

Article

A Poset-Based Approach to Curvature of Hypergraphs

Yasharth Yadav ¹, Areejit Samal ^{1,2}  and Emil Saucan ^{3,*}

¹ The Institute of Mathematical Sciences (IMSc), Chennai 600113, India; yasharthby@imsc.res.in (Y.Y.); asamal@imsc.res.in (A.S.)

² Homi Bhabha National Institute (HBNI), Mumbai 400094, India

³ Department of Applied Mathematics, ORT Braude College of Engineering, Karmiel 2161002, Israel

* Correspondence: semil@braude.ac.il

Abstract: In this contribution, we represent hypergraphs as partially ordered sets or posets, and provide a geometric framework based on posets to compute the Forman–Ricci curvature of vertices as well as hyperedges in hypergraphs. Specifically, we first provide a canonical method to construct a two-dimensional simplicial complex associated with a hypergraph, such that the vertices of the simplicial complex represent the vertices and hyperedges of the original hypergraph. We then define the Forman–Ricci curvature of the vertices and the hyperedges as the scalar curvature of the associated vertices in the simplicial complex. Remarkably, Forman–Ricci curvature has a simple combinatorial expression and it can effectively capture the variation in symmetry or asymmetry over a hypergraph. Finally, we perform an empirical study involving computation and analysis of the Forman–Ricci curvature of hyperedges in several real-world hypergraphs. We find that Forman–Ricci curvature shows a moderate to high absolute correlation with standard hypergraph measures such as eigenvector centrality and cardinality. Our results suggest that the notion of Forman–Ricci curvature extended to hypergraphs in this work can be used to gain novel insights on the organization of higher-order interactions in real-world hypernetworks.

Keywords: hypergraph; poset; simplicial complex; Forman–Ricci curvature; higher-order interactions



Citation: Yadav, Y.; Samal, A.; Saucan, E. A Poset-Based Approach to Curvature of Hypergraphs. *Symmetry* **2022**, *14*, 420. <https://doi.org/10.3390/sym14020420>

Academic Editors: Serge Lawrencenko and Alice Miller

Received: 1 January 2022

Accepted: 17 February 2022

Published: 20 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Networks provide a powerful framework for modelling interactions within systems composed of a large number of components [1–3]. A network description of any real-world system consists of a collection of vertices or nodes representing the constituents of the system, and edges or links indicating the presence of interactions between pairs of vertices or nodes. For certain systems, the underlying network could also provide additional information in the form of directed edges, weighted edges or signed edges. However, a fundamental limitation of all network representations is that they can only be used to model pairwise interactions within a system, whereas many real-world complex systems display higher-order interactions between three or more components [4,5]. One possible framework to model such higher-order interactions within complex systems is through the use of hypergraphs [6].

In the past two decades it has become trendy to establish discrete analogs of classical theorems in geometry [7,8]. Especially in recent years there has been an increasing interest in characterizing networks as discrete geometric objects, which has led to significant developments in the field of network geometry [9]. Various complementary approaches to network geometry are currently being explored, including hyperbolic network geometry and emergent network geometry [10], and topological data analysis [11]. A notable direction of research in network geometry involves the study of discrete Ricci curvatures for complex networks [12]. Forman–Ricci curvature and Ollivier–Ricci curvature are the two widely-used notions of discrete Ricci curvature that have also found several practical applications, including detection of crashes and bubbles in financial markets [13] and

community identification in complex networks [14]. While the geometric and topological properties of complex networks are being extensively analyzed, studies involving the geometric and topological characterisation of hypergraphs are rare. Nevertheless, there have been a few attempts at extending the notion of discrete Ricci curvature to the framework of hypergraphs [15,16].

The present work is dedicated to the proposition that hypergraphs can be naturally construed as partially ordered sets or posets, which in turn, have an innate interpretation as simplicial complexes. Posets exhibit a strict partial order relation, which is a homogeneous, transitive and asymmetric relation. However, it is important to note that the Hasse diagrams representing posets can be quite symmetrical, even though they are directed graphs. Following the above mentioned properties of hypergraphs and posets, we provide a canonical method to construct two-dimensional simplicial complexes associated with hypergraphs. Two-dimensional simplicial complexes are endowed with intrinsic topological and geometric properties, which allow us to easily compute the Euler characteristic of hypergraphs. Notably, the interpretation of hypergraphs as two-dimensional simplicial complexes also allows us to define the Forman–Ricci curvature for hypergraphs. Subsequently, we performed an empirical study of real-world hypergraphs, where we computed and analyzed the Forman–Ricci curvatures of hyperedges in several real-world hypergraphs. Specifically, we compared the Forman–Ricci curvature of hyperedges with two standard hyperedge measures, namely eigenvector centrality [17] and cardinality.

2. Preliminaries

A hypergraph is defined as a pair $H = (V, A)$, where V is a finite set of elements called vertices or nodes and A is a set of non-empty subsets of V called hyperedges. Note that hypergraphs are a generalization of graphs or networks. In contrast to graphs where an edge can connect only two vertices, hyperedges in hypergraphs can connect a group (or a set) of two or more vertices together.

2.1. Posets

We briefly summarize here the minimal definitions and properties of posets that we have used in this work. Further, we presume the reader is familiar with the basic definition of posets.

Definition 1 (Coverings). Let $(\mathcal{P}, <)$ be a poset, where $<$ denotes the partial order relation on \mathcal{P} , and let p, q be elements of \mathcal{P} . We say that p covers q if $q < p$ and there does not exist $r \in \mathcal{P}$, such that $q < r < p$. We denote the fact that p covers q by $q \prec p$.

Definition 2 (Ranked posets). Given a poset $(\mathcal{P}, <)$, a rank function for \mathcal{P} is a function $\rho: \mathcal{P} \rightarrow \mathbb{N}$ such that

1. If q is a minimal element of \mathcal{P} , then $\rho(q) = 0$;
2. If $q \prec p$, then $\rho(p) = \rho(q) + 1$. A poset \mathcal{P} is called ranked if there exists a rank function for \mathcal{P} . The maximal value of $\rho(p)$, $p \in \mathcal{P}$ is called the rank of \mathcal{P} , and it is denoted by $r(\mathcal{P})$.

Note that for the definition of ranked posets we essentially (but not strictly) follow the one provided by Bloch [18] and, while other terminologies exist [19,20], we prefer the one above for the sake of clarity and concordance with Bloch's paper. Let us also note that if a poset is ranked, then the rank function is unique. Furthermore, if \mathcal{P} is a ranked poset of rank r , and if $j \in \{0, \dots, r\}$, we denote $\mathcal{P}_j = \{p \in \mathcal{P} \mid \rho(p) = j\}$, and by F_j the cardinality of \mathcal{P}_j , i.e., $F_j = |\mathcal{P}_j|$.

While a variety of examples of posets pervade mathematics, a basic example is that of the set of subsets, i.e., the power set 2^X of a given set X . Here, the partial order relation $<$ is defined by the inclusion relation \subset . Note that the hypergraph $H = (V, A)$ defined on the vertex set V is essentially a subset of 2^V , and therefore, can also be regarded as a

poset. As for the case of posets in general, $(2^X, \subset)$ represents the archetypal example of ranked posets. Thus, hypergraphs represent, in essence, ranked posets, which is essential for the sequel. Many of the hypergraphs arising in real-world datasets are actually directed in nature [15,21]. For such hypergraphs, the poset structure is even more evident, as the order relation is emphasized by the directionality. Moreover, if there are no loops in such a hypergraph, the resulting poset is also ranked.

2.2. Simplicial Complexes and the Euler Characteristic

For this work on posets, rather than providing the full technical definition of a simplicial complex, we refer the reader to classic references (see e.g., [22,23]).

Given a poset \mathcal{P} , there exists a canonical way of producing an associated simplicial complex $\Delta(\mathcal{P})$, by considering a vertex for each element $p \in \mathcal{P}$ and an m -simplex for each chain $p_0 \prec p_1 \prec \dots \prec p_m$ of elements of \mathcal{P} .

If the poset \mathcal{P} is finite, we can define its Euler characteristic as being equal to that of the associated simplicial complex $\Delta(\mathcal{P})$, i.e.,

$$\chi(\mathcal{P}) = \chi(\Delta(\mathcal{P})). \tag{1}$$

Note that the hypergraph $H = (V, A)$ defined previously is a finite hypergraph. Thus, we can define the Euler characteristic of the simplicial complex $\Delta(H)$ associated with H .

Equation (1) enables us to define the Euler characteristic of any poset, even if it is not ranked, due to the fact that the associated simplicial complex is naturally ranked by the dimension of the simplices (faces). However, if \mathcal{P} is itself ranked—as indeed it is in our setting—then there exists a direct, purely combinatorial way of defining the Euler characteristic of \mathcal{P} that emulates the classical one, in the following manner:

$$\chi_g(\mathcal{P}) = \sum_{j=0}^r (-1)^j F_j. \tag{2}$$

While in general $\chi(\mathcal{P})$ and $\chi_g(\mathcal{P})$ do not coincide, they are identical in the case of CW complexes, thus in particular for polyhedral complexes, hence a fortiori for simplicial complexes. In particular, we shall obtain the same Euler characteristic irrespective of the model of hypergraph that we choose to operate with: the poset model \mathcal{P} , its associated complex $\Delta(\mathcal{P})$, the geometric view of posets as simplicial complexes attached to each subset of cardinality k (i.e., to each vertex a k -simplex), or the more general polyhedral model that we considered in Ref. [15]. Thus, the Euler characteristic of a hypergraph is a well defined invariant, independent of the chosen hypergraph model, and as such captures the essential topological structure of the network.

In this work, we provide a slightly modified procedure for obtaining a simplicial complex from a poset \mathcal{P} . Specifically, we considered a vertex for each element in \mathcal{P} and a 2-simplex (triangle) for chains consisting of at most 3 elements. This allows us to produce a two-dimensional simplicial complex associated with \mathcal{P} , and we denote such a complex as $\Delta^2(\mathcal{P})$. Figure 1 is a graphical illustration summarizing the construction of $\Delta(\mathcal{P})$ and $\Delta^2(\mathcal{P})$ for an example poset. The Euler characteristic χ of $\Delta^2(\mathcal{P})$ is given by the following simple expression:

$$\chi = |V| - |E| + |F| \tag{3}$$

where $|V|$, $|E|$ and $|F|$ are the number of vertices, number of edges and number of triangular faces, respectively, in $\Delta^2(\mathcal{P})$. In the succeeding sections, we shall show that $\chi(\Delta^2(\mathcal{P}))$ is not just a topological invariant, but it is also closely related to the geometry of $\Delta^2(\mathcal{P})$.

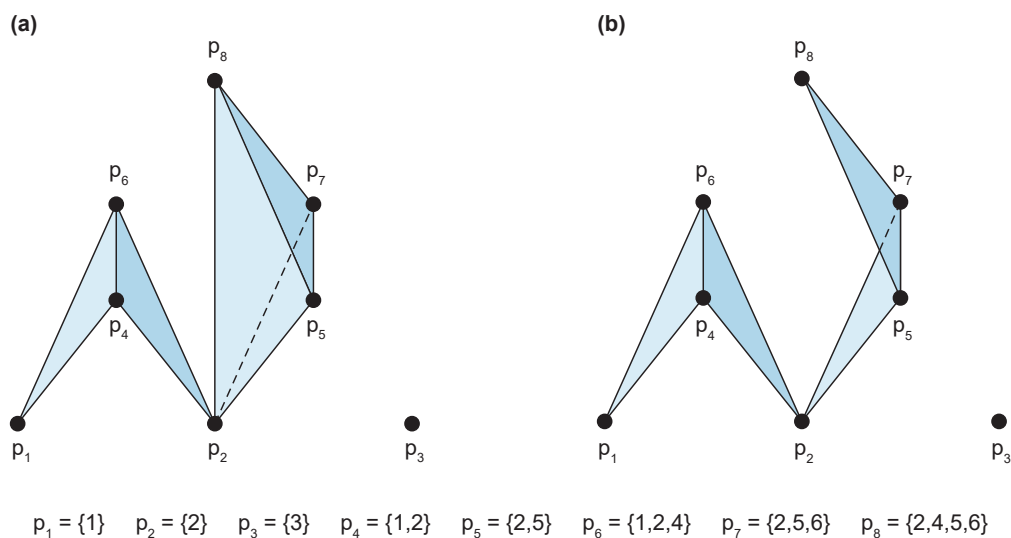


Figure 1. Schematic figure illustrating the construction of simplicial complexes associated with the poset $\mathcal{P} = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7\}$. (a) Construction of $\Delta(\mathcal{P})$, where simplices of any dimension may appear. (b) Construction of $\Delta^2(\mathcal{P})$, where only simplices of dimension ≤ 2 can appear.

2.3. Forman–Ricci Curvature

Forman introduced a discretization of the notion of Ricci curvature [24], by adapting the Bochner–Weizenböck formula [25] to the quite general setting of CW complexes. Some of us have expatiated on the geometric content of Forman’s discretization of Ricci curvature elsewhere [12]. Therefore, to avoid repetition, we refer the reader to [12]. Note that in [12], Forman’s original notion of Ricci curvature is referred to as the augmented Forman–Ricci curvature. Moreover, we had introduced a one-dimensional notion of Forman–Ricci curvature and employed it for the study of complex networks in [26].

While Forman–Ricci curvature applies to both vertex- and edge-weighted complexes, (a feature which makes it suitable for broad range of applications), we concentrate here on the combinatorial case, wherein all vertex as well as edge weights are equal to 1. In this combinatorial case, Forman–Ricci curvature has the following simple and appealing form:

$$\text{Ric}_F(e) = \#\{t^2 : t^2 \supset e\} - \#\{\hat{e} : \hat{e} \parallel e\} + 2. \tag{4}$$

Here, # denotes the number of elements in a set, t^2 denotes triangles and e denotes edges, while $\{t^2 : t^2 \supset e\}$ denotes the set of triangles that consist of the edge e . Moreover, “ \parallel ” denotes parallelism, where two faces of the same dimension (e.g., edges) are said to be parallel if they share a common parent (i.e., a higher dimensional face containing the edges, in this case, a triangle), or a common child (i.e., a lower dimensional face, in this case a vertex), but not both.

Note that for “shallow” real-world hypergraphs, like the metabolic networks considered in [21], both Forman–Ricci curvature and the Euler characteristic are readily computable.

2.4. The Gauss–Bonnet Formula

In the smooth setting, there exists a strong connection between curvature and the Euler characteristic, that is captured by the classical Gauss–Bonnet formula [25]. However, the Forman–Ricci curvature as defined in [24] does not satisfy a Gauss–Bonnet type theorem, since no counterparts in dimensions 0 and 2, essential in the formulation of the Gauss–Bonnet theorem, are defined therein. Interestingly, Bloch defined these necessary curvature terms and was thus able to formulate [18] an analogue of the Gauss–Bonnet theorem in the setting of ranked posets. While in general the one-dimensional curvature term has no close classical counterpart, in the particular case of cell complexes (and thus of

simplicial complexes), Euler characteristic and Forman–Ricci curvature are intertwined in the following discrete version of the Gauss–Bonnet theorem:

$$\sum_{v \in F_0} R_0(v) - \sum_{e \in F_1} \text{Ric}_F(e) + \sum_{f \in F_2} R_2(t) = \chi(X). \tag{5}$$

Here, $F_0, F_1,$ and F_2 denote the sets of vertices, edges and faces (triangles), respectively, and R_0 and R_2 denote the respective zero-dimensional and two-dimensional curvature terms required in a Gauss–Bonnet type formula. These curvature functions are defined via a number of auxiliary functions, as follows:

$$R_0(v) = 1 + \frac{3}{2}A_0(v) - A_0^2(v), R_2(t) = 1 + 6B_2(t) - B_2^2(t), \tag{6}$$

where A_0, B_2 are the aforementioned auxiliary functions, which are defined in the following simple and combinatorially intuitive manner:

$$A_0(x) = \#\{y \in F_1, x < y\}, B_2(x) = \#\{z \in F_1, z < x\}. \tag{7}$$

Since we consider only triangular 2-faces, the formulas for the curvature functions reduce to the following very simple and intuitive ones:

$$R_0(v) = 1 + \frac{3}{2}\text{deg}(v) - \text{deg}^2(v), R_2(t) = 1 + 6 \cdot 3 + 3^2 = 28, \tag{8}$$

where $\text{deg}(v)$ denotes, conform to the canonical notation, the degree of the vertex v , i.e., the number of its adjacent vertices.

From these formulas and from the general expression of the Gauss–Bonnet formula (Equation (5)), we obtain the following combinatorial formulation of the noted formula in the setting of the two-dimensional simplicial complexes:

$$\chi(X) = \sum_{v \in F_0} \left(1 + \frac{3}{2}\text{deg}(v) - \text{deg}^2(v)\right) - \sum_{e \in F_1} \text{Ric}_F(e) + 28, \tag{9}$$

or, after taking into account also Equation (9), and some additional manipulations:

$$\chi(X) = \sum_{v \in F_0} \left(1 + \frac{3}{2}\text{deg}(v) - \text{deg}^2(v)\right) - \sum_{e \in F_1} \left(4 + 9 \cdot \#\{t > e\} + \sum_{v < e} \text{deg}(v)\right) + 28. \tag{10}$$

Note that for unity of notation throughout the paper, we preferred to write in the formulas above, F_0 rather than V , and F_2 instead of F , as commonly used. Equation (5) and its variations also allow for studying the long time behaviour of evolving (hyper-)networks via the use of prototype networks of given Euler characteristic [27].

3. Hypergraph Datasets

For our empirical study, we applied the poset-based construction on 12 real-world hypergraphs spanning multiple domains. Table 1 provides the summary statistics for these 12 hypergraphs.

First, we considered three hypergraphs derived from biological datasets, namely biogrid [28], protein-complex [29], and disease-gene [30]. In the biogrid hypergraph, vertices are chemicals and a hyperedge is a set of chemicals that interact with a given gene. In the protein-complex hypergraph, vertices are proteins and a hyperedge corresponds to a set of proteins contained in a particular complex. In the disease-gene hypergraph, vertices correspond to diseases and a hyperedge consists of a set of diseases associated with a particular gene.

Second, we acquired four hypergraphs based on affiliation networks, namely corporate-memberships [31,32], senate-committees [33], norwegian-directorate [34], and youtube-groups [35]. In the corporate-memberships hypergraph, vertices correspond to corporate

executive officers and a hyperedge consists of members of a particular social organization. The senate-committees hypergraph is derived from Congressional data [36], where vertices are members of the US Senate and hyperedges correspond to committee memberships. The norwegian-directorate hypergraph is based on an interlocking directorate among public limited companies in Norway, wherein the vertices are directors and hyperedges comprise of directors affiliated to boards. In the youtube-groups hypergraph, vertices are YouTube users and a hyperedge is a set of users in a given group.

Table 1. Summary statistics for 12 real-world hypergraphs considered in this study. For each hypergraph, the table lists the labels corresponding to the vertices (V) and hyperedges (H), as well as the number of vertices $|V|$ and hyperedges $|H|$.

Hypergraph	Vertex (V)	Hyperedge (A)	$ V $	$ A $
biogrid	chemical	gene	2138	4455
protein-complex	protein	complex	3639	2848
disease-gene	disease	gene	13,074	8948
corporate-memberships	officer	organization	25	15
senate-committees	US senate member	committee	282	315
norwegian-directorate	director	board	1495	367
youtube-groups	user	group	30,087	94,238
facebook-forum	user	forum topic	899	522
enron-email	email	sender/recipient	10,915	58,575
contact-highschool	student	students in proximity	327	7818
github-projects	user	project	120,867	56,519
cond-mat-publications	author	scientific paper	16,726	22,016

Third, we considered 2 hypergraphs based on communication networks, namely facebook-forum [37], enron-email [38], and one hypergraph contact-highschool [39] which is based on proximity data. The facebook-forum hypergraph is based on users' activity in a Facebook forum, wherein the vertices represent Facebook users and a hyperedge corresponds to the users that have participated in a given topic of the forum. In the enron-email hypergraph, vertices correspond to emails at Enron and a hyperedge consists of the sender and all the recipients of an email. The hypergraph contact-highschool is based on close-proximity human interactions [40], obtained from sensor data worn by students at a high school, wherein vertices represent students and each hyperedge is a group of students that were in close proximity of one another.

Finally, we acquired two hypergraphs, github-projects [31] and cond-mat-publications [3] that are based on collaboration networks and co-authorship networks, respectively. In the github-projects hypergraph, vertices correspond to GitHub users and a hyperedge consists of users which have contributed to a project. The cond-mat-publications hypergraph is based on preprints posted to Condensed Matter section of arXiv E-Print Archive between 1995 and 1999, wherein the vertices correspond to authors and a hyperedge represents a scientific paper, which comprises of co-authors of that paper.

4. Results

For each real-world hypergraph H , we obtained the associated two-dimensional simplicial complex $\Delta^2(H)$ by the canonical construction procedure described in Section 2.2. Note that the vertices (0-simplices) of $\Delta^2(H)$ are comprised of the vertices V and hyperedges A of the original hypergraph H , and the edges (1-simplices) and triangular faces (2-simplices) are defined by the partial order relation (see Figure 1). To investigate the topological properties of the hypergraphs, we computed the Euler characteristic of the associated complex $\chi(\Delta^2(H))$ as defined in Equation (3). The computed Euler characteristic corresponding to each of the 12 hypergraphs considered in this study is reported in Table 2.

Table 2. Euler characteristic of the two-dimensional simplicial complexes $\chi(\Delta^2(H))$ associated with the 12 real-world hypergraphs.

Hypergraph	Euler Characteristic χ
biogrid	−1806
protein-complex	−690
disease-gene	−26,305
corporate-memberships	−31
senate-committees	−4370
norwegian-directorate	48
youtube-groups	−152,325
facebook-forum	−5546
enron-email	−9977
contact-highschool	−3184
github-projects	−210,323
cond-mat-publications	−6560

To study the geometrical properties of the hypergraphs, we computed the Forman–Ricci curvature $\text{Ric}_F(e)$ of the edges in $\Delta^2(H)$ (see Equation (4)). For all the 12 hypergraphs, we found that the Forman–Ricci curvature was negative for most of the edges in $\Delta^2(H)$, indicating a strong prevalence of partial order relation in all the hypergraphs. Moreover, negative edge curvatures indicate a rarity of triangular faces in $\Delta^2(H)$. Recall that the Forman–Ricci curvature of an edge depends on the degrees of the anchoring vertices as well as the triangular faces in which this edge belongs.

Forman–Ricci curvature is a discrete analogue of the classical Ricci curvature [25] and is fundamentally defined on edges. However, it is also possible to define the scalar curvature for a vertex or a node as the sum of the curvatures of all edges incident to that vertex:

$$\mathbf{F}(v) = \sum_{e \sim v} \text{Ric}_F(e) . \quad (11)$$

The above definition was previously used to analyse the curvatures of vertices in complex networks [12]. In the present work, we have extended this definition to the case of two-dimensional simplicial complexes $\Delta^2(H)$ associated with hypergraphs. We found that the scalar curvature is also negative for most of the vertices in $\Delta^2(H)$. Figure 2 shows the distribution of the scalar curvatures of those vertices in $\Delta^2(H)$ that correspond to the hyperedges A of the original hypergraph.

While edges and triangular faces in the poset transformation $\Delta^2(H)$ starting from the original hypergraph do not have direct correspondence to real-world data, the vertices in the transformed poset or two-dimensional simplicial complex $\Delta^2(H)$ represent the vertices and hyperedges of the original hypergraph. Hence, the notion of scalar curvature makes it possible to extend the definition of Forman–Ricci curvature to vertices as well as hyperedges in hypergraphs, and also allows for a comparative analysis with existing hypergraph measures. In the present work, we compute the correlation between Forman–Ricci curvature of hyperedges and two standard hyperedge measures, namely eigenvector centrality and cardinality. A brief description of both these measures is provided in Appendix A.

Eigenvector centrality was first introduced by Bonacich [17] as a measure to quantify the important individuals in the context of social networks. Since then, it has become a standard tool for network analysis with a wide range of applications including Google’s PageRank [41] and as a means to predict neuron’s firing rate [42]. It is based on the idea that the centrality score of a vertex depends on the centrality score of its neighbours. Moreover, Bonacich [17] generalised this concept to two-mode data such as hypergraphs, which allows one to compute the centralities of not only the vertices, but also the hyperedges. It follows from the definition of eigenvector centrality (see Appendix A) that central vertices or nodes belong to central hyperedges and central hyperedges contain central nodes.

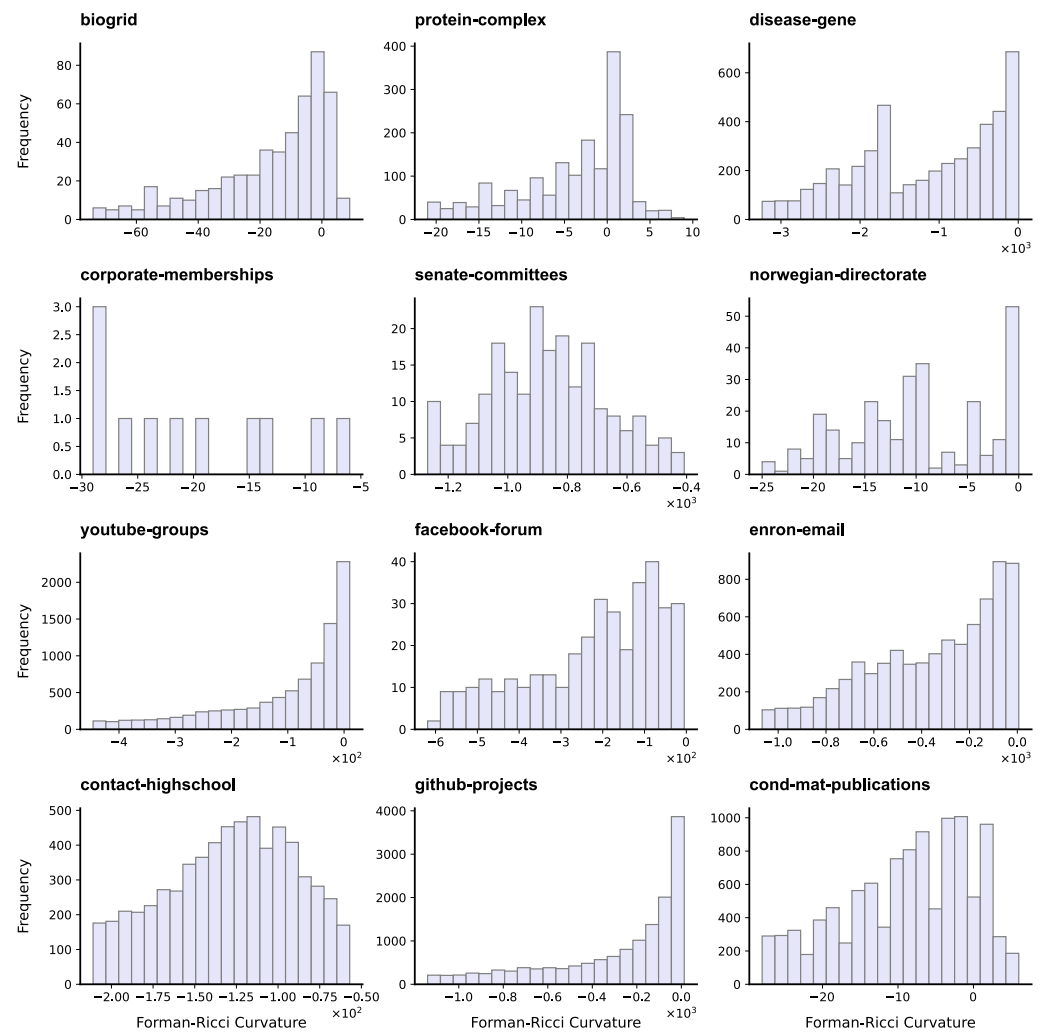


Figure 2. Distribution of the scalar curvature of the vertices in the two-dimensional simplicial complexes $\Delta^2(H)$ associated with the 12 real-world hypergraphs. The plots display the scalar curvatures of only those nodes in $\Delta^2(H)$ that correspond to hyperedges in the original hypergraph H .

Figure 3 is a scatter plot between Forman–Ricci curvature and eigenvector centrality of the hyperedges for the 12 hypergraphs considered in this study. Table 3 shows the Pearson correlation between Forman–Ricci curvature and eigenvector centrality of the hyperedges for the hypergraphs. We found that most real hypergraphs show a moderate to high negative correlation between Forman–Ricci curvature and eigenvector centrality, regardless of the domain that they belong to. However, we should emphasize that 3 hypergraphs, namely norwegian-directorate, enron-email and cond-mat-publications display a weak correlation between Forman–Ricci curvature and eigenvector centrality. On the other hand, affiliation hypergraphs such as corporate-memberships, senate-committees, and youtube-groups have a very high absolute correlation (≥ 0.9) between the two measures. Additionally, we found that hypergraphs representing biological datasets show moderate correlation between Forman–Ricci curvature and eigenvector centrality.

Figure 4 is a scatter plot between Forman–Ricci curvature and cardinality of the hyperedges for the 12 hypergraphs considered in this study. Table 3 shows the Pearson correlation between Forman–Ricci curvature and cardinality of the hyperedges for the hypergraphs. The Forman–Ricci curvature of hyperedges in all the hypergraphs display a high negative correlation with their cardinalities. Similar to the results of correlation with eigenvector centrality, we found that affiliation hypergraphs such as hypergraphs of senate-committees, corporate-memberships and norwegian-directorate show a very high absolute

correlation of 0.94, 0.95 and 0.95, respectively. The affiliation hypergraph youtube-groups also shows a high absolute correlation of 0.81. Moreover, we found that hypergraphs representing biological datasets display a moderate to high absolute correlation between the two measures.

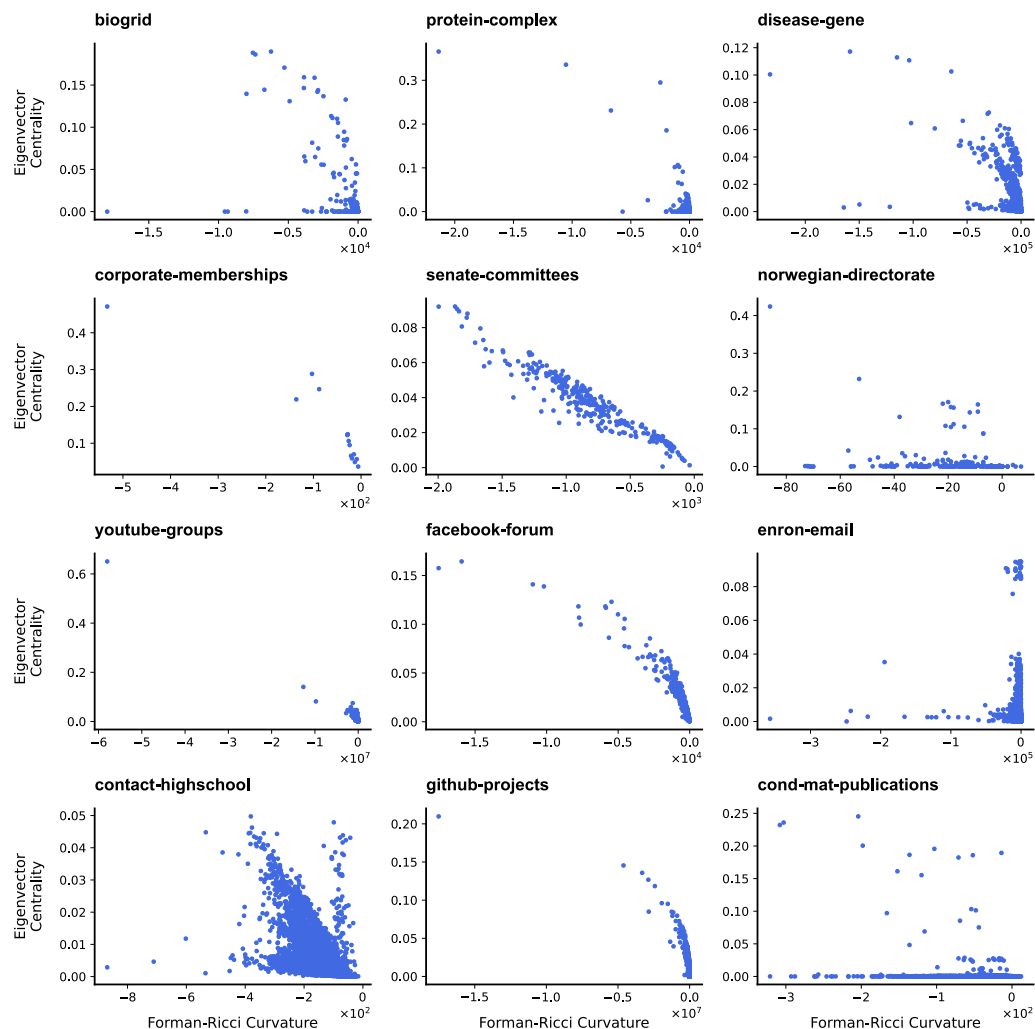


Figure 3. Scatter plots displaying the correlation between Forman–Ricci curvature and eigenvector centrality of the hyperedges for the 12 real-world hypergraphs. The Forman–Ricci curvature of an hyperedge is equal to the scalar curvature of the representative vertex in the two-dimensional simplicial complex $\Delta^2(H)$ associated with the hypergraph H .

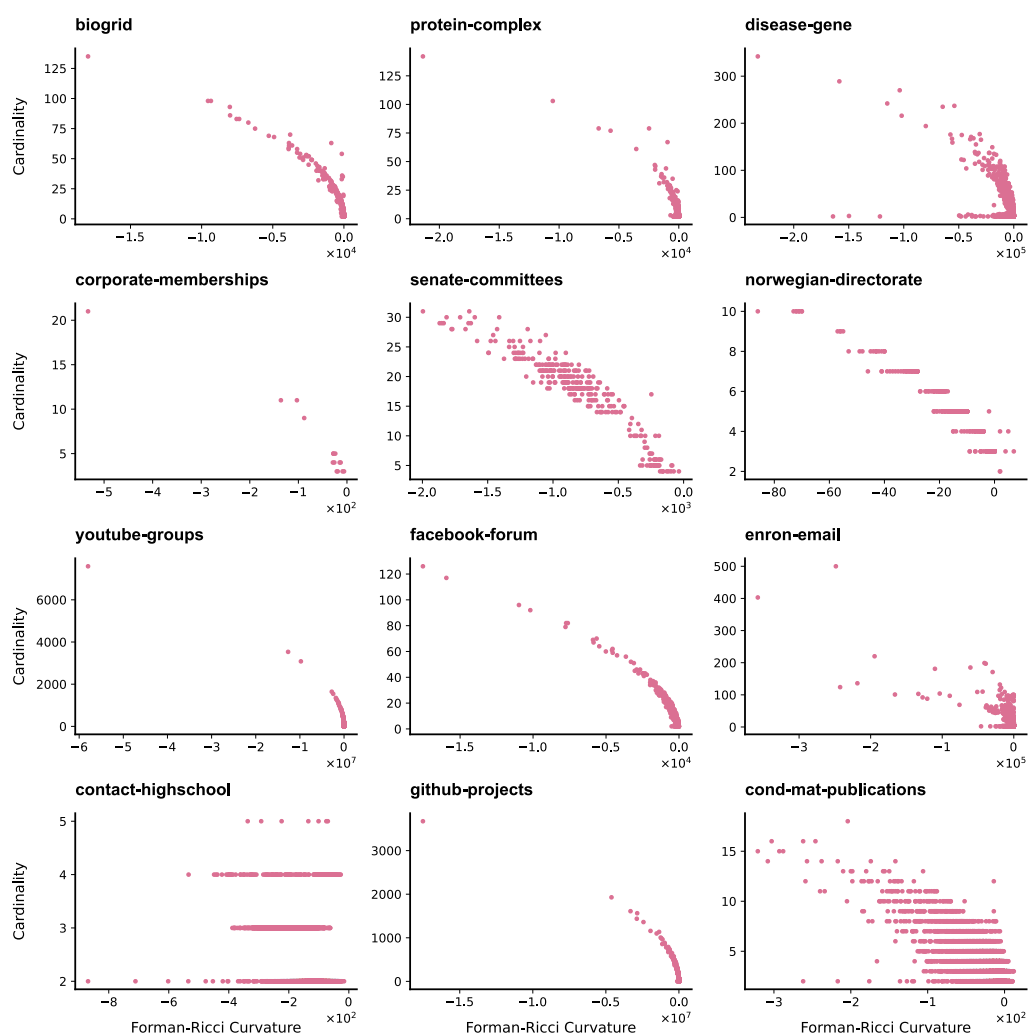


Figure 4. Scatter plots displaying the correlation between Forman–Ricci curvature and cardinality of the hyperedges for the 12 real-world hypergraphs. The Forman–Ricci curvature of an hyperedge is equal to the scalar curvature of the representative vertex in the two-dimensional simplicial complex $\Delta^2(H)$ associated with the hypergraph H .

Table 3. Pearson correlation between Forman–Ricci curvature (FRC) and two traditional hyperedge measures, namely eigenvector centrality (EVC) and cardinality.

Hypergraph	Correlation between FRC and	
	EVC	Cardinality
biogrid	−0.55	−0.88
protein-complex	−0.79	−0.78
disease-gene	−0.57	−0.63
corporate-memberships	−0.9	−0.95
senate-committees	−0.95	−0.94
norwegian-directorate	−0.25	−0.95
youtube-groups	−0.97	−0.81
facebook-forum	−0.88	−0.92
enron-email	−0.08	−0.69
contact-highschool	−0.49	−0.51
github-projects	−0.6	−0.73
cond-mat-publications	−0.18	−0.71

5. Discussion

In this contribution, we present a canonical method to construct two-dimensional simplicial complexes from hypergraphs. Such a construction allows for an easy investigation of the topological and geometric properties of hypergraphs. Specifically, by employing Gauss–Bonnet formula, we have shown that the Euler characteristic of the transformed poset or two-dimensional simplicial complex is mathematically related to the Forman–Ricci curvature of its edges. Additionally, representing hypergraphs as two-dimensional simplicial complexes allows us define Forman–Ricci curvature for vertices or nodes as well as hyperedges. Finally, representing hypergraphs as two-dimensional simplicial complexes could help extend our previously proposed method for persistent homology of complex networks [43–45] to the case of hypergraphs.

Subsequently, we applied the poset transformation to 12 real-world hypergraphs and computed the Forman–Ricci curvature of the hyperedges in the real-world hypergraphs. We have also compared Forman–Ricci curvature of the hypergraphs with two standard hyperedge measures, namely eigenvector centrality and cardinality. We found that Forman–Ricci curvature shows a moderate to high negative correlation with both eigenvector centrality and cardinality. Notably, eigenvector centrality was introduced to quantify the importance of individuals as well as groups in the context of social networks [17]. We found that Forman–Ricci curvature shows a high negative correlation with eigenvector centrality for three affiliation hypergraphs considered in our empirical analysis. These results suggest that Forman–Ricci curvature could also be used to identify important vertices or hyperedges in hypergraphs.

Author Contributions: Y.Y., A.S. and E.S. designed and performed research. Y.Y. carried out the computations and generated the figures. Y.Y., A.S. and E.S. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: A.S. received support from the Max Planck Society, Germany via a Max Planck Partner Group in Mathematical Biology. E.S. received support from the German-Israeli Foundation (GIF) grant number I-1514-304.6/2019.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All datasets generated or analyzed during this study are included in this article or is available upon request from the corresponding author.

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

Appendix A. Traditional Hypergraph Measures

Appendix A.1. Eigenvector Centrality

For the definition of eigenvector centrality of hypergraphs, we follow the one provided by Bonacich [17].

Consider a hypergraph H with n vertices and m hyperedges. Let \mathbf{A} be a $n \times m$ matrix such that $A_{ij} = 1$ if vertex i belongs to hyperedge j and $A_{ij} = 0$ otherwise. Let \mathbf{p} and \mathbf{g} be n and m dimensional vectors of vertex and hyperedge centrality scores, respectively. Then the following matrix equations hold true, where λ is a singular value of the matrix \mathbf{A} :

$$\mathbf{A}\mathbf{g} = \lambda\mathbf{p}, \quad (\text{A1})$$

$$\mathbf{A}^T\mathbf{p} = \lambda\mathbf{g}. \quad (\text{A2})$$

Alternatively, we can consider the hypergraph H as a bipartite graph G with $n + m$ vertices, where the n vertices and m hyperedges of H are the two sets of vertices in G , and define an edge between node i and hyperedge j if $a_{ij} = 1$. We can also define a $n \times m$ adjacency matrix for G . Equations (A1) and (A2) are then equivalent to the following matrix equation:

$$\lambda \begin{pmatrix} \mathbf{p} \\ \mathbf{g} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{p} \\ \mathbf{g} \end{pmatrix}. \quad (\text{A3})$$

We remark that Bonacich has also provided a means to correct for the eigenvector centrality depending on the size of the hypergraphs [17]. However, the need to correct for hypergraph size is not always evident. Therefore, in order to keep our analysis simple, we restrict ourselves to the standard form of eigenvector centrality as described by Equations (A1)–(A3).

Appendix A.2. Cardinality

Consider a hypergraph $H = (V, A)$ and a hyperedge $e \in A$. The cardinality of e is defined as the number of elements in the set e . A hypergraph $H = (V, A)$ is called k -uniform if all hyperedges have cardinality k . Thus, a graph is equivalent to a 2-uniform hypergraph.

References

1. Watts, D.J.; Strogatz, S.H. Collective dynamics of ‘small-world’ networks. *Nature* **1998**, *393*, 440–442. [[CrossRef](#)]
2. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [[CrossRef](#)]
3. Newman, M.E.J. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 404–409. [[CrossRef](#)]
4. Bianconi, G. *Higher-Order Networks; Elements in Structure and Dynamics of Complex Networks*; Cambridge University Press: Cambridge, UK, 2021.
5. Battiston, F.; Cencetti, G.; Iacopini, I.; Latora, V.; Lucas, M.; Patania, A.; Young, J.G.; Petri, G. Networks beyond pairwise interactions: Structure and dynamics. *Phys. Rep.* **2020**, *874*, 1–92. [[CrossRef](#)]
6. Torres, L.; Blevins, A.S.; Bassett, D.; Eliassi-Rad, T. The why, how, and when of representations for complex systems. *SIAM Rev.* **2021**, *63*, 435–485. [[CrossRef](#)]
7. Lawrencenko, S.; Plummer, M.D.; Zha, X. Isoperimetric constants of infinite plane graphs. *Discret. Comput. Geom.* **2002**, *28*, 313–330. [[CrossRef](#)]
8. Higuchi, Y.; Shirai, T. Isoperimetric constants of (d, f) -regular planar graphs. *Interdiscip. Inf. Sci.* **2003**, *9*, 221–228. [[CrossRef](#)]
9. Boguna, M.; Bonamassa, I.; De Domenico, M.; Havlin, S.; Krioukov, D.; Serrano, M.Á. Network geometry. *Nat. Rev. Phys.* **2021**, *3*, 114–135. [[CrossRef](#)]
10. Bianconi, G.; Rahmede, C. Emergent hyperbolic network geometry. *Sci. Rep.* **2017**, *7*, 41974. [[CrossRef](#)] [[PubMed](#)]
11. Patania, A.; Vaccarino, F.; Petri, G. Topological analysis of data. *EPJ Data Sci.* **2017**, *6*, 7. [[CrossRef](#)]
12. Samal, A.; Sreejith, R.P.; Gu, J.; Liu, S.; Saucan, E.; Jost, J. Comparative analysis of two discretizations of Ricci curvature for complex networks. *Sci. Rep.* **2018**, *8*, 8650. [[CrossRef](#)] [[PubMed](#)]
13. Samal, A.; Pharasi, H.K.; Ramaia, S.J.; Kannan, H.; Saucan, E.; Jost, J.; Chakraborti, A. Network geometry and market instability. *R. Soc. Open Sci.* **2021**, *8*, 201734. [[CrossRef](#)]
14. Sia, J.; Jonckheere, E.; Bogdan, P. Ollivier-ricci curvature-based method to community detection in complex networks. *Sci. Rep.* **2019**, *9*, 9800. [[CrossRef](#)]
15. Saucan, E.; Weber, M. Forman’s Ricci curvature—From networks to hypernetworks. In *International Conference on Complex Networks and Their Applications*; Springer, Cham, Switzerland, 2019; Volume 812, pp. 706–717.
16. Eidi, M.; Jost, J. Ollivier ricci curvature of directed hypergraphs. *Sci. Rep.* **2020**, *10*, 12466. [[CrossRef](#)] [[PubMed](#)]
17. Bonacich, P. Simultaneous group and individual centralities. *Soc. Netw.* **1991**, *13*, 155–168. [[CrossRef](#)]
18. Bloch, E. Combinatorial Ricci Curvature for Polyhedral Surfaces and Posets. *arXiv* **2014**, arXiv:math.CO/1406.4598.
19. Rosen, K.H.; Michaels, J.G.; Gross, J.L.; Grossman, J.W.; Shier, D.R. (Eds.) *Handbook of Discrete and Combinatorial Mathematics*; CRC Press: Boca Raton, FL, USA, 1999.
20. Stanley, R. *Enumerative Combinatorics*; Cambridge University: Cambridge, UK, 1997; Volume 1.
21. Saucan, E.; Samal, A.; Weber, M.; Jost, J. Discrete Curvatures and Network Analysis. *MATCH* **2018**, *80*, 605–622.
22. Hudson, J.F.P. *Piecewise Linear Topology*; W. A. Benjamin: Amsterdam, NY, USA, 1969.
23. Rourke, C.P.; Sanderson, B.J. *Introduction to Piecewise-Linear Topology*; Springer: Berlin, Germany, 1972.
24. Forman, R. Bochner’s method for cell complexes and combinatorial Ricci curvature. *Discret. Comput. Geom.* **2003**, *29*, 323–374. [[CrossRef](#)]
25. Jost, J. *Riemannian Geometry and Geometric Analysis*, 7th ed.; Springer: Berlin/Heidelberg, Germany, 2017.
26. Sreejith, R.P.; Mohanraj, K.; Jost, J.; Saucan, E.; Samal, A. Forman curvature for complex networks. *J. Stat. Mech. Theory Exp.* **2016**, *2016*, P063206. [[CrossRef](#)]
27. Weber, M.; Saucan, E.; Jost, J. Coarse geometry of evolving networks. *J. Complex Netw.* **2018**, *6*, 706–732. [[CrossRef](#)]
28. Stark, C.; Breitkreutz, B.J.; Reguly, T.; Boucher, L.; Breitkreutz, A.; Tyers, M. BioGRID: A general repository for interaction datasets. *Nucleic Acids Res.* **2006**, *34*, D535–D539. [[CrossRef](#)] [[PubMed](#)]

29. Giurgiu, M.; Reinhard, J.; Brauner, B.; Dunger-Kaltenbach, I.; Fobo, G.; Frishman, G.; Montrone, C.; Ruepp, A. CORUM: The comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Res.* **2018**, *47*, D559–D563. [[CrossRef](#)]
30. Piñero, J.; Ramírez-Anguita, J.M.; Saüch-Pitarch, J.; Ronzano, F.; Centeno, E.; Sanz, F.; Furlong, L.I. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* **2019**, *48*, D845–D855. [[CrossRef](#)] [[PubMed](#)]
31. Kunegis, J. KONECT: The Koblenz Network Collection. In Proceedings of the 22nd International Conference on World Wide Web (WWW '13 Companion), Rio de Janeiro, Brazil, 13–17 May 2013; pp. 1343–1350.
32. Faust, K. Centrality in affiliation networks. *Soc. Netw.* **1997**, *19*, 157–191. [[CrossRef](#)]
33. Benson, A.R.; Abebe, R.; Schaub, M.T.; Jadbabaie, A.; Kleinberg, J. Simplicial closure and higher-order link prediction. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E11221–E11230. [[CrossRef](#)] [[PubMed](#)]
34. Seierstad, C.; Opsahl, T. For the few not the many? The effects of affirmative action on presence, prominence, and social capital of women directors in Norway. *Scand. J. Manag.* **2011**, *27*, 44–54. [[CrossRef](#)]
35. Mislove, A.; Marcon, M.; Gummadi, K.P.; Druschel, P.; Bhattacharjee, B. Measurement and Analysis of Online Social Networks. In Proceedings of the 5th ACM/Usenix Internet Measurement Conference (IMC'07), San Diego, CA, USA, 24–26 October 2007.
36. Stewart, C.; Woon, J. Congressional Committee Assignments, 103rd to 114th Congresses, 1993–2017. Available online: http://web.mit.edu/17.251/www/data_page.html (accessed on 10 February 2022) .
37. Opsahl, T. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Soc. Netw.* **2013**, *35*, 159–167. [[CrossRef](#)]
38. Klimt, B.; Yang, Y. The Enron Corpus: A New Dataset for Email Classification Research. In *Machine Learning: ECML 2004*; Boulicaut, J.F., Esposito, F., Giannotti, F., Pedreschi, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; pp. 217–226.
39. Mastrandrea, R.; Fournet, J.; Barrat, A. Contact Patterns in a High School: A Comparison between Data Collected Using Wearable Sensors, Contact Diaries and Friendship Surveys. *PLoS ONE* **2015**, *10*, e0136497. [[CrossRef](#)]
40. Chodrow, P.S.; Veldt, N.; Benson, A.R. Generative hypergraph clustering: From blockmodels to modularity. *Sci. Adv.* **2021**, *7*, eabh1303. [[CrossRef](#)]
41. Langville, A.N.; Meyer, C.D. Deeper Inside PageRank. *Internet Math.* **2003**, *1*, 335–380. [[CrossRef](#)]
42. Fletcher, J.M.; Wennekers, T. From Structure to Activity: Using Centrality Measures to Predict Neuronal Activity. *Int. J. Neural Syst.* **2018**, *28*, 1750013. [[CrossRef](#)] [[PubMed](#)]
43. Kannan, H.; Saucan, E.; Roy, I.; Samal, A. Persistent homology of unweighted complex networks via discrete Morse theory. *Sci. Rep.* **2019**, *9*, 1–18. [[CrossRef](#)] [[PubMed](#)]
44. Roy, I.; Vijayaraghavan, S.; Ramaia, S.J.; Samal, A. Forman–Ricci curvature and persistent homology of unweighted complex networks. *Chaos Solitons Fractals* **2020**, *140*, 110260. [[CrossRef](#)]
45. Saucan, E. Discrete Morse Theory, Persistent Homology and Forman–Ricci Curvature. *arXiv* **2020**, arXiv:math.DG/2003.03844.