Ollivier-Ricci Flow for Community Detection

Lewis D.L. Hart[1]

[1] *Mathematical Institute, University of Oxford*, OX1 3LB, United Kingdom

This manuscript was compiled on 20/04/2022

Network geometry is a blossoming new research direction of network science, whereas community detection has long-been considered a taproot of the field. At the forefront of new research in Ollivier-Ricci curvature for community detection (1–3) is a novel algorithm: the Ollivier-Ricci flow, created in 2019 by Ni et al. in (1), and grounded in network geometry. We make the case for its inclusion into the current arsenal of community detection algorithms and, through the lens of Ollivier's unique discretisation in (4), explore its foundations in optimal transport theory. However, the implementation of Ollivier-Ricci flow in (1) heavily relies on incorporating modularity-based optimisation methods to achieve accurate results. Using the Ollivier-Ricci flow defined in (1), and our own original numerical simulations, we investigate the efficacy of this community detection method and the extent to which it is independent from modularity. In our discourse, we also reexamine the resolution limit in modularity optimisation (5).

Complex Networks | Network Geometry | Community Detection | Ricci Flow | Optimal Transport | Resolution Limit

By modelling data as a complex network, the nature of many systems in the real-world can be sufficiently, or entirely, captured. Complex networks have been used to model the acquaintances of persons (a social network) in (6, 7), the pages of the internet (an information network) in (8, 9), and cellular functions of molecules (a biological network) in (10, 11). These few examples demonstrate studying complex networks has significant practical applications in the real-world, and in many interdisciplinary contexts. Complex networks are discrete mathematical objects which capture connections (termed *edges*) between points of inter-dependent data (termed *vertices*, or *nodes*). The focus of this article surrounds a scalar quantity of network geometry, the Ollivier-Ricci curvature, and its novel applications to community detection.

The study of communities contained within networks is a linchpin of network science and is frequently discussed in the literature (12–16). Many real-world networks constructed from empirical data have inherent community structures, but it is difficult to ascertain exactly which nodes belong to which communities, since a measure which exposits community belonging is usually obscure, or is infeasible to empirically quantify. For instance, it is relatively simple to gather the required adjacency data of who knows whom in a social network, but much harder to identify and measure which are the groups of people in which everyone knows each other.

Community detection is an umbrella term and refers to the collection of algorithms which take a graph as input (possibly alongside some chosen parameters, depending on the algorithm) and output the communities within a graph. Generally speaking, communities are locally-dense connected subgraphs within a network. Community detection algorithms exist to elucidate obscure community structures, and allow network scientists to uncover communities with methods which rely only on the adjacency data observed. By recognising community structures, we can isolate the unique functional components in a real-world network which contrast to its ubiquitous features. This could have naturally efficacious utilities, for example, in contexts of isolating communities of people in disease-modelling networks to stunt contagion, targeting communication in social networks to individuals whom ideate disparate consensuses, and aggregating clusters of categoric knowledge in information networks which are encyclopaedic.

Curvature is a property introduced by Gauss and Riemann, and is often confined to Eulerian or Riemannian geometry. In general, curvature measures how much a surface curves around a point, *i.e.* the local deviation from flatness. Regardless of dimension, the value of curvature is always scalar. This leads to technical disadvantages: the formulation must be adapted to the shape, and the precise interpretation understood in different dimensions of the space in which it exists. For example, the curvature at any point in a plane is zero, because it is flat, and the curvature at any saddle point on a surface is negative, because the area around a saddle point grows faster than would corresponding area in a plane. However, this scalar nature also leads to conceptual advantages: curvature understands symmetries between different shapes which belong to different dimensions. For example, a surface with constant positive curvature is a sphere, whereas a line with constant positive curvature is a circle (in fact, the latter is true for any non-zero curvature). In the nomenclature, the curvature as described above takes its namesake, but curvature has been animated in many variations to extend the notion to other geometries. These extensions have posited fundamental results which *au fait* the very foundations of their underlying geometry.

The Ricci curvature is one such extension of curvature which informs wide-reaching geometric characteristics of shapes, but

Significance Statement

Whilst community structures in complex networks have been widely observed, identifying their constituent communities algorithmically remains abstruse. Many community detection methods have been developed, and have evolved from very different mathematical foundations. However, every known community detection method is vulnerable to some limitation (17), and their computational complexities vastly vary (18). Thus, community detection remains a significant research objective in network science. Presently, very few methods are based in network geometry, and so this article serves to informs the reader of one novel geometrical approach to community detection: the Ollivier-Ricci flow. We build the necessary foundations to apply and understand the Ollivier-Ricci flow as defined by Ni et al. in (1), and bridge the gap between a first course in network science and novel research in network geometry.

There are no competing interests to declare. Lewis D.L. Hart contributed the entirety of this work.
Correspondence should be addressed to **E-mail:** hart@maths.ox.ac.uk

is a more complicated notion to exemplify. Ricci curvature encapsulates many of the geometrical properties which uniquely define the geometry of a shape within its space, such as the divergence of geodesics, and the volume of balls. In this article we discuss the Ricci curvature of graphs (analogous to shapes) in spaces of network geometry. Providing explicit definitions of all necessary constituents, we formalise the Ricci curvature and apply the Ricci curvature in the setting of complex networks. As Ricci curvature was first defined with applications to classical geometry in mind, it was defined in a continuous sense. Thus, since complex networks are discrete objects, we require a discretised notion of Ricci curvature. There have been many attempts to discretise the Ricci curvature, and this has resulted in many non-equivalent discrete analogues (see (19, 20), and appendix).

In the sequel we study one analogue, termed the *Ollivier-Ricci curvature*, which was originally discretised by Ollivier in (4). Owing to its roots in probability measures from optimal transport theory, Ollivier's definition excellently captures stochastic properties of networks. Building on recent work in this field (see (1, 2, 21)), we examine the usefulness of Ollivier-Ricci curvature in applications to community detection using an algorithmic process called Ollivier-Ricci flow (1), and demonstrate original numerical simulations of Ollivier-Ricci flow on a selection of synthetic graphs over the course of two experiments.

Community Detection

Community detection algorithms aim to recognise dense clusters in a graph. According to Barabási in (22), the fundamental hypothesis of community detection is that the community structure within a network is uniquely encoded in the adjacency matrix A. Thus, all communities which are detectable algorithmically should only depend on inspecting A.

Qualitatively, communities within a graph are defined as partitions of nodes with the property that nodes in any one community are more densely connected with each other, and less densely connected with nodes in other communities. Thus, different communities should be, relatively, sparsely connected. When working with small networks, visualisations can help to determine whether the output communities of a community detection algorithm are sensible, but this quickly becomes ineffective for giant networks.

Since there is some freedom in choosing an appropriate measure of this edge density, the literature does not widely-agree upon a single quantitative definition of a community (23, 24). For this reason, a quantitative definition is, generally, not implemented directly in community detection algorithms. This does lead to difficulties in comparing the efficacy of community detection algorithms, especially when additional non-topological information does not exist in the data-set (23). One example of non-topological information is ground-truth communities metadata. The ground-truth communities in a network are pre-specified partitions of nodes, which in real-world networks are obtained in data collection, and which in synthetic networks are encoded by a model.

Community detection algorithms abound in the literature (12, 25–27), and are categorised by the theme of their underlying mathematics. Ollivier-Ricci flow belongs to the category of geometric community detection algorithms. Notable other categories include, but are not limited to, randomised algorithms (*e.g.* optimising random walks (26)), optimising centrality (*e.g.* optimising betweenness centrality (28)), and optimising modularity (e.g. the Louvain method (12)). Many community detection algorithms work by defining a desirable quantity that can be measured in the network, and then manipulating the network over several iterations (*e.g.* removing edges, removing nodes, or changing edge-weights), with the aim of optimising this quantity. Ollivier-Ricci flow manipulates edge-weights and removes edges with the aim of optimising the Ollivier-Ricci curvature to zero everywhere, which effectively "flattens" the network.

Modularity, introduced by Newman and Girvan in (25), is well-studied in the literature (5, 12, 29), and adheres to the fundamental hypothesis above. The calculation of modularity requires incorporating a null model which randomly generates artificial graphs. The null model obeys the same degree distribution as a given network, but distributes edges uniformly at random. Essentially, within subsets of nodes in a given network, modularity counts the number of edges and subtracts the expected number of edges generated by the null model.

Resolution Limit. No "perfect" community detection algorithm is known to exist. Each approach suffers some shortcoming, and the imperfections due to each method are often unique (17). Modularity, for example, suffers from a resolution limit. This resolution limit is prevalent when optimising modularity directly, and causes groups of smaller communities to sometimes be detected as single communities, *i.e.* merged together. Consider for example, a network with L total edges consisting of n identical cliques, each clique with m nodes, and wherein each clique is connected minimally to exactly two others, *i.e.* between two cliques there is exactly one edge. Such a network is termed a *ring of cliques* (*e.g.* fig. 6). Naturally, each clique in a ring of cliques is a community, since each clique is locally dense. The work of Fortunato and Barthélemy in (5) proved that, if $n > \sqrt{L}$, with n even, then modularity optimisation leads to incorrect community detection. Specifically, two or more cliques are detected as single communities. See fig. 5 for an illustration.

Curvature in Network Geometry

Geometry is routinely described as the concern of spaces and the shapes defined therein. The scope and facets of classical geometry can be extended to complex networks, and this field of study is known as *network geometry*. In the sense that each node in a network occupies a natural arrangement in space, complex networks possess geometry. In this space, different nodes have different positions relative to others, and this distance is determined by a distance metric. Network geometry is properly constructed by pairing the adjacency matrix, A, which encodes all the information in a network, with a sufficient distance metric, d, which takes a pair of 2 nodes as its argument. For vertices x, y in a simple graph, we define the distance metric $d(x, y)$ as the edge-count in any shortest path from x to y.

Discretising Geodesics and Balls. Geodesics and balls are fundamental notions in geometry. On a surface, a geodesic is a locally length-minimising curve, *i.e.* a curve which represents the shortest path between two points. So in the context of graphs, the discrete analogue of a geodesic is the shortest path

between two nodes x, y, and the length of this geodesic is $d(x, y)$. In the context of a graph G, the discrete analogue of a (closed) ball with centre $x \in V$ and radius $r > 0$, is the set $N_r(x)$ defined $N_r(x) = \{y \in V | d(x, y) < r\}$. Thus, the ball $N_r(x)$ consists of all vertices in V which are a path of length at most $r - 1$ edges away from x. The Ricci curvature controls the rate of divergence of geodesics, as well as the growth rate of the volume of balls.

The Ollivier-Ricci discretisation of Ricci curvature in (4) is based in optimal transport theory. In 1781, Monge proposed the optimal transport problem. The problem was originally formulated in practical terms: moving ore from mines to factories. Each mine produces ore, each factory consumes ore, and between each mine and factory there is a certain distance. The total production of ore is equal to the total consumption of factories. The *transportation distance* is equal to the distance the ore is transported multiplied by the mass of the ore (in physics, this is analogous to the work done). The problem then, is to find an optimal transport map wherein all available ore is entirely distributed from all mines to all factories, and which is optimal insofar as the total transportation distance is minimal.

The Ricci curvature determines the average rate at which geodesics starting from the same point diverge. Equivalently, the Ricci curvature controls the rate at which the volume of balls grow via a function of the radius. In cases when geodesics diverge relatively slowly everywhere, or the volume of balls grow relatively slowly everywhere, the Ricci curvature has a small lower bound globally. In terms of two metric balls, the Ricci curvature also controls how much these balls overlap via a function which takes the radii of the balls, and the distance between the centres of these balls, as arguments.

Ollivier's method to port Ricci curvature to the discrete setting in (4) was to compare the distance between two balls with the distance between their centres. Ollivier proved that the overlapping volume of two balls is directly related to the transportation cost of moving mass from one ball to the other. When the volume of two balls is mostly overlapping, this corresponds to a lower transportation cost of moving mass between these two balls.

In 1930, Kantorovich reformulated the problem using probability measures (30), which transformed the optimal transport problem into a linear optimisation problem. The metric of optimisation in Kantorovich's formulation is called the Wasserstein Distance, and it measures the distance between two probability distributions.

Ollivier-Ricci Curvature. We now formally define the Ollivier-Ricci curvature (ORC), κ. Let $G = (V, E)$ be an undirected, unweighted, simple graph. There is a probability measure m_x, which distributes mass to every vertex in the neighbourhood of x and to x itself. For an edge $xy \in E$, the value of κ_{xy} captures the optimal transportation cost between m_x and m_y. In previous work by Lin, Lu and Yau in (21), m_x was chosen so as to distribute mass uniformly to the neighbours of x:

$$m_x^\alpha(x_i) = \begin{cases} \alpha & \text{if } x_i = x, \\ \frac{1-\alpha}{|\pi(x)|} & \text{if } x_i \in \pi(x), \\ 0 & \text{otherwise.} \end{cases}$$

where $\pi(x) = N_1(x) = \{y \in V | xy \in E\}$ is the neighbourhood of x. The parameter $\alpha \in [0, 1]$ is the proportion of mass which is left on node x.

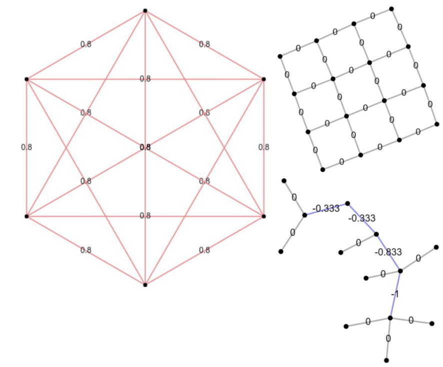

Fig. 1. Numerical values of ORC plotted on each edge using $\alpha = 0.5$. **(Left)** A clique graph with 6 nodes. All edges have positive ORC. **(Top Right)** A snapshot of an infinitely sized lattice graph, with node arrangement and edge connectivity analogous to a 2x2 grid. All edges have zero ORC. **(Bottom Right)** A tree graph. Edges connected to leaves have zero ORC and all other edges have negative ORC. **Credit:** Petrulionyte in (31).

A discrete transport plan is a map $A : V \times V \to [0, 1]$, wherefore the amount of mass from u to be moved to vertex v is $A(u, v)$. Since the total mass demanded by all vertices must be equal to the total mass supplied by all vertices, the transport plan satisfies $\sum_{v' \in V} A(u, v') = m_x(u)$ and $\sum_{u' \in V} A(u', v) = m_y(v)$ for all $xy \in E$. The (discretised) Wasserstein distance W represents the minimum total weighted travel distance to move m_x to m_y, and is defined $W(m_x, m_y) = \inf\{\sum_{u,v \in V} A(u, v) d(u, v)\}$.

We now have all the constituents to define the ORC of the edge $xy \in E$:

$$\kappa_{xy} = 1 - \frac{W(m_x, m_y)}{d(x, y)},$$

where $d(x, y)$ is the same distance metric as before, namely the shortest path (geodesic) distance between x and y. Fig.1 demonstrates calculations of κ_{xy} in 3 different graphs. On comparing values of κ_{xy}, we see each graph is in stark contrast to every other.

With κ_{xy} defined, we can portray some intuition as to how the ORC operates in community detection by examining the relationship between W and d. Let $x, y \in V$. Then, in the optimal transport of m_x to m_y, the majority of mass moved by the transport map A from x to y should be along edges of shortest paths from x to y. Suppose x, y belong to different communities. Then, the nodes in their neighbourhoods, $N_1(x)$ and $N_1(y)$, should have few common neighbours, so we expect there should be very few shortest paths from x to y. Thus, we expect the majority of mass to be moved along these paths. Therefore, the Wasserstein distance $W(m_x, m_y)$ should be larger than $1 \leq d(x, y)$, which would imply $\kappa_{xy} < 0$. So intra-community edges should have negative ORC. On the other hand, if we suppose x, y belong to the same community, then there should be many edges which connect $N_1(x)$ to $N_1(y)$, so there should also exist many equivalently short paths from x to y involved in the optimal transport to move m_x to m_y. Therefore, the Wasserstein distance $W(m_x, m_y)$ should be less than $1 \leq d(x, y)$, which would imply $\kappa_{xy} > 0$. So inter-community edges should have positive ORC.

Ollivier-Ricci Flow

In classical geometry, the Ricci flow transforms a Riemannian manifold via a diffusion process similar to heat diffusion. This diffusion process morphs a manifold by essentially "smoothing it out". Regions of space which contain mostly points with large positive curvature are shrunk, whereas regions of space which contain mostly points with large negative curvature are stretched. In the setting of graphs, the ORC is a reliable medium to analogously capture these diffusion dynamics in a process known as Ollivier-Ricci flow (ORF). In this section, we explain how ORF can be used algorithmically to detect communities as a geometric decomposition of the network via an evolving, iterative process.

Let $G = (V, E)$ be an unweighted, undirected, simple graph. The ORF begins by calculating the ORC values for all edges. Then, the initial edge-weights are formed by assigning these Ollivier-Curvature values to their respective edges. Then, in each iteration, all edge weights are updated simultaneously via the relation:

$$w_{xy}^{(i+1)} = d^{(i)}(x, y) - \kappa_{xy}^{(i)}.d^{(i)}(x, y)$$

where $^{(i)}$ indicates the i-th iteration of each variable. Let $uv \in E$. Then, in the graph induced by the weights $w_{xy}^{(i)}$, and for the edge xy which corresponds to uv in this graph, $w_{xy}^{(i)}$ is the edge weight of xy at the i-th iteration, $\kappa_{xy}^{(i)}$ is the ORC of xy at the i-th iteration, and $d^{(i)}(x, y)$ is the weighted shortest path distance between x and y. Initially $w_{xy}^{(0)} = w_{xy}$ and $d_{xy}^{(0)} = d(x, y)$. As subsequent iterations involve weighted graphs, the distance metric d has to be adapted to incorporate these edge-weights (see appendix). Note, since ORF is a novel algorithm, there is presently minimal evidence (1) to guide the optimal choice in the value of α for defining the ORC.

The dynamics of ORF involve stretching and shrinking edges. Between the i-th and $(i+1)$-th iteration we say that an edge xy is shrunk if $w_{xy}^{(i+1)} < w_{xy}^{i}$, and that xy is stretched if $w_{xy}^{(i+1)} > w_{xy}^{i}$. Iteratively, the ORF stretches edges with negative ORC, and shrinks edges with positive ORC. After many iterations of ORF, intra-community edges become very shrunk, and inter-community edges become very stretched. Then, a thresholding procedure, termed surgery, is used to separate the different communities. The surgery procedure removes all edges with edge-weights greater than a sufficiently large value, termed the cut-threshold. In theory, edges with weights above this cut-threshold are likely to be inter-community edges, and so the surgery disconnects the graph into components, wherein each component should either be itself a community, or should strictly contain two or more communities. If a network contains a hierarchical community structure several rounds of surgery may be required, as there could exist several cut-thresholds which vary in size.

Programming Ollivier-Ricci Flow. ORF can be programmed with relative ease because, owing to optimal transport theory, the program essentially only needs to solve a linear optimisation problem. Using modern programming languages the program is quite simple to write, but such programs generally require significant computational resources. This is why ORF is very slow in comparison to other community detection methods, such as the Louvain method (12) (a greedy-based modularity optimisation algorithm).

After initialising (*i.e.* receiving all necessary inputs), a properly defined algorithm must terminate without manual intervention. In general, ORF is less efficacious when programmed as an algorithm, as was observed by Ni et al. in (1), wherein deterministic predictions of cut-threshold were shown to frequently identify sub-optimal cuts for surgery, or apply too many cuts in surgery. Fig.7 demonstrates an example of ORF implemented as a deterministic algorithm on Zachary's (32) Karate Club, and we see after the second surgery the community output is sub-optimal. Therefore, the approach of Ni et al in many of the networks analysed in (1) was to incorporate ORF into an algorithmic *procedure* instead.

This means, before each surgery the computations are interrupted to await an input either confirming the deterministically predicted cut-threshold, or accepting a a different cut-threshold via manual override (see appendix). In the case of the confirmation input, the algorithmic prediction is heavily guided by modularity-based optimisation methods. In the ORF algorithm in (1) the deterministic prediction of cut-threshold, in the absence of ground-truth data, is effectively programmed as the smallest cut-threshold that yields the maximum modularity (see supplementary information in (1) and (33) for code). To adapt the programme of a deterministic algorithm so that it may presciently identify the optimal cut and number of cuts is an ongoing research objective. So whilst ORF is indeed a community detection method based in network geometry, we remark that for practical purposes, its accuracy and computational complexity (see appendix) is, presently, heavily tied to modularity optimisation theory.

Theoretical Results

In 2019 in (1), Ni et al. investigated whether ORF can accurately identify communities in certain families of graphs. Let K_n denote the complete graph on n vertices and let $a \geq 2$ and $b \geq 2$. Graphs in the family $G(a, b)$ are defined:

1. Take K_{b+1} with vertices labelled $p_1, ..., p_{b+1}$.
2. Take $b + 1$ copies of K_{a+1} and denote these copies $C_1, ..., C_{b+1}$. For each C_i the vertices are labelled $u_1^i, ..., u_{a+1}^i$.
3. For each i coalesce the vertex u_1^i with p_i (*i.e.* set $u_1^i = p_i$). The resulting graph is $G(a, b)$.

Now let $a > b$ and suppose a and b are fixed. Then the graph $G(a, b)$ is highly symmetrical and possesses clear community structure: each $C_1, ..., C_{b+1}$ is a community. This observation is justified because each C_i is a copy of K_{a+1} (so the edge distribution within each C_i is dense), and for each C_i there are $\frac{1}{2}a(a + 1)$ intra-community edges and $b - 1$ inter-community edges.

On consideration, we see many graphs in $G(a, b)$ look similar to, but are not quite the same, as rings of cliques. For example, $G(3, 2)$ bears similar community structure to a ring of 3 cliques, each of size 4 (see fig. 6). A more accurate description is that $G(a, b)$ is a *clique of cliques*. The structural similarity of graphs in $G(a, b)$ to rings of cliques, and the following decisive theorem, were the inspiration for our second experiment, wherein we investigated whether a resolution limit applies to ORF in rings of cliques.

Theorem 1. *Let $a > b$. Let K^0 denote the ORC defined by $\alpha = 0$. Then, for all graphs in the family $G(a, b)$ the ORF associated to the ORC K^0 accurately detects all communities.*

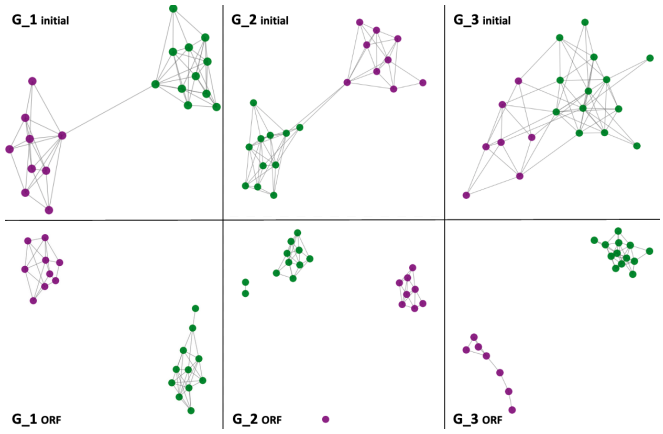

Fig. 2. Top: the artificially generated graphs G_1, G_2 and G_3. Colours indicate the communities. **Bottom:** the output communities of G_1, G_2 and G_3 after surgery on the graphs induced from 50 iterations of ORF using $\alpha = 0.5$.

This theorem was rigorously proven by Ni et al. in (1). In the proof of this theorem it was shown that, for any graph belonging to $G(a, b)$ (with $a > b$), at the n-th iteration of the ORF associated to K^0, the ORC can be explicitly calculated for all edges in the induced graph. Then, it was further shown that the weights of intra-community edges shrink asymptotically faster than the weights of inter-community edges. This asymptotic relation makes it very simple to determine an adequate cut-threshold. Then, using this cut-threshold, only one round of surgery is required, because the community structure is homogeneous. This surgery removes all edges in K_{b+1}, thus separating the graph into components. Each component is a copy of K_{a+1}, and therefore the ORF associated to K^0 detects each community with perfect accuracy.

Materials and Methods

We conducted original numerical simulations over two experiments. ORF is a new approach to community detection, first published in (1) only 3 years ago, whereas modularity has been studied often in the literature for many years (5, 12, 29). The aim of experiment 1 was to compare the efficacy of ORF with modularity-based optimisation methods. The aim of experiment 2 was to investigate the distribution of edge-weights induced by many iterations of ORF in graphs which are susceptible to the resolution limit in (direct) modularity optimisation.

Experiment 1. Our first experiment involved artificially generated graphs, each beholding a clear community structure. The Louvain method (12) and the Clauset-Newman-Moore greedy modularity maximization method (abbrev. *greedy method*) (13) were chosen as our modularity-based methods to compare the efficacy of community detection with ORF.

Our randomised model artificially generated three synthetic graphs, labelled G_1, G_2, and G_3. Each graph is composed of 20 vertices and 60 edges, and exposits a binary community structure: G_1 has a very strong community structure, G_2 has a slightly weaker community structure, and G_3 has the weakest community structure. Upon examining the visualisations provided in the top section of fig.2, this depreciating community structure as described is convincing, at least qualitatively.

In our model, each node is assigned to one of two possible communities uniformly at random. Then, the edges between nodes are distributed in a stochastic process determined by a parameter: the expected ratio of iter-community to intra-community edges (*abbrev. inter:intra*). The value of *inter:intra* encodes the strength of the community structure. For example, if *inter:intra* is equal to 0.1, then nodes in the same community are 10 times more likely to be connected with each other, and for every 10 inter-community edges we expect 1 intra-community edge.

The two benchmark modularity-based community detection algorithms were applied to each graph and their outputs were recorded. Using $\alpha = 0.5$, ORF was also applied to each graph for 50 iterations. Then, the algorithm deterministically decided the cut-threshold and surgery disconnected the graph, outputting the predicted communities as components. For each method, and for each graph, an accuracy statistic was collected by calculating the number of nodes assigned to their communities correctly and dividing by the total number of nodes. The accuracy statistics gathered are recorded in table 1.

Table 1. Results of experiment 1

Graph	*inter:intra*	Louvain acc.	greedy acc.	ORF acc.
1. G_1	0.01	1.0	1.0	1.0
2. G_2	0.05	1.0	1.0	0.85
3. G_3	0.1	1.0	0.80	1.0

Key: Accuracy=acc., greedy = Clauset-Newman-Moore greedy modularity maximization, ORF = ORF.

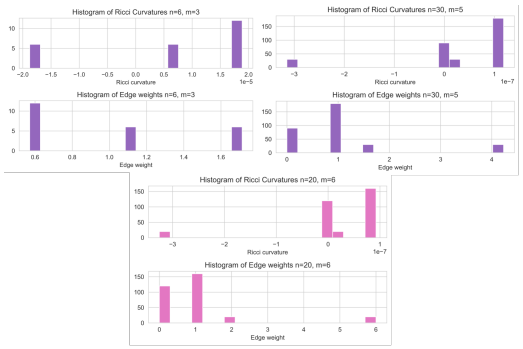

Fig. 3. The histograms of edge-weights and ORCs for 3 rings of cliques after 50 iterations of ORF using $\alpha = 0.5$.

Experiment 2. The aim of our second investigation was to contrast the efficacy of ORF with modularity-based optimisation methods in a more extreme setting: rings of cliques where the resolution limit is known to inhibit accurate community detection when directly optimising modularity.

Rings of cliques, each consisting of n copies of K_m, were generated (so each ring of cliques had, in total, $N = nm$ nodes and $L = \frac{1}{2}nm(m-1) + 2$ edges). We used $n = 6, m = 3$, $n = 20, m = 6$, and $n = 30, m = 5$ to construct three rings of cliques. In each ring of cliques we ensured $n > \sqrt{L}$ since, owing to (5), it is known that rings of cliques with $n > \sqrt{L}$ and n even, are susceptible to the resolution limit.

Using $\alpha = 0.5$, ORF was applied to each ring of cliques over 50 iterations. The resulting distributions of edge weights and ORCs of the induced graphs are organised into the histograms in fig 3.

Discussion

Experiment 1 Results. We acknowledge that the rudimentary stochastic machinery in the graph-generating model, the fixed number of nodes and edges, and the small sample size restrict us from drawing very general conclusions. Nonetheless, we resolve to cogitate what scientific merit we can garner, and we argue, as per our brief in the significance statement, our methods are of pedagogical value: they are a demonstration an application of ORF on a scale that can be understood by inspection, with graphs involving easily visualisable community structures.

Examining table 1, we see that all methods detected the community structure in G_1 with perfect accuracy. This would be expected, since G_1 has an explicit and obvious community structure. In G_2, the Louvain and greedy methods detected both communities with perfect accuracy, whereas the ORF method assigned 3 out of 20 nodes to incorrect communities (see fig. 2). This result was surprising, and is due to a cut-threshold which was too low, and which thus resulted in four output communities rather than two. Observe that the second distinct choice of cut-threshold was 2.1, as can be seen in fig. 9, and a manual-override using this cut-threshold would have detected the communities in G_2 with perfect accuracy. Therefore, we can attribute the accuracy drop-off to an incorrect choice of cut-threshold by the deterministic algorithm, rather than the dynamics of ORF. Furthermore, since the deterministic choice of cut-threshold in G_2 significantly stunts the accuracy of the ORF method, this motivates further research to improve the underlying program.

In G_3, the ORF method and the Louvain method achieved perfect accuracy, whereas the greedy method achieved 0.75 accuracy. These results evidence that, in practice, ORF is competitive with modularity-based community detection methods, and the results of G_2 emphasise that, in at least one case, ORF can perform better.

Experiment 2 Results. Examining the histograms in fig. 3, we see clearly that in each ring of cliques, after 50 iterations of ORF the edges with greatest edge-weight are distinctly disparate from the others. By targeting only the edges with these greatest edge-weights, for example, by choosing cut-thresholds for surgeries equal to 1.6 in $n = 6, m = 3$, to 5 in $n = 20, m = 6$, and to 4 in $n = 30, m = 5$, every community in each ring of cliques was detected with perfect accuracy. This is one piece of evidence that, in theory, and when restricted to rings of cliques, a clear cut-threshold for surgery in ORF can be found deterministically, which does not incorporate modularity-based optimisation techniques. For instance, the cut-thresholds could have been determined by a naive algorithm which simply chose any cut-threshold between the two greatest, non-identical, edge-weights. This adaption bears similarity to a different community detection approach using ORC by Sia et al. in (2). Our suggestion would not be an overarching improvement to the modularity-based method used by Ni in (1), but does reveal that in some cases, modularity-optimisation is not absolutely necessary to determine the cut-threshold.

In the limited purview of our 3 graphs, the results show that ORF outperforms direct modularity optimisation. In the restricted case of all rings of cliques with $n > \sqrt{L}$, the results lend plausibility to a conjecture, our own, that ORF does not suffer from the resolution limit in the same sense as modularity. Our results provide meaningful motivation to probe this conjecture with further research. To investigate this, we suggest the next step would be to conduct numerical simulations involving graphs with multi-layered community structures. For example, graphs consisting many rings of cliques, each of varying clique sizes, and with each ring of cliques sparingly connected to others.

ACKNOWLEDGMENTS. Thank you to Professor Lambiotte, Sofia Medina and Alison Peard for all your enlightening discussions.

1. CC Ni, YY Lin, F Luo, J Gao, Community Detection on Networks with Ricci Flow. *Sci. Reports* **9** (2019).
2. J Sia, E Jonckheere, P Bogdan, Ollivier-Ricci Curvature-Based Method to Community Detection in Complex Networks. *Sci. Reports 2019 9:1* **9**, 1–12 (2019).
3. NG Ananov, et al., Using discrete Ricci curvatures to infer COVID-19 epidemic network fragility and systemic risk. *J. Stat. Mech. Theory Exp.* **2021**, 053501 (2021).
4. Y Ollivier, Ricci curvature of Markov chains on metric spaces. *J. Funct. Analysis* **256**, 810–864 (2007).
5. S Fortunato, M Barthélemy, Resolution limit in community detection. *Proc. Natl. Acad. Sci. United States Am.* **104**, 36–41 (2007).
6. C Stadtfeld, K Takács, A Vörös, The Emergence and Stability of Groups in Social Networks. *Soc. Networks* **60**, 129–145 (2020).
7. ACM Brito, FN Silva, DR Amancio, A complex network approach to political analysis: application to the Brazilian Chamber of Deputies. (2019).
8. FB Gonzaga, VC Barbosa, GB Xexéo, The network structure of mathematical knowledge according to the Wikipedia, MathWorld, and DLMF online libraries. *Netw. Sci.* **2**, 367–386 (2012).
9. Z Wu, T Li, R Roman, Analyzing Wikipedia Membership Dataset and PredictingUnconnected Nodes in the Signed Networks. (2021).
10. E Ravasz, AL Somera, DA Mongru, ZN Oltvai, AL Barabási, Hierarchical organization of modularity in metabolic networks. *Sci. (New York, N.Y.)* **297**, 1551–1555 (2002).
11. LH Hartwell, JJ Hopfield, S Leibler, AW Murray, From molecular to modular cell biology. *Nat. 1999 402:6761* **402**, C47–C52 (1999).
12. LV Nguyen, et al., Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
13. A Clauset, ME Newman, C Moore, Finding community structure in very large networks. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* **70**, 6 (2004).
14. S Fortunato, D Hric, Community detection in networks: A user guide. *Phys. Reports* **659**, 1–44 (2016).
15. H Cherifi, G Palla, BK Szymanski, X Lu, On community structure in complex networks: challenges and opportunities. *Appl. Netw. Sci.* **4** (2019).
16. W Guo, R Chen, YC Chen, AG Banerjee, Efficient Community Detection in Large-Scale Dynamic Networks Using Topological Data Analysis. (2022).
17. R Aldecoa, I Marín, Exploring the limits of community detection strategies in complex networks. *Sci. Reports* **3** (2013).
18. Z Yang, R Algesheimer, CJ Tessone, A Comparative Analysis of Community Detection Algorithms on Artificial Networks. *Sci. Reports 2016 6:1* **6**, 1–18 (2016).
19. K Devriendt, R Lambiotte, Discrete curvature on graphs from the effective resistance. (2022).
20. RR Forman, Bochner's method for cell complexes and combinatorial Ricci curvature. *Discret. Comput. Geom.* **29**, 323–374 (2003).
21. Y Lin, L Lu, ST Yau, Ricci curvature of graphs. *Tohoku Math. J.* **63**, 605–627 (2011).
22. Chapter 9 – Network Science by Albert-László Barabási (year?).
23. F Radicchi, C Castellano, F Cecconi, V Loreto, D Paris, Defining and identifying communities in networks. *Proc. Natl. Acad. Sci. United States Am.* **101**, 2658–2663 (2004).
24. H Lu, A Nayak, S Member, A General Definition of Network Communities and the Corresponding Detection Algorithm. (year?).
25. ME Newman, M Girvan, Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).
26. M Rosvall, CT Bergstrom, Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. United States Am.* **105**, 1118–1123 (2008).
27. L Donetti, MA Muñoz, Detecting Network Communities: a new systematic and efficient algorithm. *J. Stat. Mech. Theory Exp.* (2004).
28. M Girvan, ME Newman, Community structure in social and biological networks. *Proc. Natl. Acad. Sci. United States Am.* **99**, 7821–7826 (2002).
29. ME Newman, Modularity and community structure in networks. *Proc. Natl. Acad. Sci. United States Am.* **103**, 8577–8582 (2006).
30. AM Vershik, Long History of the Monge-Kantorovich Transportation Problem. *The Math. Intell. 2013 35:4* **35**, 1–9 (2013).
31. G Petrulionyte, Ricci Curvature in Network Embedding and Clustering. (year?).
32. WW Zachary, An Information Flow Model for Conflict and Fission in Small Groups. *Source: J. Anthropol. Res.* **33**, 452–473 (1977).
33. Tutorial: GraphRicciCurvature — GraphRicciCurvature 0.5.1 documentation (year?).
34. J Han, W Li, W Deng, Multi-resolution community detection in massive networks. *Sci. Reports 2016 6:1* **6**, 1–12 (2016).
35. X Lu, B Cross, BK Szymanski, Asymptotic resolution bounds of generalized modularity and multi-scale community detection. *Inf. Sci.* **525**, 54–66 (2020).
36. How to Use our Data - Johns Hopkins Coronavirus Resource Center (year?).
37. GitHub - gintepe/RicciEmbeddings: Y4 Final Project (year?).
38. U Brandes, et al., Maximizing Modularity is hard. (2006).
39. Louvain method for community detection (year?).

40. J Van Den Brand, A Deterministic Linear Program Solver in Current Matrix Multiplication Time. (2020).
41. A Samal, et al., Comparative analysis of two discretizations of Ricci curvature for complex networks. *Sci. Reports* **8**, 8650 (2018).
42. E Heintze, H Karcher, A general comparison theorem with applications to volume estimates for submanifolds. *Annales scientifiques de l'École Norm. Supérieure* **11**, 451–470 (1978).
43. F Dorfler, JW Simpson-Porco, F Bullo, Electrical Networks and Algebraic Graph Theory: Models, Properties, and Applications. *Proc. IEEE* **106**, 977–1005 (2018).

Supporting Information Appendix

The code which accompanies this article is available on Github at https://github.com/cannon-complexity/Ollivier-Ricci-Flow-Experiments.

Summary. In this report we promoted new research in network geometry (main credit to Ni et al. in (1)) which uses Ollivier-Ricci curvature (ORC) for community detection via the Ollivier-Ricci flow (ORF) process. We justified some of the theoretical merits of ORF, demonstrated implementations of ORF in our own numerical simulations, and based on our experimental results, suggested avenues for further research.

Working from the definitions of geodesics and balls in classical geometry, we discussed some key properties of Ricci curvature, namely how the Ricci curvature controls the rate of divergence of geodesics and the overlapping volume of two balls, and introduced their discrete analogues in the context of network geometry. After explaining the intuition behind Ollivier's method to connect these discretised geometric properties with optimal transport, we explicitly defined ORC and ORF, and explained why ORF can be used for community detection. Then, we examined algorithmic implementations of ORF and discussed why determining the cut-threshold for surgery is, in many instances, dependent on modularity optimisation.

Over the course of two experiments, we implemented ORF in our own numerical simulations. In the first experiment we showed ORF can successfully detect communities in some graphs with binary community structures, which were artificially generated by our model. Our results showed that, in this small sample, ORF performed with competitive accuracy when compared to the Louvain and Clauset-Newman-Moore greedy modularity maximization methods. In the second experiment, we carefully selected a sample of rings of cliques, for which the resolution limit is prohibitive to community detection via direct modularity optimisation in each, and discussed the extent to which modularity optimisation is necessary to determine the cut-threshold. Our results showed that ORF can detect communities in these rings of cliques with perfect accuracy, which lends weight to our conjecture that ORF is not susceptible to the resolution limit (in the same sense as of direct modularity optimisation) in the restricted case of rings of cliques.

Weighted Shortest Path Distance Metric. This definition of d given here is more refined than the one given in the main body. However, note that this definition of the *weighted* shortest path distance, and of the shortest path distance already given both adhere to the relationship described between inter/intra-community edges and parity of ORC explained in the section on ORC, as well as to the relationship described between inter/intra-community edges and shrinking/stretching of edge-weights explained in the section on ORF.

Let $G = (V, E, w)$ be an edge-weighted graph such that w_{ij} is the weight of the edge $ij \in E$, and for all $i, j \in V$ let $w_{ij} > 0$. For two nodes $i, j \in V$, the notation $i\tilde{j}$ denotes that $ij \in E$. The weighted shortest path distance d induces a metric on V by

$$d(v, v') = \min\{\sum_{i=0}^{n} w_{k_i k_{i+1}} | k_i \in V, k_0 = v, k_{n+1} = v', k\tilde{k}_{i+1}\}$$

where the minimum is taken over all edge paths from v to v'.

Note, we require all edge-weights to be strictly positive in order for $d : V \times V \to \mathbf{R}$ to properly define an induced metric on V. Specifically, we require the following axioms to hold

1. $d(x, y) = 0 \iff x = y$ (identity of indiscernibles)
2. $d(x, y) = d(y, x)$ (symmetry)
3. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality)

for all $x, y, z \in V$.

The identity of indiscernibles axiom requires edge-weights be non-zero. The symmetry axiom combined together with the triangle inequality, requires edge-weights be strictly positive. For brevity and simplicity we omitted this extended definition of the shortest path distance d in the main text. In the redacted definition it is self-evident that the metric axioms are satisfied, and in the case of an unweighted graph, the weighted shortest path distance is equivalent to the shortest path distance used in the prequel.

A Note on the Resolution Limit. Ever since Fortunato and Barthélemy discovered the resolution limit in (5) it has been a major imposition to network scientists working with giant real-world complex networks. Despite the more recent incorporation of a resolution parameter to mollify the affects of the resolution limit, in networks which have hierarchical community structures there is still tendency for modularity optimisation to lead to aggregated communities nearer the top or bottom of this hierarchy (34).

There is presently no convincing adaptation of modularity optimisation which incorporates a global parameter to avoid the resolution limit entirely. However, in 2020 there was progress by Lu et al. in (35) to mitigate the resolution limit through the implementation of a novel, progressive, agglomerative heuristic which gradually increases the resolution parameter. Albeit, their method is not simple and to increment the resolution parameter iteratively necessarily requires substantially more computational resources than direct modularity optimisation. Thus, dilemmas arising from the resolution limit remain a present point of concern in network science.

Real-World Use-Cases of Ollivier-Ricci Curvature for Community Detection. ORC has recently been used to detect communities in real-world networks which incise contextual scientific merit. For example, in the *Facebook Ego-network* in the work of Ni et al. (1), in *drug-drug interaction networks* in the work of Sia et al. (2), and in John Hopkin's COVID-19 data (36) in the work of Souza et al. (3). We refer the reader to (33, 37) for recent Python3 implementations of ORF.

A Note on the Computational Complexity of Ollivier-Ricci Flow. Exact modularity optimization is a known **NP**-hard problem (38), so is not practical to directly implement into an ORF algorithm. Therefore, the Louvain method (12) is commonly used to speed up these calculations. The Louvain method is widely believed to run in $O(N.logN)$ time, where N is the number of nodes in the network (39). Hitherto, there is no known research into the computational complexity of ORF. We suggest a good starting point would be recent work from 2019 by Brand in (40) on the complexity of linear optimisation problems.

A note on Manual Override in Ollivier-Ricci Flow Algorithmic Processes. To guide a manual override, one approach is to construct an xy-line plot as a helping tool. In this line, the domain x-coordinates are all possible cut-thresholds, and each image y-coordinate is the maximum modularity (from optimisation) after surgery using each possible cut (these modularities are averaged and normalised according to the size of each component resulting from surgery). Figures 8, 9 and 10 demonstrate these line plots from our own numerical simulations in experiment 1. Note that the computations to plot this helper-tool are already necessary processes, albeit are implicit, in the deterministic ORF algorithm of Ni et al. in (1).

Ricci Curvature in Classical Geometry. As the focus of this article is on the subject of network geometry, it was ineloquent to present the definition of Ricci curvature in the context of Riemmanian manifolds and formulate all the necessary constituent measures. Instead, we opted to narrate an outlay of the most fundamental geometrical properties of a manifold which depend on Ricci curvature.

Here, we briefly take the opportunity to elucidate these mainstay properties of Ricci curvature in Riemmanian geometry more clearly.

For a a vector $\mathbf{v}$ on an n-dimensional manifold, the Ricci curvature $Ric(\mathbf{v})$ quantifies the local deviation of the manifold from being locally Euclidean (around $\mathbf{v}$). The average dispersion of geodesics around $\mathbf{v}$, as well as the growth of balls and spheres around $\mathbf{v}$ both depend on the Ricci curvature (topologically, a "ball" refers to the

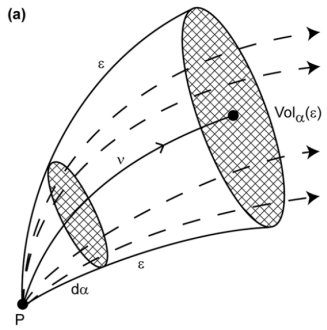

Fig. 4. Sketch of the relationship between the Ricci curvature, the rate of growth of the volume of balls, and the divergence of geodesics. **Credit:** Samal et al. in (41)

space inside a sphere; "sphere" refers to only the surface). These two derivative qualities are acquiesced within the formula (41, 42):

$$\mathrm{Vol}_\alpha(\epsilon) = d\alpha\epsilon^{n-1}(1 - \frac{Ric(\mathbf{v})}{3}\epsilon^2 + o(\epsilon^2))$$

where $\mathrm{Vol}_\alpha(\epsilon)$ is the $(n-1)$-dimensional volume generated within an n-solid angle $d\alpha$ by geodesics of length ϵ in the direction of the vector $\mathbf{v}$. Fig.4 encompasses this relation.

Other Discretisations of Ricci Curvature. The Forman-Ricci curvature (defined by Forman in (20)) is an alternate discretisation of Ricci curvature, and has been discussed in parallel discourse with Ollivier-Ricci curvature in many works (we refer the reader to two choice papers (41) and (19)). Ollivier and Forman developed their discretisations independently, and each definition intuits different interpretations of the notion of smoothness (41). Each discretisation has led to different results, understanding the structure and behaviour of networks through different lenses. The advantage of Forman-Ricci curvature is that its mechanism is local within the network topology. In consequence, algorithms which use Forman-Ricci curvature are efficient to compute, and of significantly lower computational complexity than those which use ORC. Interestingly, Forman's and Ollivier's Ricci curvatures are highly correlated for many real-world networks (41), which is a surprising result given their involved computational demands, which are in stark contrast to each other.

We refer special mention to the reader of recent work by Devriendt and Lambiotte: in January 2022 in (19), Devriendt and Lambiotte introduced a new discretisation of curvature which utilises the concept of effective resistance (inspired by electrical circuits, wherein the resistance exerted by the whole network can be captured by current flowing between any two nodes (43)) and gave strong theoretical evidence for a number of relations between Ollivier-Ricci curvature and Forman-Ricci curvature.

Additional Figures.

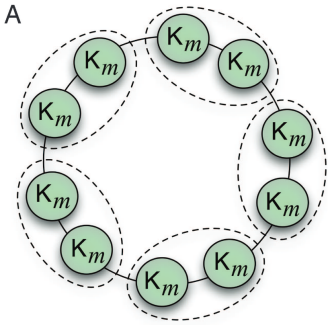

Fig. 5. A ring of K_m cliques. Dotted lines represent communities which are aggregated into single communities due to the resolution limit when $n\sqrt{L}$ with n even. See main body for details. **Credit:** Fortunato and Barthélemy in (5).

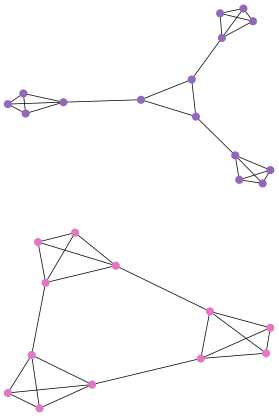

Fig. 6. Top: The graph $G(3,2)$ belonging to the family $G(a,b)$. **Bottom:** A ring of 3 cliques, each of size 4.

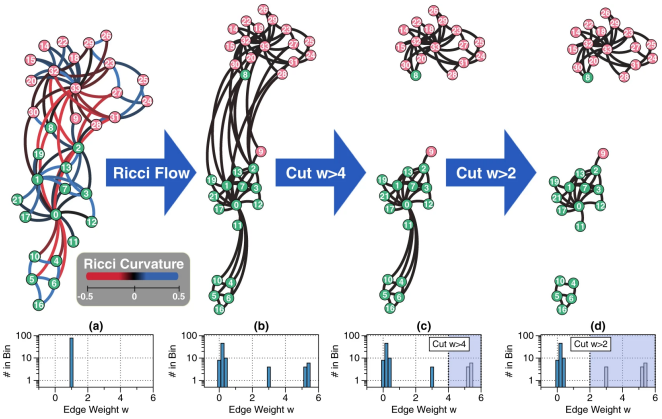

Fig. 7. (a): Zachary's famous Karate Club Network from (32). Red and green colours on the nodes represent the ground-truth communities, each are centred around the "Officer" node and the "Mr Hi" node. Initially, all edge-weights are equal to 1. **(b):** The resulting graph induced by 100 iterations of ORF. Observe that the majority of inter-community edges are stretched, and the majority of intra-community edges are shrunk. **(c):** The resulting graph after surgery using a cut-threshold equal to 4. Observe the community detection is largely accurate, adhering to the ground-truth communities in all nodes, save for 2. **(d):** The resulting graph after a second surgery using a cut-threshold equal to 2. We see this disconnects the graph into two too many communities, which yields a significantly stunted community detection result in comparison to the first surgery. **Credit:** Ni et al. in (1)

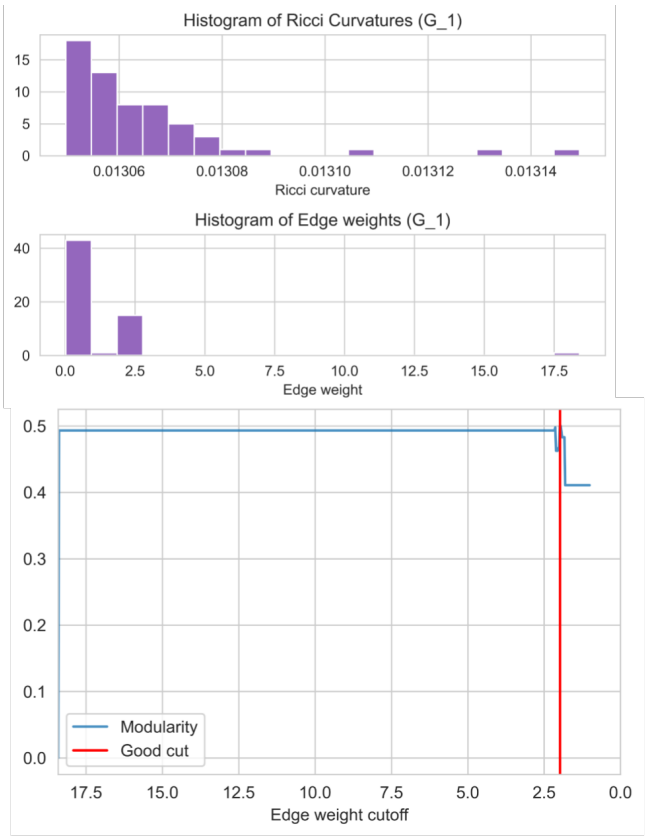

Fig. 8. ORF applied to graph G_1 in experiment 1. **Top:** Distribution of edge weights after 50 iterations of ORF using $\alpha = 0.5$. **Bottom:** Modularity line-plot "helper-tool" for all possible cut-thresholds with the red line indicating the deterministic cut-threshold prediction.

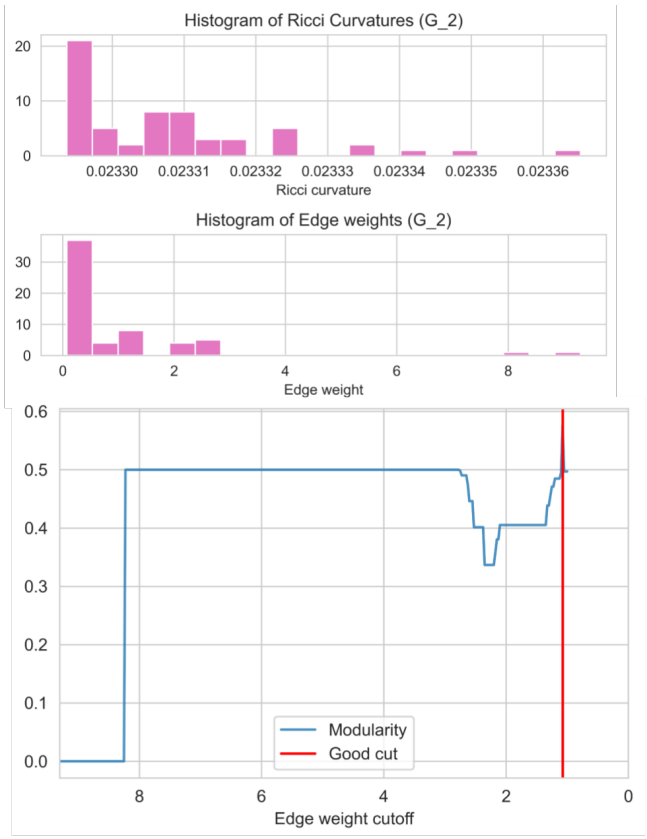

Fig. 9. ORF applied to graph G_2 in experiment 1. **Top:** Distribution of edge weights after 50 iterations of ORF using $\alpha = 0.5$. **Bottom:** Modularity line-plot "helper-tool" for all possible cut-thresholds with the red line indicating the deterministic cut-threshold prediction.

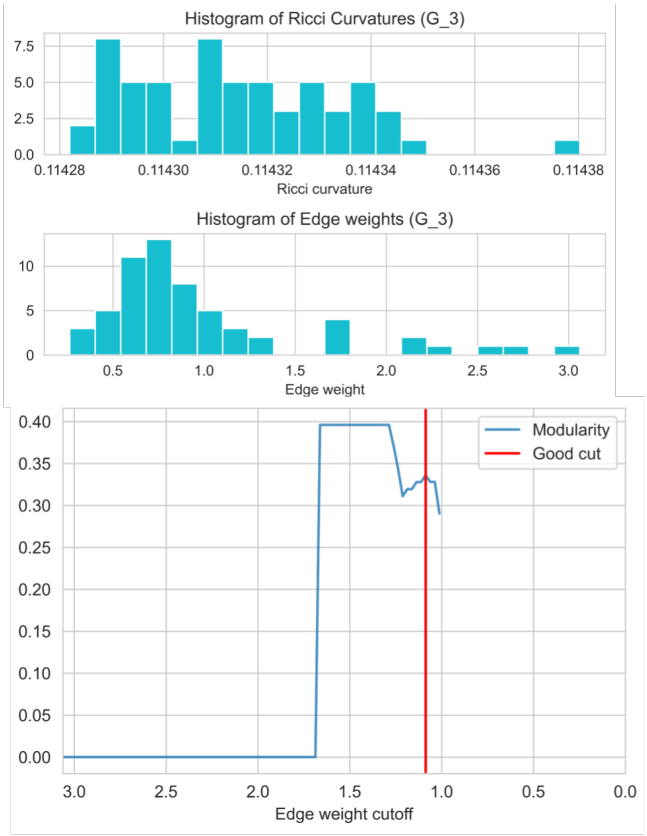

Fig. 10. ORF applied to graph G_3 in experiment 1. **Top:** Distribution of edge weights after 50 iterations of ORF using $\alpha = 0.5$. **Bottom:** Modularity line-plot "helper-tool" for all possible cut-thresholds with the red line indicating the deterministic cut-threshold prediction.