

# Counting simplicial pairs in hypergraphs

Jordan Barrett\*, Paweł Prałat†, Aaron Smith‡, François Thériberge§

October 16, 2024

## Abstract

We present two ways to measure the simplicial nature of a hypergraph: the simplicial ratio and the simplicial matrix. We show that the simplicial ratio captures the frequency, as well as the rarity, of simplicial interactions in a hypergraph while the simplicial matrix provides more fine-grained details. We then compute the simplicial ratio, as well as the simplicial matrix, for 10 real-world hypergraphs and, from the data collected, hypothesize that simplicial interactions are more and more *deliberate* as edge size increases. We then present a new Chung-Lu model that includes a parameter controlling (in expectation) the frequency of simplicial interactions. We use this new model, as well as the real-world hypergraphs, to show that multiple stochastic processes exhibit different behaviour when performed on simplicial hypergraphs vs. non-simplicial hypergraphs.

## 1 Introduction

Many datasets that are typically represented as graphs would be more accurately represented as hypergraphs. For example, in the graph representation of a collaboration dataset, authors are represented as vertices and an edge exists between two vertices if the corresponding authors wrote a paper together [26]. Using this representation, it is impossible to distinguish between a three-author paper and three separate two-author papers. In contrast, when we represent a collaboration dataset as a hypergraph we can clearly distinguish between a three-author paper (a single hyperedge) and three separate two-author papers (three distinct hyperedges). Hypergraph representations have proven to be useful for studying collaboration datasets [12], protein complexes and metabolic reactions [9], and many other datasets that

---

\*Department of Mathematics, Toronto Metropolitan University, Toronto, ON, Canada; e-mail: [jordan.barrett@torontomu.ca](mailto:jordan.barrett@torontomu.ca)

†Department of Mathematics, Toronto Metropolitan University, Toronto, ON, Canada; e-mail: [pralat@torontomu.ca](mailto:pralat@torontomu.ca)

‡Department of Mathematics and Statistics, University of Ottawa, Ottawa, ON, Canada; e-mail: [asmi28@uottawa.ca](mailto:asmi28@uottawa.ca)

§Tutte Institute for Mathematics and Computing, Ottawa, ON, Canada; email: [theberge@ieee.org](mailto:theberge@ieee.org)

are traditionally represented as graphs [24]. Moreover, after many years of intense research using graph theory in modelling and mining complex networks [8, 11, 16, 25], hypergraph theory has started to gain considerable traction [3, 4, 5, 6, 18, 14, 17]. It is becoming clear to both researchers and practitioners that higher-order representations are needed to study datasets involving higher-order interactions [5, 20, 28, 24].

Similar to hypergraph representations, simplicial complexes provide another way to represent datasets with higher-order interactions and, in some cases, it is not clear what the better model is for a given dataset [19, 29, 31]. The notion of *simpliciality* was first introduced by Landry, Young and Eikmeier in [22] as a way of describing how closely a hypergraph resembles its simplicial closure. In their work, they discover that many hypergraphs built from real-world data, although not actually simplicial complexes, resemble their simplicial closures more closely than random hypergraphs. In a similar but distinct study, LaRock and Lambiotte in [23] find that real-world hypergraphs often contain more instances of hyperedges contained in other hyperedges than in random hypergraphs. The results found in these two papers suggest that real-world hypergraphs are organized in a way where many of the small hyperedges live inside larger hyperedges. In our work, we pursue this idea further and define a ratio and a matrix for hypergraphs, which we call the *simplicial ratio* and *simplicial matrix* respectively, based on the number of instances of hyperedges inside other hyperedges compared to that of a null model.

The remainder of the paper is organized as follows. In Sections 1.1 and 1.2 we discuss notation as well as the measures for simpliciality given in [22]. Next, we define the simplicial ratio in Section 2.1, the simplicial matrix in Section 2.2, and temporal variants in Section 2.3. Then, in Section 3.1 we compute the simplicial ratio and simplicial matrix of the same 10 real-world hypergraphs that were studied in [22] and then analyse this data in Section 3.2. In Section 4 we present a new random graph model that allows for more or less instances of hyperedges inside other hyperedges depending on an input parameter  $q \in [0, 1]$ . In Section 5 we experiment with four stochastic processes, comparing the processes on real-world hypergraphs and on our proposed model for varying  $q$ . We conclude and suggest further research in Section 6. Finally, let us mention that this paper is a (long) journal version of the (short) proceeding paper published in [2].

## 1.1 Notation

In this paper, we use the terms graph and edge in lieu of hypergraph and hyperedge.

A graph  $G$  is a pair  $(V(G), E(G))$  where  $V(G)$  is a set of vertices and  $E(G)$  is a collection of edges, i.e., a collection of subsets of vertices. We insist that  $\emptyset \notin E(G)$  for any graph  $G$ . In general, for a graph  $G$  and edge  $e \in E(G)$ , it is acceptable that  $|e| = 1$ . In this paper, however, we forbid such edges and consider only edges of size at least 2. We write  $[n] := \{1, \dots, n\}$  and typically label the vertices in  $G$  as  $[n]$ . A subgraph of a graph  $G$  is any graph  $H = (V(H), E(H))$  with  $V(H) \subseteq V(G)$  and  $E(H) \subseteq E(G)$  (note that, as  $H$  is itself a graph, any edge  $e \in E(H)$  contains only vertices in  $V(H)$ ). For  $e \in E(G)$ , write  $|e|$  for the

size of  $e$  and, for each positive integer  $k$ , define

$$E_k(G) := \{e \in E(G), |e| = k\}.$$

If  $E_k(G) = E(G)$  for some  $k > 0$ , then we call  $G$  a  $k$ -uniform graph. Note that, for any graph  $G$ , the graph  $G_k := (V(G), E_k(G))$  is a  $k$ -uniform subgraph of  $G$ , and

$$G = \bigcup_{k>0} G_k,$$

and thus every graph is the edge-disjoint union of uniform subgraphs.

A *multigraph*  $G$  is a graph that allows edges  $e \in E(G)$  with more than one instance of the same vertex (multiset edges) and allows multiple edges  $e_1, \dots, e_k \in E(G)$  that are identical (parallel edges); a graph  $G$  is *simple* if it contains no multiset edges or parallel edges. Note that all simple graphs are multigraphs. For a multigraph  $G$  and a vertex  $v$ , writing  $m_G(v, e)$  for the number of instances of  $v$  in  $e$ , the *degree of  $v$  in  $G$* , denoted  $\deg_G(v)$ , is defined as

$$\deg_G(v) := \sum_{e \in E(G)} m_G(v, e).$$

If  $G$  is simple, we equivalently have

$$\deg_G(v) = \left| \{e \in E(G) \mid v \in e\} \right|.$$

All graphs in this paper are simple except for the random graphs generated by Algorithm 2 and Algorithm 4.

We use standard notation for probability, i.e.,  $\mathbb{P}(\cdot)$  for probability,  $\mathbb{E}[\cdot]$  for expectation. We write  $X \sim \mathcal{U}$  to mean  $X$  is sampled from distribution  $\mathcal{U}$  and write  $X_1, \dots, X_k \stackrel{i.i.d.}{\sim} \mathcal{U}$  to mean  $X_1, \dots, X_k$  are sampled independently and identically from distribution  $\mathcal{U}$ . For a set  $S$ , we write  $X \in_u S$  to mean that  $X$  is chosen uniformly at random from  $S$ .

## 1.2 Measures for simpliciality

In [22], Landry, Young and Eikmeier establish three distinct measures quantifying how close a graph is to a simplicial complex. The first measure they establish is the *simplicial fraction*. Given a graph  $G$ , let  $S \subseteq E(G)$  be the set of edges such that  $e \in S$  if and only if  $|e| \geq 3$  and, for all  $f \subseteq e$  with  $|f| \geq 2$ ,  $f \in E(G)$ . Then the *simplicial fraction* of  $G$ , written  $\sigma_{\text{SF}}(G)$ , is defined as

$$\sigma_{\text{SF}}(G) := \frac{|S|}{\left| \bigcup_{k \geq 3} E_k(G) \right|}.$$

In words,  $\sigma_{\text{SF}}(G)$  is the proportion of edges of size at least 3 in  $E(G)$  that satisfy downward closure.

The second and third measures that Landry, Young and Eikmeier establish are the *edit simpliciality* and the *face edit simpliciality*, respectively. For a graph  $G$ , define the  $k$ -closure, written  $\overline{G}_k$ , as the graph  $(V(\overline{G}_k), E(\overline{G}_k))$  where

$$\begin{aligned} V(\overline{G}_k) &= V(G), \\ E(\overline{G}_k) &= \left\{ e \subseteq V(G) \mid |e| \geq k \text{ and } e \subseteq f \text{ for some } f \in E(G) \right\}. \end{aligned}$$

Then the *edit simpliciality* of  $G$ , written  $\sigma_{\text{ES}}(G)$ , is defined as

$$\sigma_{\text{ES}}(G) := \frac{|E(G)|}{|E(\overline{G}_2)|}.$$

Thus,  $1 - \sigma_{\text{ES}}(G)$  is the (normalized) number of additional edges needed to turn  $G$  into its 2-closure. Similarly, the *face edit simpliciality* of  $G$ , written  $\sigma_{\text{FES}}(G)$ , is the average edit simpliciality across all induced subgraphs defined by maximal edges (edges not contained in other edges) in  $\bigcup_{k \geq 3} E_k(G)$ .

Using the three measures defined above, Landry, Young and Eikmeier show that real-world graphs are significantly more simplicial than graphs sampled from random models. However, they also note some unique short-comings of each measure. In the following two examples, we show some additional short-comings that are shared among all three measures. The first example shows that none of the measures properly capture the *types* of simplicial relationships in a graph.

*Example 1.1.* Fix  $n, k$  with  $5 \leq k$  and  $3k \leq n$ . Let  $G_1$  be a graph on the vertex set  $[n]$  and with three edges  $\{1, \dots, k\}, \{k+1, \dots, 2k\}, \{2k+1, \dots, 3k\}$  of size  $k$  and three edges  $\{1, 2, 3\}, \{k+1, k+2, k+3\}, \{2k+1, 2k+2, 2k+3\}$  of size 3. Let  $G_2$  be a graph on the same vertex set and with the same three edges  $\{1, \dots, k\}, \{k+1, \dots, 2k\}, \{2k+1, \dots, 3k\}$  of size  $k$ , but now with three edges  $\{1, \dots, k-1\}, \{k+1, \dots, 2k-1\}, \{2k+1, \dots, 3k-1\}$  of size  $k-1$ . See Figure 1 for an illustration of  $G_1$  and  $G_2$  with  $n = 18$  and  $k = 6$ .

With  $G_1$  and  $G_2$  as defined above, we have

$$\begin{aligned} \sigma_{\text{SF}}(G_1) &= \sigma_{\text{SF}}(G_2) = 0, \\ \sigma_{\text{ES}}(G_1) &= \sigma_{\text{ES}}(G_2) = \frac{2 \cdot 3}{(2^k - k - 1) \cdot 3} = \frac{2}{2^k - k - 1}, \text{ and} \\ \sigma_{\text{FES}}(G_1) &= \sigma_{\text{FES}}(G_2) = \frac{2}{2^k - k - 1}, \end{aligned}$$

the value  $2^k - k - 1$  coming from the fact that there are  $2^k$  subsets,  $k$  of which are subsets of size 1, and 1 of which is the empty set. Thus, by all three measures,  $G_1$  and  $G_2$  are equally simplicial. However, qualitatively the simplicial relationships in  $G_1$  are different than in  $G_2$ . Consider, for example, edges  $e_3, e_5, e_6$  in an Erdős-Rényi random graph on  $n$  vertices with  $|e_3| = 3, |e_5| = 5$  and  $|e_6| = 6$ . Then, the probability of  $e_3 \subseteq e_6$  (as in  $G_1$ ) is of order  $n^{-3}$ , whereas the probability of  $e_5 \subseteq e_6$  (as in  $G_2$ ) is of order  $n^{-5}$ .

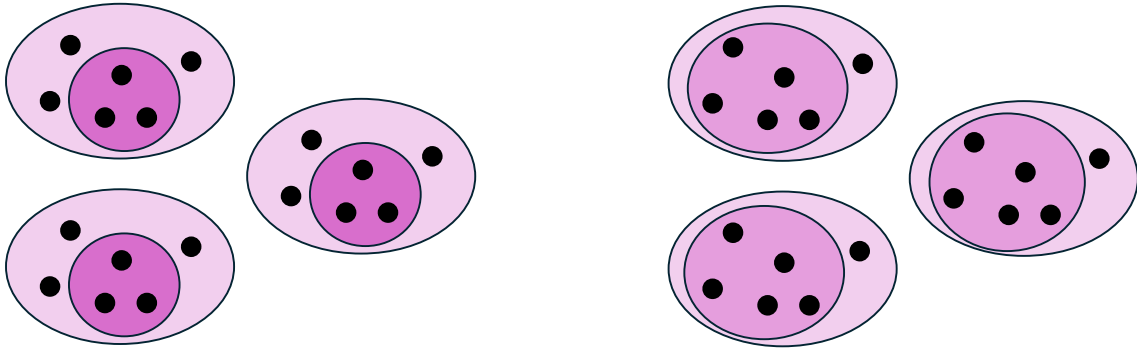


Figure 1: (left) a graph  $G_1$  with 18 vertices, 3 edges of size 6, and 3 edges of size 3, and (right) a graph  $G_2$  with 18 vertices, 3 edges of size 6, and 3 edges of size 5. We have  $\sigma_{\text{SF}}(G_1) = \sigma_{\text{SF}}(G_2) = 0$ ,  $\sigma_{\text{ES}}(G_1) = \sigma_{\text{ES}}(G_2) = 2/57$ , and  $\sigma_{\text{FES}}(G_1) = \sigma_{\text{FES}}(G_2) = 2/57$ .

The second example shows that, while the three measures are good indicators of how close a graph is to its 2-closure, none of the measures are good indicators of how common it is to see edges inside of other edges in the graph.

*Example 1.2.* Let  $G_1$  and  $G_2$  be as shown in Figure 2. There is a clear, strong simplicial structure in  $G_1$ , and there is clearly no simplicial structure in  $G_2$ . However, in both graphs, the simplicial fraction is 0 (none of the edges satisfy downward closure). Moreover, the edit simpliciality of  $G_1$  is  $4/57 \approx 0.07$  and of  $G_2$  is  $3/41 \approx 0.07$ . Likewise, the face edit simpliciality of  $G_1$  is  $4/57 \approx 0.07$  and of  $G_2$  is

$$\frac{1}{3} \left( \frac{1}{26} + \frac{1}{11} + \frac{1}{4} \right) \approx 0.13.$$

Thus,  $G_1$  and  $G_2$  are equally simplicial according to the simplicial fraction and the edit simpliciality and, more strikingly,  $G_1$  is *less* simplicial than  $G_2$  according to the face edit simpliciality.

As mentioned previously, Examples 1.1 and 1.2 are not issues when we treat the simplicial fraction, edit simpliciality, and face edit simpliciality as measures of how close a graph is to its 2-closure (as was their intended purpose). Instead, these examples suggest that if we want to understand the extent to which edges sit inside other edges in real-world networks then we need a new type of scoring system.

## 2 A new approach to simpliciality

We aim to quantify a graph based on the frequency and rarity of edges inside other edges when compared to a null model. The metrics we present focus on the regime where data is “slightly” more simplicial than random (and so nearly-complete large simplices are extremely rare), while previous metrics focus on the regime where data is “almost completely”

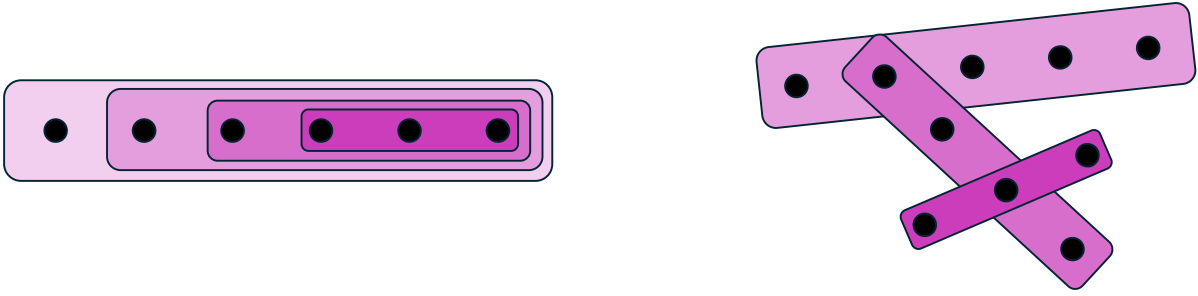


Figure 2: (left) a graph  $G_1$  with 6 vertices and 4 edges, and (right) a graph  $G_2$  with 10 vertices and 3 edges. We have  $\sigma_{\text{SF}}(G_1) = 0$ ,  $\sigma_{\text{ES}}(G_1) \approx 0.07$ ,  $\sigma_{\text{FES}}(G_1) \approx 0.07$ , and  $\sigma_{\text{SF}}(G_2) = 0$ ,  $\sigma_{\text{ES}}(G_2) \approx 0.07$ ,  $\sigma_{\text{FES}}(G_2) \approx 0.13$ .

simplicial. The motivation behind these metrics is that the former regime is often more appropriate in real networks.

### The hypergraph Chung-Lu model

In the material to come, we frequently reference the hypergraph Chung-Lu model. The original model was defined for graphs [7] and has been extensively studied since then. More recently, the model was generalized to other structures, including geometric graphs [15, 13] (both undirected and directed variants) as well as hypergraphs [14]. We give an algorithm for building the hypergraph model, conditioned on the number of edges, and point the reader to [14] for a full description of the model.

Let  $(d_1, \dots, d_n)$  be a degree sequence on vertex set  $[n]$  and let  $(m_{k_{\min}}, \dots, m_{k_{\max}})$  be a sequence of edge sizes where  $m_k$  represents the number of edges of size  $k$ . Then, writing  $p(\cdot)$  for the probability distribution with  $p(v) = d_v / \sum_{u \in [n]} d_u$  for all  $v \in [n]$ , we first give the algorithm that generates a Chung-Lu edge of a given size.

---

**Algorithm 1** Chung-Lu edge.

---

**Require:**  $(d_1, \dots, d_n)$ ,  $k$

- 1: Sample  $e[1], \dots, e[k] \stackrel{i.i.d.}{\sim} p(\cdot)$ .
  - 2: Return  $\{e[1], \dots, e[k]\}$
- 

We now give the algorithm that generates a Chung-Lu graph.

---

**Algorithm 2** Chung-Lu Model.

---

**Require:**  $(d_1, \dots, d_n), (m_{k_{\min}}, \dots, m_{k_{\max}})$ .

- 1: Initialize edge list  $E = \{\}$ .
  - 2: **for**  $k \in \{k_{\min}, \dots, k_{\max}\}$  **do**
  - 3:     **for**  $i \in [m_k]$  **do**
  - 4:         sample  $e \sim \mathbf{Algorithm\ 1}((d_1, \dots, d_n), k)$ .
  - 5:         Set  $E = E \cup \{e\}$ .
  - 6:     **end for**
  - 7: **end for**
  - 8: Return  $G = ([n], E)$ .
- 

For a graph  $G$  with degree sequence  $\mathbf{d} = (d_1, \dots, d_n)$  and edge size sequence  $\mathbf{m} = (m_{k_{\min}}, \dots, m_{k_{\max}})$ , we write  $\hat{G} \sim \text{CL}(G)$  to mean  $\hat{G} \sim \text{CL}(\mathbf{d}, \mathbf{m})$ , where  $\text{CL}(\mathbf{d}, \mathbf{m})$  is the random graph returned by **Algorithm 2**. A key feature of the Chung-Lu model is that the degree sequence is preserved in expectation.

**Lemma 2.1.** *Let  $\hat{G} \sim \text{CL}(G)$  for some graph  $G$ . Then*

$$\mathbb{E}[\text{deg}_{\hat{G}}(v)] = \text{deg}_G(v)$$

for all  $v \in [n]$ .

*Proof.* Let  $d_v := \text{deg}_G(v)$  for all  $v \in [n]$ . First, note that every vertex in every edge of  $\hat{G}$  is sampled independently with probability  $p$ , where  $p(v) = \frac{d_v}{\sum_{u \in [n]} d_u}$ . Thus, the expected total occurrence of  $v$  in  $E(\hat{G})$  is

$$p(v) \sum_{e \in E(\hat{G})} |e| = \left( \frac{d_v}{\sum_{u \in [n]} d_u} \right) \sum_{e \in E(\hat{G})} |e| = \left( \frac{d_v}{\sum_{u \in [n]} d_u} \right) \sum_{u \in [n]} d_u = d_v,$$

the second equality coming from the hypergraph counterpart of the hand-shaking lemma. Given that the total occurrence of  $v$  in  $E(\hat{G})$  is precisely  $\text{deg}_{\hat{G}}(v)$ , the lemma follows.  $\square$

## 2.1 The simplicial ratio

For a graph  $G$ , a *simplicial pair in  $G$*  is a pair of distinct edges  $e_1, e_2 \in E(G)$  with  $e_1 \subset e_2$ . Let  $\text{sp}(G)$  be the number of simplicial pairs in  $G$ .

Let  $G$  be a graph and let  $\hat{G} \sim \text{CL}(G)$  conditioned on  $\hat{G}$  having no multiset edges. Then the *simplicial ratio*, denoted by  $\sigma_{\text{SR}}(G)$ , is defined as

$$\sigma_{\text{SR}}(G) := \frac{\text{sp}(G)}{\mathbb{E}[\text{sp}(\hat{G})]},$$

if  $\mathbb{E}[\text{sp}(\hat{G})] > 0$ , and  $\sigma_{\text{SR}}(G) := 1$  otherwise. In words,  $\sigma_{\text{SR}}(G)$  is the ratio of the number of simplicial pairs to the expected number of simplicial pairs.

*Remark 2.2.* If  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right] = 0$  then it is necessarily the case that  $\text{sp}(G) = 0$ , since it is always true that  $\mathbb{P} \left( \hat{G} = G \right) > 0$ . Moreover, if  $\text{sp}(G) = 0$  and  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right] = 0$  then the number of simplicial pairs is as expected and so we define  $\sigma_{\text{SR}}(G) = 1$ .

*Remark 2.3.* We have mentioned already that the sizes of the edges in a simplicial pair are important. For this reason, we condition on  $\hat{G} \sim \text{CL}(G)$  having no multiset edges.

*Remark 2.4.* Our choice of the Chung-Lu model is not necessary for defining the simplicial ratio. One could equivalently define the simplicial ratio by taking expectations with respect to any model: the configuration model, Erdős-Rényi model, Stochastic Block Model, ABCD model, etc. We choose to use the Chung-Lu model as, in our opinion, it achieves the best balance of (a) retaining important features of a graph and (b) allowing for fast approximations of  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right]$ .

*Remark 2.5.* As mentioned in the previous remark, we *approximate*  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right]$  rather than compute this expectation exactly. For a graph  $G$ , computing  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right]$  is quite difficult as we discuss in the open problems presented in Section 6.1. We approximate using a Monte Carlo estimator which is detailed in Appendix B.

## Examples

Let us revisit Examples 1.1 and 1.2.

Starting with Example 1.1, the number of simplicial pairs in both graphs is 3. However, in  $G_1$  the expected number of simplicial pairs is  $\approx 0.3$ , and in  $G_2$  this expectation is  $\approx 0.008$ . Thus,  $\sigma_{\text{SR}}(G_1) \approx 10$ , whereas  $\sigma_{\text{SR}}(G_2) \approx 380$ , suggesting that the number of simplicial relationships in  $G_2$  is far more surprising than in  $G_1$ . This result confirms that the simplicial ratio weighs different types of simplicial pairs differently.

Continuing with Example 1.2, we have that  $\text{sp}(G_1) = 6$  and  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right] \approx 4.3$ , meaning  $\sigma_{\text{SR}}(G_1) \approx 1.4$ , whereas  $\text{sp}(G_2) = 0$  and  $\mathbb{E} \left[ \text{sp} \left( \hat{G}_2 \right) \right] \approx 0.2 > 0$ , meaning  $\sigma_{\text{SR}}(G_2) = 0$ . Thus, the simplicial ratio can clearly distinguish  $G_1$  and  $G_2$ .

By computing the simplicial ratio of the graphs in Examples 1.1 and 1.2, we see a clear distinction between the three measures given in [22] and the simplicial ratio that we present here: the simplicial fraction, edit simpliciality, and face edit simpliciality are all ways of measuring how close a graph is to its induced simplicial complex, whereas the simplicial ratio is a way to measure how *surprisingly simplicial* a graph is.

## 2.2 The simplicial matrix

For a graph  $G$ , write  $\text{sp}(G, i, j)$  for the number of simplicial pairs  $(e_1, e_2)$  in  $G$  with  $|e_1| = i$  and  $|e_2| = j$  with  $i < j$ . Then, letting  $\hat{G} \sim \text{CL}(G)$  conditioned on having no multiset edges,



the simplicial matrix of  $G$ , denoted by  $M_{\text{SR}}(G)$ , is the partial matrix with cell  $(i, j)$  equalling

$$M_{\text{SR}}(G, i, j) := \frac{\text{sp}(G, i, j)}{\mathbb{E} \left[ \text{sp}(\hat{G}, i, j) \right]}$$

whenever  $i < j$  and  $G$  contains edges of size  $i$  and of size  $j$  (and substituting 0 if there are no simplicial pairs of this type), and with cell  $(i, j)$  being empty otherwise.

*Remark 2.6.* We once again approximate  $\mathbb{E} \left[ \text{sp}(\hat{G}, i, j) \right]$  via the sampling technique found in Appendix B.

Intuitively, the simplicial matrix “unpacks” the simplicial ratio and shows how powerful the simplicial interactions between edges of all different sizes are. More formally, the simplicial matrix and simplicial ratio of  $G$  satisfy the following weighted sum.

$$\sigma_{\text{SR}}(G) = \sum_{i < j} w_{i,j} \cdot M_{\text{SR}}(G, i, j)$$

where

$$w_{i,j} := \frac{\mathbb{E} \left[ \text{sp}(\hat{G}, i, j) \right]}{\mathbb{E} \left[ \text{sp}(\hat{G}) \right]}, \quad \sum_{i < j} w_{i,j} = 1.$$

We will see in Section 3 that the simplicial matrix reveals information about real-world graphs that the simplicial ratio alone does not. In particular, a hypothesis we make in this paper, as suggest by these matrices, is that *the composition of an edge in a real-world network becomes more dependent on simpliciality as the edge size increases.*

## Examples

We again revisit Examples 1.1 and 1.2. In Example 1.1,  $M_{\text{SR}}(G_1)$  contains one non-empty cell,  $(3, 6)$ , with value  $\approx 10$ , and  $M_{\text{SR}}(G_2)$  contains one non-empty cell,  $(5, 6)$ , with value  $\approx 380$ .

Example 1.2 is more interesting as  $G_1$  contains simplicial pairs of various types. For  $G_1$ , we have

$$M_{\text{SR}}(G_1) \approx \begin{bmatrix} \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \mathbf{3.8} & \mathbf{1.7} & \mathbf{1} \\ \emptyset & \emptyset & \emptyset & \emptyset & \mathbf{2.4} & \mathbf{1} \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \mathbf{1} \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \end{bmatrix},$$

and for  $G_2$  we have

$$M_{\text{SR}}(G_2) \approx \begin{bmatrix} \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \\ \emptyset & \emptyset & \emptyset & \mathbf{0} & \mathbf{0} \\ \emptyset & \emptyset & \emptyset & \emptyset & \mathbf{0} \\ \emptyset & \emptyset & \emptyset & \emptyset & \emptyset \end{bmatrix}.$$

The simplicial matrix for  $G_1$  unpacks the information about its simplicial interactions. Indeed, the simplicial ratio simply tells us that the number of simplicial pairs is 1.4 times more than expected. On the other hand, the simplicial matrix tells us that all 3 simplicial pairs involving the edge of size 6 are to be expected, whereas the other three simplicial pairs are at least somewhat surprising. We can also see that the existence of the (3, 4) pair in  $G_1$  is more surprising than the existence of the (3, 5) pair, which is in turn more surprising than the existence of the (3, 6) pair. In general, given a graph  $G$  and distinct edge sizes  $i < j < k$ , if  $G$  has the property that  $|E_j(G)| \leq |E_k(G)|$  then it follows from the sampling process in **Algorithm 1** that  $\mathbb{E}[\text{sp}(G, i, j)] \leq \mathbb{E}[\text{sp}(G, i, k)]$ . In the case of Example 1.2, we have that  $|E_4(G_1)| = |E_5(G_1)| = |E_6(G_1)| = 1$  and  $\mathbb{E}[\text{sp}(G_1, 3, 4)] \approx 0.26$ ,  $\mathbb{E}[\text{sp}(G_1, 3, 5)] \approx 0.59$ , and  $\mathbb{E}[\text{sp}(G_1, 3, 6)] = 1$ .

## 2.3 Including a temporal element

Many networks (both real and synthetic) are not merely static graphs, but rather evolving process with edges forming over time. In these evolving processes, there are two distinct formations of a simplicial pair: either a small edge could form first, followed by a larger (superset) edge, or a large edge could form first, followed by a smaller (subset) edge. In the context of a collaboration graph, a “bottom-up” formation is a group of collaborators who invite more people for a future collaboration, whereas a “top-down” formation is a group who exclude some people for a future collaboration. At least in this context, there is a substantial difference between bottom-up simplicial pairs and top-down simplicial pairs, and we would ultimately like to know how different networks bias towards or against the two types of simplicial formations. For this reason, we include a version of the simplicial ratio and of the simplicial matrix that accounts for time-stamped edges. In the definitions to come, we assume that no two edges are born at the exact same time.

Let  $G$  be an evolving graph with time-stamped edges  $E(G) = (e_1, \dots, e_m)$  such that  $e_i$  was generated before  $e_{i+1}$  for all  $1 \leq i < m$ . Next, let  $\text{sp}^\nearrow(G)$  be the number of simplicial pairs  $(e_i, e_j)$  in  $G$  with  $i < j$  and  $|e_i| < |e_j|$ , and let  $\text{sp}^\searrow(G)$  be the number of simplicial pairs  $(e_i, e_j)$  with  $i > j$  and  $|e_i| < |e_j|$ . Finally, let  $\hat{G} \sim \text{CL}(G)$  and assign a uniformly random ordering to the edges of  $\hat{G}$ . Then the bottom-up simplicial ratio and top-down simplicial ratio of  $G$ , denoted  $\sigma_{\text{SR}}^\nearrow(G)$  and  $\sigma_{\text{SR}}^\searrow(G)$  respectively, are defined as

$$\sigma_{\text{SR}}^\nearrow(G) := \frac{\text{sp}^\nearrow(G)}{\mathbb{E}[\text{sp}^\nearrow(\hat{G})]} \quad \text{and} \quad \sigma_{\text{SR}}^\searrow(G) := \frac{\text{sp}^\searrow(G)}{\mathbb{E}[\text{sp}^\searrow(\hat{G})]}.$$

*Remark 2.7.* By symmetry, we have that  $\mathbb{E}[\text{sp}^\nearrow(\hat{G})] = \mathbb{E}[\text{sp}^\searrow(\hat{G})] = \frac{1}{2} \cdot \mathbb{E}[\text{sp}(\hat{G})]$ . Thus, we can equivalently define the bottom-up simplicial ratio and top-down simplicial ratio respectively as

$$\frac{2 \cdot \text{sp}^\nearrow(G)}{\mathbb{E}[\text{sp}(\hat{G})]} \quad \text{and} \quad \frac{2 \cdot \text{sp}^\searrow(G)}{\mathbb{E}[\text{sp}(\hat{G})]}.$$

For the temporal version of the simplicial matrix we distinguish between bottom-up and top-down simplicial pairs by their location in the matrix. For a temporal graph  $G$  with edge ordering  $E(G) = (e_1, \dots, e_m)$  and for  $k < \ell$ , write  $\text{sp}^\rightarrow(G, k, \ell)$  for the number of simplicial pairs  $(e_i, e_j)$  such that  $i < j$ ,  $|e_i| = k$ , and  $|e_j| = \ell$ . Likewise, write  $\text{sp}^\leftarrow(G, k, \ell)$  for the number of simplicial pairs  $(e_i, e_j)$  such that  $i > j$ ,  $|e_i| = k$  and  $|e_j| = \ell$ . Then the temporal simplicial matrix, denoted  $M_{\text{SR}}^\rightarrow(G)$ , is the partial matrix with cell  $(k, \ell)$  equalling

$$M_{\text{SR}}^\rightarrow(G, k, \ell) := \frac{\text{sp}^\rightarrow(G, k, \ell)}{\mathbb{E} \left[ \text{sp}^\rightarrow(\hat{G}, k, \ell) \right]},$$

cell  $(\ell, k)$  equalling

$$M_{\text{SR}}^\leftarrow(G, \ell, k) := \frac{\text{sp}^\leftarrow(G, k, \ell)}{\mathbb{E} \left[ \text{sp}^\leftarrow(\hat{G}, k, \ell) \right]},$$

for all valid  $k < \ell$ , and cells  $(k, \ell)$  and  $(\ell, k)$  being empty otherwise.

### 3 Empirical results

In this section, we show the simplicial ratio and simplicial matrix, both with and without a temporal element where applicable, for the same 10 graphs that were analysed in [22]. We then comment on the data and build some hypotheses about the simplicial nature of real networks.

The 10 graphs are all taken from [21] and full descriptions can be found there. We paraphrase and summarize the descriptions below.

**contact-primary-school:** a temporal graph where nodes are primary students and edges are instances of contact (physical proximity) between students.

**contact-high-school:** the same as contact-primary-school except with high-school students.

**hospital-lyon:** the same as contact-primary-school and contact-high-school except with patients and health-care workers in a hospital.

**email-enron:** a temporal graph where nodes are email-addresses and edges comprise the sender and receivers of emails.

**email-eu:** the same as email-enron except built from a different organization.

**diseasome:** a static (non-temporal) graph where nodes are diseases and edges are collections of diseases with a common gene.

**disgenenet:** a static graph where nodes are genes and edges are collections of genes found in a disease. In other words, disgenenet is precisely the line-graph of diseasome.

**ndc-substances:** a static graph where nodes are substances and edges are collections of substances that make up various drugs.

**congress-bills:** a temporal graph where nodes are US Congresspersons and edges comprise the sponsor and co-sponsors of legislative bills put forth in both the House of Representatives and the Senate.

**tags-ask-ubuntu:** a temporal graph where nodes are tags and edges are collections of tags applied to questions on the website `askubuntu.com`.

For each graph, we restrict to edges of sizes 2 through 11, as is the case in [22]. We throw away multi-edges, only keeping the first occurrence of each edge in the case of temporal graphs. We approximate  $\mathbb{E}[\hat{G}]$  using the Monte Carlo method presented in Appendix B.

### 3.1 The data

We first show a table which includes the ratios, as well as useful information about each graph.

| $G$             | $ V(G) $ | $ E(G) $ | $[ E_2 ,  E_3 ,  E_4 ,  E_{\geq 5} ]$ | $\sigma_{\text{SR}}(G)$ | $\sigma_{\text{SR}}^{\uparrow}(G)$ | $\sigma_{\text{SR}}^{\downarrow}(G)$ |
|-----------------|----------|----------|---------------------------------------|-------------------------|------------------------------------|--------------------------------------|
| disgenenet      | 1982     | 760      | [157, 139, 93, 371]                   | 28.81                   | n.a.                               | n.a.                                 |
| contact-h.s.    | 327      | 7818     | [5498, 2091, 222, 7]                  | 6.68                    | 11.19                              | 2.17                                 |
| diseasome       | 516      | 314      | [153, 92, 26, 43]                     | 6.49                    | n.a.                               | n.a.                                 |
| email-eu        | 967      | 23729    | [13k, 5k, 2k, 4k]                     | 5.19                    | 5.77                               | 3.72                                 |
| email-enron     | 143      | 1442     | [809, 317, 138, 178]                  | 4.96                    | 6.98                               | 2.94                                 |
| congress-bills  | 1715     | 58788    | [14k, 10k, 8k, 27k]                   | 4.46                    | 5.23                               | 3.69                                 |
| ndc-substances  | 2740     | 4754     | [1130, 745, 535, 2344]                | 4.22                    | n.a.                               | n.a.                                 |
| contact-p.s.    | 242      | 12704    | [7748, 4600, 347, 9]                  | 2.74                    | 4.82                               | 0.66                                 |
| hospital-lyon   | 75       | 1824     | [1107, 657, 58, 2]                    | 0.94                    | 1.71                               | 0.17                                 |
| tags-ask-ubuntu | 3021     | 145053   | [28k, 52k, 39k, 25k]                  | 0.69                    | 1.09                               | 0.29                                 |

Table 1: The simplicial ratio of 10 real networks and the corresponding bottom-up simplicial ratio and top-down simplicial ratio for the 7 temporal networks. The graphs are ordered according to  $\sigma_{\text{SR}}(G)$ , from largest to smallest.

In Figure 4 we show the simplicial matrices of these graphs and in Figure 5 we show the temporal matrices of the 7 temporal graphs. For readability we show only the non-empty cells of the partial matrices and omit cells involving edges of size greater than 5. Figure 3 shows the simplified presentation of the simplicial matrix of  $G_1$  from Example 1.2.

$G_1$

|   | 4   | 5   | 6   |
|---|-----|-----|-----|
| 3 | 3.8 | 1.7 | 1.0 |
| 4 |     | 2.4 | 1.0 |
| 5 |     |     | 1.0 |

Figure 3: The simplicial matrix of  $G_1$  from Example 1.2, presented in a simplified manner.

| <p><b>disgenenet</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>53.7</td> <td>21.5</td> <td>19.7</td> </tr> <tr> <th>3</th> <td></td> <td>&gt;1k</td> <td>&gt;1k</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>&gt;1k</td> </tr> </tbody> </table> |      | 3    | 4    | 5 | 2 | 53.7 | 21.5 | 19.7 | 3 |  | >1k  | >1k  | 4 |  |  | >1k | <p><b>contact-h.s.</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>6.1</td> <td>6.0</td> <td>6.1</td> </tr> <tr> <th>3</th> <td></td> <td>966</td> <td>&gt;1k</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>&gt;1k</td> </tr> </tbody> </table> |  | 3 | 4 | 5 | 2 | 6.1 | 6.0 | 6.1 | 3 |  | 966 | >1k | 4 |  |  | >1k | <p><b>diseasome</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>14.2</td> <td>5.2</td> <td>6.0</td> </tr> <tr> <th>3</th> <td></td> <td>267</td> <td>0</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>0</td> </tr> </tbody> </table>                |  | 3 | 4 | 5 | 2 | 14.2 | 5.2 | 6.0 | 3 |  | 267 | 0   | 4 |  |  | 0   | <p><b>email-eu</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>4.0</td> <td>3.5</td> <td>3.5</td> </tr> <tr> <th>3</th> <td></td> <td>505</td> <td>442</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>&gt;1k</td> </tr> </tbody> </table>     |  | 3 | 4 | 5 | 2 | 4.0 | 3.5 | 3.5 | 3 |  | 505 | 442 | 4 |  |  | >1k |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|------|------|---|---|------|------|------|---|--|------|------|---|--|--|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|---|---|---|---|-----|-----|-----|---|--|-----|-----|---|--|--|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|---|---|---|---|------|-----|-----|---|--|-----|-----|---|--|--|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|---|---|---|---|-----|-----|-----|---|--|-----|-----|---|--|--|-----|
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 53.7 | 21.5 | 19.7 |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | >1k  | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 6.1  | 6.0  | 6.1  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 966  | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 14.2 | 5.2  | 6.0  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 267  | 0    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | 0    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 4.0  | 3.5  | 3.5  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 505  | 442  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| <p><b>email-enron</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>4.3</td> <td>4.0</td> <td>3.8</td> </tr> <tr> <th>3</th> <td></td> <td>154</td> <td>125</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>&gt;1k</td> </tr> </tbody> </table>         |      | 3    | 4    | 5 | 2 | 4.3  | 4.0  | 3.8  | 3 |  | 154  | 125  | 4 |  |  | >1k | <p><b>congress-bills</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>5.3</td> <td>4.6</td> <td>4.4</td> </tr> <tr> <th>3</th> <td></td> <td>236</td> <td>167</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>&gt;1k</td> </tr> </tbody> </table>  |  | 3 | 4 | 5 | 2 | 5.3 | 4.6 | 4.4 | 3 |  | 236 | 167 | 4 |  |  | >1k | <p><b>ndc-substances</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>13.7</td> <td>8.2</td> <td>6.3</td> </tr> <tr> <th>3</th> <td></td> <td>&gt;1k</td> <td>630</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>&gt;1k</td> </tr> </tbody> </table> |  | 3 | 4 | 5 | 2 | 13.7 | 8.2 | 6.3 | 3 |  | >1k | 630 | 4 |  |  | >1k | <p><b>contact-p.s.</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>2.5</td> <td>2.6</td> <td>2.5</td> </tr> <tr> <th>3</th> <td></td> <td>187</td> <td>171</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>&gt;1k</td> </tr> </tbody> </table> |  | 3 | 4 | 5 | 2 | 2.5 | 2.6 | 2.5 | 3 |  | 187 | 171 | 4 |  |  | >1k |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 4.3  | 4.0  | 3.8  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 154  | 125  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 5.3  | 4.6  | 4.4  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 236  | 167  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 13.7 | 8.2  | 6.3  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | >1k  | 630  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 2.5  | 2.6  | 2.5  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 187  | 171  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| <p><b>hospital-lyon</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>0.9</td> <td>0.9</td> <td>0.9</td> </tr> <tr> <th>3</th> <td></td> <td>19.3</td> <td>13.4</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>0</td> </tr> </tbody> </table>          |      | 3    | 4    | 5 | 2 | 0.9  | 0.9  | 0.9  | 3 |  | 19.3 | 13.4 | 4 |  |  | 0   | <p><b>tags-ask-ubuntu</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>0.5</td> <td>0.5</td> <td>0.5</td> </tr> <tr> <th>3</th> <td></td> <td>8.8</td> <td>9.7</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>295</td> </tr> </tbody> </table>    |  | 3 | 4 | 5 | 2 | 0.5 | 0.5 | 0.5 | 3 |  | 8.8 | 9.7 | 4 |  |  | 295 | <p><b>average</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>2</th> <td>10.5</td> <td>5.7</td> <td>5.4</td> </tr> <tr> <th>3</th> <td></td> <td>434</td> <td>356</td> </tr> <tr> <th>4</th> <td></td> <td></td> <td>&gt;1k</td> </tr> </tbody> </table>           |  | 3 | 4 | 5 | 2 | 10.5 | 5.7 | 5.4 | 3 |  | 434 | 356 | 4 |  |  | >1k |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 0.9  | 0.9  | 0.9  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 19.3 | 13.4 |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | 0    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 0.5  | 0.5  | 0.5  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 8.8  | 9.7  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | 295  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
|                                                                                                                                                                                                                                                                                                                                                                                       | 3    | 4    | 5    |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                     | 10.5 | 5.7  | 5.4  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                     |      | 434  | 356  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                     |      |      | >1k  |   |   |      |      |      |   |  |      |      |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                   |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                      |  |   |   |   |   |      |     |     |   |  |     |     |   |  |  |     |                                                                                                                                                                                                                                                                                                                                                                                |  |   |   |   |   |     |     |     |   |  |     |     |   |  |  |     |

Figure 4: The simplicial matrix of 10 real networks, as well as the cell-wise average matrix. For each graph  $G$ , only non-empty cells of  $M_{SR}(G)$  are shown, and cells involving edges of size greater than 5 are omitted. The value of a cell is replaced with “> 1k” whenever the value is above 1000.

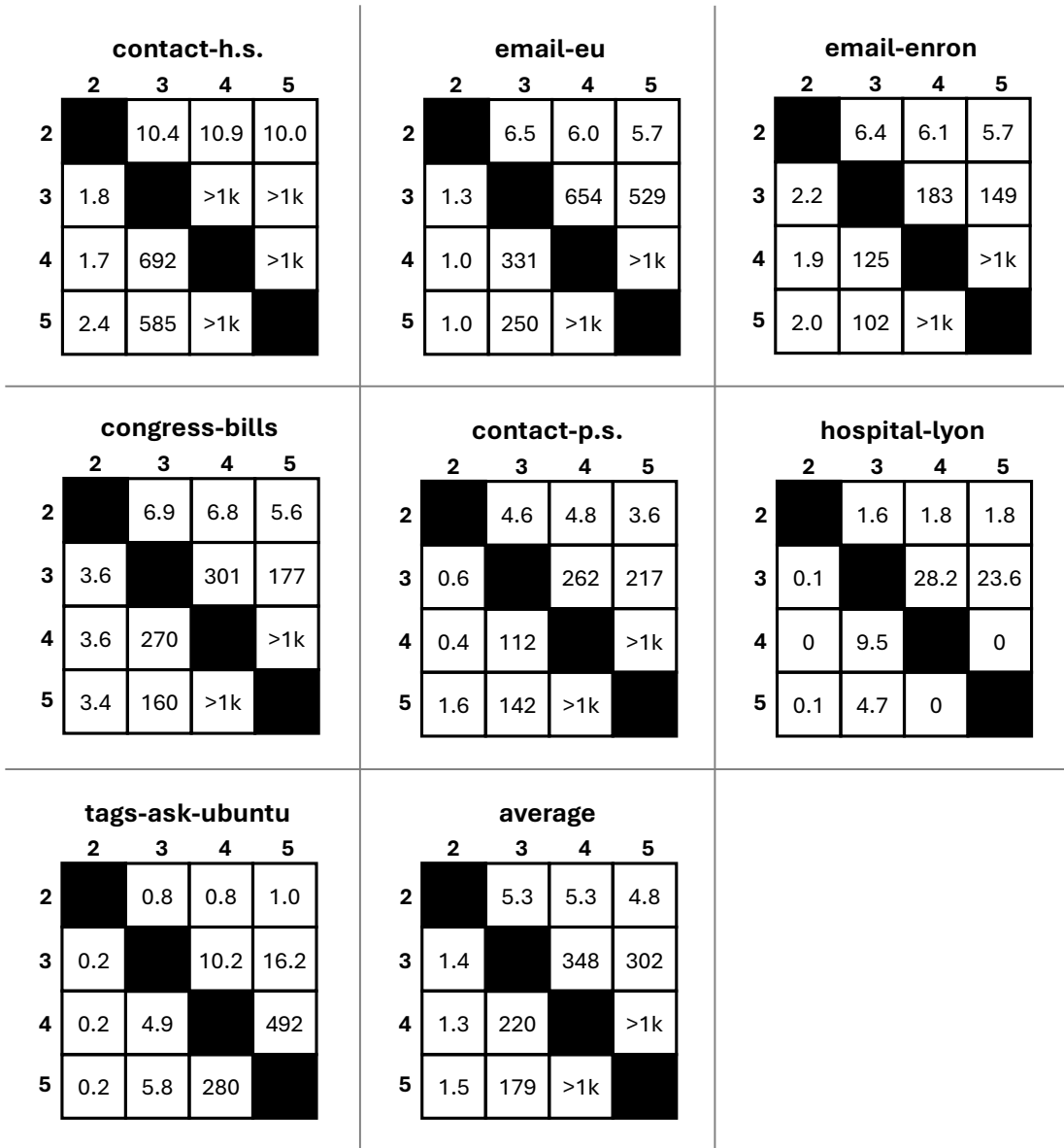


Figure 5: The temporal simplicial matrix of 7 real networks, as well as the cell-wise average matrix. For each graph  $G$ , only non-empty cells of  $M_{\text{SR}}^{\rightarrow}(G)$  are shown, and cells involving edges of size greater than 5 are omitted. The value of a cell is replaced with “> 1k” whenever the value is above 1000.

## 3.2 Analysis

### Simplicial ratio

Based on our results, we see that that biology networks are, on average, more surprisingly simplicial than contact-based networks and email networks. In contrast, it was shown in [22] that contact-based networks are the closest to their simplicial closures and biological networks

are furthest from theirs. In fact, comparing the ranks of the 3 existing measures (sf, es, fes) and the ranks from our simplicial ratio (sr), we get the following Kendall correlation values.

|     | sf     | es     | fes    | sr     |
|-----|--------|--------|--------|--------|
| sf  | 1.000  | 0.706  | 0.989  | -0.270 |
| es  | 0.706  | 1.000  | 0.722  | -0.256 |
| fes | 0.989  | 0.722  | 1.000  | -0.289 |
| sr  | -0.270 | -0.256 | -0.289 | 1.000  |

These values show that our ranking system is negatively correlated with the ranking systems in [22]. A partial explanation for this correlation is that (a) the measures behave differently under different regimes of edge density and (b) the 10 datasets cover a wide range of edge density.

### Bottom-up and top-down simplicial ratios

In our testing, we find that every temporal graph contains more bottom-up simplicial pairs than top-down simplicial pairs. This suggests that, in general for many real networks, a small edge leading to a larger (superset) edge is more common than a large edge leading to a smaller (subset) edge. However, this result is *heavily* biased on our choice of keeping only the first instance of an edge. To see this bias, let  $G$  be a temporal graph with edges  $e_1, e_2 \in E(G)$  such that  $e_1 \subset e_2$  and suppose that  $e_1$  appears with multiplicity 5 and that  $e_2$  appears with multiplicity 1. Then there are 6 possible birth orderings for  $e_2$  and the 5 copies of  $e_1$ , and only 1 such ordering sees  $e_2$  born before  $e_1$ . In most of the temporal networks analysed, the highest frequency of multi-edges are indeed 2-edges, and hence this bottom-up trend is at least partly explained by the above discussion. The topic of temporal simpliciality is one that we intend on exploring further in future works.

### Simplicial matrix

Arguably the most immediate take-away from these matrices is that simplicial interactions become more surprising as edge size increases. Although this feature is interesting, there is at least a partial explanation for this phenomenon that we explore in the following example.

*Example 3.1.* Let  $n \in \mathbb{N}$ ,  $\mathbf{d}$  be a uniform degree sequence, and let  $\mathbf{m} = (m_2, \dots, m_5)$  be a sequence of edge sizes with  $m_2 = m_3 = m_4 = m_5 = n$ . Now let  $G \sim \text{CL}(n, \mathbf{p}, \mathbf{m})$ , and let  $e_2, e_3, e_4, e_5 \in E(G)$  be chosen uniformly at random conditioned on  $|e_i| = i$  for each  $i \in \{2, 3, 4, 5\}$ . Then, writing  $X_{i,j}$  for the indicator variable which is 1 if  $e_i \subset e_j$ , we have

$$\begin{array}{c|c|c}
 \mathbb{E}[X_{2,3}] \propto n^{-2} & \mathbb{E}[X_{2,4}] \propto n^{-2} & \mathbb{E}[X_{2,5}] \propto n^{-2} \\
 \hline
 & \mathbb{E}[X_{3,4}] \propto n^{-3} & \mathbb{E}[X_{3,5}] \propto n^{-3} \\
 \hline
 & & \mathbb{E}[X_{4,5}] \propto n^{-4}
 \end{array}$$

which implies

$$\begin{array}{c|c|c} \mathbb{E} [\text{sp}(G, 2, 3)] \propto 1 & \mathbb{E} [\text{sp}(G, 2, 4)] \propto 1 & \mathbb{E} [\text{sp}(G, 2, 5)] \propto 1 \\ \hline & \mathbb{E} [\text{sp}(G, 3, 4)] \propto 1/n & \mathbb{E} [\text{sp}(G, 3, 5)] \propto 1/n \\ \hline & & \mathbb{E} [\text{sp}(G, 4, 5)] \propto 1/n^2 \end{array}$$

Now, let  $H$  be a graph with degree sequence  $\mathbf{d}$  and edge-size sequence  $\mathbf{m}$ , and suppose  $H$  has one simplicial pair of each type. Then, based on the above calculations, we get that

$$\begin{array}{c|c|c} \sigma_{\text{SR}}(H, 2, 3) \propto 1 & \sigma_{\text{SR}}(H, 2, 4) \propto 1 & \sigma_{\text{SR}}(H, 2, 5) \propto 1 \\ \hline & \sigma_{\text{SR}}(H, 3, 4) \propto n & \sigma_{\text{SR}}(H, 3, 5) \propto n \\ \hline & & \sigma_{\text{SR}}(H, 4, 5) \propto n^2 \end{array}$$

Thus, the above matrix acts as a loose, point-wise lower-bound on the simplicial matrix for sparse graphs with at least one simplicial pair of each type. For many of the graphs analysed, this rough sketch of a simplicial matrix is a good approximation of the actual matrices. In summary, what the simplicial matrix is capturing, above all else, is that (a) real graphs contain simplicial pairs of all types, and (b) synthetic (sparse) models very rarely generate simplicial pairs other than pairs containing 2-edges.

### Temporal simplicial matrix

Here, the bias towards bottom-up simplicial pairs is consistent with the cell-wise comparisons. This suggests that the bias is independent, or at least not heavily dependent, on edge size.

## 4 A new model that incorporates simpliciality

In this section, we define a random graph model, called the *simplicial Chung-Lu model*, that generalizes the Chung-Lu hypergraph model defined in [14]. We begin with the algorithm that generates a simplicial edge.

Let  $(d_1, \dots, d_n)$  be a degree sequence,  $k$  be an edge size,  $E$  be a set of existing edges, and  $E_k \subseteq E$  be a set of existing edges that are of size  $k$ . Recalling that  $p(\cdot)$  is the probability distribution governed by  $(d_1, \dots, d_n)$ , writing  $\binom{S}{k}$  for the collection of  $k$ -subsets of  $S$ , and recalling that  $x \in_u X$  means  $x$  is sampled uniformly from  $X$ , the algorithm to generate a simplicial edge is as follows.



---

**Algorithm 3** Simplicial edge.

---

**Require:**  $(d_1, \dots, d_n)$ ,  $k$ ,  $E$ .

```
1: if  $E \setminus E_k = \emptyset$  then
2:   Sample  $e \sim$  Algorithm 1 $((d_1, \dots, d_n), k)$ 
3: else
4:   Sample  $e' \in_u E \setminus E_k$ .
5:   if  $|e'| < k$  then
6:     Sample  $e'' \sim$  Algorithm 1 $((d_1, \dots, d_n), k - |e'|)$ .
7:     Set  $e = e' \cup e''$ 
8:   else
9:     Sample  $e \in_u \binom{e'}{k}$ 
10:  end if
11: end if
12: Return  $e$ 
```

---

In words, we first check if there is at least one edge in  $E$  *not* of size  $k$  to pair  $e$  with. If there is no such edge, we return a Chung-Lu edge. Otherwise, we choose an existing edge  $e'$  uniformly at random from the set of edges *not* of size  $k$  and construct our edge  $e$  from  $e'$  in one of two ways: if  $k < |e'|$  we set  $e$  to be a uniform  $k$ -subset of  $e'$ , whereas if  $k > |e'|$  we build  $e$  by combining  $e'$  with a Chung-Lu edge of size  $k - |e'|$ .

We now give the algorithm to generate a simplicial Chung-Lu graph. Let  $(d_1, \dots, d_n)$  be a degree sequence,  $(m_{k_{\min}}, \dots, m_{k_{\max}})$  be a sequence of edge sizes, and  $S = (s_1, \dots, s_\ell)$  be a random permutation of all available sizes for an edge, i.e.,  $S$  contains  $m_k$  copies of  $k$  for each edge size  $k$  in some random order. Additionally, let  $q \in [0, 1]$  be a parameter governing the number of simplicial edges created during the process.

---

**Algorithm 4** Simplicial Chung Lu model.

---

**Require:**  $(d_1, \dots, d_n)$ ,  $(m_{k_{\min}}, \dots, m_{k_{\max}})$ ,  $q$ .

```
1: Initialize edge list  $E = \{\}$  and random edge-size list  $S$ .
2: for  $k \in S$  do
3:   Sample  $X \sim \text{Bernoulli}(q)$ .
4:   if  $X = 1$  then
5:     Sample  $e \sim$  Algorithm 3 $((d_1, \dots, d_n), k, E)$ 
6:   else
7:     Sample  $e \sim$  Algorithm 1 $((d_1, \dots, d_n), k)$ 
8:   end if
9:   Set  $E = E \cup \{e\}$ .
10: end for
11: Return  $G = ([n], E)$ 
```

---

Note that, if  $q = 0$ , the simplicial Chung-Lu model yields a Chung-Lu model, ensuring

that this new model is indeed a generalized Chung-Lu model. Moreover, the following lemma shows that the main feature of the Chung-Lu model is still present in this new model.

**Lemma 4.1.** *Let  $G$  be a random graph generated as a simplicial Chung-Lu model with input parameters  $(d_1, \dots, d_n)$ ,  $(m_{k_{\min}}, \dots, m_{k_{\max}})$ , and  $q \in [0, 1]$ . Then, for all  $v \in [n]$ ,*

$$\mathbb{E}[\deg_G(v)] = d_v.$$

*Proof.* Let us generate a random edge-size list  $S$  that will be used to create the simplicial Chung-Lu graph  $G$ . We will first prove (by induction on  $i$ ) the following claim.

**Claim:** Each vertex  $v$  of the  $i$ 'th edge  $e_i$  formed during the construction process of  $G$  satisfies

$$\mathbb{P}(v = u) = p(u) \text{ for all } u \in [n].$$

Note that edges of  $G$  are not generated independently; the graph has rich dependence structure. The distribution of  $e_i$  is affected by edges generated earlier. It is important to keep in mind that the claim applies to the edge formed at time  $i$  but without exposing information about earlier edges.

Firstly, if  $i = 1$ , then  $e_1$  is necessarily generated via **Algorithm 1** and the claim follows immediately. Now fix  $i > 1$  and consider the formation of  $e_i$ . On the one hand, if  $e_i$  was generated via **Algorithm 1** then the claim is once again immediate. Otherwise,  $e_i$  was generated via **Algorithm 3**, i.e., generated constructively from another edge  $e_j$  with  $j < i$ . In this case, if  $|e_i| < |e_j|$  then  $e_i \in_u \binom{e_j}{|e_i|}$  and, regardless which subset of  $e_j$  is selected to form  $e_i$ , the claim holds by induction. Otherwise, if  $|e_i| > |e_j|$ , then  $e_i$  is the union of  $e_j$  and another edge  $e''$  generated via **Algorithm 1**: the claim holds immediately for vertices in  $e''$ , and for vertices in  $e_j$ , the claim holds by induction.

Thus, for any  $e \in E(G)$ ,  $v \in e$ , and  $u \in [n]$ , we have that  $\mathbb{P}(v = u) = p(u)$ . Summing over all vertices in all edges, we get that

$$\mathbb{E}[\deg_G(u)] = \left( \sum_{e \in E(G)} \sum_{v \in e} \mathbb{P}(v = u) \right) = \left( p(u) \sum_{e \in E(G)} |e| \right) = \left( p(u) \sum_{v \in [n]} d_v \right) = d_u,$$

the first equality following from linearity of expectation, and the third equality following from the generalized handshaking lemma.  $\square$

The simplicial Chung Lu model does in fact generate more simplicial pairs as  $q$  increases. Figure 6 shows the expected number of simplicial pairs (approximated via 1000 samples) for graphs generated via **Algorithm 4** with  $q$  varying from 0 to 1 in 0.1 increments.

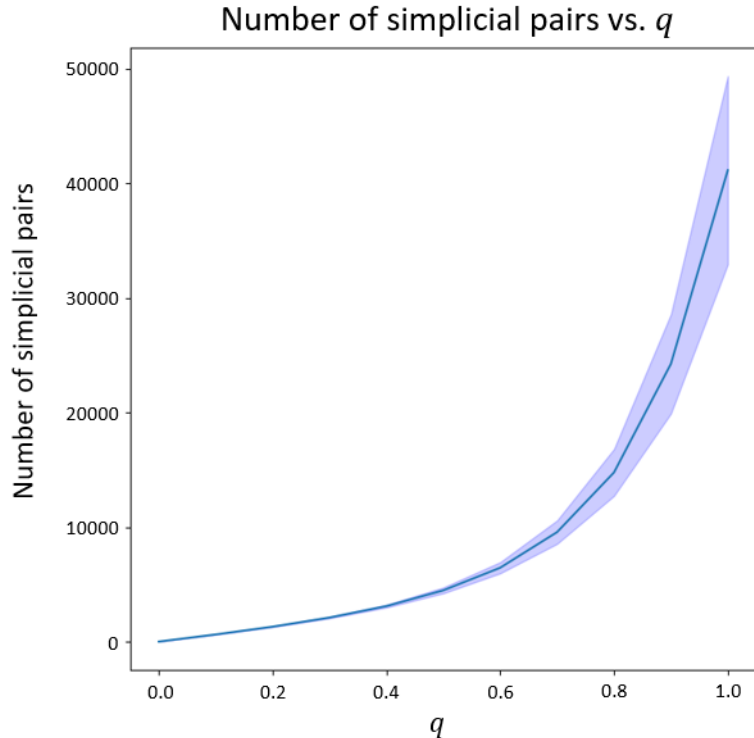


Figure 6: The average number of simplicial pairs (taken over 1000 samples) for simplicial Chung Lu graphs with varying  $q$ . For each  $q \in [0, 0.1, \dots, 1]$ ,  $G_q$  is a simplicial Chung Lu graph with  $n = 1000$ ,  $\mathbf{d}$  a uniform degree sequence, and  $[|E_2|, |E_3|, |E_4|, |E_5|] = [5000, 1000, 100, 10]$ . The shaded region represents the standard deviation over the 1000 samples.

## 5 Experiments

One reason to study simpliciality is that it likely has an impact on the evolution of stochastic processes on the associated graphs. We illustrate this potential impact via two toy processes with varying parameters. The first process is component growth which is a standard way to measure the robustness of a network (see, for example, Chapter 8 in [1]). The second type is information diffusion which simulates how quickly a substance (e.g., a disease, a rumour) spreads through a network. Intuitively, both of these processes should be affected by a graph containing a large number of simplicial pairs: in the case of component growth the smaller edge of a simplicial pair does not contribute to component size, and in the case of information diffusion a simplicial pair transfers information less efficiently than two non-overlapping edges.

## 5.1 Descriptions of the experiments

We perform four experiments (two experiments for each of the two types of stochastic processes) on the real networks and on the corresponding simplicial Chung-Lu graphs for varying  $q \in \{0, 0.5, 1\}$ .

**Giant component growth with random edge selection:** We choose a uniform random order for  $E(G)$  and track the size of the largest component as edges are added to  $G$  according to a random ordering. We plot the growth up to the point where  $\min\{|E(G)|, |V(G)|\}$  edges have been added. We perform this experiment independently 100 times on the real graphs, meaning we shuffle the edge ordering and track the growth 100 times. For the simplicial Chung-Lu models we (a) sample the graph, (b) shuffle the edges, and (c) track the growth, performing steps (a), (b), and (c) independently 100 times.

**Giant component growth with adversarial edge selection:** We order  $E(G)$  in ascending order of betweenness (breaking ties randomly) and track the size of the largest component as edges are added to  $G$  according to this adversarial ordering. Note that the betweenness of an edge  $e$  in a hypergraph is equivalent to the betweenness of its corresponding vertex  $v_e$  in the line graph (see [10], or any textbook on network science such as [16], for a definition of betweenness for graphs). For the real graphs, we run the experiment only once (the results will be the same every time), and for the Chung-Lu models we sample and track growth 10 times. We sample significantly less here than in the other three experiments due to the time complexity of calculating betweenness.

**Information diffusion from a single source:** We initialize a function  $w_0 : V(G) \rightarrow [0, 1]$  with  $w_0(v) = 0$  for all vertices, except for one randomly chosen vertex  $v^*$  which has  $w_0(v^*) = 1$ . Then, in round  $i + 1$ , we choose a random edge  $e$  and, for each  $v \in e$ , set  $w_{i+1}(v) = w(e)/|e|$ , where  $w(e) = \sum_{u \in e} w(u)$  (keeping  $w_{i+1}(v) = w_i(v)$  for all  $v \notin e$ ). We track the Wasserstein-1 distance (also known as the “earth mover’s distance” [27]) between  $w_i$  and the uniform distribution  $w_\infty : V(G) \rightarrow 1/|V(G)|$ . We run the experiment 100 times, re-rolling the Chung-Lu model every time.

**Information diffusion from  $|V(G)|/10$  sources:** This experiment is identical to the previous experiment, except that  $w_0(v^*) = 1$  for 10% of the vertices chosen at random, and that  $w_\infty : V(G) \rightarrow 1/10$ .

### Insisting on connected graphs

These experiments, and in particular the two diffusion experiments, are highly dependent on connectivity. The real graphs are restricted to their largest component, and so we insist that the random graphs are also connected. To achieve this, we modify the simplicial Chung-Lu model and insist that incoming edges must connect disjoint components, until the point the graph is connected when we continue generating edges as normal. A full description of this algorithm is presented in Appendix B.

## 5.2 The results

Here, we will show the results for the two graphs: **hospital-lyon** and **disgenenet**. Recall that the **hospital-lyon** graph has a simplicial ratio of approximately 0.94, whereas the **disgenenet** graph has a ratio of approximately 28.81. The full collection of results can be found in Appendix A <sup>1</sup> and the sampling technique can be found in Appendix B.

### Experiment 1: random growth

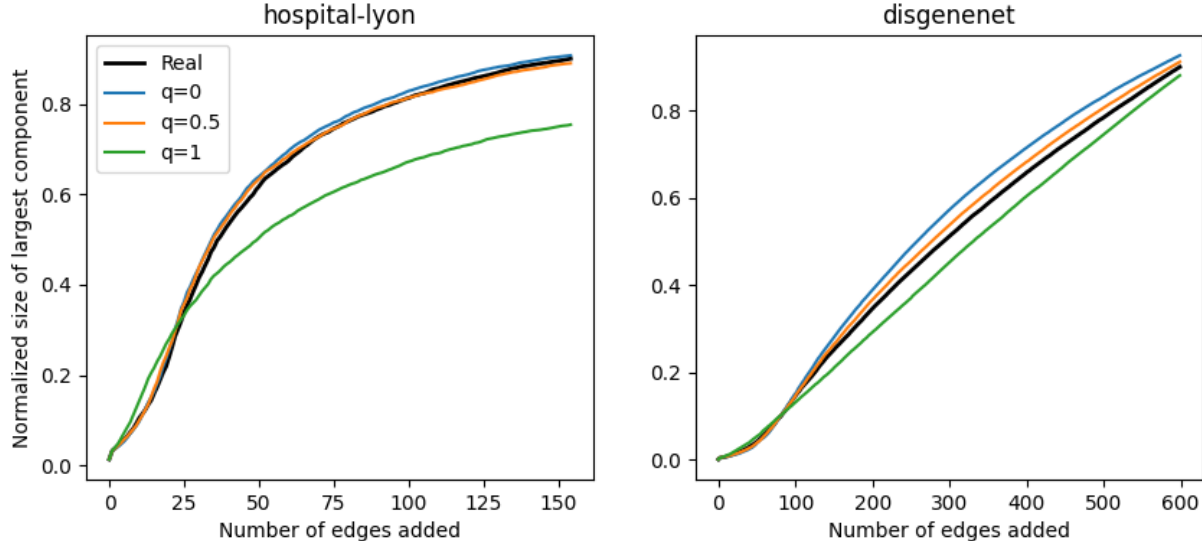


Figure 7: Giant component size (normalized by the number of vertices) vs. number of edges added in the random growth process on the **hospital-lyon** graph (left) and the **disgenenet** graph (right). The curve is the point-wise average across 100 independent experiments: for the real graph the edges are resampled each time, and for the random models the entire graphs are resampled each time.

In this first experiment we see the following. For **hospital-lyon** the real graph grows in a near identical way to the random model with  $q = 0$  and  $q = 0.5$ , whereas the random model with  $q = 1$  grows much slower. In contrast, for **disgenenet** the real graph grows at a rate somewhere between the random model with  $q = 0.5$  and  $q = 1$ . Of course, these graphs have very different growth behaviour due to the difference in edge densities. Nevertheless, this result suggests that the high simplicial ratio of **disgenenet** plays a role in slowing down the growth of the graph, whereas the low simplicial ratio of **hospital-lyon** leads it to grow as quickly as a classical Chung-Lu model.

<sup>1</sup>Due to computational limitations, we do not present the adversarial growth experiment for all 10 networks.

## Experiment 2: adversarial growth

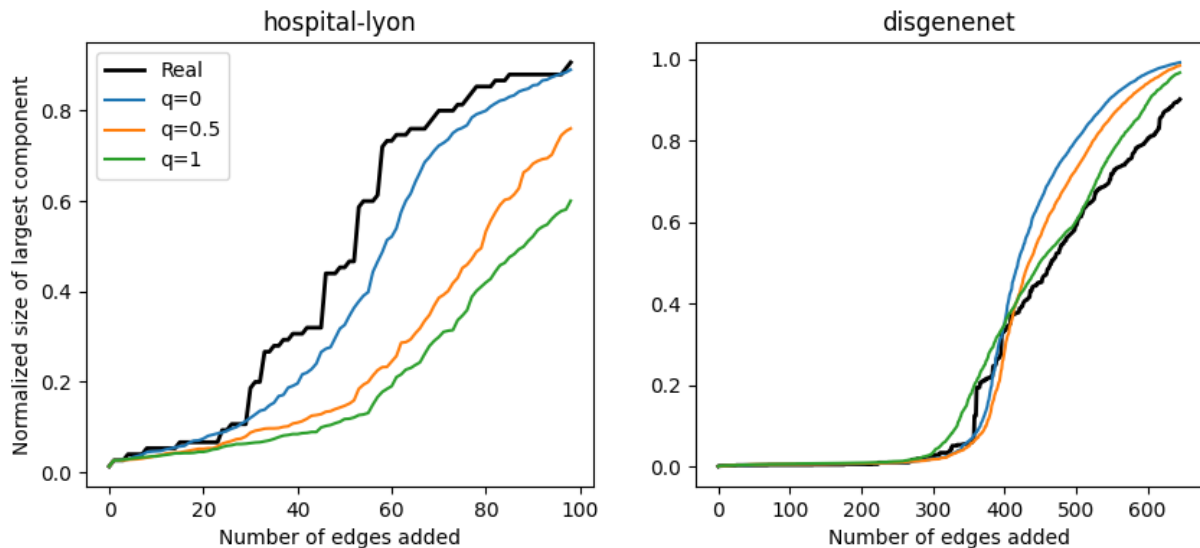


Figure 8: Giant component size vs. number of edges added in the adversarial growth process on the **hospital-lyon** graph (left) and the **disgenenet** graph (right). The curve is the point-wise average across 10 independent experiments: for the real graph the experiment is performed only once as the result will always be the same, and for the random models the graphs are resampled each time.

The results of this second experiment, adversarial growth, are less clear due to the fact that we averaged over 10 samples instead of 100. Nonetheless, there is still a clear distinction between the real growth vs. the synthetic growth for these two graphs. On the left, we see that the real graph grows faster than all the random models, whereas on the right the real graph grows slower than all the models.

### Experiment 3: single-source diffusion

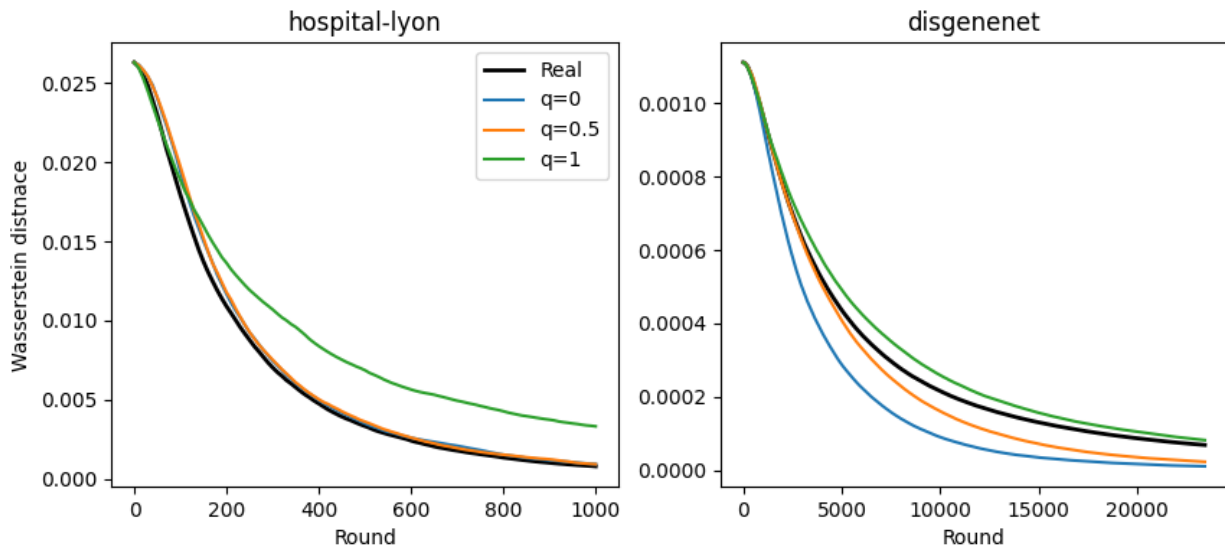


Figure 9: Wasserstein distance to uniform vs. number of rounds in the single-source diffusion process on the **hospital-lyon** graph (left) and the **disgenenet** graph (right). The curve is the point-wise average across 100 independent experiments: for the real graph the chosen edges per round, as well as the location of the initial vertex with weight 1, are resampled each time, and for the random models the entire graphs are resampled each time.

This experiment is perhaps the most substantial in showing the effect of simpliciality on a random process, namely, that information diffusion is slower on highly simplicial graphs vs. non-simplicial graphs.

## Experiment 4: 10% diffusion

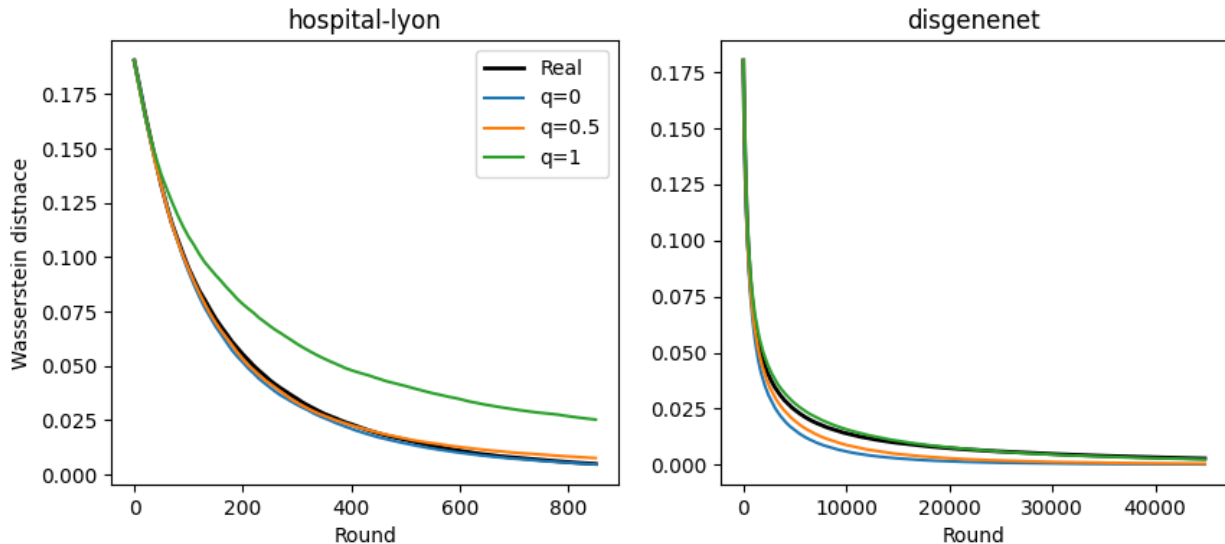


Figure 10: Wasserstein distance to uniform vs. number of rounds in the 10% sprinkled diffusion process on the **hospital-lyon** graph (left) and the **disgenenet** graph (right). The curve is the point-wise average across 100 independent experiments: for the real graph the chosen edges per round, as well as the location of the initial 10% of vertices with weight 1, are resampled each time, and for the random models the entire graphs are resampled each time.

The result here mirrors the result in the previous experiment, except of course that the diffusion is much faster.

## 6 Conclusion

The phenomenon of edges inside of other edges is a feature of hypergraphs not present in graphs and, based on our results and on the preceding results of Landry, Young and Eikmeier, it is clear that this phenomenon is a key feature of real-world networks with multi-way interactions. The simplicial ratio captures the strength of simplicial interactions in a graph and, from the collection of 10 real-world networks analysed, we have showed that (a) the simplicial ratio is not at all consistent across the graphs, (b) the simplicial ratio varies significantly even for graphs of a similar type (e.g., **contact high-school**, **contact primary-school**, and **hospital-lyon**), (c) the number of simplicial interactions involving edges of size  $k, \ell > 2$  is not at all captured by the Chung Lu model, and (d) the simplicial ratio can affect the outcome of random growth, adversarial growth, and information diffusion. We hope that our work continues to motivate research into the phenomenon of edges inside edges, and we discuss some potential follow ups to this research.



## 6.1 Further research

The simplicial ratio involves the parameter  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right]$  where  $\hat{G} \sim \text{CL}(G)$ . Instead of approximating  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right]$  as we do, one could compute  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right]$  explicitly. For example, given a uniform degree sequence  $\mathbf{d}$  and edge-size sequence  $(m_{k_{\min}}, \dots, m_{k_{\max}})$ , and conditioning on  $\hat{G}$  containing no multiset edges, the probability that  $e_1, e_2$  form a simplicial pair is

$$\frac{\binom{|e_2|}{|e_1|}}{\binom{n}{|e_1|}}.$$

Thus, by linearity of expectation, conditioning on the event that  $\hat{G}$  has no multiset edges, we have

$$\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right] = \sum_{k=k_{\min}}^{k_{\max}-1} \sum_{\ell=k+1}^{k_{\max}} m_k m_\ell \binom{\ell}{k} / \binom{n}{k}.$$

Thus, for a uniform degree sequence,  $\mathbb{E} \left[ \text{sp} \left( \hat{G} \right) \right]$  is relatively straightforward to compute. However, trying to compute this expectation if the degree sequence is not uniform is significantly harder. Finding a closed form for this expectation, or even a closed form approximation, would allow for a significantly faster algorithm for computing  $\sigma_{\text{SR}}(G)$ . Such a result would also allow for a better understanding of the nature of the simplicial matrix for both sparse and dense graphs.

Understanding the degree to which edges form simplicial pairs could aid in predicting the composition of future edges, especially large edges, in temporal networks. If a graph has a high simplicial ratio, then a potential new edge should be given more weight based on the number of new simplicial pairs it creates, as well as on the size of the smaller edge in each pairs. For example, when considering the location for a new edge of size 5 in a highly simplicial graph  $G$ , a location that creates many  $(2, 5)$  pairs should be given more weight, but perhaps a location that creates a single  $(4, 5)$  pair should be given *even more* weight. In any case, incorporating simpliciality in the link prediction problem should improve existing algorithms, at least for highly simplicial graphs.

Along with the simplicial ratio and simplicial matrix, we introduce temporal variants. In our experiments where only the first instance of an edge is kept in a temporal network, we find that, typically, more bottom-up pairs are generated than top-down pairs, in part because there are more small multi-edges than large multi-edges. There are of course other ways to measure the difference in frequency between bottom-up pairs and top-down pairs. For example, we could insist that a simplicial pair  $e_k, e_\ell$  is “temporally relevant” if and only if both  $e_k$  and  $e_\ell$  were born within the same  $\epsilon$ -window of time. In this case, we could measure the frequency of  $e_k$  pairs followed shortly by  $e_\ell$  pairs, and vice versa. The temporal formation of simplicial pairs could once again be valuable for the task of link prediction.

## Acknowledgements

This research is supported in part by the Tutte Institute for Mathematics and Computing (TIMC), a division of the Communications Security Establishment (CSE). TIMC does not specifically endorse the contents of this work, or any other work by the authors. Any opinions or positions represented herein do not represent the official position of CSE or the Government of Canada.

## References

- [1] Albert-László Barabási and Márton Pósfai. *Network science*. Cambridge University Press, Cambridge, 2016. URL: <http://barabasi.com/networksciencebook/>.
- [2] Jordan Barret, Paweł Prałat, Aaron Smith, and François Théberge. Counting simplicial pairs in hypergraphs. In *International Conference on Complex Networks and Their Applications*, page in press. Springer, 2024.
- [3] Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri. Networks beyond pairwise interactions: structure and dynamics. *Physics Reports*, 874:1–92, 2020.
- [4] Austin R Benson, Rediet Abebe, Michael T Schaub, Ali Jadbabaie, and Jon Kleinberg. Simplicial closure and higher-order link prediction. *Proceedings of the National Academy of Sciences*, 115(48):E11221–E11230, 2018.
- [5] Austin R Benson, David F Gleich, and Desmond J Higham. Higher-order network analysis takes off, fueled by classical ideas and new data. *arXiv preprint arXiv:2103.05031*, 2021.
- [6] Austin R Benson, David F Gleich, and Jure Leskovec. Higher-order organization of complex networks. *Science*, 353(6295):163–166, 2016.
- [7] Fan RK Chung and Linyuan Lu. *Complex graphs and networks*. Number 107. American Mathematical Soc., 2006.
- [8] David Easley and Jon Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge university press, 2010.
- [9] Song Feng, Emily Heath, Brett Jefferson, Cliff Joslyn, Henry Kvinge, Hugh D Mitchell, Brenda Praggastis, Amie J Einfeld, Amy C Sims, Larissa B Thackray, et al. Hypergraph models of biological networks to identify genes critical to pathogenic viral response. *BMC bioinformatics*, 22(1):287, 2021.
- [10] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

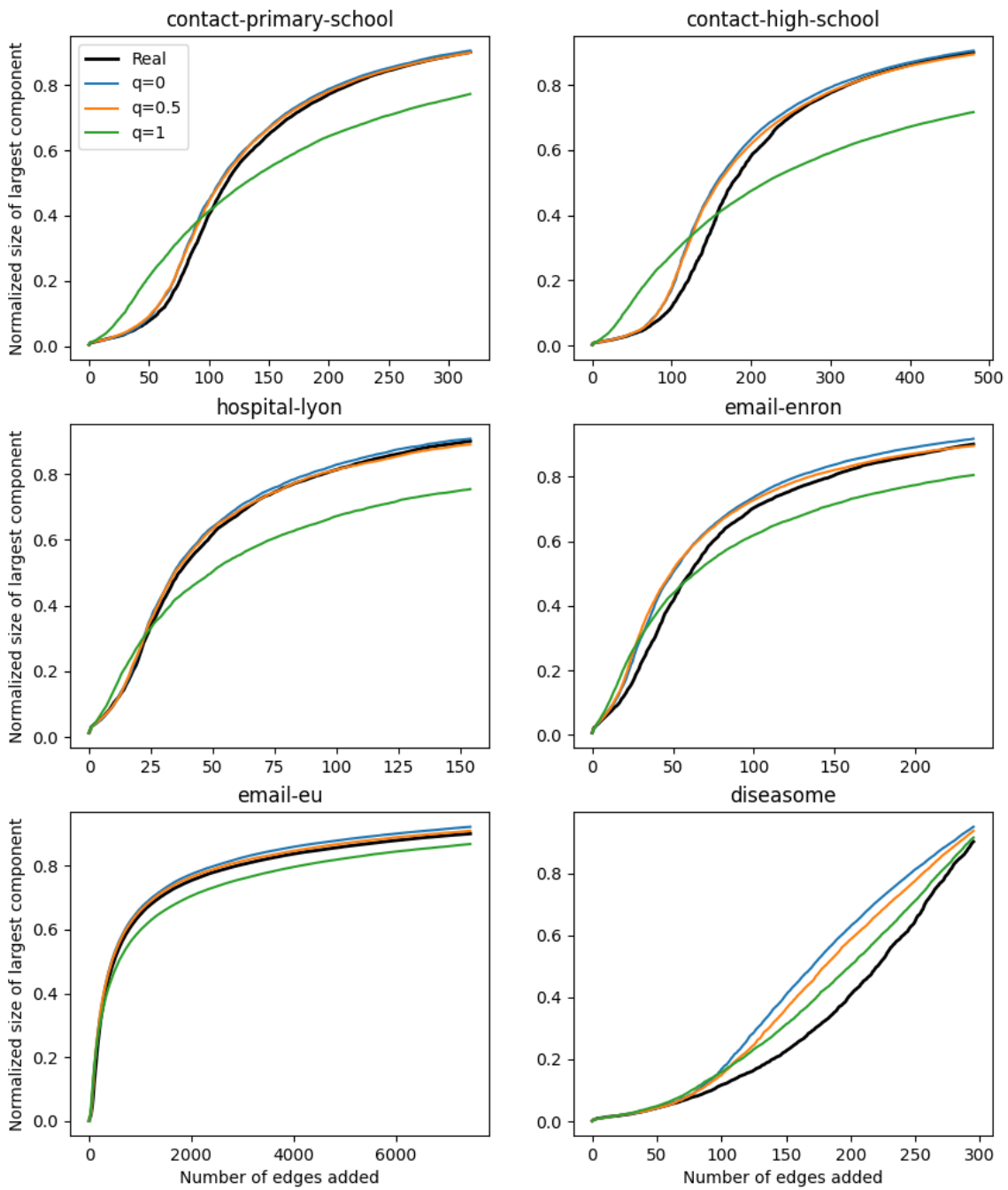
- [11] Matthew O Jackson. *Social and economic networks*. Princeton university press, 2010.
- [12] Jonas L Juul, Austin R Benson, and Jon Kleinberg. Hypergraph patterns and collaboration structure. *Frontiers in Physics*, 11:1301994, 2024.
- [13] Bogumił Kamiński, Łukasz Kraiński, Paweł Prałat, and François Théberge. A multi-purposed unsupervised framework for comparing embeddings of undirected and directed graphs. *Network Science*, 10(4):323–346, 2022.
- [14] Bogumił Kamiński, Valérie Poulin, Paweł Prałat, Przemysław Szufel, and François Théberge. Clustering via hypergraph modularity. *PLoS one*, 14(11):e0224307, 2019.
- [15] Bogumił Kamiński, Paweł Prałat, and François Théberge. An unsupervised framework for comparing graph embeddings. *Journal of Complex Networks*, 8(5):cnz043, 2020.
- [16] Bogumil Kaminski, Pawel Prałat, and François Théberge. *Mining complex networks*. Chapman and Hall/CRC, 2021.
- [17] Bogumił Kamiński, Paweł Prałat, and François Théberge. Hypergraph artificial benchmark for community detection (h-abcd). *Journal of Complex Networks*, 11(4):cnad028, 2023.
- [18] Bogumił Kamiński, Paweł Misiorek, Paweł Prałat, and François Théberge. Modularity based community detection in hypergraphs, 2024. URL: <https://arxiv.org/abs/2406.17556>, arXiv:2406.17556.
- [19] Jihye Kim, Deok-Sun Lee, and K.-I. Goh. Contagion dynamics on hypergraphs with nested hyperedges. *Physical Review E*, 108(3), 2023. URL: <http://dx.doi.org/10.1103/PhysRevE.108.034313>, doi:10.1103/physreve.108.034313.
- [20] Renaud Lambiotte, Martin Rosvall, and Ingo Scholtes. Understanding complex systems: From networks to optimal higher-order models. *arXiv preprint arXiv:1806.05977*, 2018.
- [21] Nicholas W. Landry, Maxime Lucas, Iacopo Iacopini, Giovanni Petri, Alice Schwarze, Alice Patania, and Leo Torres. XGI: A Python package for higher-order interaction networks. *Journal of Open Source Software*, 8(85):5162, May 2023. URL: <https://joss.theoj.org/papers/10.21105/joss.05162>, doi:10.21105/joss.05162.
- [22] Nicholas W Landry, Jean-Gabriel Young, and Nicole Eikmeier. The simpliciality of higher-order networks. *EPJ Data Science*, 13(1):17, 2024.
- [23] Timothy LaRock and Renaud Lambiotte. Encapsulation structure and dynamics in hypergraphs. *Journal of Physics: Complexity*, 4(4):045007, nov 2023. URL: <https://dx.doi.org/10.1088/2632-072X/ad0b39>, doi:10.1088/2632-072X/ad0b39.
- [24] Geon Lee, Fanchen Bu, Tina Eliassi-Rad, and Kijung Shin. A survey on hypergraph mining: Patterns, tools, and generators, 2024. URL: <https://arxiv.org/abs/2401.08878>, arXiv:2401.08878.

- [25] Mark Newman. *Networks*. Oxford university press, 2018.
- [26] Tom Odda. On properties of a well-known graph or what is your ramsey number. *Annals of the New York Academy of Sciences*, 328:166 – 172, 12 2006. doi:10.1111/j.1749-6632.1979.tb17777.x.
- [27] Y. Rubner, C. Tomasi, and L.J. Guibas. A metric for distributions with applications to image databases. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pages 59–66, 1998. doi:10.1109/ICCV.1998.710701.
- [28] Hao Tian and Reza Zafarani. Higher-order networks representation and learning: A survey. *ACM SIGKDD Explorations Newsletter*, 26(1):1–18, 2024.
- [29] Leo Torres, Ann S. Blevins, Danielle Bassett, and Tina Eliassi-Rad. The why, how, and when of representations for complex systems. *SIAM Review*, 63(3):435–485, 2021. doi:10.1137/20M1355896.
- [30] Dominic Yeo. Multiplicative coalescence. URL: <https://api.semanticscholar.org/CorpusID:2555801>.
- [31] Yuanzhao Zhang, Maxime Lucas, and Federico Battiston. Higher-order interactions shape collective dynamics differently in hypergraphs and simplicial complexes. *Nature Communications*, 14(1), 2023. doi:10.1038/s41467-023-37190-9.

## A All experiments

Here we show the results of the random growth, single-source diffusion, and 10% diffusion experiments. Due to the time complexity of computing edge-betweenness, we are unable to perform the adversarial growth experiment for all 10 graphs. Note that **ubuntu (edge-chopped)** is the subgraph of **tags-ask-ubuntu** containing only the first 20000 edges. The simplicial ratio of this edge-chopped graph is  $\approx 0.37$  and so this subgraph is even less simplicial than the whole graph.

The experiments are presented in the the following order: random growth, single-source diffusion, and 10% diffusion. Each of the three figures are presented on two separate pages.



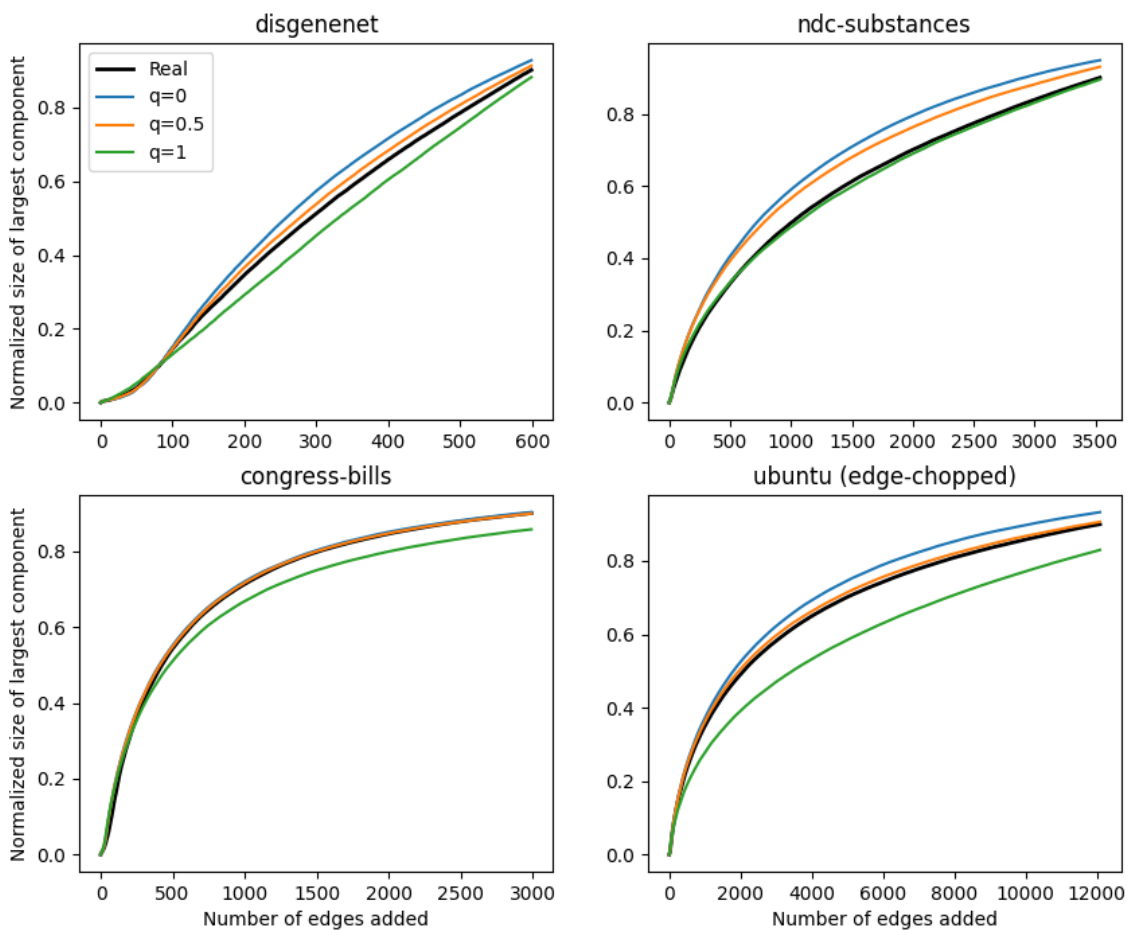
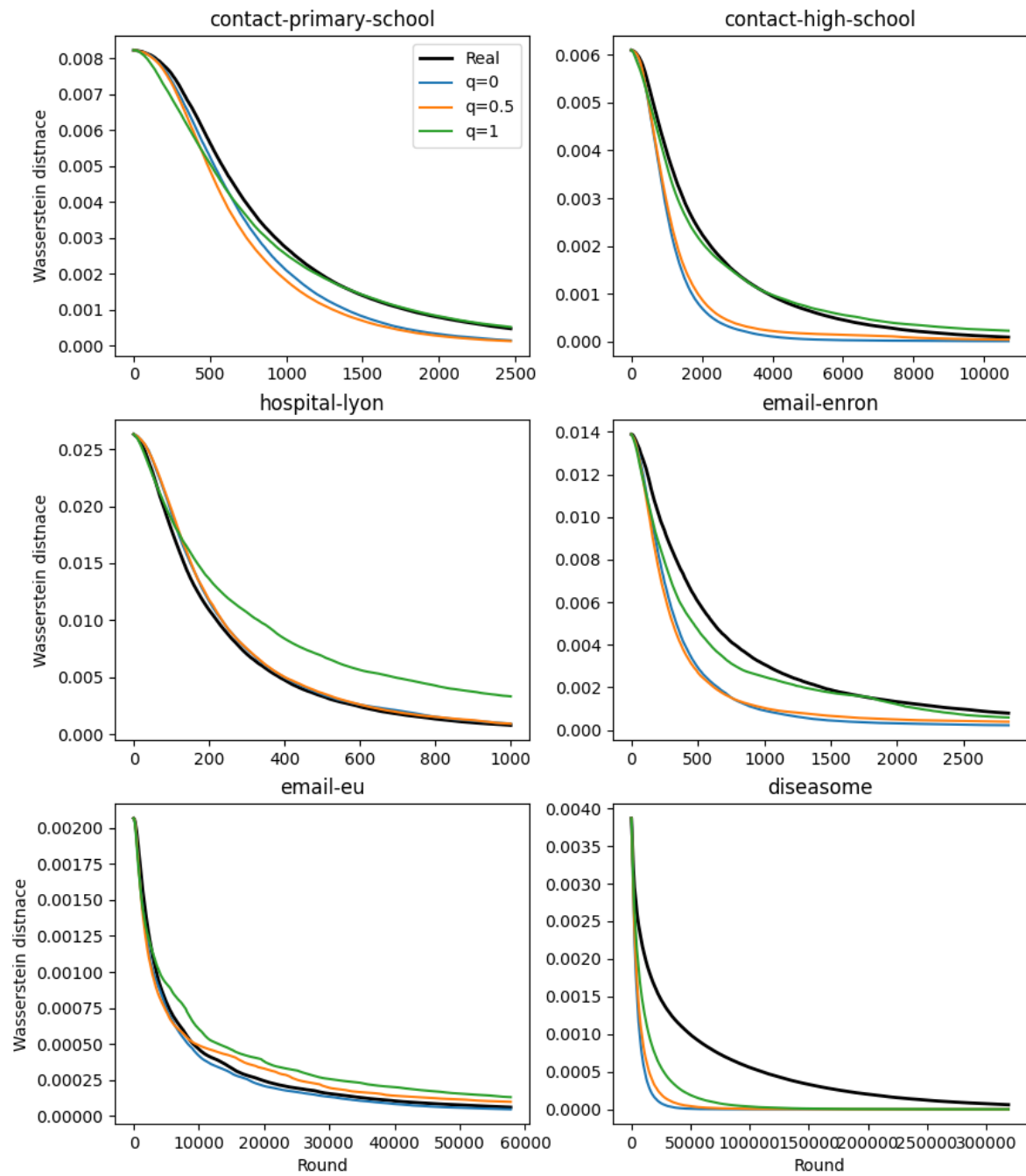


Figure 11: Giant component size (normalized by the number of vertices) vs. number of edges added in the random growth process for all 10 graphs. The curve is the point-wise average across 100 independent experiments: for the real graph the edges are resampled each time, and for the random models the entire graphs are resampled each time.



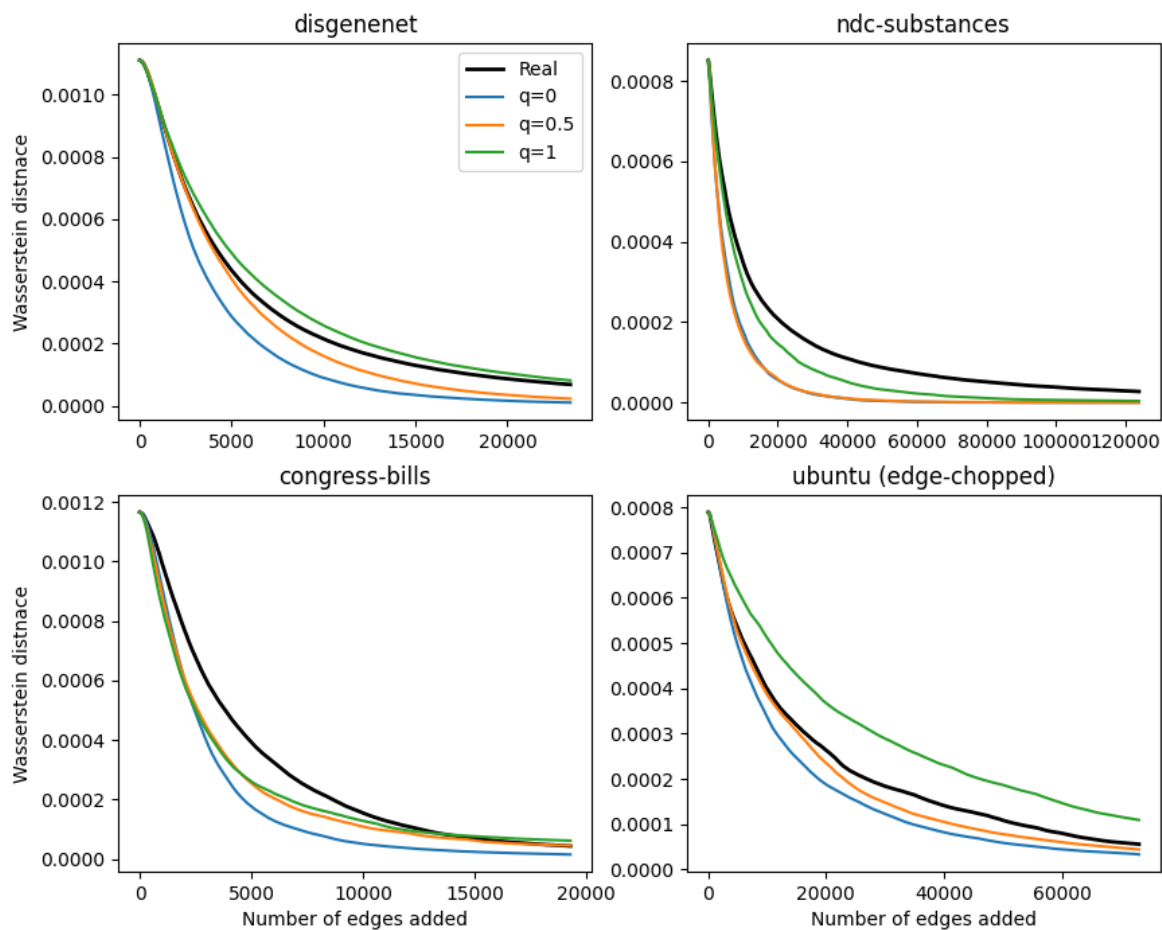
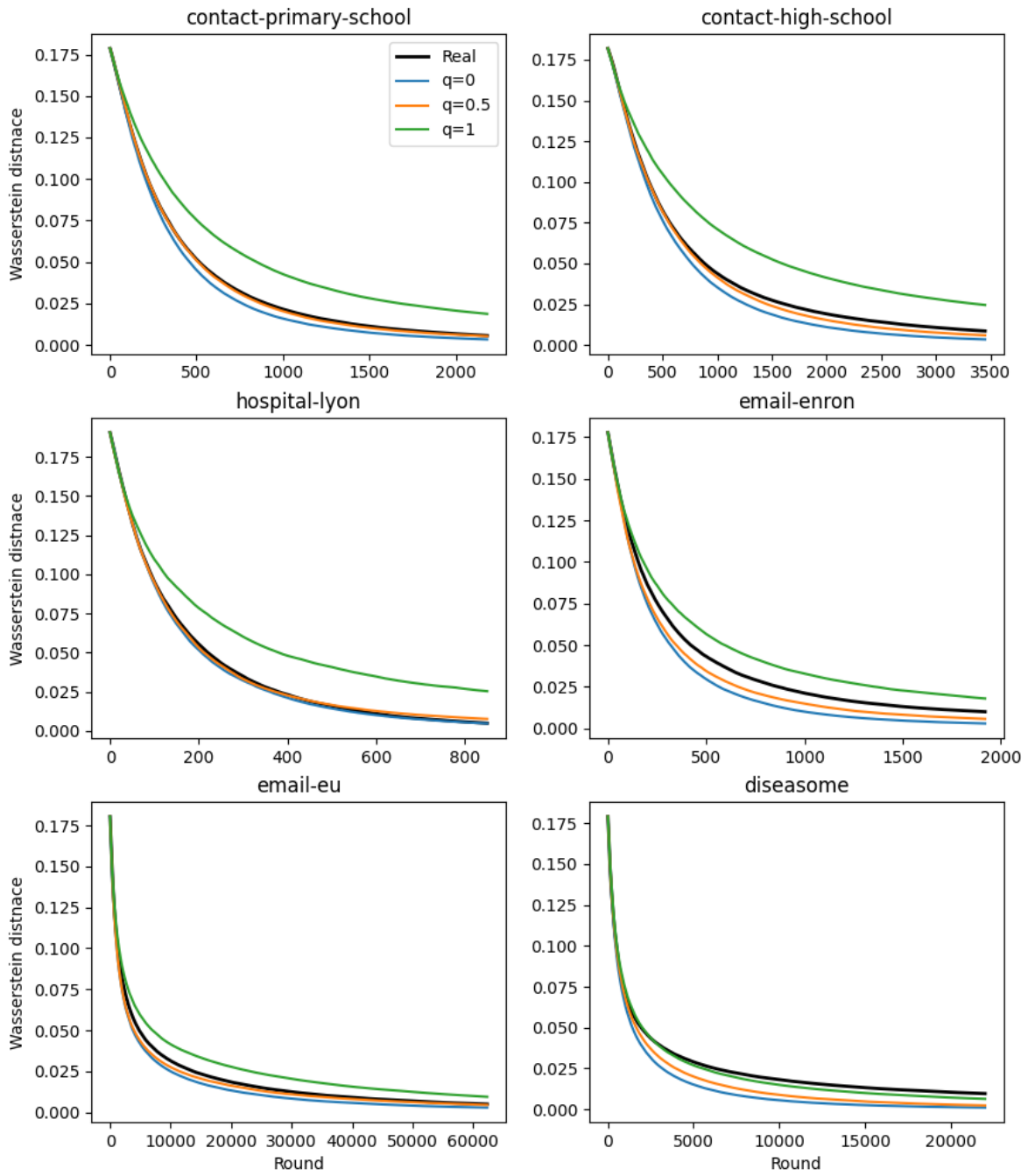


Figure 12: Wasserstein distance to uniform vs. number of rounds in the single-source diffusion process for all 10 graphs. The curve is the point-wise average across 100 independent experiments: for the real graph the chosen edges per round, as well as the location of the initial vertex with weight 1, are resampled each time, and for the random models the entire graphs are resampled each time.





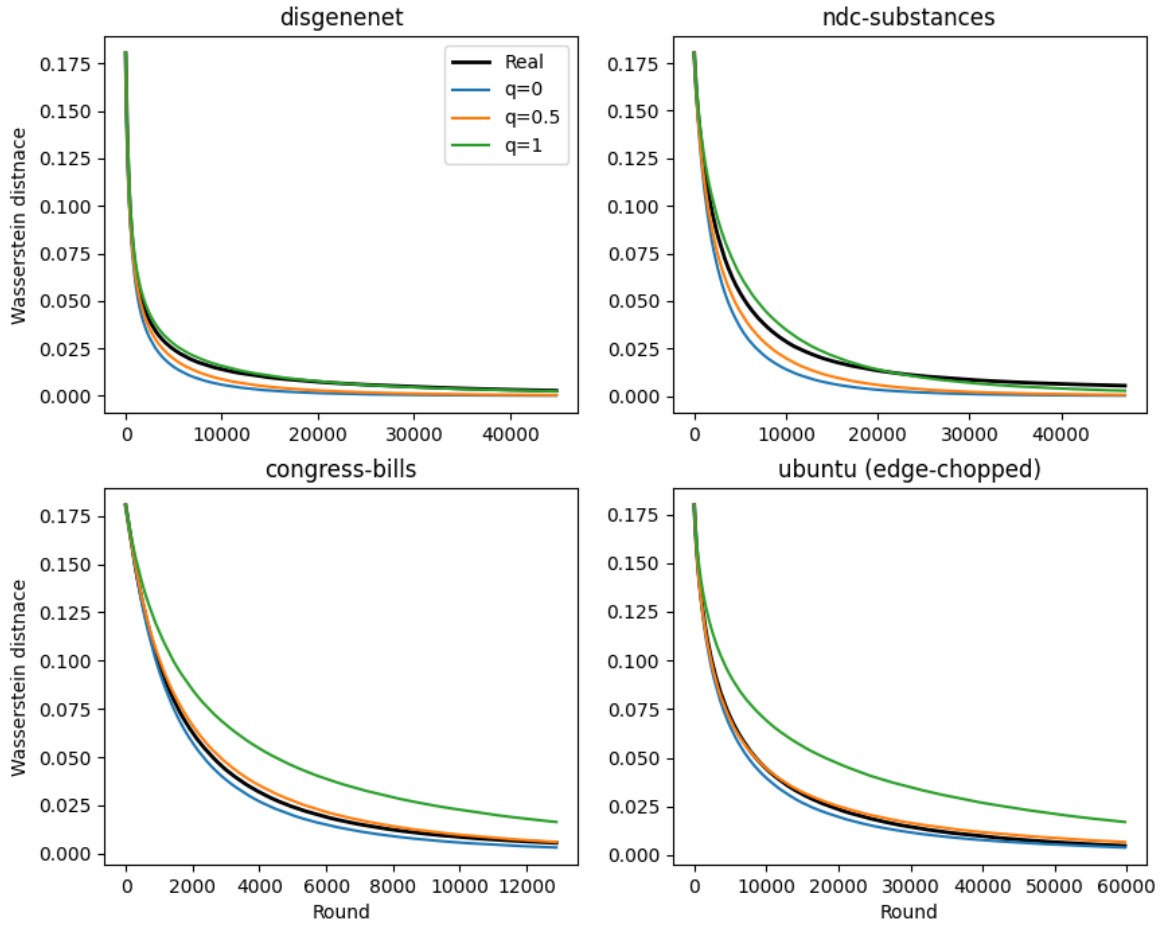


Figure 13: Wasserstein distance to uniform vs. number of rounds in the 10% sprinkled diffusion process for all 10 graphs. The curve is the point-wise average across 100 independent experiments: for the real graph the chosen edges per round, as well as the location of the initial 10% of vertices with weight 1, are resampled each time, and for the random models the entire graphs are resampled each time.

## B Algorithms

### B.1 Estimating the expected number of simplicial pairs

To compute the simplicial ratio of a graph  $G$ , we must first compute the expected number of simplicial pairs in  $\hat{G} \sim \text{CL}(G)$ . As discussed in Section 6, computing this expectation is quite difficult. In this section, we outline a Monte Carlo approximate technique for this expectation.

For a degree sequence  $\mathbf{d} = (d_1, \dots, d_n)$  and an edge size  $k$ , write  $\mathbb{P}(\text{simple} \mid \mathbf{d}, k)$  for the probability that **Algorithm 1** generates a simple edge when given inputs  $\mathbf{d}$  and  $k$ . For a graph  $G$  with degree sequence  $\mathbf{d}$ , we first approximate  $\mathbb{P}(\text{simple} \mid \mathbf{d}, k)$  for all edge sizes  $k$  in  $E(G)$ . To do this, we chose a number of samples  $s$ , sample  $s$  edges independently as **Algorithm 1**( $\mathbf{d}, k$ ), and compute the ratio  $x/s$  where  $x$  is the number of simple edges generated. In all experiments performed for the paper, we use  $s = 1000$ .

With  $\mathbb{P}(\text{simple} \mid \mathbf{d}, k)$  approximated for all edge sizes  $k$ , we can now approximate the number of simplicial pairs. We will show the algorithm for computing the expected number of  $(3, 5)$ -pairs here, as the generalization is straightforward to interpret but difficult to notate. Write  $|\mathbf{d}| := \sum_{i \in [n]} d_i$ . For an edge  $e = \{v_1, \dots, v_5\}$ , the probability that an edge  $e'$  of size 3 generated by **Algorithm 1** is (a) simple and (b) satisfies  $e' \subset e$  is given by

$$\sum_{1 \leq a < b < c \leq 5} \frac{3! d_{v_a} d_{v_b} d_{v_c}}{(|\mathbf{d}|)^3 \mathbb{P}(\text{simple} \mid \mathbf{d}, 3)}. \quad (1)$$

To break this down, consider only the probability that  $e' = \{v_1, v_2, v_3\}$ . **Algorithm 1** can generate this edge in  $3!$  different orders, and the probability of generating the edge in each case is

$$\frac{d_{v_1} d_{v_2} d_{v_3}}{|\mathbf{d}|^3}.$$

It can also happen that **Algorithm 1** generates a multi-edge, requiring us to sample again. Thus, the probability of *eventually* sampling the edge  $e' = \{v_1, v_2, v_3\}$  is

$$\begin{aligned} \sum_{i \geq 0} (1 - \mathbb{P}(\text{simple} \mid \mathbf{d}, 3))^i \frac{3! d_{v_1} d_{v_2} d_{v_3}}{|\mathbf{d}|^3} &= \frac{3! d_{v_1} d_{v_2} d_{v_3}}{|\mathbf{d}|^3} \sum_{i \geq 0} (1 - \mathbb{P}(\text{simple} \mid \mathbf{d}, 3))^i \\ &= \frac{3! d_{v_1} d_{v_2} d_{v_3}}{|\mathbf{d}|^3} \left( \frac{1}{1 - (1 - \mathbb{P}(\text{simple} \mid \mathbf{d}, 3))} \right) \\ &= \frac{3! d_{v_1} d_{v_2} d_{v_3}}{(|\mathbf{d}|^3) \mathbb{P}(\text{simple} \mid \mathbf{d}, 3)}. \end{aligned}$$

Summing over all  $\binom{5}{3}$  possible 3-edges inside  $e$  gives us (1).

We now approximate the number of  $(k, \ell)$  simplicial pairs as follows.

1. Choose some sampling number  $s$ . Then, sample  $s$  independent edges via **Algorithm 1**( $\mathbf{d}, \ell$ ).
2. For each edge, compute the probability of generating a  $(k, \ell)$  simplicial pair.

3. Compute the average and multiply this result by  $m_k m_l$ , where  $m_k$  is the number of edges of size  $k$ , and similarly for  $m_l$ .

## B.2 Constructing a connected skeleton of a random graph

We will generate a connected skeleton for our random graph via multiplicative coalescence. In short, multiplicative coalescence is a process in which particles in a space join together at a rate proportional to the product of their masses. We point the reader to [30] for an overview on the multiplicative coalescence process. In the context of generating random graphs, multiplicative coalescence is the process where new edges joining disjoint components are chosen with probability proportional to the product of the weights of the components.

We will describe **Algorithm 5** in words before presenting it as pseudo-code. Let  $\mathbf{d} := (d_1, \dots, d_n)$  be a degree sequence and  $\mathbf{m} := (m_{k_{\min}}, \dots, m_{k_{\max}})$  be a sequence of edge sizes. We construct the skeleton of our graph as follows.

1. Initially, we have an empty edge list  $E = \{\}$  and a collection of components, one for each vertex. For component  $C = \{v\}$ , define the weight of  $C$ , written  $w(C)$ , as  $w(C) := d_v$ .
2. We generate a random edge-size list  $S$  as per **Algorithm 4**, i.e., a uniform permutation containing  $m_k$  copies of  $k$  for each edge size  $k$ .
3. Iteratively until the graph is connected, we do the following.
  - (a) Choose a size  $k$  from  $S$  (iteratively).
  - (b) Sample  $k$  components independently, each component  $C$  being chosen with probability proportional to  $w(C)$ . If the chosen components  $C_1, \dots, C_k$  are not all unique, discard them all and sample again (repeating until we have a collection of distinct components).
  - (c) For each component  $C$  chosen in the previous step, randomly sample a designated vertex for  $C$ ; for  $v \in C$ , choose  $v$  as the designated vertex for  $C$  with probability  $d_v / \sum_{u \in C} d_u$ .
  - (d) Construct the edge  $e$  consisting of all the designated vertices. Add  $e$  to  $E$ , remove the chosen components  $C_1, \dots, C_k$ , and create a new component  $C = \cup_{j \in [k]} C_j$  with  $w(C) = \sum_{i \in [k]} w(C_i)$ .

If, just before the graph is fully connected, the chosen size  $k$  is greater than the number of components  $c$ , we generate the last edge of the connected skeleton by connecting the final  $c$  components as per step 3 (with  $k$  replaced by  $c$ ) and sampling the remaining  $k - c$  vertices as per the usual Chung-Lu sampling technique, i.e., using **Algorithm 1**. We note that, other than potentially the last edge constructed, an edge constructed in step 3 is equivalent to an edge generated by **Algorithm 1** conditioned on this edge joining  $k$  distinct components. We use this observation to simplify **Algorithm 5**. We will simplify **Algorithm 5** by writing “update [collection of components]” after generating an edge.

---

**Algorithm 5** Connected skeleton.

---

**Require:**  $(d_1, \dots, d_n), (m_{k_{\min}}, \dots, m_{k_{\max}})$ 

- 1: Initialize edge list  $E = \{\}$ , a random edge-size list  $S$  as per **Algorithm 4**, and a collection of components  $\mathcal{C} = \{C_v := \{v\} | v \in [n]\}$ .
  - 2: **for**  $k \in S$  **do**
  - 3:     **if**  $k \leq |\mathcal{C}|$  **then**
  - 4:         **repeat**
  - 5:             Sample  $e \sim$  **Algorithm 1**  $\left((d_1, \dots, d_n), k\right)$ .
  - 6:             **until**  $|e \cap C| \leq 1$  for all  $C \in \mathcal{C}$
  - 7:             Set  $E = E \cup e$  and update  $\mathcal{C}$ .
  - 8:         **else**
  - 9:             Set  $c = |\mathcal{C}|$ .
  - 10:         **repeat**
  - 11:             Sample  $e' \sim$  **Algorithm 1**  $\left((d_1, \dots, d_n), c\right)$ .
  - 12:             **until**  $|e' \cap C| \leq 1$  for all  $C \in \mathcal{C}$
  - 13:             Sample  $e'' \sim$  **Algorithm 1**  $\left((d_1, \dots, d_n), k - c\right)$ .
  - 14:             Set  $E = E \cup \{e' \cup e''\}$  and update  $\mathcal{C}$ .
  - 15:         **end if**
  - 16:         **if**  $|\mathcal{C}| = 1$  **then**
  - 17:             Return  $E$
  - 18:         **end if**
  - 19:     **end for**
  - 20: Return  $E$
- 

Once we generate a connected skeleton via **Algorithm 5**, we then update the parameter  $(m_{k_{\min}}, \dots, m_{k_{\max}})$  (by subtracting, from  $m_k$ , the number of edges of size  $k$  that were generated for each  $k$ ) and generate the rest of the simplicial Chung-Lu graph via **Algorithm 4** with updated parameter  $(m_{k_{\min}}, \dots, m_{k_{\max}})$  and initial (non-empty) edge list  $E$ .