Towards Analysis of Information Content in Dynamic Networks*

Wojciech Szpankowski
Center for Science of Information
Purdue University

June 22, 2019

Center for Science of Information
NSF Science & Technology Center

AofA, France, 2019

Dedicated to **P.Flajolet** and **D. E. Knuth**

*Joint work with Y. Choi, A. Grama, T. Luczak, A. Magner, and J. Sreedharan

Outline

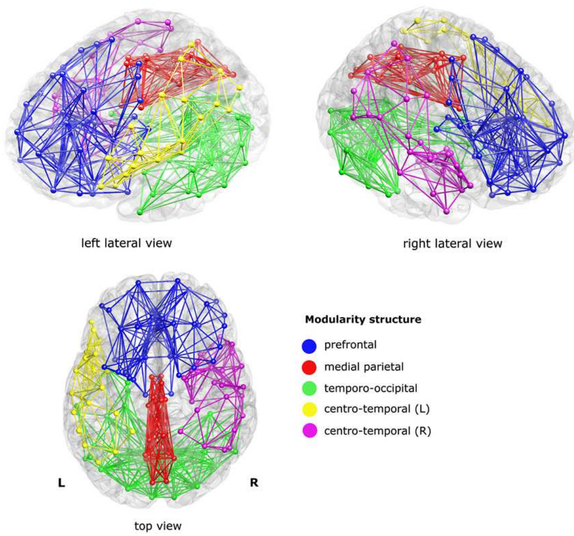
1. Dynamic Networks: Motivation and Challenges

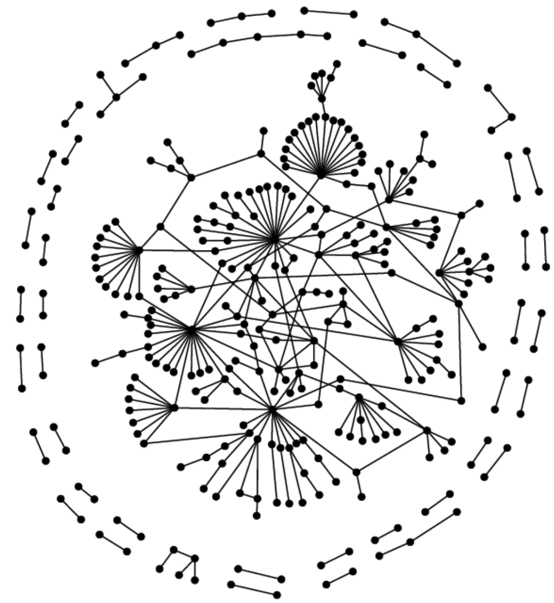
2. Models of Dynamic Networks

3. Structural Compression
   - Structural Entropy vs Graph Entropy
   - Preferential Attachment Graphs: Algorithms and Analysis
   - Duplication Model (preliminary results)
4. **TIMES**: **T**emporal **I**nformation **M**aximally **E**xtracted from **S**tructure
   - Problem Formulation
   - Precision vs Density Optimal Curve
   - Peeling Algorithm

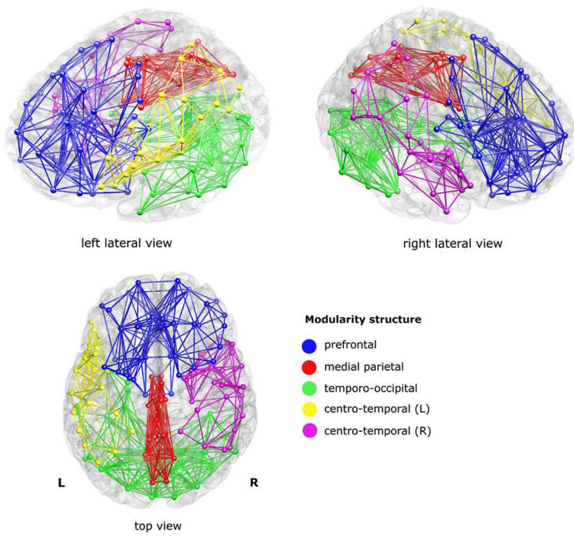
5. Application

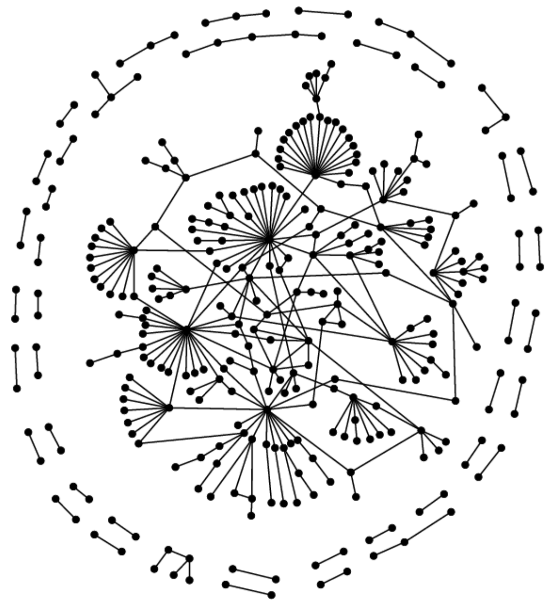
Dynamic Networks

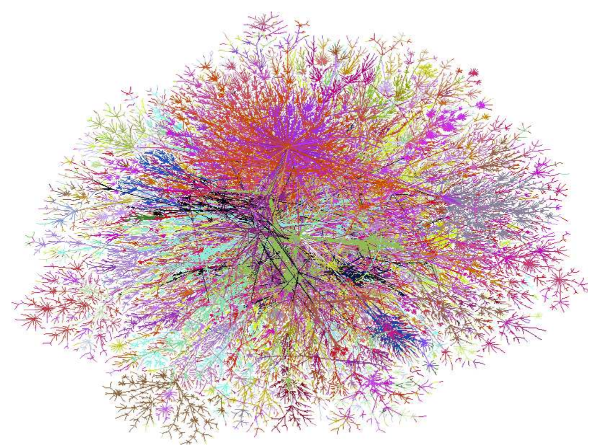




Dynamic Networks



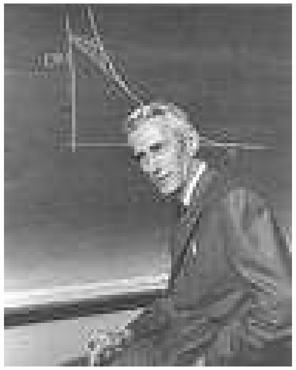




Dynamic Network Challenges

- **Inferring underlying dynamic processes** governing network evolution.
- Determine **minimum # bits to describe dynamic networks**.
- **Infer spatio-temporal properties** (e.g., node arrivals, network features including modularity, centrality, and clustering coefficient).
- Derive **control strategies** that optimize given objectives.
- Predict **long term forward evolution** of the network state

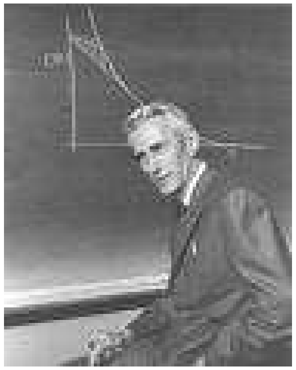
Goal: Fundamental limits and efficient algorithms that achieves these limits.



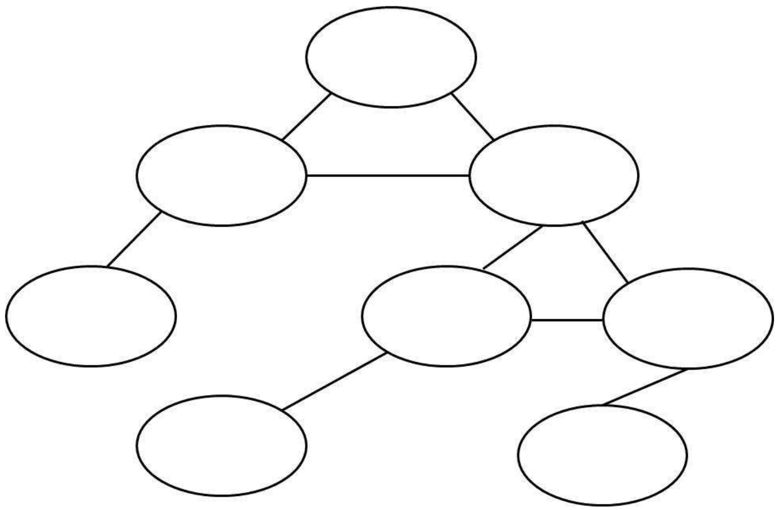
Dynamic Network Challenges

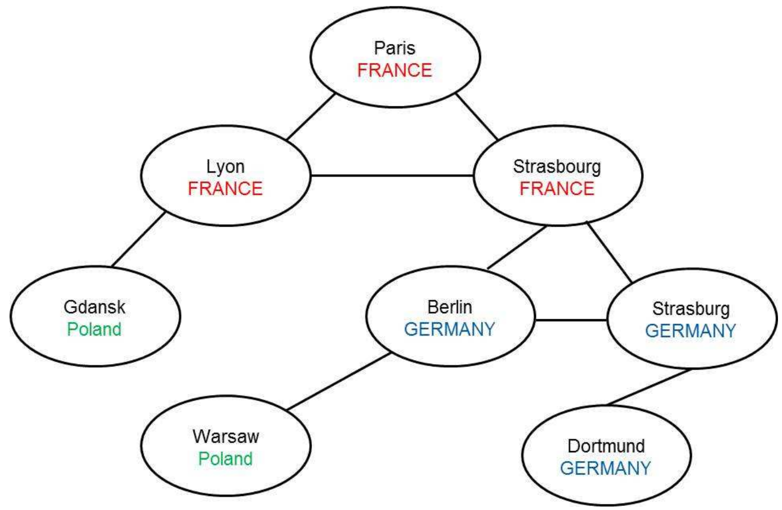
-
- Determine **minimum # bits to describe dynamic networks**.
-
-
-

Goal: Fundamental limits and efficient algorithms that achieves these limits.



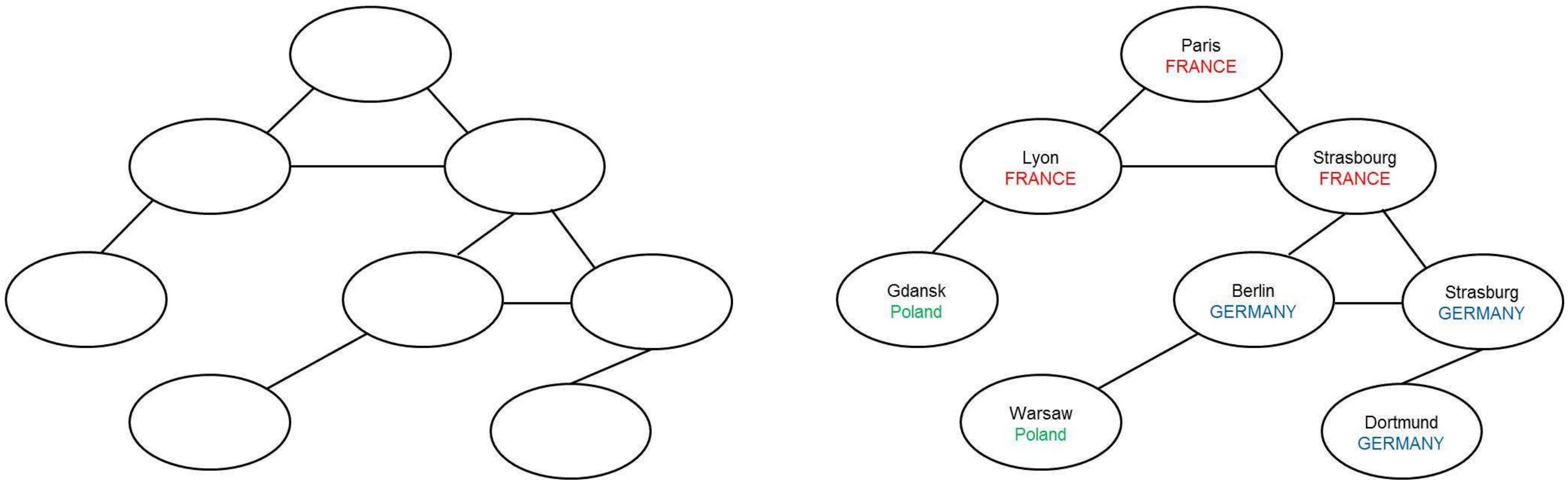
Motivation: Aim#1: Structural Compression





How many bits are required to describe the **unlabeled graph**, and how many additional bits one needs to represent the correlated labels?

Motivation: Aim#1: Structural Compression

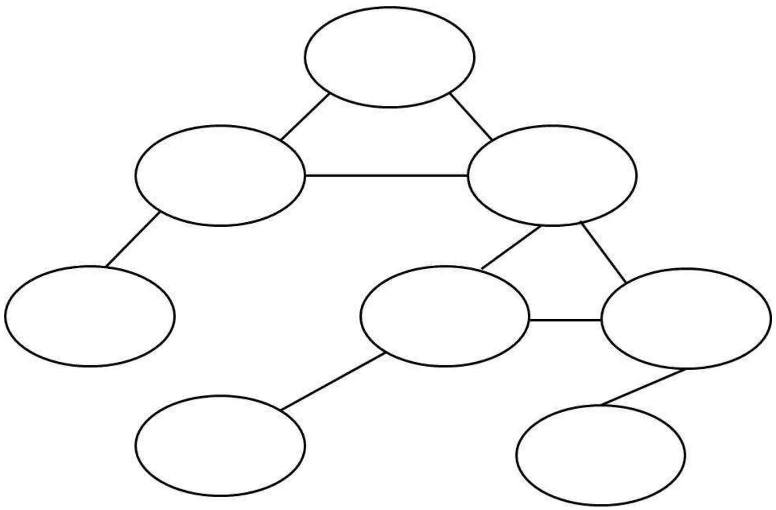


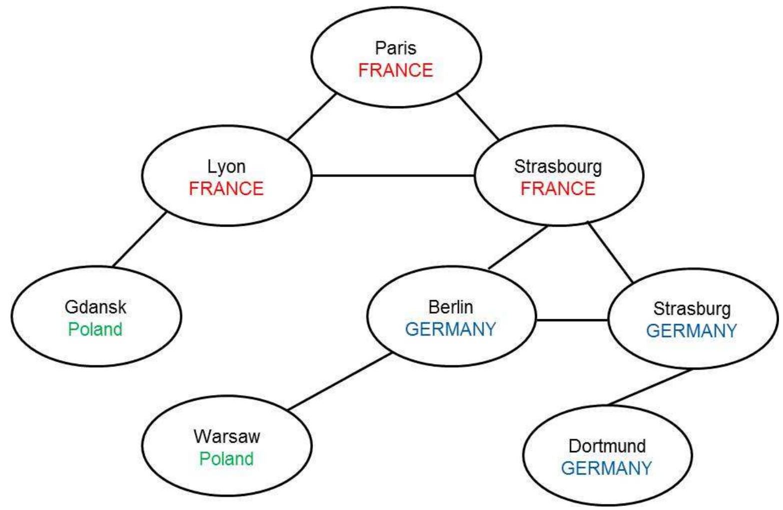
How many bits are required to describe the **unlabeled graph**, and how many additional bits one needs to represent the correlated labels?

One snapshot versus many snapshots of a dynamic graph.

Mulimodal and mixed dynamic networks.

Motivation: Aim#1: Structural Compression





How many bits are required to describe the **unlabeled graph**, and how many additional bits one needs to represent the correlated labels?

One snapshot versus many snapshots of a dynamic graph.

Mulimodal and mixed dynamic networks.

F. Brooks, jr. (JACM, 2003, "Three Great Challenges . . ."):

"We have **no theory** that gives us a metric for the Information embodied in **structure** . . . this is the most fundamental gap in the theoretical underpinning of Information and computer science".

Outline Update

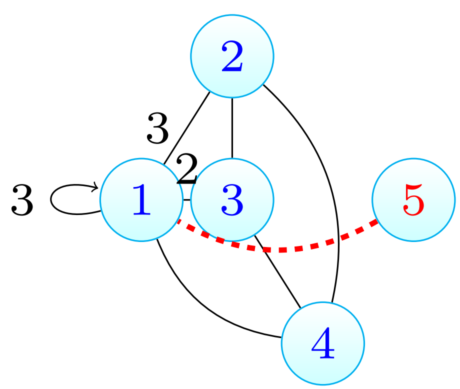
1. Dynamic Networks: Motivation and Challenges
   Goal: efficient algorithms to infer properties and to control and predict network evolution
2. Models of Dynamic Networks
3. Structural Compression

Three Graph Models

Erdős-Rényi Model $G(n, p)$:
Every pair of nodes receives an edge independently with probability p.
If $G(n, p)$ has k edges, then

$$P(G) = p^k (1 - p)^{\binom{n}{2} - k}.$$

Preferential Attachment Graph Model $\mathrm{PA}(n, m)$:
1: Vertex 1 arrives and has m self-edges.
2: $1 < t \leq n$: Vertex t makes m choices of connections, with probability of a vertex choice proportional to vertex current degree.



$$\Pr[t \to v | G_{t-1}] = \frac{\deg_{t-1}(v)}{2m(t-1)}.$$

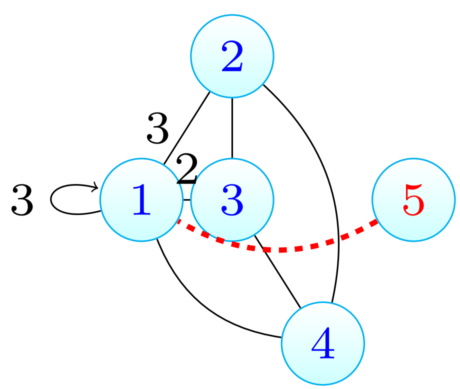
v	$\deg_4(v)$
1	9
2	5
3	4
4	3

Three Graph Models

Erdős-Rényi Model $G(n, p)$:
Every pair of nodes receives an edge independently with probability p.
If $G(n, p)$ has k edges, then

$$P(G) = p^k (1-p)^{\binom{n}{2}-k}.$$

Preferential Attachment Graph Model $\mathrm{PA}(n, m)$:
1: Vertex 1 arrives and has m self-edges.
2: $1 < t \leq n$: Vertex t makes m choices of connections, with probability of a vertex choice proportional to vertex current degree.



$$\Pr[t \to v | G_{t-1}] = \frac{\deg_{t-1}(v)}{2m(t-1)}.$$

v	$\deg_4(v)$
1	9
2	5
3	4
4	3

Duplication Model $\mathrm{DM(p)}$:
In this model a new node v copies/connects with probability p nodes in the neighbor of a uniformly selected existing node.

Outline Update

1. Dynamic Networks: Motivation and Challenges

   Goal: efficient algorithms to infer properties and to control and predict network evolution

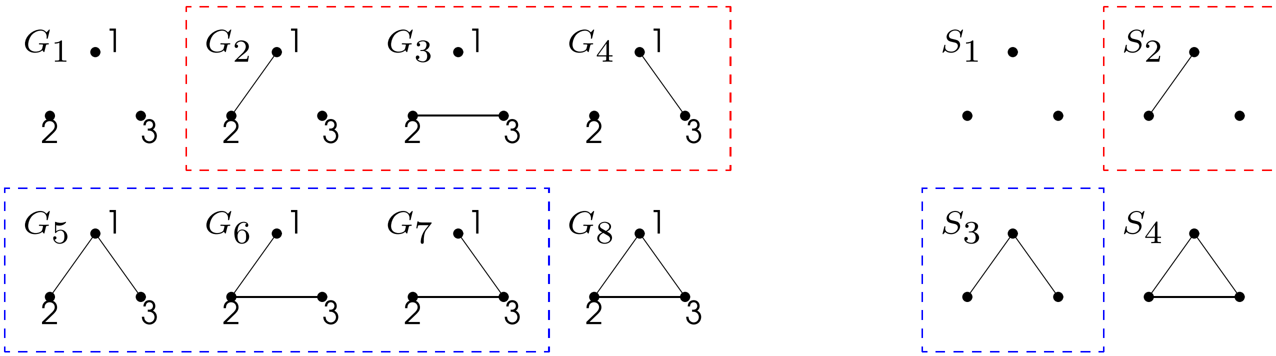
2. Models of Dynamic Networks

   Erdős-Rényi, preferential attachment, duplication

3. Structural Compression
   - Preferential Attachment Graphs
   - Duplication Model

Graphs vs Structure

A structure model S of a graph G is defined as an unlabeled version of G.



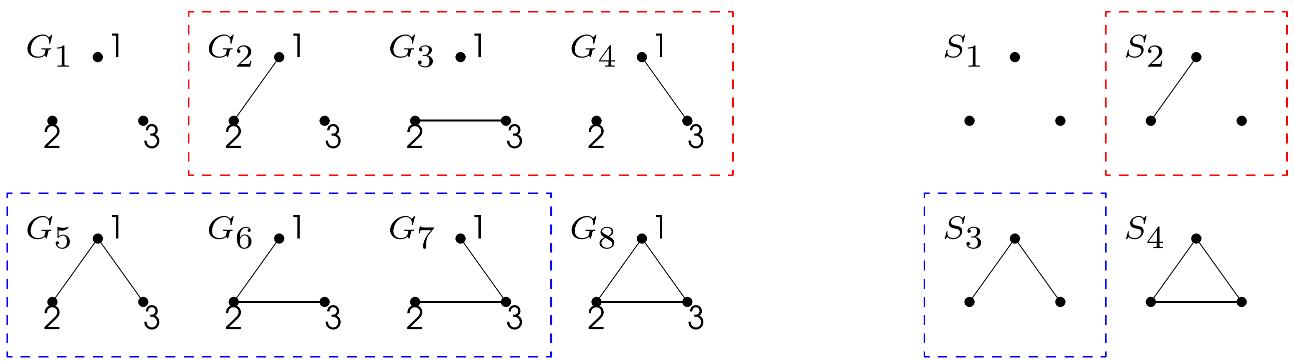
Graph Entropy vs **Structural Entropy**:

$$H(G) = \mathbb{E}[-\log P(G)] = -\sum_{G \in \mathcal{G}} P(G) \log P(G), \quad \text{graph entropy}$$

$$H(S(G)) = \mathbb{E}[-\log P(S)] = -\sum_{S \in \mathcal{S}} P(S) \log P(S) \quad \text{structural entropy}$$

Graphs vs Structure

A structure model S of a graph G is defined as an unlabeled version of G.



Graph Entropy vs **Structural Entropy**:

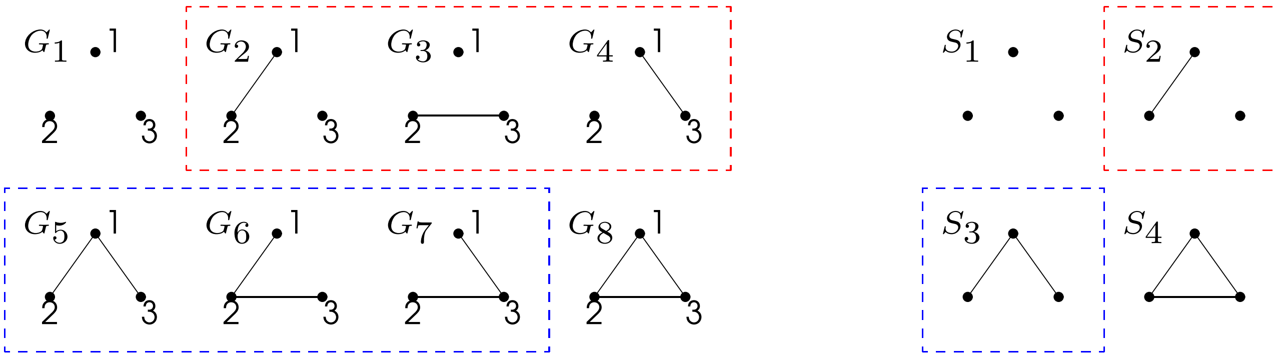
$$H(G) = \mathbb{E}[-\log P(G)] = -\sum_{G \in \mathcal{G}} P(G) \log P(G), \quad \text{graph entropy}$$

$$H(S(G)) = \mathbb{E}[-\log P(S)] = -\sum_{S \in \mathcal{S}} P(S) \log P(S) \quad \text{structural entropy}$$

How $H(G)$ and $H(S)$ are related?

Graphs vs Structure

A structure model S of a graph G is defined as an unlabeled version of G.



Graph Entropy vs **Structural Entropy**:

$$H(G) = \mathbb{E}[-\log P(G)] = -\sum_{G \in \mathcal{G}} P(G) \log P(G), \quad \text{graph entropy}$$

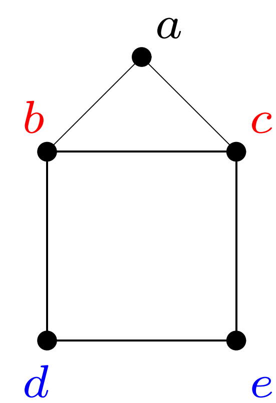
$$H(S(G)) = \mathbb{E}[-\log P(S)] = -\sum_{S \in \mathcal{S}} P(S) \log P(S) \quad \text{structural entropy}$$

How $H(G)$ and $H(S)$ are related?

Most papers deal with algorithmic issues of **labeled graph compression**.
Provably optimal: Aldous & Ross, 2014, Anantharam & Delgosha, 2017, Suresh et al, 2015, (for trees Kieffer & Yang, 2005).

Symmetries, Admissibility, and Feasibility

Graph Automorphism: An automorphism of G is adjacency preserving permutation of vertices of G (i.e., a form of symmetry). The collection $\mathrm{Aut}(G)$ of all automorphism of G is called *the automorphism group* of G.



Set of feasible permutations $\Gamma(G)$: $\sigma \in S_n$ such that $\sigma(G)$ has positive probability (of generation):

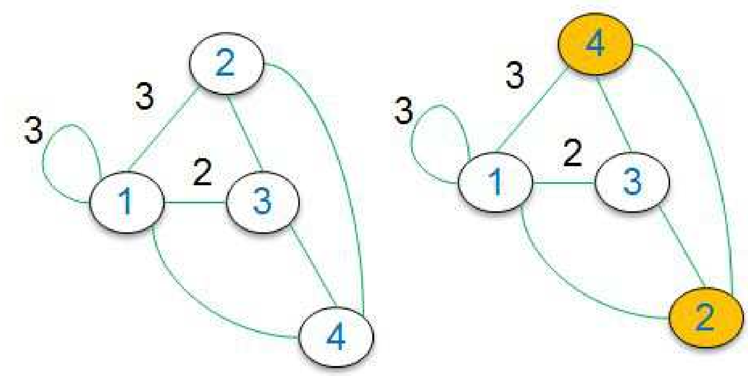
$\Gamma(G) \subseteq S_n$, where S_n is the set of all permutations.

Set of admissible graphs:

$$\mathrm{Adm}(G) := \{\sigma(G) : \ \sigma \in \Gamma(G)\}.$$

Note that by orbit-stabilizer like property we have

$$|\mathrm{Adm}(G)| = \frac{|\Gamma(G)|}{|\mathrm{Aut}(G)|}.$$



$H(G)$ versus $H(S)$

Lemma 1. *Assume that all positive probability labeled graphs $G(S)$ of the same structure $S(G)$* **have the same probability**. *Then*

$$H(G) = H(S) + \mathbb{E}[\log |\mathrm{Adm}(G)|] = H(S) + \mathbb{E}[\log |\Gamma(G)|] - \mathbb{E}[\log |\mathrm{Aut}(G)|].$$

$H(G)$ versus $H(S)$

Lemma 1. *Assume that all positive probability labeled graphs $G(S)$ of the same structure $S(G)$* **have the same probability***. Then*

$$H(G) = H(S) + \mathbb{E}[\log|\mathrm{Adm}(G)|] = H(S) + \mathbb{E}[\log|\Gamma(G)|] - \mathbb{E}[\log|\mathrm{Aut}(G)|].$$

Proof. Consider the following implications

$$H(G) = H(G, S) = H(S) + H(G|S).$$

Under our uniformity assumption we have

$$P(G|S) = \frac{1}{|\mathrm{Adm(G)}|} = \frac{|\mathrm{Aut}(G)|}{|\Gamma(G)|},$$

hence $H(G|S) = \mathbb{E}[\log|\mathrm{Adm}(G)|]$.

Remark. Lemma 1 does not hold for the duplication model.

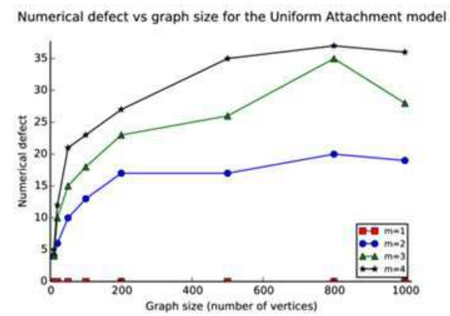
Outline Update

1. Dynamic Networks: Motivation and Challenges
   Goal: efficient algorithms to infer properties and to control and predict network evolution
2. Models of Dynamic Networks
   Erdős-Rényi, preferential attachment, duplication
3. Structural Compression
   - Preferential Attachment Graphs
   - Duplication Model

Symmetry Behavior of $\mathrm{PA}(n, m)$

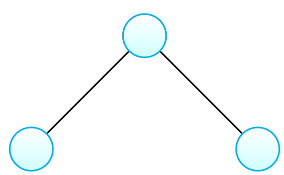
Theorem 1 (Symmetry behavior for $m = 1, 2$: Janson, Magner, Kollias, W.S., EJC 2014). *Consider* $G \sim \mathrm{PA}(n, m)$ *for* $\boxed{m = 1}$. *Then*

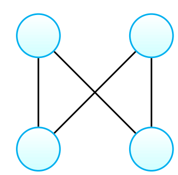
$$P(|\mathrm{Aut}(G)| > \Omega(n)] \to 1$$



For $\boxed{m = 2}$, *there is some* $c > 0$ *for which:* $P(|\mathrm{Aut}(G)| > 1) > c.$

Proof: By counting certain subgraphs (Pólya urns + birthday paradox):

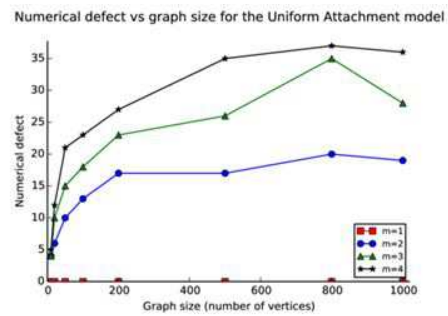




Symmetry Behavior of $\mathrm{PA}(n, m)$

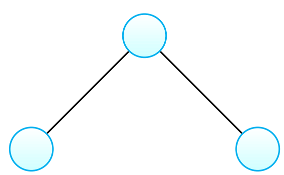
Theorem 1 (Symmetry behavior for $m = 1, 2$: Janson, Magner, Kollias, W.S., EJC 2014). *Consider $G \sim \mathrm{PA}(n, m)$ for $m = 1$. Then*

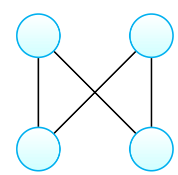
$$P(|\mathrm{Aut}(G)| > \Omega(n)] \to 1$$



For $m = 2$, there is some $c > 0$ for which: $P(|\mathrm{Aut}(G)| > 1) > c.$

Proof: By counting certain subgraphs (Pólya urns + birthday paradox):





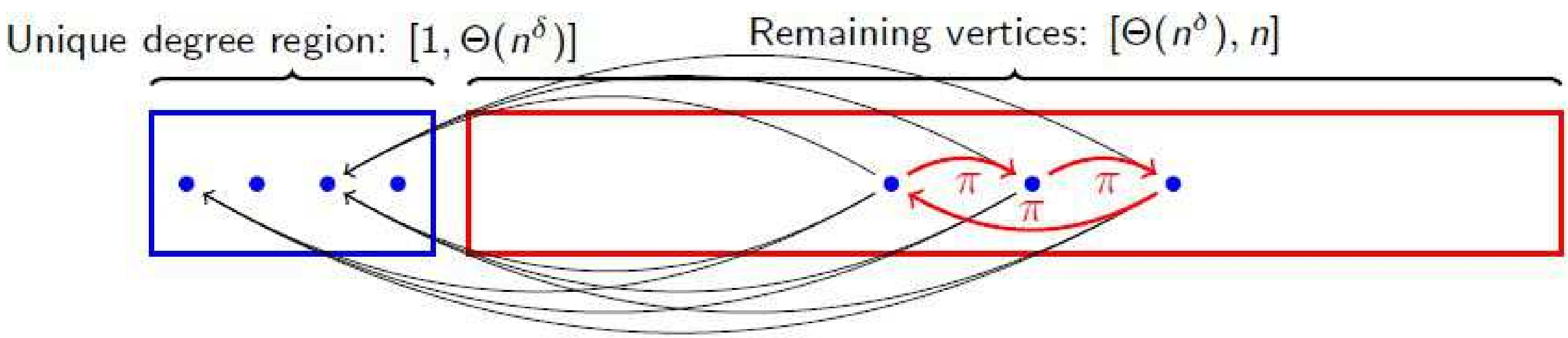
Theorem 2 (Asymmetry for $m \geq 3$: Luczak, Magner, W.S., RSA 2019). *Consider $G \sim \mathrm{PA}(n, m)$ for $m \geq 3$. Then, for some fixed $\delta > 0$,*

$$\Pr[|\mathrm{Aut}(G)| > 1] = O(n^{-\delta}).$$

Proof of Asymmetry of $\mathrm{PA}(n, m)$ for $m \geq 3$

Let $k = n^{\delta}$. Define two properties:
($\mathfrak{A}$) No two vertices t_1, t_2, where $k < t_1 < t_2$, are adjacent to the same m neighbors from the set $[t_1 - 1]$.

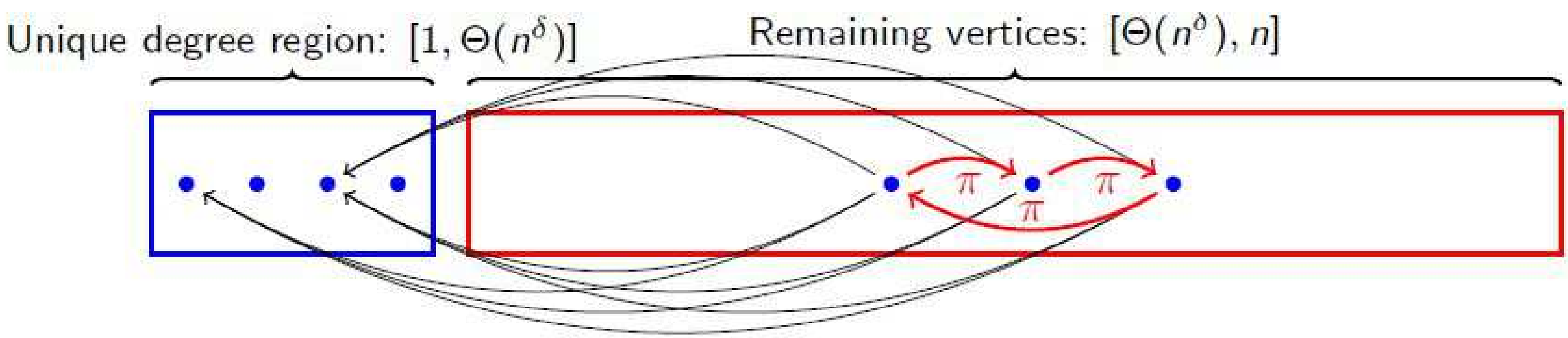
($\mathfrak{B}$) The degree of every vertex $s \leq k$ is unique, that is, $\deg_n(s) \neq \deg_n(s')$, $s < k$ and $s' < n$.



Proof of Asymmetry of $\mathrm{PA}(n, m)$ for $m \geq 3$

Let $k = n^{\delta}$. Define two properties:
($\mathfrak{A}$) No two vertices t_1, t_2, where $k < t_1 < t_2$, are adjacent to the same m neighbors from the set $[t_1 - 1]$.

($\mathfrak{B}$) The degree of every vertex $s \leq k$ is unique, that is, $\deg_n(s) \neq \deg_n(s')$, $s < k$ and $s' < n$.



Easy to see that if $\mathfrak{A}$ and $\mathfrak{B}$ hold, then $G_n(m) := \mathrm{PA}(n, m)$ is asymmetric:

$$P(|\mathrm{Aut}(G_n(m))| = 1) \geq P(G_n(m) \in \mathfrak{A} \cap \mathfrak{B}).$$

Indeed, let $G_n(m)$ satisfies $\mathfrak{A}$ and $\mathfrak{B}$ and $\sigma \in \mathrm{Aut}(G_n(m))$ which is not identity of $\mathrm{Aut}(G)$. Then:
(a) Let t_1 be smallest vertex such that $t_2 = \sigma(t_1) \neq t_1 > k$;
(b) By $\mathfrak{B}$ for $s \in [k]$ we have $\sigma(s) = s$, that is, σ is fixed for $s < k$.
(c) By $\mathfrak{A}$, we know t_1 and t_2 have different neighbourhoods in $[k]$ with fixed σ.
(d) This contradiction show that σ is identity, so that $|\mathrm{Aut}(G)| = 1$.

Property $\mathfrak{A}$ holds with small probability

1. Observe that for $k < t_1 < t_2$ we have

$$P(G_n(m) \notin \mathfrak{A}) \leq \sum_{k<t_1<t_2} \sum_{1\leq r_1 \leq r_2 \ldots \leq r_m < t_1} P(t_1, t_2 \text{ choose } r_1, \ldots, r_m)\ .$$

2. We now assume uniform attachment ($P(t \to v | G_{t-1}) = 1/t$). The

$$P(t_1, t_2 \text{ choose } r_1, \ldots, r_m) = \frac{1}{t_1^m} \frac{1}{t_2^m}.$$

3. Now we proceed as follows:

$$\begin{aligned}
P(G_n(m) \notin \mathfrak{A}) &\leq \sum_{k<t_1<t_2} \sum_{1\leq r_1 \leq r_2 \ldots \leq r_m < t_1} \frac{1}{t_1^m} \frac{1}{t_2^m} \\
&\leq \sum_{t_1=k}^{\infty} \sum_{t_2=t_1}^{\infty} \binom{t_1}{m} \frac{1}{t_1^m} \frac{1}{t_2^m} \\
&\leq \sum_{t_1=k}^{\infty} \frac{1}{m!} \sum_{t_2=t_1}^{\infty} \frac{1}{t_2^m} \\
&\leq \sum_{t_1=k}^{\infty} \frac{1}{m!} \frac{1}{t_1^{m-1}} \leq \frac{1}{k^{m-2}} \to 0 \quad m \geq 3.
\end{aligned}$$

Theorem 3 (Structural entropy; Luczak, Magner, W.S., 2019). *Let* $m \geq 3$ *be fixed. Consider* $G \sim \mathrm{PA}(n, m)$. *We have (see Sauerhoff, 2016)*

$$H(G) = mn \log n + m(\log 2m - 1 - \log m! - A(m))n + o(n)$$

where $A(m) = \sum_{d=m}^{\infty} \frac{\log d}{(d+1)(d+2)}$. *Furthermore:*

$$\boxed{\mathbb{E}[\log |\Gamma(G)|] = n \log n + O(n \log \log n)}$$ *which implies*

$$H(S(G)) = H(G) - \mathbb{E}[\log |\Gamma(G)|] + o(n) = (m-1)n \log n + O(n \log \log n),$$

Graph Entropy and Structural Entropy

Theorem 3 (Structural entropy; Luczak, Magner, W.S., 2019). *Let* $m \geq 3$ *be fixed. Consider* $G \sim \mathrm{PA}(n, m)$. *We have (see Sauerhoff, 2016)*

$$H(G) = mn \log n + m(\log 2m - 1 - \log m! - A(m))n + o(n)$$

where $A(m) = \sum_{d=m}^{\infty} \frac{\log d}{(d+1)(d+2)}$. *Furthermore:*

$$\boxed{\mathbb{E}[\log |\Gamma(G)|] = n \log n + O(n \log \log n)}$$ *which implies*

$$H(S(G)) = H(G) - \mathbb{E}[\log |\Gamma(G)|] + o(n) = (m-1)n \log n + O(n \log \log n),$$

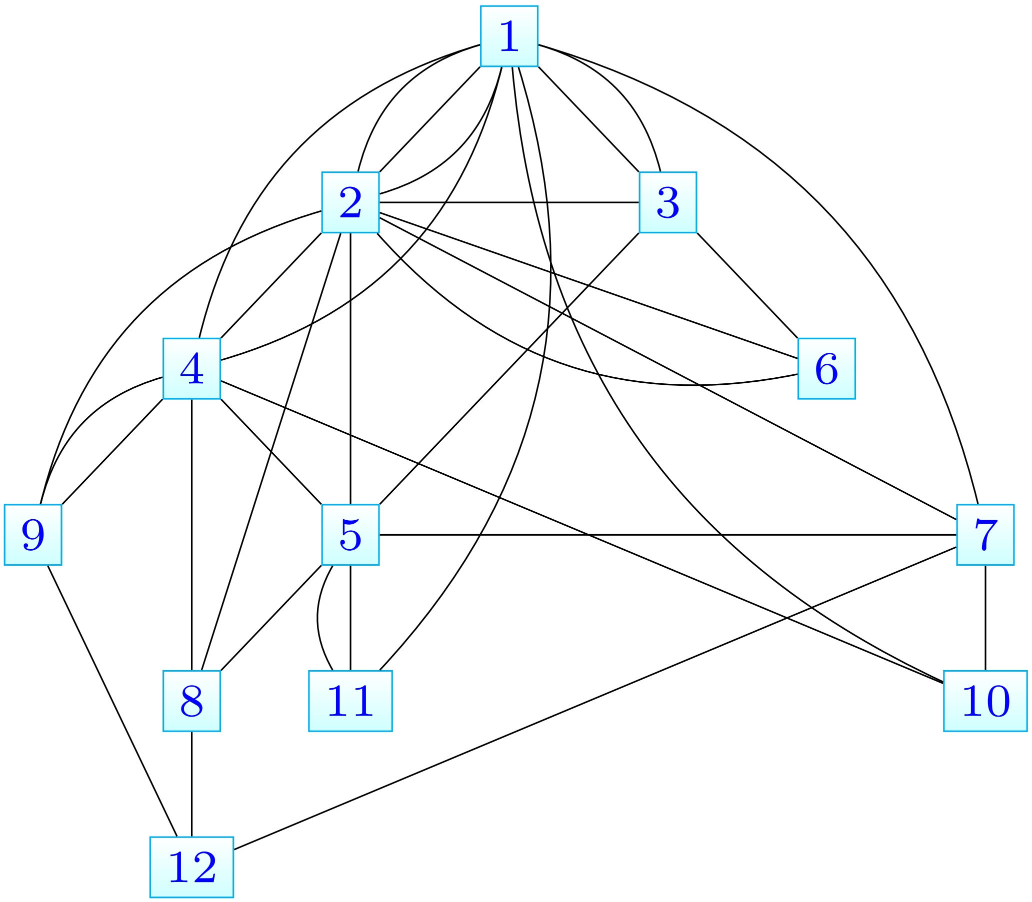
Proof. Observe that $H(G_n) = \sum_{t=1}^{n} H(v_{t+1}|G_t)$, and

$$H(v_{t+1}|G_t) = m \sum_{d=m}^{t} \mathbb{E}[N_{t,d}] p_{t,d} \log(1/p_{t,d})$$

where $\mathbb{E}[N_{t,d}(G)]$ denotes the expected number of vertices of degree d in the fixed graph G, and $p_{t,d} = \frac{d}{2mt}$. But

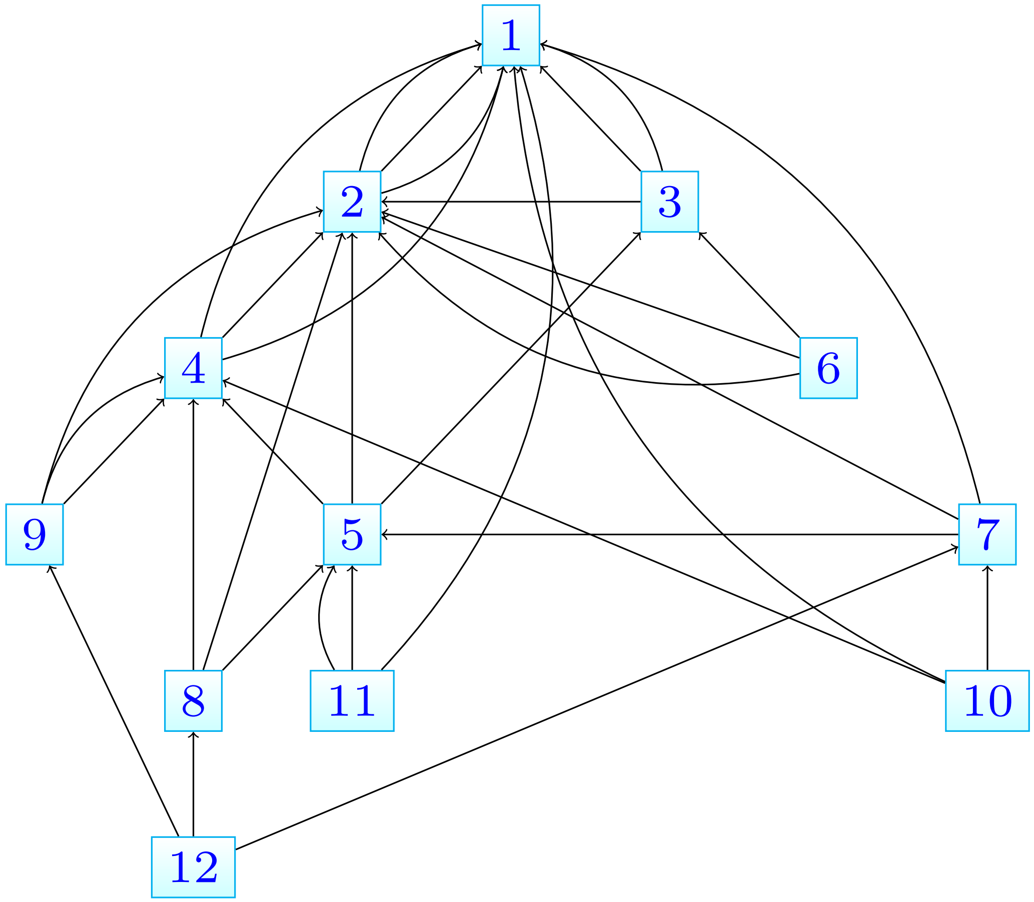
$$\left| \mathbb{E}[N_{t,d}] - \frac{2m(m+1)t}{d(d+1)(d+2)} \right| \leq C.$$

Structural Compression Algorithm



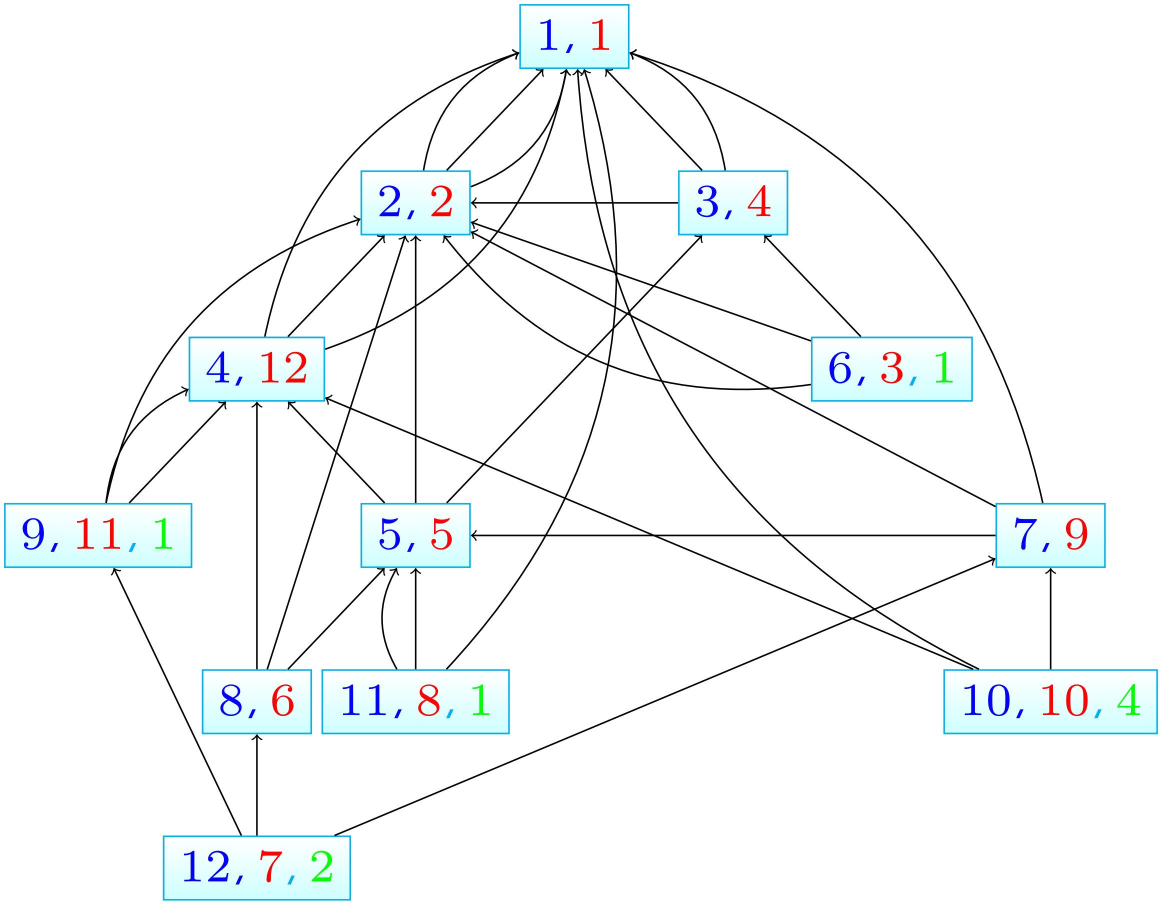
1. Recover $\mathrm{DAG}(G)$ from the input graph.
2. Label nodes with directed DFS and backtracking numbers B_j = # of nodes to pop immediately after visiting node j.

Algorithm



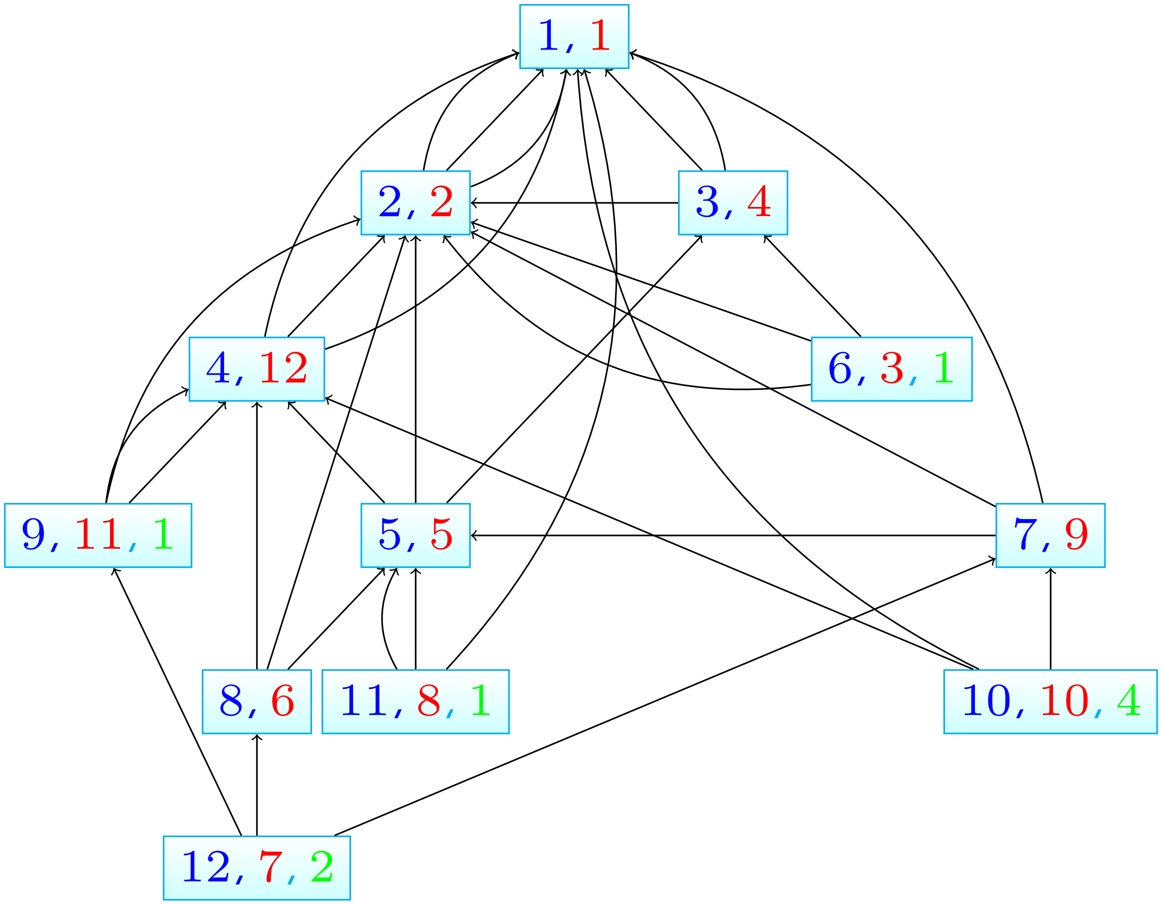
1. **Recover $\mathrm{DAG}(G)$ from the input graph.**
2. Label nodes with directed DFS and backtracking numbers B_j = # of nodes to pop immediately after visiting node j.

Algorithm



1. Recover $\mathrm{DAG}(G)$ from the input graph.
2. **Label nodes with directed DFS and backtracking numbers B_j = # of nodes to pop immediately after visiting node j.**

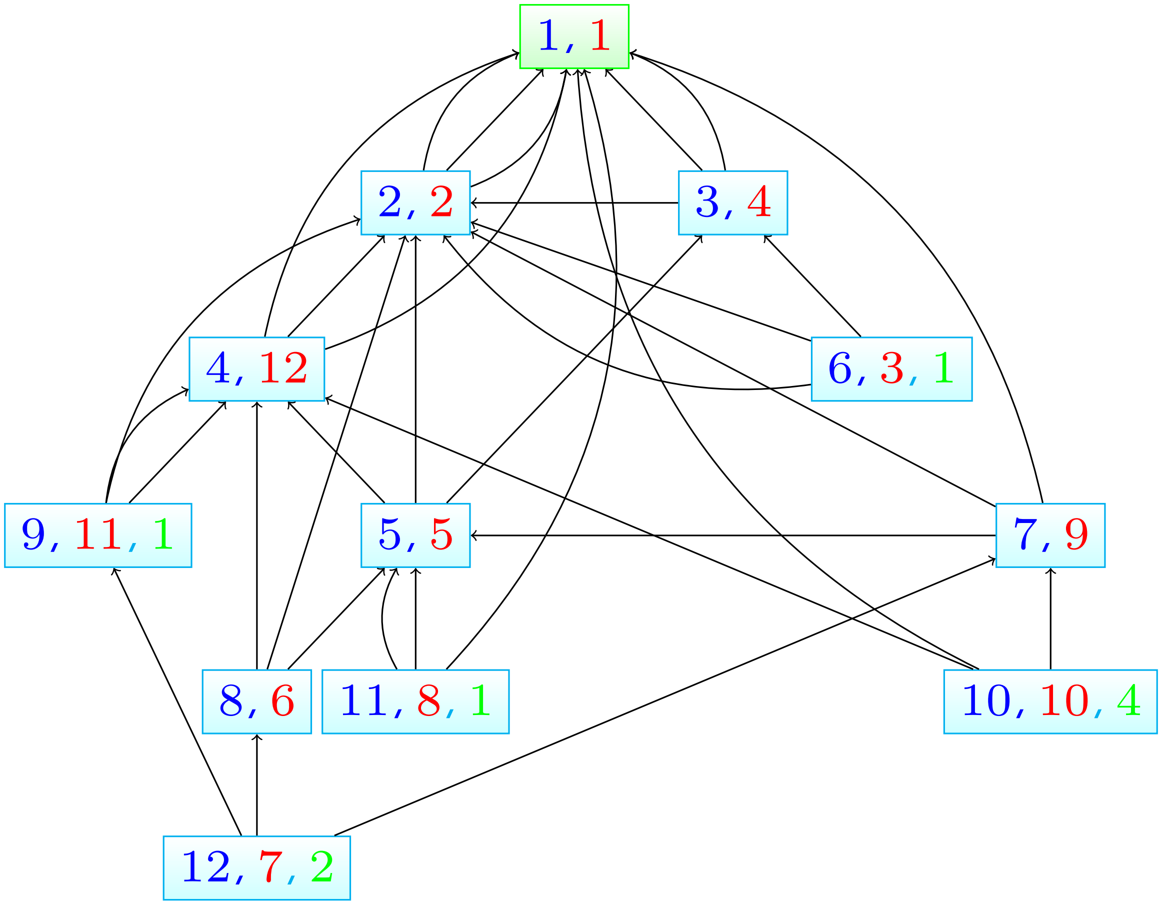
Algorithm



Output: (Bookkeeping info)

1. Empirical distribution for backtracking B_j: $[\frac{7}{12}, \frac{3}{12}, \frac{1}{12}, \frac{0}{12}, \frac{1}{12}]$. **Construct Huffman code $C(x)$.**
2. Conduct DFS again in the same order. For each visited node, encode destinations of all out-edges besides its predecessor, plus encoding of its backtracking number.

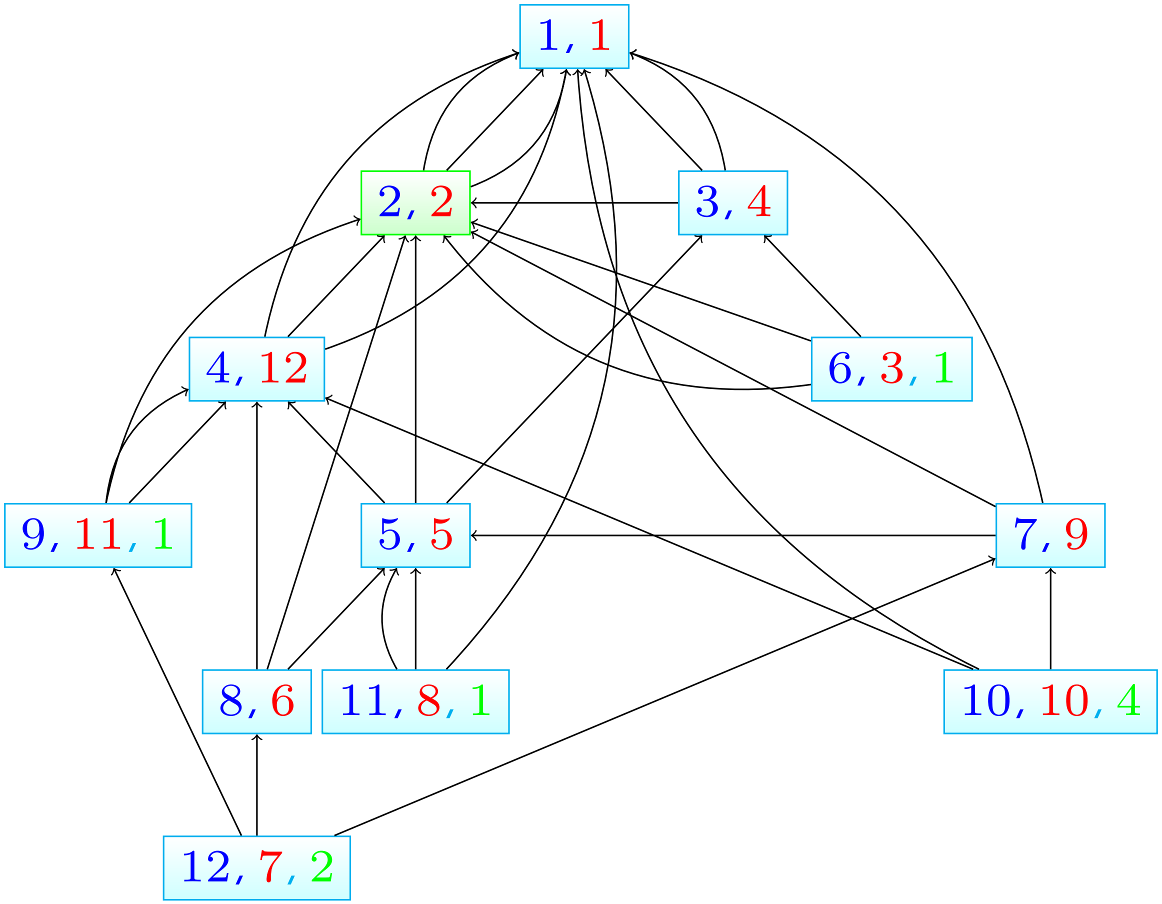
Algorithm



1. Empirical distribution for backtracking B_j: $[\frac{7}{12}, \frac{3}{12}, \frac{1}{12}, \frac{0}{12}, \frac{1}{12}]$. Construct Huffman code $C(x)$.
2. **Conduct DFS again in the same order. For each visited node, encode destinations of all out-edges besides its predecessor, plus encoding of its backtracking number.**

Output: (Bookkeeping info)

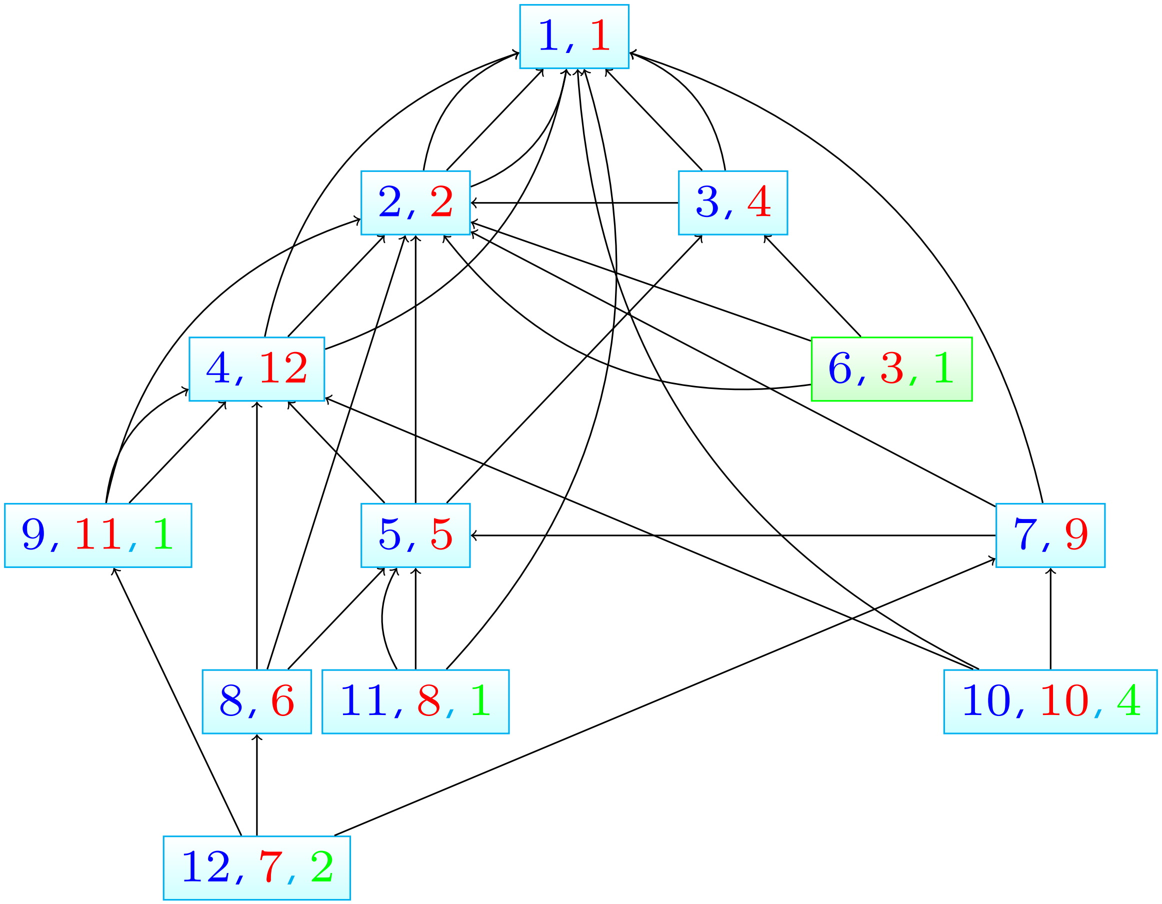
Algorithm



1. Empirical distribution for backtracking B_j: $[\frac{7}{12}, \frac{3}{12}, \frac{1}{12}, \frac{0}{12}, \frac{1}{12}]$. Construct Huffman code $C(x)$.
2. **Conduct DFS again in the same order. For each visited node, encode destinations of all out-edges besides its predecessor, plus encoding of its backtracking number.**

Output: (Bookkeeping info) $|(1)_2(1)_2C(0)$

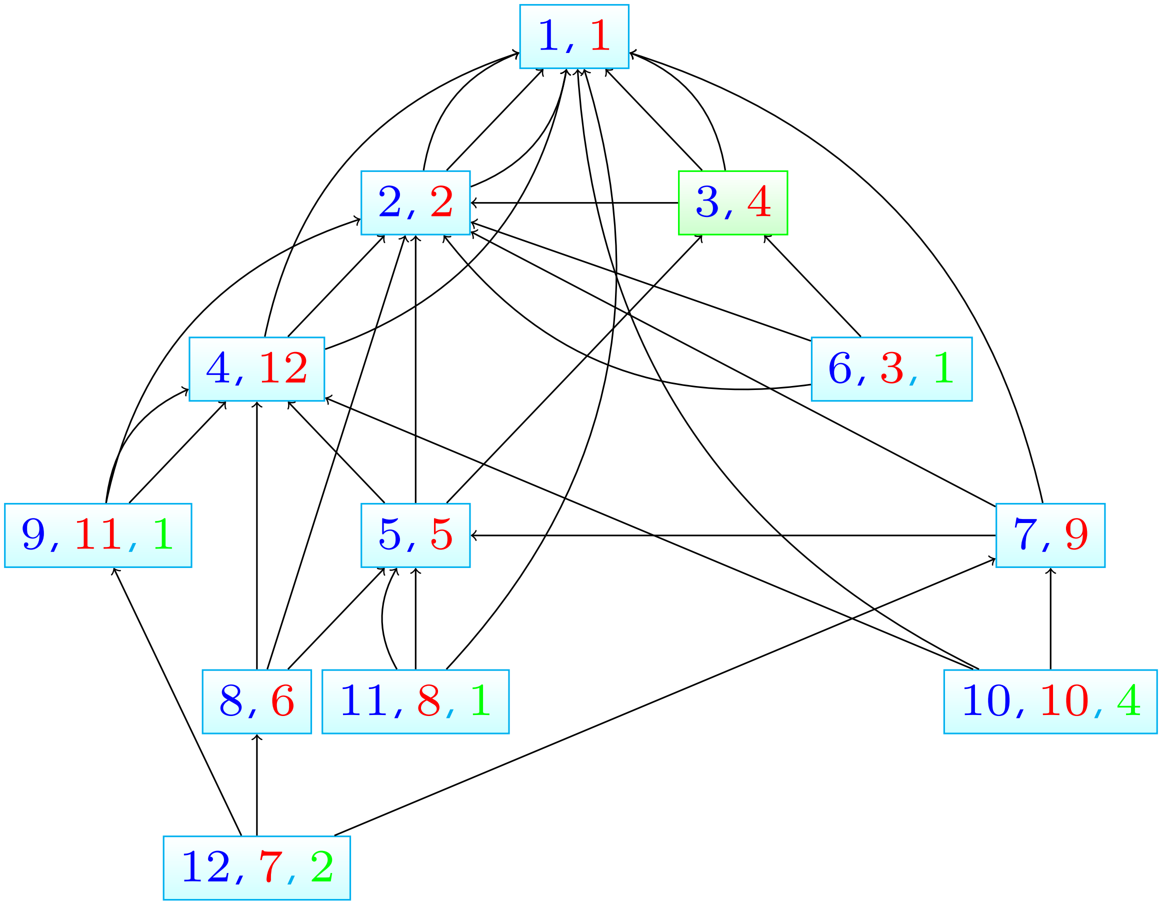
Algorithm



1. Empirical distribution for backtracking B_j: $[\frac{7}{12}, \frac{3}{12}, \frac{1}{12}, \frac{0}{12}, \frac{1}{12}]$. Construct Huffman code $C(x)$.
2. **Conduct DFS again in the same order. For each visited node, encode destinations of all out-edges besides its predecessor, plus encoding of its backtracking number.**

Output: (Bookkeeping info) $|(1)_2(1)_2C(0)(2)_2(4)_2C(1)$

Algorithm



1. Empirical distribution for backtracking B_j: $[\frac{7}{12}, \frac{3}{12}, \frac{1}{12}, \frac{0}{12}, \frac{1}{12}]$. Construct Huffman code $C(x)$.
2. **Conduct DFS again in the same order. For each visited node, encode destinations of all out-edges besides its predecessor, plus encoding of its backtracking number.**

Output: (Bookkeeping info) $|(1)_2(1)_2C(0)(2)_2(4)_2C(1)(1)_2(1)_2C(0)...$

Analysis of the Algorithm

Average code length:

- Bookkeeping: $\Theta(n)$ bits.
- Encoding of neighbors: $(m-1)n\log n + O(n)$ bits.
- How much encoding of backtracking numbers cost?

$$\frac{1}{n}\sum_{j=1}^{n}|C(B_j)| \approx H(B) \le \log(\max_j B_j),$$

and $\max_j B_j$ is upper bounded by the height of the DAG.

Lemma 2 (Luczak, Magner, W.S., RSA 2019). *The height of* $\mathrm{DAG}(G)$ *is* $\Theta(\log n)$ *with high probability.*

End result: expected code length is $(m-1)\log n + O(n\log\log n)$ bits.

Asymptotically optimal!

Outline

1. Dynamic Networks: Motivation and Challenges
2. Models of Dynamic Networks
3. Structural Compression
   - Structural Entropy vs Graph Entropy
   - Preferential Attachment Graphs: Algorithms and Analysis
   - Duplication Model (preliminary results)

Duplication Model

Pastor-Satorras Model $\mathrm{DD}(G_{n_0}, n, p, r,)$ Given a seed graph G_{n_0} on n_0 nodes, at time $t \leq n$ do:

Duplication: select a node u uniformly at random with probability $1/t$ and connect to all neighbors of u.
Divergence: (a) New connections are deleted with probability $1 - p$.
(b) Furthermore, for all nodes not connected to a selected node, add a new edge with probability r/t where n is the number of nodes.

Duplication Model

Pastor-Satorras Model $\mathrm{DD}(G_{n_0}, n, p, r,)$ Given a seed graph G_{n_0} on n_0 nodes, at time $t \leq n$ do:

Duplication: select a node u uniformly at random with probability $1/t$ and connect to all neighbors of u.
Divergence: (a) New connections are deleted with probability $1 - p$.
(b) Furthermore, for all nodes not connected to a selected node, add a new edge with probability r/t where n is the number of nodes.

Some parameters of $\mathrm{DD}(n, p, r)$ model

Degree: $\deg_t(s)$ – degree of node $s \leq n$ at time $s \leq t \leq n$.

Mean degree

$$D_n(G) = \frac{1}{n} \sum_s \deg_n(s).$$

Relation between $D_n(G)$ and $\deg_t(t)$
Lemma 3. *For any $t \geq t_0$ it holds that*

$$\mathbb{E}[\deg_{t+1}(t+1)|G_t] = \left(p - \frac{r}{t}\right) D(G_t) + r.$$

Recurrences

We have the following two recurences:

$$\mathbb{E}[\deg_{t+1}(s)] = \mathbb{E}[\deg_t(s)] \left(1 + \frac{p}{t} - \frac{r}{t^2}\right) + \frac{r}{t}$$

and

$$\mathbb{E}[D_{t+1}(G)] = \mathbb{E}[D_t(G) \left(1 + \frac{2p-1}{t+1} - \frac{2r}{t(t+1)}\right) + \frac{2r}{t+1}.$$

Both recurrences fall under the following general recurrence

$$\mathbb{E}[f(G_{n+1})] = f(G_n)g_1(n) + g_2(n)$$

that can be solve as follows

$$\mathbb{E}f(G_n) = f(G_{n_0}) \prod_{k=n_0}^{n-1} g_1(k) + \sum_{j=n_0}^{n-1} g_2(j) \prod_{k=j+1}^{n-1} g_1(k).$$

For example, with $c_1 = \frac{p+\sqrt{p^2+4r}}{2}$, $c_2 = \frac{p-\sqrt{p^2+4r}}{2}$

$$\mathbb{E}[\deg_t(s)] = \mathbb{E}[\deg_s(s)] \prod_{k=s}^{t-1} \left(1 + \frac{p}{k} - \frac{r}{k^2}\right) + \sum_{j=s}^{t-1} \frac{r}{j} \prod_{k=j+1}^{t-1} \left(1 + \frac{p}{k} - \frac{r}{k^2}\right)$$

$$= \frac{\Gamma(t+c_1)\Gamma(t+c_2)}{\Gamma(t)^2} \left(\mathbb{E}[\deg_s(s)] \frac{\Gamma(s)^2}{\Gamma(s+c_1)\Gamma(s+c_2)} + \sum_{j=s}^{t-1} \frac{r}{j} \frac{\Gamma(j)^2}{\Gamma(j+c_1)\Gamma(j+c_2)} \right)$$

Asymptotics for $\mathbb{E}[D_n(G)]$

Using standard asymptotics of gamma function we prove the following.

Theorem 4. *For all $t \geq t_0$ we have*

$$\mathbb{E}[D(G_t)] = \frac{\Gamma(t+c_3)\Gamma(t+c_4)}{\Gamma(t)\Gamma(t+1)} \left(D(G_{t_0}) \frac{\Gamma(t_0)\Gamma(t_0+1)}{\Gamma(t_0+c_3)\Gamma(t_0+c_4)} + \sum_{j=t_0}^{t-1} \frac{2r}{j+1} \frac{\Gamma(j)\Gamma(j+1)}{\Gamma(j+c_3)\Gamma(j+c_4)} \right),$$

where $c_3 = p + \sqrt{p^2+2r}$, $c_4 = p - \sqrt{p^2+2r}$, and $\Gamma(z)$ is the Euler gamma function.

Furthermore, asymptotically as $t \to \infty$ we find

$$\mathbb{E}[D(G_t)] = \begin{cases} \frac{2r}{1-2p}(1+o(1)) & \text{if } p < \frac{1}{2}, \\ 2r \ln t \, (1+o(1)) & \text{if } p = \frac{1}{2}, \\ \frac{\Gamma(t_0)\Gamma(t_0+1)}{\Gamma(t_0+c_3)\Gamma(t_0+c_4)} t^{2p-1} \left(D(G_{t_0}) + \frac{2r}{t_0+1} \, {}_4F_3\left[{t_0, t_0+1, t_0+1, 1 \atop t_0+2, t_0+c_3, t_0+c_4} ; 1 \right] \right) (1+o(1)) & \text{if } p > \frac{1}{2} \end{cases}$$

where $D(G_{t_0}$ is the mean degree of the initial graph G_{t_0} and ${}_4F_3\left[{t_0, t_0+1, t_0+1, 1 \atop t_0+2, t_0+c_3, t_0+c_4} ; 1 \right]$ is the hypergeometric series.

Asymptotics of $\mathbb{E}[\deg_t(s)]$

We only present asymptotics for $s = O(1)$ and $s = O(t)$ for $\mathbb{E}[\deg_t(s)]$.

Theorem 5. *Asymptotically as $t \to \infty$:*
(i) *for* $s = O(1)$

$$\mathbb{E}[\deg_t(s)] = t^p \Bigg[\left(p - \frac{r}{s-1}\right) \frac{\Gamma(s+c_3)\Gamma(s+c_4)}{s\Gamma(s+c_1)\Gamma(s+c_2)} \left(D(G_{t_0}) \frac{\Gamma(t_0)\Gamma(t_0+1)}{\Gamma(t_0+c_3)\Gamma(t_0+c_4)} + \sum_{j=t_0}^{s-1} \frac{2r}{j+1} \frac{\Gamma(j)\Gamma(j+1)}{\Gamma(j+c_3)\Gamma(j+c_4)} \right) + \frac{r\Gamma(s)^2}{\Gamma(s+c_1)\Gamma(s+c_2)} \left(1 + {}_4F_3\left[{s,s,s,1 \atop s+1,s+c_1,s+c_2}; 1\right] s^{-1}\right) \Bigg] (1+o(1)).$$

(ii) *for* $s \to \infty$

$$\mathbb{E}[\deg_t(s)] = \begin{cases} \frac{2rp}{1-2p} \left(\frac{t}{s}\right)^p (1+o(1)) & \text{if } p < \frac{1}{2} \\ 2rp\sqrt{\frac{t}{s}} \log s \, (1+o(1)) & \text{if } p = \frac{1}{2} \\ \frac{p\Gamma(t_0)\Gamma(t_0+1)}{\Gamma(t_0+c_3)\Gamma(t_0+c_4)} \left(D(G_{t_0}) + \frac{2r}{t_0+1} {}_4F_3\left[{t_0,t_0+1,t_0+1,1 \atop t_0+2,t_0+c_3,t_0+c_4}; 1\right]\right) & \\ \qquad \left(\frac{t}{s}\right)^p s^{2p-1} (1+o(1)) & \text{if } p > \frac{1}{2}. \end{cases}$$

Thank you!