensional objects. The topological association itself can be carried out several ways. Although this type of curvature originated relatively recently, there are already a number of papers investigating properties of these measures and applying them to real-world networks, e.g., see [8,15–18].

The network curvature discussed in this paper is another discretization of Ricci curvature, namely Ollivier's discretization [19–22], henceforth dubbed as the "Ollivier–Ricci curvature". Both Ollivier–Ricci curvature and Forman-Ricci curvature assign measures that assign a number to each edge of the given network, but the numbers are calculated in quite different ways in these two curvatures since they capture different metric properties of a Riemannian manifold. The reader is referred to the paper by [15] for a comparative analysis of these two measures. In addition to the network curvatures measures discussed above, researchers have also explored other notions of curvature, such as the one based on circle packings by Chow and Luo [23].

Basic Notations and Terminologies

To simplify exposition, we assume in this paper that the given network (In this paper the terms "graph" and "network" will be used interchangeably.) $G = (V, E)$ is an undirected unweighted connected graph; generalization of the corresponding definitions and concepts to the case of non-negative edge weights is mostly straightforward. The following notations will be used in the rest of this paper.

- ▷ For a node $v \in V$, $\mathsf{Nbr}(v) = \{u \mid \{v, u\} \in E\}$ denotes the set of neighbors of v, and $\deg(v) = |\mathsf{Nbr}(v)|$ denotes the degree of v.
- ▷ $\mathrm{dist}_G(u, v)$ (or simply $\mathrm{dist}(u, v)$) denote the distance (i.e., number of edges in a shortest path) between the nodes u and v in G.

2. Ollivier–Ricci Curvature: Motivation, Definition and Illustration

In this section, we provide the formal definition of the Ollivier–Ricci curvature. First, we need to define the so-called Earth Mover's Distance (EMD) (also known as the L_1 transportation distance, the L_1 Wasserstein distance and the Monge-Kantorovich-Rubinstein distance) [24–27]. For the purpose of this paper, it suffices to define the distance in the discrete setting of a network as follows. Suppose that we have two probability distributions $\mathbb{P}_1$ and $\mathbb{P}_2$ on a subset $\emptyset \subset V' \subseteq V$ of nodes, i.e., two real numbers $0 \leq \mathbb{P}_1(v), \mathbb{P}_2(v) \leq 1$ for every node $v \in V'$ with $\sum_{v \in V'} \mathbb{P}_1(v) = \sum_{v \in V'} \mathbb{P}_2(v) = 1$. We can think of every number $\mathbb{P}_1(v)$ as the maximum total amount of "earth" (dirt) at node v that can be moved to other nodes, and every number $\mathbb{P}_2(v)$ as the maximum total amount of earth node v can store in its storage. The cost of transporting one unit of earth from node u to node v is $\mathrm{dist}_G(u, v)$, and the goal is to satisfy the storage requirement of all nodes by moving earths as needed while minimizing the total transportation cost. Letting the variable $z_{u,v} \in [0, 1]$ denote the amount of shipment from node u to node v in an optimal solution, EMD for the two probability distributions $\mathbb{P}_1$ and $\mathbb{P}_2$ on V' can be formulated as the linear programming (LP) problem shown in Figure 1 which can be solved in polynomial time. One can also think of the EMD solution as the distance between two probability distributions $\mathbb{P}_1$ and $\mathbb{P}_2$ on the set of nodes V' based on the shortest-path metric on G. We will use the notation $\mathrm{EMD}(V', \mathbb{P}_1, \mathbb{P}_2)$ to denote the value of the objective function in an optimal solution of the LP in Figure 1.

variables: $z_{u,v}$ for every pair of nodes $u, v \in V'$

minimize $\sum_{u \in V'} \sum_{v \in V'} \text{dist}(u,v)\, z_{u,v}$ (* minimize total transportation cost *)

subject to

$\sum_{v \in V'} z_{u,v} = \mathbb{P}_1(u)$, for each $u \in V'$ (* take from u as much as it has *)

$\sum_{u \in V'} z_{u,v} = \mathbb{P}_2(v)$, for each $v \in V'$ (* ship to v as much as it needs *)

$z_{u,v} \geq 0$, for all $u, v \in V'$

Figure 1. LP-formulation for EMD on the set of nodes $|V'|$ with $|V'|^2$ variables. Comments are enclosed by (* and *). Note that the constraints $z_{u,v} \leq 1$ are unnecessary and therefore omitted.

For an intuitive understanding of the connection of EMD to Ollivier–Ricci curvature for networks, we informally recall one way of defining Ricci curvature measure for a smooth Riemannian manifold. The Ricci curvature at a point x in the manifold along a direction can be thought of transporting a small ball centered at x along that direction and measuring the "distortion" of that ball. The role of the direction is captured by the edge $\{u, v\}$, the roles of the balls at the two nodes are played by the distributions $\mathbb{P}_1$ and $\mathbb{P}_2$, and the role of the distortion due to transportation is captured by the EMD measure. More precisely, given our input graph $G = (V, E)$ and an edge $\{u, v\} \in E$, the paper [20] uses the EMD measure to define the "course Ricci curvature" $\text{RIC}(u, v)$ along the edge $\{u, v\}$ in the following manner (see Figure 2 for an illustration):

- Let V' be the set of nodes $V_{u,v} \stackrel{\text{def}}{=} \{u, v\} \cup \mathsf{Nbr}(u) \cup \mathsf{Nbr}(v)$.
- Let the probability distributions $\mathbb{P}_1$ and $\mathbb{P}_2$ be uniform distributions (If the given graph is non-negative node weights then another option is to normalize the restrictions of these node weights to the sub-graph $H_{u,v}$ and use them for the distributions $\mathbb{P}_1$ and $\mathbb{P}_2$.) $\mathbb{P}_u$ and $\mathbb{P}_v$, respectively, over the nodes in $\{u\} \cup \mathsf{Nbr}(u)$ and $\{v\} \cup \mathsf{Nbr}(v)$, respectively, i.e.,

$$\mathbb{P}_u(x) \stackrel{\text{def}}{=} \mathbb{P}_1(x) = \begin{cases} \frac{1}{|\{u\} \cup \mathsf{Nbr}(u)|}, & \text{if } x \in \{u\} \cup \mathsf{Nbr}(u) \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbb{P}_v(x) \stackrel{\text{def}}{=} \mathbb{P}_2(x) = \begin{cases} \frac{1}{|\{v\} \cup \mathsf{Nbr}(v)|}, & \text{if } x \in \{v\} \cup \mathsf{Nbr}(v) \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

- Remembering that $\text{dist}_G(u, v) = 1$ for an edge $\{u, v\} \in E$, we can then define the course Ricci curvature as (*cf.* [20] (Definition 3)):

$$\text{RIC}(u,v) = 1 - \frac{\text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v)}{\text{dist}_G(u,v)} \equiv \text{RIC}(u,v) = 1 - \text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v) \tag{2}$$

The measure can easily be extended for graphs with non-negative edge weights; redefine $\text{dist}(u, v)$ to be minimum total weight over all possible paths between u and v and use the equation:

$$\text{RIC}(u,v) = 1 - \frac{\text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v)}{\text{dist}_G(u,v)}$$

Some authors also define the discrete Ricci curvature $\text{RIC}(u)$ for a node $u \in V$ by taking the average of the discrete Ricci curvarure over all edges incident on u, e.g., by letting $\text{RIC}(u) = \frac{\sum_{\{u,v\} \in E} \text{RIC}(u,v)}{\deg(u)}$.

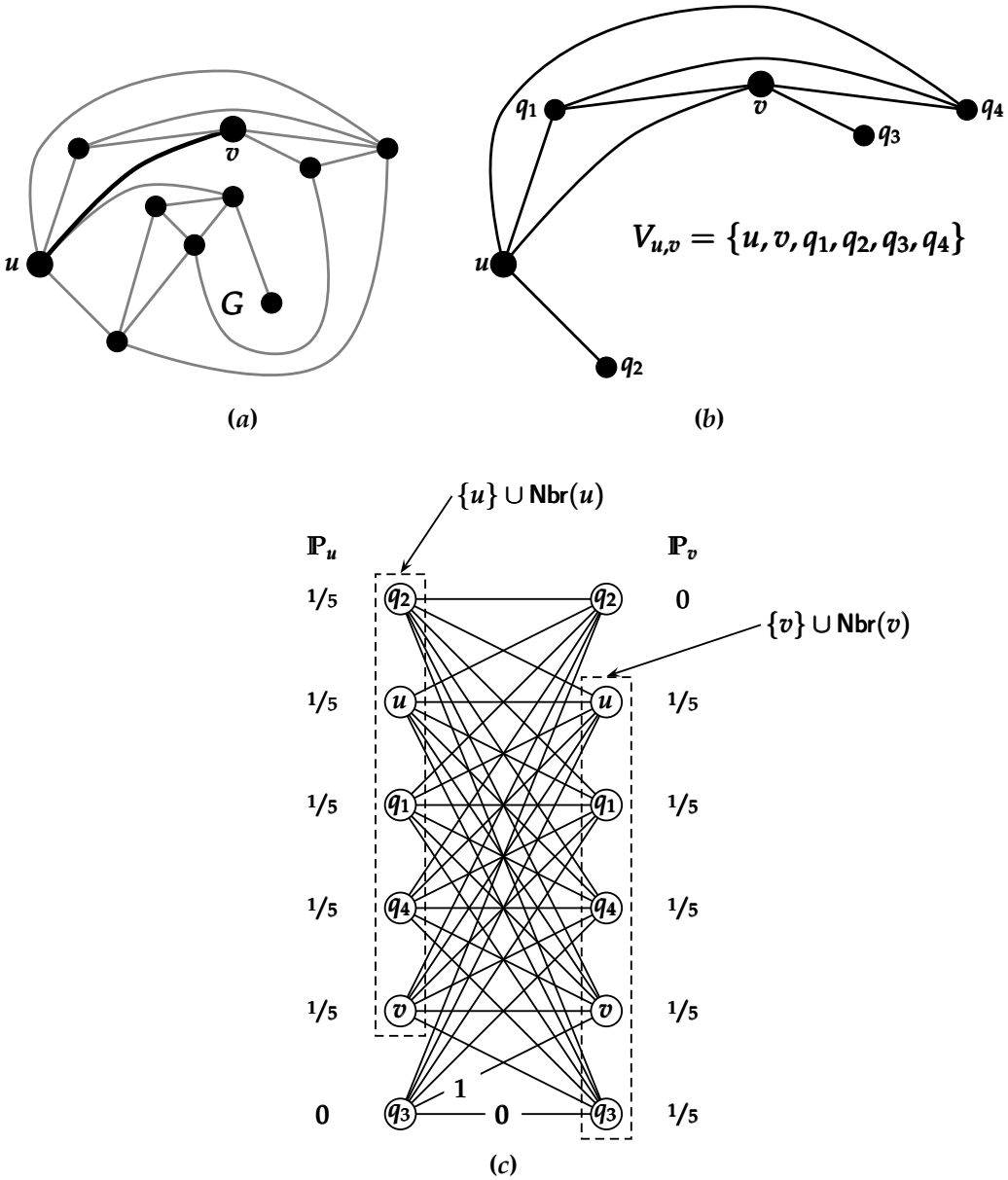

Figure 2. A pictorial illustration of calculation of $\textsc{Ric}(u, v)$. (**a**) The given graph G; (**b**) The subset of nodes $V_{u,v}$; (**c**) The distributions $\mathbb{P}_u$ and $\mathbb{P}_v$. For visual clarity, only two distances $\mathrm{dist}(q_3, q_3) = 0$ and $\mathrm{dist}(v, q_3) = 1$ are shown.

An Illustration of Computing the Value of $\textsc{Ric}(u, v)$ *For a Two-dimensional Grid*

Consider an infinite two-dimensional grid on the plane and any edge $\{u, v\}$ of the grid as shown in Figure 3. Note that any node of the grid has exactly 4 neighbors, thus $\mathbb{P}_u(x) = \begin{cases} 1/5, & \text{if } x \in \{u\} \cup \mathsf{Nbr}(u) \\ 0, & \text{otherwise} \end{cases}$ and $\mathbb{P}_v(x) = \begin{cases} 1/5, & \text{if } x \in \{v\} \cup \mathsf{Nbr}(v) \\ 0, & \text{otherwise} \end{cases}$. Moreover, the set of nodes $\mathsf{Nbr}(u) \setminus \{v\}$ and $\mathsf{Nbr}(v) \setminus \{u\}$ are disjoint, thus it is easy to see that $\textsc{Emd}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v) = 1$ (see Figure 3). Using (2) we therefore get $\textsc{Ric}(u, v) = 0$.

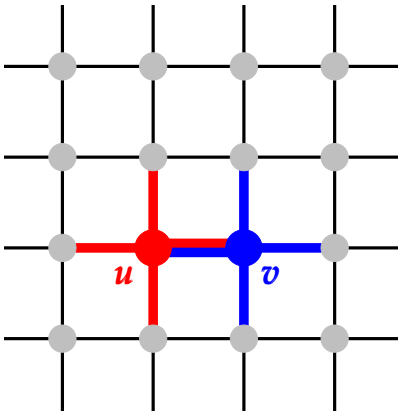

Figure 3. A pictorial illustration of calculation of $\textsc{Ric}(u,v)$ for a two-dimensional grid. The blue edges, when shifted to the left by one unit, coincide with the red edges, giving $\textsc{Emd}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v) \leq 1$. It can also be argued that $\textsc{Emd}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v) \geq 1$ (e.g., see [20] (Example 5) with $N=2$), thus giving $\textsc{Emd}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v) = 1$.

3. Exact and Approximate Computation of $\mathbf{Ric}(u,v)$

Note that any node $x \in V_{u,v}$ with either $\mathbb{P}_u(x) = 0$ or $\mathbb{P}_v(x) = 0$ can be ignored in the calculation of $\textsc{Emd}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v)$. Thus, a straightforward calculation of $\textsc{Ric}(u,v)$ requires the following two steps:

- ▷ Find the pair-wise distances between the nodes in $\mathsf{Nbr}(u)$ and $\mathsf{Nbr}(v)$. This can be done in $O(n^\omega \log n)$ using Seidel's algorithm [28] where n is the number of nodes and ω be the value such that two $n \times n$ matrices can be multiplied in $O(n^\omega)$ time; the smallest current value of ω is slightly less than 2.373 [29].
- ▷ Solve an LP with $O(\deg(u)\deg(v))$ variables and $O(\deg(u)\deg(v))$ constraints via standard LP solvers such as the interior-point method. Alternatively, the LP can be solved by minimum-cost network flow algorithms by viewing it as a transportation problem, e.g., see [30].

However, the calculation of $\textsc{Emd}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v)$ (and therefore $\textsc{Ric}(u,v)$) can be further simplified if we make some more observations.

Consider a pair of nodes $u' \in \mathsf{Nbr}(u)$ and $v' \in \mathsf{Nbr}(v)$ for an edge $\{u,v\} \in E$. Note that there are only four possible values of $\text{dist}_G(u',v')$: $\text{dist}_G(u',v') = 0$ if $u' = v'$, $\text{dist}_G(u',v') = 1$ if $\{u',v'\} \in E$, $\text{dist}_G(u',v') = 2$ if there is a path of length 2 between u' and v', and $\text{dist}_G(u',v') = 3$ for all other cases. Thus, to to find all pair-wise distances between the nodes in $\mathsf{Nbr}(u)$ and $\mathsf{Nbr}(v)$ we only need to check for paths up to length 3, which can be done faster in $O(n^\omega)$ time using Seidel's algorithm [28] again.

For further discussion, consider the total variation distance (TVD) between the two distributions $\mathbb{P}_u$ and $\mathbb{P}_v$ on the set of nodes in $V_{u,v}$:

$$||\,\mathbb{P}_u - \mathbb{P}_v\,||_{\text{TVD}} \stackrel{\text{def}}{=} \frac{1}{2}\left(\sum_{v' \in V_{u,v}} \left(\,|\mathbb{P}_u(v') - \mathbb{P}_v(v')|\,\right)\right)$$

Note that $||\,\mathbb{P}_u - \mathbb{P}_v\,||_{\text{TVD}}$ can be trivially computed in $O(\deg(u) + \deg(v))$ time.

Proposition 1. $1 - 3||\,\mathbb{P}_u - \mathbb{P}_v\,||_{\text{TVD}} \leq \textsc{Ric}(u,v) \leq 1 - ||\,\mathbb{P}_u - \mathbb{P}_v\,||_{\text{TVD}}$.

Proof. Since every pair of non-identical nodes $u', v' \in V_{u,v}$ satisfy $1 \leq \text{dist}_G(u',v') \leq 3$, we have $||\,\mathbb{P}_u - \mathbb{P}_v\,||_{\text{TVD}} \leq \textsc{Emd}(V_{u,v},\mathbb{P}_u,\mathbb{P}_v) \leq 3||\,\mathbb{P}_u - \mathbb{P}_v\,||_{\text{TVD}}$ which imply the claimed result via definition of $\textsc{Ric}(u,v)$. □

The bound in Proposition 1 may not necessarily be a tight approximation for $\textsc{Ric}(u,v)$; for example, for the grid in Figure 3 we get $||\,\mathbb{P}_u - \mathbb{P}_v\,||_{\text{TVD}} = 3/5$ giving $-4/5 \leq \textsc{Ric}(u,v) \leq 2/5$ as an approximation to the actual value of $\textsc{Ric}(u,v) = 0$.

For development of further bounds, consider the edge $\{u, v\} \in E$. Assume without loss of generality that $\deg(u) \leq \deg(v)$ and G has 4 or more nodes, thus $\deg(v) \geq 2$. Suppose that u and v have $0 \leq \ell \leq \deg(u)$ common neighbour nodes as shown pictorially below:

$$\mathsf{Nbr}(u) = \{ \overbrace{p_1, p_2, \ldots, p_k, q_1, q_2, \ldots, q_\ell}^{k+\ell=\deg(u)-1 \geq \ell+1 \text{ nodes}} \}$$

$$\{ \underbrace{q_1, q_2, \ldots, q_\ell}_{\substack{\ell \geq 0 \text{ common} \\ \text{neighbours}}}, \; r_1, r_2, \ldots, r_m \} = \mathsf{Nbr}(v)$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{m+\ell=\deg(v)-1 \geq \ell+1 \text{ nodes}}$$

Note that the two probability vectors $\mathbb{P}_u$ and $\mathbb{P}_v$ for the edge $\{u, v\}$ are as shown below:

$$\begin{array}{cccccccccccc} & p_1 & \ldots & p_k & q_1 & \ldots & q_\ell & u & r_1 & \ldots & r_m & v \\ \mathbb{P}_u = \Big(& \frac{1}{\deg(u)+1} & \cdots & \frac{1}{\deg(u)+1} & \frac{1}{\deg(u)+1} & \cdots & \frac{1}{\deg(u)+1} & \frac{1}{\deg(u)+1} & 0 & \ldots & 0 & \frac{1}{\deg(u)+1} \Big) \\ \mathbb{P}_v = \Big(& 0 & \ldots & 0 & \frac{1}{\deg(v)+1} & \cdots & \frac{1}{\deg(v)+1} & \frac{1}{\deg(v)+1} & \frac{1}{\deg(v)+1} & \cdots & \frac{1}{\deg(v)+1} & \frac{1}{\deg(v)+1} \Big) \end{array}$$

By our assumption $\frac{1}{\deg(u)+1} \geq \frac{1}{\deg(v)+1}$, and thus a straightforward calculation gives the following value for $|| \mathbb{P}_u - \mathbb{P}_v ||_{\text{TVD}}$:

$$\begin{aligned} || \mathbb{P}_u - \mathbb{P}_v ||_{\text{TVD}} &= \frac{1}{2} \times \left(\frac{k}{\deg(u)+1} + \frac{m}{\deg(v)+1} + (\ell+2) \times \left(\frac{1}{\deg(u)+1} - \frac{1}{\deg(v)+1} \right) \right) \\ &= \frac{\frac{k+\ell}{2}+1}{\deg(u)+1} + \frac{\frac{m-\ell}{2}-1}{\deg(v)+1} = \frac{1}{2} + \frac{(\deg(v)+1) - 2(\ell+2)}{2(\deg(v)+1)} = 1 - \frac{\ell+2}{\deg(v)+1} \end{aligned} \quad (3)$$

Proposition 2. $-2 + 3\frac{\ell+2}{\deg(v)+1} \leq \text{RIC}(u, v) \leq \frac{\ell+2}{\deg(v)+1}$, *and in particular it always holds that* $-2 < \text{RIC}(u, v) \leq 1$.

Proof. Plugging the bound (3) in Proposition 1 proves the first claim. To prove the second claim, note that $0 < \frac{\ell+2}{\deg(v)+1} \leq 1$. □

For further bounds, suppose that there exists a $\gamma \in \{1, 2, 3\}$ such that for any two distinct nodes $u' \in \mathsf{Nbr}(u)$ and $v' \in \mathsf{Nbr}(v)$ we have $\text{dist}(u', v')$ is exactly γ. In that case, it follows that

$$\text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v) = \gamma \times || \mathbb{P}_u - \mathbb{P}_v ||_{\text{TVD}} \Rightarrow \text{RIC}(u, v) = 1 - \gamma \times || \mathbb{P}_u - \mathbb{P}_v ||_{\text{TVD}} = 1 - \gamma + \frac{\gamma(\ell+2)}{\deg(v)+1}$$

Now, suppose that G has no cycles of 5 of fewer edges containing the edge $\{u, v\}$ (a tree is a trivial example of such a graph). This implies $\gamma = 3$ and $\ell = 0$, giving the following bound.

Proposition 3. *If G has no cycles of* 5 *of fewer edges containing the edge* $\{u, v\}$ *then* $\text{RIC}(u, v)$ *is precisely* $-2 + \frac{6}{\deg(v)+1} \leq 0$ *and can be computed in* $O(\deg(u) + \deg(v))$ *time.*

4. Review of Efficient Approximate Computation of $\text{Ric}(u, v)$ via Linear Sketching

It is clear that a crucial bottleneck in computing $\text{RIC}(u, v)$ for an arbitrary graph $G = (V, E)$ is the computation of $\text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v)$ since it seems to require solving a linear program with $O(\deg(u) \deg(v))$ variables and $O(\deg(u) \deg(v))$ constraints (note that in the worst case $\deg(u) \deg(v)$ can be as large as $\Theta(n^2)$ when n is the number of nodes of G). In this section we review

a non-trivial approach for computing $\text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v)$ provided we settle for a *slightly* non-optimal solution for $\text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v)$.

Linear sketching is a popular method to perform approximate computations on large data sets using dimensionality reduction [31]. The general (informal) intuition behind linear sketching is to take linear projections of the given data set and then use these projections to provide solutions to the original problem. Significant research has been done on the problem of estimating EMD using linear sketches for general metric spaces [32–36]. In this section, we discuss the results by McGregor and Stubbs [37] to approximately estimate EMD on a graph metric (i.e., metric induced by inter-node distances in a graph, as is the case for computing $\text{RIC}(u, v)$). Recall that our bottleneck is the computation of $\text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v)$ for the given graph G.

The first step is to transform the problem of computing $\text{EMD}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v)$ by standard techniques to the following equivalent problem which will be denoted by EMD_{d}. Given two multi-sets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{X}$ over a ground set $\mathcal{X}$ with $|\mathcal{A}| = |\mathcal{B}| = k$, and a metric $\mathsf{d} : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}^+$ on $\mathcal{X}$, compute the minimum-cost of perfect matching between $\mathcal{A}$ and $\mathcal{B}$, i.e., using $\pi_{\mathcal{A},\mathcal{B}}$ to denote a 1-1 mapping from $\mathcal{A}$ to $\mathcal{B}$, we need to compute

$$\text{EMD}_{\mathsf{d}}(\mathcal{A}, \mathcal{B}) = \min_{\pi_{\mathcal{A},\mathcal{B}}} \Big\{ \sum_{a \in \mathcal{A}} \mathsf{d}(a, \pi_{\mathcal{A},\mathcal{B}}(a)) \Big\}$$

For the purpose of measuring approximation quality, we say that an algorithm is an (ϵ, δ)-algorithm for computing a quantity of value Q if the value Q' returned by the algorithm satisfies $\Pr[\,|Q - Q'| < \epsilon Q\,] \geq 1 - \delta$.

The basic approach of McGregor and Stubbs in [37] is to define two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{|E|}$ corresponding to the set $\mathcal{A}$ and $\mathcal{B}$. We then estimate $\text{EMD}_{\mathsf{d}}(\mathcal{A}, \mathcal{B})$ by posing it as a ℓ_1-regression problem using the vectors $\mathbf{x}, \mathbf{y}$ and a set of other vectors defined by the structure of the underlying graph. The idea is take some random projections of these vectors to a smaller dimensional space and then perform ℓ_1-regression on these projections to save space and time. The following result by Kane et al. [38] is crucial to the analysis of this approach (the notation $\Pr_{M \sim \nu}$ is the standard notation for denoting that the entries of M are drawn from the distribution ν):

> $(\star)$ There exists a distribution ("q-dimensional sketch") ν over linear maps from $\mathbb{R}^n \mapsto \mathbb{R}^q$ where $q = O(\varepsilon^{-2} \log n \log \delta^{-1})$ and a "post-processing" function $f : \mathbb{R}^q \mapsto \mathbb{R}$ such that for any $\mathbf{x} \in \mathbb{R}^n$ with polynomially-bounded entries, it holds that
>
> $$\Pr_{M \sim \nu} \big[\, \big|\, \| \mathbf{x} \|_1 - f(M\mathbf{x}) \,\big| \leq \varepsilon \| \mathbf{x} \|_1 \,\big] \geq 1 - \delta$$

To understand how the above result relates to the calculation of $\text{EMD}_{\mathsf{d}}(\mathcal{A}, \mathcal{B})$, first consider the case when the given instance of $\text{EMD}_{\mathsf{d}}(\mathcal{A}, \mathcal{B})$ is one dimensional, i.e., let $G = (V, E)$ be a path with n nodes $V = \{1, \ldots, n\}$ and $n - 1$ edges $E = \{e_1, \ldots, e_{n-1}\}$ where $e_i = \{i, i+1\}$, let $A, B \subseteq V$, and let $\mathsf{d}(i, j) = \text{dist}_G(i, j)$ for all $i, j \in V$. Then we can associate computation of $\text{EMD}_{\mathsf{d}}(\mathcal{A}, \mathcal{B})$ to a norm estimation problem in the following manner. Assume that we have vectors $\mathbf{x} = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$ and $\mathbf{y} = (y_1, \ldots, y_{n-1}) \in \mathbb{R}^{n-1}$ such that for all $i \in \{0, 1, n-1\}$ the following assertions hold: $x_i = |\{a \in \mathcal{A} \,|\, i \geq a\}|$ and $y_i = |\{b \in \mathcal{B} \,|\, i \geq b\}|$. Then, it can be shown that $\text{EMD}_{\mathsf{d}}(\mathcal{A}, \mathcal{B}) = \| \mathbf{x} - \mathbf{y} \|_1$ and thus we can use the result of Kane et al. [38] as stated in $(\star)$ directly.

As a second illustration of the above point, suppose that the graph G in the previous example is now a cycle of n nodes $V = \{1, \ldots, n\}$ and n edges $E = \{e_1, \ldots, e_n\}$ where $e_i = \{i, i+1\}$ for $i \in \{1, \ldots, n-1\}$ and $e_n = \{n, 1\}$. Suppose that we simply ignore the last edge e_n so that the graph becomes a path and we can apply the previous approach. However, this omission of e_n changes the distance between the nodes $i \in \mathcal{A}$ and $j \in \mathcal{B}$ from $\mathsf{d}(i, j) = \min\{|i - j|, |i - n| + 1 + |1 - j|, |i - 1| + 1 + |n - j|\}$ to a new distance $\mathsf{d}'(i, j) = |i - j|$. To resolve this issue, we make a sequence of guesses for the number of pairs of nodes that will be joined using the edge e_n. More precisely,

for $\lambda \in \{-k, -k+1, \ldots, k-1, k\}$ let $\mathcal{C}_\lambda$ be the multi-set consisting of λ copies of "1" if $\lambda > 0$ and $|\lambda|$ copies of "n" if $\lambda < 0$. Then, one can show that

$$\textsc{Emd}_{\mathsf{d}}(\mathcal{A}, \mathcal{B}) \leq |\lambda| + \textsc{Emd}_{\mathsf{d}'}(\mathcal{A} \uplus \mathcal{C}_\lambda, \mathcal{B} \uplus \mathcal{C}_{-\lambda})$$

with equality for some $\lambda \in \{-k, -k+1, \ldots, k-1, k\}$, where $\uplus$ denotes the union for multi-sets. Thus, we can use the result in $(\star)$ in the following manner. First define two vectors $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ where $x_i = |\{a \in \mathcal{A} \mid i \geq a\}|$ and $y_i = |\{b \in \mathcal{B} \mid i \geq b\}|$ for $i \in \{1, \ldots, n-1\}$, and $x_n = y_n = 0$. Let $\mathbf{z} = \mathbf{x} - \mathbf{y}$ and $\mathbf{c} = (1, \ldots, 1) \in \mathbb{R}^n$. Then, it follows that

$$\textsc{Emd}_{\mathsf{d}}(\mathcal{A}, \mathcal{B}) = \min_{\lambda \in \{-k, -k+1, \ldots, k-1, k\}} \left\{ \parallel \mathbf{z} + \lambda \mathbf{c} \parallel_1 \right\}$$

Define the function $f : \mathbb{R} \mapsto \mathbb{R}$ as $f(\lambda) = \parallel \mathbf{z} + \lambda \mathbf{c} \parallel_1$; clearly $\textsc{Emd}_{\mathsf{d}}(\mathcal{A}, \mathcal{B}) = \min_{\lambda \in \{-k, -k+1, \ldots, k-1, k\}} \{f(\lambda)\}$. For a specific $\lambda \in \{-k, -k+1, \ldots, k-1, k\}$, we can use $(\star)$ to find an approximation $\widetilde{f_\lambda}$ of f_λ using a $O(\varepsilon^{-2} \log n \log(k\delta^{-1}))$-dimensional sketch of $\mathbf{z}$ such that $\Pr\left[\, | \, \widetilde{f_\lambda} - f(\lambda) \, | > \varepsilon f(\lambda) \right] \,] < \frac{\delta}{2k+1}$. Iterating the process $2k+1$ times and using the union bound for probabilities, we get

$$\begin{aligned} \Pr\left[\forall \lambda \in \{-k, \ldots, k\} : |\widetilde{f_\lambda} - f(\lambda)| \leq \varepsilon f(\lambda)\right] &\geq 1 - \sum_{\lambda=-k}^{k} \Pr\left[\, | \, \widetilde{f_\lambda} - f(\lambda) \, | > \varepsilon f(\lambda) \right] \,] \\ &> 1 - (2k+1) \times \frac{\delta}{2k+1} = 1 - \delta \end{aligned}$$

It is possible to design a more careful approach that iterates only $O(\log k)$ times instead of $2k+1$ times. The ideas behind this approach as described above can be extended to trees with some non-trivial effort.

Finally the approach can indeed be generalized to the case when G is an arbitrary graph (which applies to computing $\textsc{Ric}(u, v)$) in the following manner. The basic idea to calculate $\textsc{Emd}_{\mathsf{d}}(\mathcal{A}, \mathcal{B})$ for an arbitrary graph G is to reduce it in an approximate sense to that of computing $\textsc{Emd}$ for a tree. Let $T = (V, E_T)$ be an arbitrary spanning tree of G, and let $F = E \setminus E_T$. The tree T defines a natural tree metric d' where $\mathsf{d}'(a, b)$ is the length of the shortest path between a and b in T for all $a, b \in V$. One can then express $\textsc{Emd}_{\mathsf{d}}(\mathcal{A}, \mathcal{B})$ in terms of $\textsc{Emd}_{\mathsf{d}'}(\mathcal{A}', \mathcal{B}')$ for some $\mathcal{A}' \supseteq \mathcal{A}$ and $\mathcal{B}' \supseteq \mathcal{B}$ in the following manner. For $f = (u, v) \in F$ and $\lambda_f \in \{-k, -k+1, \ldots, k-1, k\}$, let $\mathcal{C}^f_{\lambda_f}$ be the multi-set consisting of λ_f copies "u" if $\lambda_f > 0$ and $|\lambda_f|$ copies of "v" if $\lambda_f < 0$. Then the following bound holds:

$$\textsc{Emd}_{\mathsf{d}}(\mathcal{A}, \mathcal{B}) \leq \sum_{f \in F} |\lambda_f| + \textsc{Emd}_{\mathsf{d}'}\left(\mathcal{A} \uplus \sum_{f \in F} \mathcal{C}^f_{\lambda_f}, \, \mathcal{B} \uplus \sum_{f \in F} \mathcal{C}^f_{-\lambda_f}\right)$$

The above inequality leads to the following approach. Fix an arbitrary node $r \in V$ as the root of the spanning tree T, and let $\mathcal{P}_T(u, v)$ denote the set of edges in the unique path in T between nodes u and v. Define the two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{|E|}$ as follows (x_e and y_e denote the component of $\mathbf{x}$ and $\mathbf{y}$, respectively, indexed by the edge $e \in E$):

$$x_e = \begin{cases} |\{a \in \mathcal{A} \mid e \in \mathcal{P}_T(a, r)\}|, & \text{if } e \in E_T \\ 0, & \text{otherwise} \end{cases} \qquad y_e = \begin{cases} |\{b \in \mathcal{B} \mid e \in \mathcal{P}_T(b, r)\}|, & \text{if } e \in E_T \\ 0, & \text{otherwise} \end{cases}$$

and let $\mathbf{z} = \mathbf{x} - \mathbf{y}$. For each $f = (u, v) \in F$, define a vector $\mathbf{c}^f \in \mathbb{R}^{|E|}$ where the component $\mathbf{c}^f_e$ of $\mathbf{c}^f$ indexed by the edge $e \in E$ is given by:

$$\mathbf{c}_e^f = \begin{cases} 1, & \text{if } e \in \mathcal{P}_T(u,r) \setminus \mathcal{P}_T(v,r) \\ -1, & \text{if } e \in \mathcal{P}_T(v,r) \setminus \mathcal{P}_T(u,r) \\ 1, & \text{if } e = f \\ 0, & \text{otherwise} \end{cases}$$

This leads to the following optimization problem:

$$\textsc{Emd}_{\mathbf{d}}(\mathcal{A},\mathcal{B}) = \min_{\forall f \in F: \lambda_f \in \{-k,-k+1,\ldots,k-1,k\}} \| \mathbf{z} + \sum_{f \in F} \lambda_f \mathbf{c}^f \|_1$$

The above optimization problem can be solved using several approaches, e.g., using a recursive regression algorithm that exploits the convexity of f or using some recent results on robust regression via sub-space embeddings [39,40].

5. Discussion

In this paper we have reviewed some computational aspects of the Ollivier–Ricci curvature for networks, and shown a few simple computational bounds. As already mentioned in Section 1, there are other notions of network curvature that is also used by researchers and therefore this review should not be viewed as championing the Ollivier–Ricci curvature over other curvatures. We hope that this review will motivate further research on the exciting interplay between notions of curvatures from network and non-network domains. Some applications of network curvatures for real-world networks appear in references such as [11,13,15,16,18].

We conclude our article by mentioning an interesting application of the Ollivier–Ricci curvature for Markov chains for graph coloring and other problems (recise technical descriptions of these results are beyond the scope of this introductory review). The probability distributions on nodes used to compute EMD in the Ollivier–Ricci curvature can be naturally associated with a Markov process on the given graph (as a very simplified illustration, one can use a "normalized version" of $\textsc{Emd}(V_{u,v}, \mathbb{P}_u, \mathbb{P}_v)$ as the probability of transition between the states corresponding to nodes u and v). Such associations have a long history in the Markov chain literature under various names such as path coupling [41] and the values of $\textsc{Ric}(u,v)$'s have been used (explicitly or implicitly) to prove useful properties of the Markov chain, such as fast convergence to its stationary distribution, in many settings such as graph colouring [41] and sampling of paths with constraints [42].

Author Contributions: The author contributions are as follows: Conceptualization, N.A., P.S. and B.D.; methodology, N.A., P.S. and B.D.; software, N.A., P.S. and B.D.; validation, N.A., P.S. and B.D.; formal analysis, N.A., P.S. and B.D.; investigation, N.A., P.S. and B.D.; resources, N.A., P.S. and B.D.; data curation, N.A., P.S. and B.D.; writing–original draft preparation, N.A., P.S. and B.D.; writing–review and editing, N.A., P.S. and B.D.; visualization, N.A., P.S. and B.D.; supervision, B.D.; project administration, B.D.; funding acquisition, B.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NSF grant number IIS-1814931.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EMD	Earth Mover's Distance
RIC	Ricci curvature

References

1. Berger, M. *A Panoramic View of Riemannian Geometry*; Springer: Berlin/Heidelberg, Germany, 2012.
2. Bridson, M.R.; Haefliger, A. *Metric Spaces of Non-Positive Curvature*; Springer: Berlin/Heidelberg, Germany, 1999.
3. Schutz, B.F. *A First Course in General Relativity*; Cambridge University Press: Cambridge, UK, 1990.
4. Gromov, M. Hyperbolic groups. *Essays Group Theory* **1987**, *8*, 75–263.
5. Benjamini, I. Expanders are not hyperbolic. *Isr. J. Math.* **1998**, *108*, 33–36. [CrossRef]
6. Chalopin, J.; Chepoi, V.; Dragan, F.F.; Ducoffe, G.; Mohammed, A.; Vaxès, Y. Fast approximation and exact computation of negative curvature parameters of graphs. In Proceedings of the 34th International Symposium on Computational Geometry, Budapest, Hungary, 11–14 June 2018.
7. Chepoi, V.; Dragan, F.F.; Estellon, B.; Habib, M.; Vaxès, Y. Diameters, centers, and approximating trees of δ-hyperbolic geodesic spaces and graphs. In Proceedings of the 24th Annual Symposium on Computational Geometry, College Park, MD, USA, 9–11 June 2008; pp. 59–68.
8. DasGupta, B.; Janardhanan, M.V.; Yahyanejad, F. How did the shape of your network change? (On detecting network anomalies via non-local curvatures). *Algorithmica* **2020**, *82*, 1741–1783. [CrossRef]
9. DasGupta, B.; Karpinski, M.; Mobasheri, N.; Yahyanejad, F. Effect of Gromov-hyperbolicity Parameter on Cuts and Expansions in Graphs and Some Algorithmic Implications. *Algorithmica* **2018**, *80*, 772–800. [CrossRef]
10. Fournier, H.; Ismail, A.; Vigneron, A. Computing the Gromov hyperbolicity of a discrete metric space. *Inf. Process. Lett.* **2015**, *115*, 576–579. [CrossRef]
11. Albert, R.; DasGupta, B.; Mobasheri, N. Topological implications of negative curvature for biological and social networks. *Phys. Rev. E* **2014**, *89*, 032811. [CrossRef] [PubMed]
12. Jonckheere, E.; Lou, M.; Bonahon, F.; Baryshnikov, Y. Euclidean versus hyperbolic congestion in idealized versus experimental networks. *Internet Math.* **2011**, *7*, 1–27. [CrossRef]
13. Papadopoulos, F.; Krioukov, D.; Boguna, M.; Vahdat, A. Greedy Forwarding in Dynamic Scale-Free Networks Embedded in Hyperbolic Metric Spaces. In Proceedings of the IEEE Conference on Computer Communications, San Diego, CA, USA, 15–19 March 2010; pp. 1–9.
14. Forman, R. Bochner's method for cell complexes and combinatorial ricci curvature. *Discret. Comput. Geom.* **2003**, *29*, 323–374. [CrossRef]
15. Samal, A.; Sreejith, R.P.; Gu, J.; Liu, S.; Saucan, E.; Jost, J. Comparative analysis of two discretizations of Ricci curvature for complex networks. *Sci. Rep.* **2018**, *8*, 8650. [CrossRef] [PubMed]
16. Sreejith, R.P.; Jost, J.; Saucan, E.; Samal, A. Systematic evaluation of a new combinatorial curvature for complex networks. *Chaos Solitons Fractals* **2017**, *101*, 50–67. [CrossRef]
17. Sreejith, R.P.; Mohanraj, K.; Jost, J.; Saucan, E.; Samal, A. Forman curvature for complex networks. *J. Stat. Mech. Theory Exp.* **2016**, *2016*, 063206. [CrossRef]
18. Weber, M.; Saucan, E.; Jost, J. Characterizing complex networks with Forman-Ricci curvature and associated geometric flows. *J. Complex Netw.* **2017**, *5*, 527–550. [CrossRef]
19. Ollivier, Y. Ricci curvature of metric spaces. *Comptes Rendus Math.* **2007**, *345*, 643–646. [CrossRef]
20. Ollivier, Y. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.* **2009**, *256*, 810–864. [CrossRef]
21. Ollivier, Y. A survey of Ricci curvature for metric spaces and Markov chains. In *Probabilistic Approach to Geometry*; Kotani, M., Hino, M., Kumagai, T., Eds.; Mathematical Society of Japan: Tokyo, Japan, 2010; pp. 343–381.
22. Ollivier, Y. A visual introduction to Riemannian curvatures and some discrete generalizations. In *Analysis and Geometry of Metric Measure Spaces*; Lecture Notes of the 50th Séminaire de Mathématiques Supérieures, Montréal, 2011; Dafni, G., McCann, R.J., Stancu, A., Eds.; American Mathematical Society: Providence, RI, USA, 2013; pp. 197–219.
23. Chow, B.; Luo, F. Combinatorial Ricci flows on surfaces. *J. Differ. Geom.* **2003**, *63*, 97–129. [CrossRef]
24. Mallows, L. A note on asymptotic joint normality. *Ann. Math. Stat.* **1972**, *43*, 508–515. [CrossRef]
25. Rubner, Y.; Tomasi, C.; Guibas, L.J. A metric for distributions with applications to image databases. In Proceedings of the 6th International Conference on Computer Vision, Bombay, India, 4–7 January 1998; pp. 59–66.

26. Rubner, Y.; Tomasi, C.; Guibas, L.J. The earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vis.* **2000**, *40*, 99–121. [CrossRef]
27. Villani, C. Topics in optimal transportation. In *Graduate Studies in Mathematics*; American Mathematical Society: Providence, RI, USA, 2003; p. 58
28. Seidel, R. On the All-Pairs-Shortest-Path Problem in Unweighted Undirected Graphs. *J. Comput. Syst. Sci.* **1995**, *51*, 400–403. [CrossRef]
29. Williams, V.V. Multiplying matrices faster than Coppersmith-Winograd. In Proceedings of the 44th ACM Symposium on Theory of Computing, New York, NY, USA, 20–22 May 2012; pp. 887–898.
30. Ahuja, R.K.; Magnanti, T.L.; Orlin, J.B. *Network Flows: Theory, Algorithms, and Applications*; Prentice-Hall, Inc.: Upper Saddle River, NJ, USA, 1993.
31. Cormode, G.; Garofalakis, M.N.; Haas, P.J.; Jermaine, C. Synopses for massive data: Samples, histogram, wavelets, sketches. *Found. Trends Databases* **2012**, *4*, 1–294. [CrossRef]
32. Andoni, A.; Ba, K.D.; Indyk, P.; Woodruff, D.P. Efficient sketches for earth-mover distance, with applications. In Proceedings of the 50 thAnnual IEEE Symposium on Foundations of Computer Science, Atlanta, GA, USA, 25–27 October 2009; pp. 324–330.
33. Brody, J.; Liang, H.; Sun, X. Space-efficient approximation scheme for circular earth mover distance. In *LATIN 2012*; Fernández-Baca, D., Ed.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7256; pp. 97–108
34. Indyk, P. A near linear time constant factor approximation for euclidean bichromatic matching (cost). In Proceedings of the 18th annual ACM-SIAM symposium on Discrete algorithms, New Orleans, LA, USA, 7–9 January 2007; pp. 39–42.
35. Indyk, P.; Price, E. K-median clustering, model-based compressive sensing, and sparse recovery for earth mover distance. In Proceedings of the 43rd Annual ACM Symposium on Theory of Computing, San Jose, CA, USA, 6–8 June 2011; pp. 627–636.
36. Verbin, E.; Zhang, Q. Rademacher-sketch: A dimensionality-reducing embedding for sum-product norms, with an application to earth-mover distance. In Proceedings of the International Colloquium on Automata, Languages, and Programming, Warwick, UK, 9–13 July 2012; pp. 834–845.
37. McGregor, A.; Stubbs, D. Sketching earth-mover distance on graph metrics. In *International Workshop on Approximation Algorithms for Combinatorial Optimization*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 274–286.
38. Kane, D.M.; Nelson, J.; Porat, E.; Woodruff, D.P. Fast moment estimation in data streams in optimal space. In Proceedings of the 43rd annual ACM symposium on Theory of computing, San Jose, CA, USA, 6–8 June 2011; pp. 745–754.
39. Clarkson, K.L.; Drineas, P.; Magdon-Ismail, M.; Mahoney, M.W.; Meng, X.; Woodruff, D.P. The fast cauchy transform and faster robust linear regression. In Proceedings of the 24th annual ACM-SIAM Symposium on Discrete algorithms, New Orleans, LA, USA, 6–8 January 2013; pp. 466–477.
40. Kane, D.M.; Nelson, J.; Porat, E.; Woodruff, D.P. Subspace embeddings for the ℓ_1-norm with applications. In Proceedings of the 43rd Annual ACM Symposium on Theory of Computing, San Jose, CA, USA, 6–8 June 2011; pp. 755–764.
41. Bubley, R.; Dyer, M.E. Path coupling: A technique for proving rapid mixing in Markov chains. In Proceedings of the 38th Annual Symposium on Foundations of Computer Science, Miami Beach, FL, USA, 20–22 October 1997; pp. 223–231.
42. Gerin, L. Random sampling of lattice paths with constraints, via transportation. In Proceedings of the 21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms, Vienna, Austria, 28 June–2 July 2010; pp. 317–328.