Bakry-Émery Ricci Curvature on Weighted Graphs with Applications to Biological Networks

Maryam Pouryahya, Rena Elkin, Romeil Sandhu
Sarah Tannenbaum, Tryphon Georgiou, Allen Tannenbaum

ABSTRACT

In recent years, there have been tremendous efforts to elucidate the complex mechanisms of cancer networks by investigating the interactions of different genetic and epigenetic factors. Mathematical tools can help significantly with overcoming many of these challenges and can facilitate better understanding of the complexities of the corresponding networks. This has lead to the emergence of the field of network and systems biology. The formal models employed in biological networks range from graphs as abstract representations of pairwise interactions to complicated systems of partial differential equations that try to capture all details of biological interactions. Therefore, the mathematical methods and tools employed in networks are quite diverse and heterogeneous. We propose an integrative framework to identify genetic features related to cancer networks and to distinguish them from the normal tissue networks by geometrical analysis of the networks provided by The Cancer Genome Atlas (TCGA) data. Our study is based on the analogous notion of fundamental concepts in Riemannian geometry, namely Ricci curvature, on discrete spaces.

I. INTRODUCTION

This paper describes a number of facts about graph curvature and its relation to functional network robustness with applications to studying mechanisms of robustness in cancer. Essentially, it gives all the background material for understanding [33] while proposing another notion of graph curvature valid for graphs with negative weights [23], [42].

Cancer cells exhibit extensive mutational heterogeneity, and subsequent elaborate protein interactions complicate the discovery and understanding of involved genes and pathways. In recent years, there have been tremendous efforts to elucidate the complex mechanisms of cancer networks by investigating the interactions of different genetic and epigenetic factors. Mathematical tools can help significantly with overcoming many of the these challenges and can facilitate a better understanding of the complexities of their networks. This has led to the emergence of the field of network and systems biology. The formal models employed in biological networks range from graphs as abstract representations of pairwise interactions to complicated systems of partial differential equations that try to capture all details of biological interactions. Therefore, the mathematical methods and tools employed in networks are quite diverse and heterogeneous. Furthermore, recent decades have witnessed the thriving development of new mathematical, computational and theoretical approaches to tackle fundamental issues in network analysis. We have proposed an integrative framework to identify genetic features related to cancer networks and to distinguish them from the normal tissue networks by geometrical analysis of the networks provided by The Cancer Genome Atlas (TCGA) data.

Our study is based on the analogous notion of fundamental concepts in Riemannian geometry, namely Ricci curvature, on discrete spaces. The concept of curvature was initially introduced to express the deviation of a geometric object from being flat. The Riemann curvature tensor of such a manifold encodes its geometric invariants and expresses the deviation from Euclidean, that is, flat space. In Riemannian geometry, sectional curvature is defined on two-dimensional tangent planes and it expresses the convexity property of the distance function between geodesics. The essential notion of Ricci curvature is the average of sectional curvatures of all tangent planes with some given direction. One important aspect of Ricci curvature is that it can control the eigenvalues of the Laplace-Beltrami operator [13].

Recently, several different notions of curvature have been proposed applicable to more general metric spaces, and various notions of generalized sectional curvature inequalities were proposed. The one that we are interested in here is the Γ_2-calculus of Bakry-Émery. In their paper published in 1985 [3], Bakry and Émery suggested a notion analogous to curvature in the very general framework of a Markov semigroup. The condition was based on the Bochner inequality with curvature-dimension condition of (K, N), denoted by $CD(K, N)$. These attempts lead us to define the notions of curvature on discrete structures like graphs. In fact, graphs are ideally suited for explaining the meaning of curvature and the appeal of this discrete feature is that it is fairly straightforward to be computed

M. Pouryahya, R. Elkin, R. Sandhu, and A. Tannenbaum are with the Departments of Computer Science and Applied Mathematics/Statistics, Stony Brook University, Stony Brook, NY 11794. email: arobertan@gmail.com

S. Tannenbaum is with the Department of Pediatric Cancer, Columbia Medical School, NY. email:sarah.tannenbaum@gmail.edu

T.T. Georgiou is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota MN 55455, USA; email: tryphon@umn.edu

The research was partially supported by the AFOSR under Grants FA9550-12-1-0319 and FA9550-15-1-0045, the National Center for Research Resources (P41-RR-013218), the National Institute of Biomedical Imaging and Bioengineering (P41-EB-015902) of the National Institutes of Health, and by NIH grant 1U24CA18092401A1.

on graphs in many cases. There have been several efforts to define the notion of Ricci curvature on discrete spaces [25] [34] [28] [15] . Previously, among these definitions of generalized Ricci curvature, we applied the one of Ollivier [28] on graphs, [33]. We showed that the Olivier-Ricci curvature can be useful to differentiate the cancer networks from normal ones. In this paper, we apply a notion of curvature based on Bakry-Émery theory [3] to weighted graphs following the approach of [23]. Also, of note is that this discrete version of Ricci curvature does not limit us to graphs with only positive weights, which was the case with our previous work in [33]. Here, we can define the curvature on graphs of positive/negative weights as long as we have negative semi-definite Laplacian. The criteria that guaranteed the negative semi-definiteness of graph Laplacian is provided in a work by [42] which we briefly review in section (V).

The other subject of interest is the theory of optimal mass transport and its relationship with Riemannian geometry, in particular Ricci curvature. In the Riemannian context, the L_2-Wasserstein distance is well-adapted to the study of Ricci curvature. We study an alternative definition of Wasserstein suggested by Benamou and Brenier [5] on the space of density probability. This leads us to discover the geodesics in the Riemannian probability spaces as the L_2-Wasserstein distance.

Finally, considering this Riemannian structure, via the Bochner formula we can show that there is a positive correlation between changing of the entropy and the Ricci curvature. In conjunction with the Fluctuation Theorem [12], we conclude that increasing the Ricci curvature is positively correlated with increasing the robustness, herein expressed as $\Delta \text{Ric} \times \Delta R \geq 0$. An appealing feature of this correlation is that we can quantify the difference between normal and cancer networks by a fairly straightforward computation of the Bakry-Émery-Ricci curvature on the networks.

II. Wasserstein Distance and Riemannian Geometry

Before describing the results on graphs, we will review the necessary literature on manifolds. Accordingly, let M be a Riemmannian manifold equipped with volume element dvol(M).

A. The Monge-Kantrovich Problem

The first optimal transport problem was proposed by Monge in the 1780's, a civil engineering problem which asks for the minimal transportation cost to move a pile of soil ("déblais") to an excavation ("remblais"). The places where the soil should be extracted, and the ones which should be filled, are all known. A more modern form of the problem introduced by Kantorovich in 1940's yields the so-called Monge-Kantrovich problem (MKP). The framework of the problem is as follows.

Let (X, μ) and (Y, ν) be two probability spaces and $\Pi(\mu, \nu)$ denotes the set of all joint probability measures on $X \times Y$ whose marginals are μ and ν. The optimal transport cost suggested by Kantorovich was the following linear programming problem:

$$\text{minimize} \int c(x,y)\, d\pi(x,y) \quad \text{over all } \pi \in \Pi(\mu,\nu).$$

where $c(x, y)$ is the cost for transporting one unit of mass from x to y. The cost function was originally defined in a distance form on a metric space (X, d). This leads us to the following distance function know as L^p Wasserstein:

$$W_p(\mu,\nu) = \Big(\inf_{\pi \in \Pi(\mu,\nu)} \int_{X\times X} d(x,y)^p d\pi(x,y) \Big)^{1/p},$$

Two exponents of $p = 1$ and $p = 2$ are particularly interesting. The L^1-Wasserstein distance is also known as the Kantorovich-Rubinstein distance, or Earth Mover's distance among computer scientists. In [29], Ollivier used the L^1-Wasserstein distance to define the (Ollivier-)Ricci curvature. We used this curvature in [33] to differentiate the cancer networks from normal tissue networks. In this paper we focus on the case $p = 2$. The L^2-Wasserstein distance has some very remarkable properties and can be related to fluid mechanics; the research is based on the work of Yann Brenier which was published in [5]. In this paper, Benamou and Brenier suggested the following alternative numerical method to calculate the L^2-Wasserstein distance:

$$W_2(\mu_0,\mu_1) = \Big(\inf_{\mu,v} \int \int_0^1 \mu(t,x)|v(x,t)|^2 dx dt \Big)^{1/2}, \quad (1)$$

subject to the continuity equation (conservation of mass formula):

$$\frac{\partial \mu}{\partial t} + \nabla \cdot (\mu v) = 0 \quad (2)$$

with the initial and final conditions:

$$\mu(0,.) = \mu_0, \mu(1,.) = \mu_1. \quad (3)$$

Here $\mu(t, x)$ is the density of a system of particles and $v(x, t)$ is the velocity field at time t and position x; $\nabla\cdot$ stands for the divergence operator.
The functional in (1) is the kinetic energy and it turns out that the optimal conditions satisfy:

$$v(t,x) = \nabla g(t,x),$$

where g is the Lagrange multiplier of constraints (2) and (3). Also, the Hamilton-Jacobi equation holds for this optimal g:

$$\frac{\partial g}{\partial t} + \frac{1}{2}||\nabla g||^2 = 0. \quad (4)$$

B. Riemannian Structure on Probability Densities

The Wasserstein distance defines a natural Riemannian structure on the space of probability measures, and this structure is essential in explicating the relationship of entropy and curvature [25] that we will sketch in Section VII-A.

Define

$$\mathcal{P} := \{\mu \geq 0 : \int \mu \,\text{dvol(M)} = 1\},$$

the space of probability densities. The tangent space at a given point μ may be identified with

$$T_\mu \mathcal{P} \cong \{u : \int u \,\text{dvol(M)} = 0\}.$$

Thus inspired by the Benamou and Brenier [5], given two "points", $\mu_0, \mu_1 \in \mathcal{P}$, the geodesic (Wasserstein) distance is:

$$\inf_{\mu,g}\{\int\int_0^1 \mu(t,x)\|\nabla g(t,x)\|^2\, dt\, \text{dvol(M)} \quad \text{subject to} \quad \frac{\partial \mu}{\partial t} + \text{div}(\mu \nabla g) = 0, \quad \mu(0,\cdot) = \mu_0,\ \mu(1,\cdot) = \mu_1\} \tag{5}$$

In other words, we look at all curves in $\mathcal{P}$ connecting μ_0 and μ_1, and take the shortest one with respect to the Wasserstein metric.

This leads us to give $\mathcal{P}$ a Riemannian structure, which will induce this Wasserstein distance. This idea is due to Jordan *et al.* [21], [38]. Namely, under suitable assumptions on differentiability for $\mu \in \mathcal{P}$, and $u \in T_\mu \mathcal{P}$, one solves the Poisson equation

$$u = -\text{div}(\mu \nabla g). \tag{6}$$

This allows us to identify the tangent space with functions up to an additive constant. Thus, for any given u we denote the solution of (6) by g_u. Then given, $u_1, u_2 \in T_\mu \mathcal{D}$, we can define the inner product

$$\langle u_1, u_2 \rangle_{W_2} := \int \nabla g_{u_1} \cdot \nabla g_{u_2} \mu \, \text{dvol(M)}, \tag{7}$$

which imposes a Riemannian metric on the probability measure space $P(X)$. Using integration by parts, we can see this inner product will induce exactly the Wasserstein distance defined by (5). We also have

$$\begin{aligned} \langle u, u \rangle_{W_2} &= \int \mu\, \nabla g_u \cdot \nabla g_u \, \text{dvol(M)} \\ &= -\int g_u \text{div}(\mu \nabla g_u)\, \text{dvol(M)} \\ &= \int u\, g_u \, \text{dvol(M)}. \end{aligned} \tag{8}$$

We note finally that the optimal g in the above optimization problem satisfies the Hamilton-Jacobi equation:

$$\frac{\partial g}{\partial t} + \frac{1}{2}\|\nabla g\|^2 = 0. \tag{9}$$

We also have the continuity equation:

$$\frac{\partial \mu}{\partial t} + \text{div}(\mu \nabla g) = 0, \tag{10}$$

with boundary conditions

$$\mu(0,\cdot) = \mu_0,\ \mu(1,\cdot) = \mu_1.$$

Thus with these conditions, $\mu_t(x) = \mu(t,x)$ defines a geodesic in the Wasserstein sense on $\mathcal{P}(M)$. Further, the tangent vector along the geodesic μ_t at μ_0 is exactly

$$u := \frac{\partial \mu}{\partial t}|_{t=0}. \tag{11}$$

III. Bakry-Émery and Bochner

We now review the Bochner formula [38] and Bakry-Émery curvature [3]. We begin with the Bochner formula. Since this is local, it is sufficient to review the notions in Euclidean space. One of the ideas of Bakry-Émery [3] is to formulate the Bochner formula through the Γ calculus that we sketch as well now.

Let Δ denote the Laplacian. Following [3], we define

$$\Gamma(f,g) := 1/2[\Delta(fg) - f\Delta g - g\Delta f].$$

Therefore,

$$\Gamma(f,f) = 1/2[\Delta f^2 - 2f\Delta f].$$

It is easy to compute that

$$\Delta f^2 = 2[f\Delta f + \|\nabla f\|^2].$$

Therefore, $\Gamma(f,f) = \|\nabla f\|^2$. Again following Bakry-Émery [3], we iterate the above operation as follows:

$$\Gamma_2(f,g) := 1/2[\Delta(\Gamma(f,g)) - \Gamma(f,\Delta g) - \Gamma(g,\Delta f)],$$

and a simple calculation shows that

$$\Gamma_2(f,f) = -\nabla f \cdot \nabla \Delta f + 1/2 \Delta \|\nabla f\|^2.$$

Recall the following argument in the continuous case on an n-dimensional manifold via the classical Bochner formula:

$$-\nabla f \cdot \nabla \Delta f + \frac{1}{2}\Delta \|\nabla f\|^2 = \|\nabla^2 f\|^2 + Ric(\nabla f),$$

where ∇^2 denotes the Hessian operator. Now applying the *Cauchy-Schwartz inequality* implies that

$$-\nabla f \cdot \nabla \Delta f + \frac{1}{2}\Delta \|\nabla f\|^2 \geq (\Delta f)^2/N + k\|\nabla f\|^2$$

if $n \leq N$ and $k \geq Ric$. This is the $C(k,N)$ criterion of Bakry-Emery. Taking $N = \infty$, we get

$$-\nabla f \cdot \nabla \Delta f + \frac{1}{2}\Delta \|\nabla f\|^2 \geq k\|\nabla f\|^2. \tag{12}$$

This motivates (15) to be given below.

IV. Laplacian of Graphs

Let $G = (V,E)$ be an undirected graph with positive and possibly negative weights. In order to define curvature in the Bakry-Émery sense [3], one needs to ensure the the associated combinatorial graph Laplacian $\mathcal{L}_G$ is non-negative definite. Necessary and sufficient conditions for this have been given in [42]. We briefly review the necessary theory in the present section.

Let $G = (V,E)$ be a weighted graph with n vertices (nodes). We always assume that the graph is connected. We set

$$\begin{aligned} d_x &= \sum_y w_{xy} \\ \mu_x(y) &:= \frac{w_{xy}}{d_x}, \end{aligned}$$

the sum taken over all neighbors of x where w_{xy} denotes the weight of an edge connecting x and y (it is taken as zero

if there is no connecting edge between x and y). Note that since the graph is undirected we have that $w_{xy} = w_{yx}$. We let $W = (w_{xy})_{1\leq x,y\leq n}$ be the matrix of weights, and D be the diagonal matrix with entries d_x. Then the (combinatorial) Laplacian matrix is defined as

$$\mathcal{L}_G := -(D - W).$$

If all the weights are positive, then via the Gershgorin Circle Theorem, $\mathcal{L}_G$ is negative semi-definite. In the more general case under consideration here, that under the assumption that G has no cycles with negative weights, the necessary and sufficient condition is the absolute values of the weights are bounded by the entries of the resistivity matrix on the positive part of the graph.

V. Positive Semi-definiteness of Graph Laplacians

Here, we review the necessary and sufficient conditions for the combinatorial Laplacian matrix $\mathcal{L}_G$ to be positive semi-definite. As we discussed before, this is equivalent to the negative semi-definiteness of our Laplacian matrix $\mathcal{L}_G$. From electric circuit analysis, the graph G can be interpreted as an electrical network where each edge represents a resistor with resistance equal to the reciprocal of its edge weight. The effective resistance across any two nodes $x, y \in V$ is given by

$$\mathcal{R}_{xy}(G) = \mathbf{e}_{xy}^T \mathcal{L}^\dagger \mathbf{e}_{xy}$$

where $\mathcal{L}^\dagger$ is the Moore-Penrose pseudo inverse of $\mathcal{L}$ and $\mathbf{e}_{xy} \in \mathbb{R}^{|E|}$ denotes the column vector with a "1" in the x-th position, a "-1" in the y-th position and zeros everywhere else [24].

When G has negative edge weights, both the magnitude of these weights and the respective location of their associated edges affect the positive semi-definiteness of the Laplacian matrix. For simplicity, we follow the notation of [42] and separate the positive and negative parts of the graph into $G_+ = (V, E_+)$ and $G_- = (V, E_-)$ where E_+ and E_- denote the set of edges with positive and negative weights, respectively. Clearly, $G = G_+ \cup G_-$.

Now we assume that G_+ is connected and require that no cycle in G contains two distinct edges with negative weights. Under this cyclic condition, we have the following theorem:

Theorem: Given a connected graph $G = (V, E)$ satisfying the conditions above, the combinatorial Laplacian, $\mathcal{L}_G = D - W$, is positive semidefinite if and only if the absolute value of each negative weight is bounded by its effective resistance over the positive part of the graph, i.e. $|w_{xy}| \leq \mathcal{R}_{xy}^{-1}(G_+)$ for all $(x, y) \in E_-$.

Two proofs to the above theorem are provided in [9]. This theorem helps us to identify the networks with negative semi-definite Laplacians, $\mathcal{L}_G$, for which we will define the discrete Ricci curvature.

VI. Bakry-Émery Ricci Curvature of Weighted Graphs

We will assume that all of the graphs under consideration have negative semidefinite Laplacians. Using the intuition of [3], the authors of [23] give a notion of curvature on graphs that is easily computable. Indeed, following [23] and using the Γ calculus of [3], we define the following bilinear operators for functions $f, g : V \to \mathbb{R}$ and $x \in V$:

$$\begin{aligned} \Gamma(f,g)(x) &:= 1/2[\mathcal{L}_G(f \cdot g)(x) \qquad (13)\\ &- f(x)\mathcal{L}_G g(x) - g(x)\mathcal{L}_G f(x)], \\ \Gamma_2(f,g)(x) &:= 1/2[\mathcal{L}_G(\Gamma(f,g)(x)) \qquad (14)\\ &- \Gamma(f, \mathcal{L}_G g)(x) - \Gamma(g, \mathcal{L}_G f)(x)]. \end{aligned}$$

Γ_2 is the Ricci curvature form on G.

Note from the former expression that

$$\mathcal{L}_G(f \cdot g)(x) = f(x)\mathcal{L}_G g(x) + 2\Gamma(f,g)(x) + g(x)\mathcal{L}_G f(x)].$$

Set $\Gamma(f) := \Gamma(f, f)$. Then following the logic of the Bochner formula [23], we define the *local graph Bakry-Émery-Ricci curvature* as the maximum value of $k(x)$ such that

$$\Gamma_2(f)(x) \geq k(x)\Gamma(f)(x), \ \forall f, \qquad (15)$$

where $x \in V$.

The preceding discussion only makes sense if Γ is non-negative definite at x. For graphs with negative weights, one must take a more global approach since at nodes with negative weights, this will not hold. More precisely, given $f : V \to \mathbb{R}$, we have that $\Gamma(f) : V \to \mathbb{R}$, given by

$$\Gamma(f)(x) = \sum_{y\sim x} w_{xy}(f(x) - f(y))^2.$$

Note that as is standard,

$$f'\mathcal{L}_G f = -1/2 \sum_{xy} w_{xy}(f(x) - f(y))^2,$$

where the sum is taken over all edges xy. These formulas are correct for any combinatorial Laplacian for any undirected graph with negative or positive weights. Assuming that the Zelazo conditions [42] hold, $f'\mathcal{L}_G f$ is always non-negative. For given $x \in V$, we have defined the local graph Ricci curvature as the maximum value of $k(x)$ such that

$$\Gamma_2(f)(x) \geq k(x)\Gamma(f)(x), \ \forall f,$$

where $x \in V$. This will only be applicable if

$$\sum_{y\sim x} w_{xy}(f(x) - f(y))^2 \geq 0, \quad \forall f. \qquad (16)$$

Unfortunately, this inequality cannot be guaranteed if $w_{xy_o} < 0$ for some adjacent y_o. Indeed, we can take a function f such that $f(x) = f(y) = 0$ for $y \neq y_o$ and $f(y_o) \neq 0$.

Accordingly, we need to use a ***global*** notion of Ricci curvature taken as follows: we simply take the maximum k such that

$$\sum_{x\in V} \Gamma_2(f)(x) \geq k \sum_{x\in V} \Gamma(f)(x), \quad \forall f.$$

(This is really summing over all edges modulo a factor of 1/2 since each edge is counted twice in the summation.) We will call this k the *global graph Bakry-Émery-Ricci (BER) curvature*. BER will always refer to the global notion that is well-defined whenever $\mathcal{L}_G$ is non-negative definite.

VII. Curvature, Entropy, and Robustness

In this section, we draw an interesting relationship between the BER Curvature defined in section (VI), Boltzman entropy and robustness. The idea is similar to the one we have used in [33]. Here, robustness is defined as the ability of a system to functionally adapt to changes in the environment. First, we implement the classical Bochner formula once again to establish a relationship between entropy and curvature. Then, in section (VII-B) we review the Fluctuation Theorem [12] which leads us to the correlation between entropy and robustness. Finally, we conclude that there is a positive correlation between entropy and curvature; this can help us to distinguish the more robust networks, namely, the cancer networks from the normal networks.

A. Entropy and Bochner

As noted by Lott-Villani [25], curvature and entropy are very closely related using the Riemannian structure described in Section II-B. We sketch this relationship now.

Define

$$H(\mu_t) := \int_M \mu_t \log \mu_t \,\text{dvol(M)}, \tag{17}$$

which is the negative of the classical *Boltzmann-Shannon entropy* functional $S(\mu_t) := -H(\mu_t)$.

We consider the Hessian operator $\nabla^2 H(\mu) : T_\mu P \to T_\mu P$ on the Riemannian manifold $P(M)$. We compute first the second derivative of H along along a geodesic path in $P(M)$, namely: we have that

$$\begin{aligned}\frac{d^2}{dt^2} H(\mu_t)|_{t=0} &= \frac{d}{dt}\langle \nabla H(\mu_t), u\rangle_{W_2} \\ &= \langle \nabla_u \nabla H(\mu_t), u\rangle_{W_2} + \langle \nabla H(\mu_t), \nabla_u u\rangle_{W_2} \\ &= \langle \nabla^2 H(\mu_t)\cdot u, u\rangle_{W_2}.\end{aligned}$$

since $\nabla_u u = 0$ (∇_u is the covariant derivative in the direction u). Here $u := \frac{\partial \mu_t}{\partial t}|_{t=0}$.

Next straightforward computation gives

$$\frac{d^2}{dt^2} H(\mu_t) = -\int_M \nabla g \cdot \nabla \Delta g + \frac{1}{2}\Delta(\|\nabla g\|^2)\mu_t \,\text{dvol(M)} \tag{18}$$

where μ_t and $g = g_u$ satisfy equations (9) and (10) above (μ_t is the geodesic path). Now we invoke Bochner:

$$\begin{aligned}&\langle \nabla^2 H(\mu_t)\cdot u, u\rangle_{W_2} \qquad (19)\\ &= \int_M \|\nabla^2 g_u\|^2 + \text{Ric}(\nabla g_u, \nabla g_u)\mu_0 \,\text{dvol(M)}.\end{aligned}$$

Assume $Ric \geq kI$ as quadratic forms on the manifold M. Then from equation (19), $\nabla^2 H$ is k-convex with respect to $\langle\cdot\rangle_{W_2}$. Thus, we recover the fact from [25] that

$$H(\mu_t) \leq tH(\mu_0) + (1-t)H(\mu_1) - \frac{k}{2}t(1-t)W_2(\mu_0,\mu_1)^2.$$

Therefore, the above inequality indicates the *positive correlation* between entropy and curvature which we express as:

$$\Delta S \times \Delta \text{Ric} \geq 0 \tag{20}$$

B. Fluctuation Theorem

One can understand the Fluctuation Theorem [12] as follows. Recall that if $p_\epsilon(t)$ denotes the probability that the mean deviates by more than ϵ from the original (unperturbed) value at time t, then

$$R := \lim_{t\to\infty,\epsilon\to 0}(-\frac{1}{t}\log p_\epsilon(t)).$$

On the other hand, evolutionary entropy S may be characterized in this setting as

$$S := \lim_{t\to\infty,\epsilon\to 0}(\frac{1}{t}\log q_\epsilon(t)),$$

where $q_\epsilon(t)$ denotes the minimal number of genealogies of length t whose total probability exceeds $1-\epsilon$. Thus the greater the $q_\epsilon(t)$, the smaller the $p_\epsilon(t)$ and so the larger the decay rate. The Fluctuation Theorem is an expression of this fact for networks, and can be expressed as

$$\Delta S \times \Delta R \geq 0, \tag{21}$$

Considering (20) and (21), we conclude that changes in robustness (ΔR) are also positively correlated with the network curvature, i.e:

$$\Delta R \times \Delta \text{Ric} \geq 0. \tag{22}$$

According to the work done in [12] and [40], it seems that in many cases the normal protein interaction networks possess a lower entropy than their cancerous analogues; hence they are less robust. This could be justified as the ability of oncoproteins to better respond to the changes in the cellular environment due to their disorganized arrangement which leads to possession of higher degrees of freedom. Since the curvature is positively correlated to the robustness of networks and easier to compute, it can help in quantifying the robustness in terms of the adaptability of networks.

VIII. Results

As we discussed before, our interest is in calculating the BER curvature at global and local scales. In this section, we apply BER curvature to certain cancer networks to differentiate them from normal tissue networks. In particular, our results illustrate that cancer networks exhibits a greater degree of functional robustness compared to the normal tissue networks [12], [33], [40].

A. Description of Data Sets

We have studied seven transcription networks composed of cancer specific genes provided by Memorial Sloan Kettering Cancer Center. The data consists of correlation values of gene-to-gene expression in cancerous and normal tissues which were computed across all samples within a given phenotype. The network is constructed using these correlation values as weights of the graph and the adjacency matrix of the graph is given by the underlying biological gene-to-gene interactions. Our TCGA data includes approximately 500 cancer-related genes of seven different tumor types: breast invasive carcinoma [BRCA], head and neck squamous cell

carcinoma [HNSC], kidney renal papillary cell carcinoma [KIRP], liver hepatocellular carcinoma [LIHC], lung adenocarcinoma [LUAD], prostate adenocarcinoma [PRAD], and thyroid carcinoma [THCA].

B. Local BER Curvature of Transcription Networks

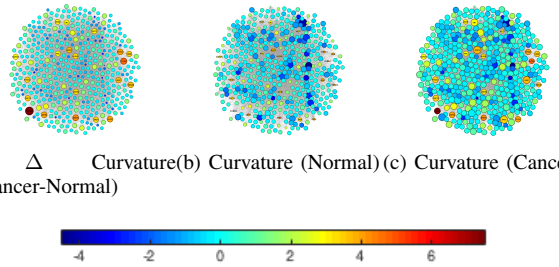
In Table I we present the difference (cancer-normal) in average local graph BER curvature of each TCGA data. As shown in the table, the average difference between cancer and normal tissue distribution is always positive. Therefore, the average local BER curvature in all seven cancer networks of our study possess a higher value than the normal corresponding tissue. Of note is that these results are consistent with our previous work in [33] where the average Ollivier-Ricci curvature was greater for the cancer networks compared to the normal one.

Cancer Type	**Δ Average Local BER Curvature**
Breast Carcinoma	0.1823
Head/Neck Carcinoma	0.1164
Kidney Carcinoma	0.2171
Liver Carcinoma	0.2272
Lung Adenocarcinoma	0.3202
Prostate Adenocarcinoma	0.1789
Thyroid Carcinoma	0.1328

TABLE I
LOCAL BER CURVATURES DEMONSTRATE A HIGHER ROBUSTNESS OF THE CANCER NETWORKS WITH RESPECT TO THE NORMAL TISSUE NETWORKS.

Additionally, since the local BER curvature measures the curvature at the nodal level, it can illuminate the genes with the most contribution to the robustness of the cancer network, as illustrated by Figure 1. To this end, we also provide the top twenty genes sorted with respect to the local BER curvature of breast carcinoma and lung adenocarcinoma in Table II and Table III, respectively.

1) Top twenty ranked genes in breast carcinoma: Taking a closer look at Table II, it is interesting to notice relationships among the top twenty genes and to consider their possible implications with respect to robustness of the breast cancer network. For instance, PIK3CA encodes the p110α protein, a catalytic subunit of the PI3K (phosphatidylinositol 3-kinase) enzyme and has been found to exhibit a 25% mutation rate in breast cancer thereby making it the most regularly mutated constituent [26]. AKT1 encodes AKT1 kinases which are involved in many signaling pathways, it is a downstream effector of PI3K, and is over-expressed in breast cancer as well [11]. Considering PIK3CA and AKT1 were both ranked among the top twenty genes seems suggestive of the PI3K/Akt signaling pathway's significance in regards to cancer robustness. In fact, the PI3K/Akt signaling pathway regulates many cell functions necessary for tumorigenesis and cancer survival such as cell growth, proliferation and apoptosis. We also notice that the MAP2K1 gene encodes for the protein kinase MEK1 which is involved in the MAPK/ERK pathway, also known as the MEK/ERK pathway. PTPN11 codes for SHP2 (Src-homology 2 domain-containing phosphatase), a known drug target in breast cancer and is required to fully activate the MAPK/ERK pathway [1]. This pathway is a signal transduction pathway involved in the promotion of cell survival, proliferation, and metastasis. Stimulation of growth-factor receptors such as EGFR, (also ranked among the top twenty genes), initiate downstream activation of the MAPK/ERK pathway. This pathway is also found to be excessively activated in many cancers including breast and lung cancer, among others, and is often caused by upstream activation from overly-expressed or aberrantly activated cell surface receptors (such as EGFR). [32] Intuitively, it makes sense that the cascading effects of mutations and abnormal activity through these pathways, which are imperative for cancer cell growth and survival, would be favorable in regards to the robustness of the network so it is not surprising that we find such functional relationships among the top ranked genes.



(a) Δ Curvature(Cancer-Normal) (b) Curvature (Normal) (c) Curvature (Cancer)

Fig. 1. Breast Carcinoma network visualization: In the above illustrations, each node represents a single gene. The color of each node remains the same in each of the figures 1(a), 1(b), 1(c) and is assigned by the difference in its local graph BER curvature (Cancer-Normal). The top twenty ranked genes are labelled and shown in red-orange colors. The only differences between the above illustrations are in the sizes of the nodes. In figure 1(a), the size of each node is scaled in the same way as its color, by the difference in its local graph BER curvature (Cancer-Normal). This is shown to highlight the top genes contributing to breast cancer network robustness. Figure 1(b) shows node sizes scaled by their local graph BER curvature in normal tissues and figure 1(c) shows node sizes scaled by their local graph BER curvature in cancerous tissues. Comparing figures 1(b) and 1(c), we see a dramatic change in the curvature of the top twenty ranked genes, where they go from demonstrating low curvature in normal tissues to demonstrating high curvature in cancerous tissues, relative to the other genes.

2) Top twenty ranked genes in lung adenocarcinoma: With regards to local BER curvature for lung cancer, we chose to highlight three genes from Table III, CDK4, NPM1, and Rac1. The CDK4 protein, encoded by the CDK4 gene, plays an important role in the cell cycle, particularly during the G1 phase. The role it plays as a significant cancer-related gene is becoming more and more clear, as its function is to drive cell-cycle progression by phosphorylating the retinoblastoma protein. Overexpression of CDK4 has been described in many tumors, including lung cancer [41]. NPM1 is an important

gene to consider with regards to human tumorigenesis and is frequently overexpressed in solid tumors of a diverse histological origin. Depending on its expression levels and gene dosage, NPM appears to function either as an oncogene or a tumour suppressor. Either partial functional loss or aberrant overexpression could lead to neoplastic transformation through distinct mechanisms. In fact, the loss of 5q35 (where NPM1 resides) is also observed in non-small-cell lung carcinoma including lung adenocarcinoma. Therefore, loss of the 5q chromosomal region could ensue genomic instability in these tumors, which could in turn favor tumor progression [19]. Rac1 is an important GTPase, encoded by the Rac1 gene that has been implicated in many cellular processes including transcriptional activation, cytoskeleton rearrangement, and cell adhesion. Furthermore, it is thought to play a role in cancer cell migration, invasion, and metastasis. Rac1 overexpression has been found in various cancers including lung adenocarcinoma. Interestingly, Rac1 overexpression also appears to be an independent predictor of adverse outcome of these carcinomas [7].

Gene Rank	Top 20 Genes	Δ Curvature (Cancer-Normal)
1	PICALM	7.4910
2	CLTCL1	4.9102
3	EPS15	4.3210
4	KIF5B	4.1284
5	CLTC	4.0657
6	PTPN11	4.0465
7	YWHAE	3.8416
8	EGFR	3.8357
9	JAK1	3.7590
10	MSN	3.6079
11	CDC73	3.5274
12	PIK3CA	3.4499
13	XPO1	3.4274
14	ALDH2	3.3854
15	SDHB	3.2626
16	GNAS	3.1372
17	AKT1	3.1279
18	MAP2K1	3.0754
19	CBL	3.0287
20	PML	3.0043

TABLE II

THE TOP 20 GENES WITH RESPECT TO LOCAL BER CURVATURE FOR BREAST CARCINOMA.

C. Global BER Curvature of Transcription Networks

We also calculate the global BER curvature for the cancer and normal networks as a whole. Since the average local curvature (Table I) is not the optimal statistic of the networks, the global BER curvature could better justify the higher robustness of the cancer network with respect to the normal network. Table IV shows five cancer networks which have a higher global curvature than their corresponding normal ones. As discussed before, this differentiates the cancer networks with greater robustness compared to the normal network.

Gene Rank	top 20 Genes	Δ Curvature (Cancer-Normal)
1	CDK4	4.5675
2	PP2R1A	4.3991
3	NPM1	4.3766
4	MAP2K2	4.2263
5	SMARCE1	3.7553
6	YWHAE	3.6548
7	RPN1	3.4337
8	NONO	3.2941
9	XPO1	3.0835
10	H3F3A	3.0566
11	RAD21	2.9930
12	RAC1	2.8890
13	CDC73	2.8444
14	MSH2	2.8199
15	SYK	2.7858
16	LCK	2.7803
17	SET	2.5233
18	CALR	2.3671
19	RPL22	2.3658
20	HMGA1	2.3211

TABLE III

THE TOP 20 GENES WITH RESPECT TO LOCAL BER CURVATURE FOR LUNG ADENOCARCINOMA.

Cancer Type	**Δ Global BER Curvature**
Breast Carcinoma	1.3881
Head/Neck Carcinoma	0.7689
Liver Carcinoma	6.0827
Lung Adenocarcinoma	6.0637
Thyroid Carcinoma	0.8727

TABLE IV

GLOBAL BER CURVATURES ILLUSTRATE THAT THE CANCER NETWORK AS A WHOLE EXHIBITS GREATER ROBUSTNESS THAN ITS CORRESPONDING NORMAL NETWORK.

IX. CONCLUSIONS AND FURTHER RESEARCH

Our definition of BER curvature has consistently demonstrated that cancerous networks characteristically exhibit higher curvature than non-cancerous networks, both globally and locally. More importantly, this characteristic is not lost when considering networks with positive and negative edge weights. As a result of the Fluctuation Theorem discussed above, BER curvature provides a way to quantify robustness of the entire network as well as the robustness of a particular gene. Such a tool has the potential to elucidate unknown key roles of proteins, suggest new promising drug targets, measure the efficacy of certain therapies and help prevent drug-resistance. While we have not yet verified that BER curvature has these capabilities, our findings are encouraging.

An immediate next step would be to repeat this study on larger networks, using TCGA data consisting of both cancer and non-cancer-related genes. Ranking the nodes by their BER curvature, we would expect to find the cancer-related

genes clustered at the top of the list and the non-cancer-related genes clustered at the bottom. This would provide additional support for our definition of BER curvature and any unexpected rankings could possibly provide some new insight.

A closely related idea for future work is to construct a directed Laplacian that is consistent with our methods thus far and also allows defining the local graph Ricci curvature for nodes connected to edges with negative weights. This would alleviate the limitation of using the combinatorial Laplacian as defined above, which is obviated by inequality (16) in section (VI).

References

[1] N. Aceto, et al., "Tyrosine phosphatase SHP2 promotes breast cancer progression and maintains tumor-initiating cells via activation of key transcription factors and a positive feedback signaling loop," *Nature Medicine* **18(4)** (2012), pp. 529-537. 6

[2] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits,* Chapman and Hall, (2006).

[3] D. Bakry and M. Émery, Diffusions hypercontractives, Séminaire de probabilités, XIX, *Lecture Notes in Math*. **1123**, Springe, Berlin, (1985), 177-206. 1, 2, 3, 4

[4] A. Barabasi, "The network takeover," *Nature Physics,* **8** (2012), pp. 14-16.

[5] J.-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem," *Numerische Mathematik* **84** (2000), pp. 375-393. 2, 3

[6] A.-I. Bonciocat, and K.-T. Sturm, "Mass transportation and rough curvature bounds for discrete spaces," *Journal of Functional Analysis.* **256** (2009), 2944-2966 .

[7] Q.-Y. Chen, L.-Q. Xu, D.-M. Jiao, Q.-H.Yao, Y.-Y. Wang, et al. "Silencing of Rac1 modifies lung cancer cell migration, invasion and actin cytoskeleton rearrangements and enhances chemosensitivity to antitumor drugs", *International Journal of Molecular Medicine* **28** (2011) pp. 769-776. 7

[8] F. Chung, "Laplacians and the Cheeger inequality for directed graphs," *Annals of Combinatorics* **9** (2005), pp. 1-19.

[9] Y. Chen, S.Z. Hong, and T. Georgiou, " On the definiteness of graph Laplacians with negative weights: Geometrical and passivity based approaches," 4

[10] S. Chow, W. Huang, Y. Li, and H. Zhou, "Fokker-Planck equations for a free energy functional or Markov process on a graph," *Archive for Rational Mechanics and Analysis,* **203** (2012), pp. 969-1008.

[11] J. Cidado and B. H. Park, "Targeting the PI3K/Akt/mTOR pathway for breast cancer therapy," *Journal of Mammary Gland Biology and Neoplasia,* **17(3-4)** (2012), pp. 205-216. 6

[12] L. Demetrius and T. Manke, "Robustness and network evolution entropic principle," *Physica A* **364** (2005), pp. 682-696. 2, 5

[13] M. DoCarmo, *Riemannian Geometry*, Birkhäuser, Boston (1993) 1

[14] H. Ebi, et al., "PI3K regulates MEK/ERK signaling in breast cancer via the Rac-GEF, P-Rex1," *Proceedings of the National Academy of Sciences* **110(52)** (2013), pp. 21124-21129.

[15] M. Erbar, J. Maas, and P. Tetali. "Discrete Ricci curvature bounds for Bernoulli-Laplace and random transposition models," *Arch. Ration. Mech. Anal.* **206 (3)** (2012), pp. 997-1038. 2

[16] L. C. Evans, "Partial differential equations and Monge–Kantorovich mass transfer, " *Current Developments in Mathematics*, International Press, Boston, MA, (1999) pp. 65-126.

[17] W. Gangbo and R. McCann, "The geometry of optimal transportation," *Acta Math.* **177** (1996), pp. 113-161.

[18] C. Givens and R. Shortt, "A class of Wasserstein metrics for probability distributions," *Michigan Mathematical Journal.* **31** (1984), pp. 231-240.

[19] S. Grisendi, C. Mecucci, B. Falini, PP. Pandolfi, "Nucleophosmin and cancer", *Nature Reviews Cancer* **6** (2006), pp. 493-505. 7

[20] L. V. Kantorovich,"On a problem of Monge," *Uspekhi Mat. Nauk.* **3** (1948), pp. 225-226.

[21] R. Jordan, D. Kinderlehrer, and F. Otto, "The variational formulation of the Fokker-Planck equation," *SIAM J. Math. Anal.* **29** (1998), pp. 1-17. 3

[22] J. Jost, *Riemannian geometry and geometric analysis,* 3rd ed., Universitext, Springer-Verlag, Berlin, (2002)

[23] B. Klartag, G. Kozma, P. Ralli, and P. Tetali, "Discrete curvature and abelian groups," http://arxiv.org/abs/1501.00516 (2015) 1, 2, 4

[24] D. Klein and M. Randić, "Resistance distance," *Journal of Mathematical Chemistry,* vol. 12, (1993), pp. 81-95. 4

[25] J. Lott and C. Villani, "Ricci curvature for metric-measure spaces via optimal transport," *Annals of Mathematics,* **169** (2009) pp. 903-991. 2, 5

[26] J. Luo, B. D. Manning and L. C. Cantley, "Targeting the PI3K-Akt pathway in human cancer: Rationale and promise," *Cancer Cell*, **4(4)** (2003), pp. 257?262. 6

[27] J. Maas, "Gradient flows of the entropy for finite Markov chains," *Journal of Functional Analysis.* **17** (2011), 2250-2292.

[28] Y. Ollivier, "Ricci curvature of metric spaces," *C. R. Math. Acad. Sci. Paris.* **345** (2007), pp. 643-646. 2

[29] Y. Ollivier, "Ricci curvature of Markov chains on metric spaces," *Journal of Functional Analysis*. **256** (2009), pp. 810-864. 2

[30] F. Otto,"The geometry of dissipative evolution equation: the porous medium equation," *Comm. Partial Differential Equations* **26** (2001), pp. 101-174.

[31] S. T. Rachev and L. Rüschendorf, *Mass Transportation Problems,* Vol. I and II, Springer-Verlag, (1998).

[32] P. J. Roberts and C. J. Der, "Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer" *Oncogene* **26(22)** (2007) pp. 3291-3310. 6

[33] R. Sandhu, T. Georgiou, E. Reznik, L. Zhu, I. Kolesov, Y. Senbabaoglu, and A. Tannenbaum, "Graph curvature for differentiating cancer networks." *Scientific Reports.* **6**, (2015). 1, 2, 5, 6

[34] K.-T. Sturm, "On the geometry of metric measure spaces," I and II. *Acta Math.*, **196(1)** (2006), pp. 65-131 & 133-177. 2

[35] K.-T. Sturm, "Probability measures on metric spaces of non-positive curvature," *Contemporary Mathematics,* **338** (2003), pp. 1-34.

[36] E. Tannenbaum, T. Georgiou, and A. Tannenbaum, "Signals and control aspects of optimal mass transport and the Boltzmann entropy, " in 49th IEEE *Conference on Decision and Control,* Atlanta, GA. December 15-17, (2010) pp. 1885-1890.

[37] S. R. S. Varadhan, *Large Deviations and Applications*, SIAM, (1984)

[38] C. Villani, *Optimal Transport, Old and New,* Springer-Verlag, (2008). 3

[39] C. Villani, *Topics in Optimal Transportation,* American Mathematical Society Publications, (2003).

[40] J. West, G. Bianconi, S. Severini, and A. Teschendorff, "Differential network entropy reveals cancer system hallmarks," *Scientific Reports.* **2**, (2012). 5

[41] A. Wu, B. Wu, J. Guo, W. Luo, D. Wu, et al. "Elevated expression of CDK4 in lung cancer," *Journal of Translational Medicine* **9:38**, (2011). 6

[42] D. Zelazo and M. Burger, "On the definiteness of the weighted Laplacian and its connection to effective resistance," *Proc. 53rd IEEE Conf. Decision Control.* (2014), pp. 2895-2900. 1, 2, 3, 4