Nestedness in Networks: A Theoretical Model and Some Applications

By Michael D. König (ETH Zurich),
Claudio J. Tessone (ETH Zurich),
Yves Zenou (Stockholm University)

Nestedness

- Nestedness is an important aspect of real-world networks.
- The organization of the New York garment industry [Uzzi, 1996]
- the Fedwire bank network [May et al., 2008; Soramaki et al., 2007] are nested in the sense that their organization is strongly hierarchical.

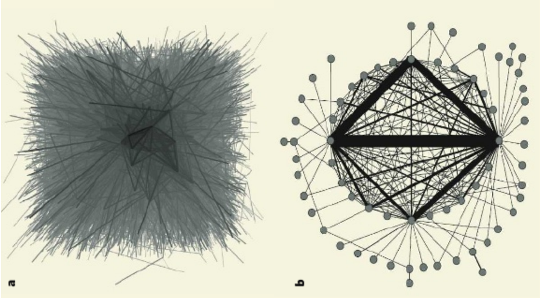
Nestedness

- **Fedwire bank network**: interbank payments transferred between commercial banks over the Fedwires Funds Service.
- Participants use Fedwire to process large-value, time-critical payments, such as payments for the settlement of interbank purchases and sales of federal funds; the purchase, sale, and financing of securities transactions; the disbursement or repayment of loans; and the settlement of real estate transactions.

Nestedness

- Network of banks: the topology of interbank payment flows within the US Fedwire service is clearly nested.

- **weighted network**: links represent transaction volumes
- **existence of a backbone**: involves small number of nodes



Example: Fedwire interbank payment network (K. Soramäki *et al.* Physica A **379** (2007) 317-333)

(left) Thousands of banks and tens of thousands of links representing USD 1.2 $\times 10^{12}$ in daily transactions

(right) Core of the network: 66 banks accounting for 75 % of transfers, 25 banks being completely connected.

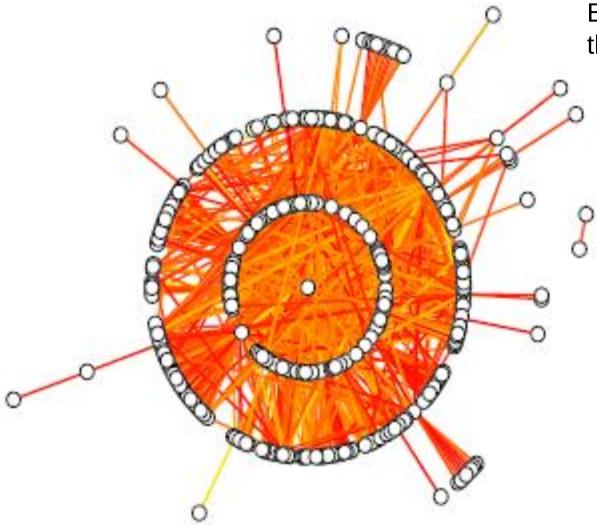
Nestedness

- The sample from this network amounted to around 700, 000 transfer funds, with just over 5, 000 banks involved on an average day.
- The authors [Soramaki et al., 2007] find that this network is characterized by a relatively small number of strong flows (many transfers) between nodes, with the vast majority of linkages being weak to non-existing (few to no flows).

- The topology of this Fedwire network is highly dissortative, since large banks are disproportionately connected to small banks, and vice versa; the average bank was connected to 15 others.
- Most banks have only a few connections while a small number of “hubs” have thousands.

Visualizing financial networks

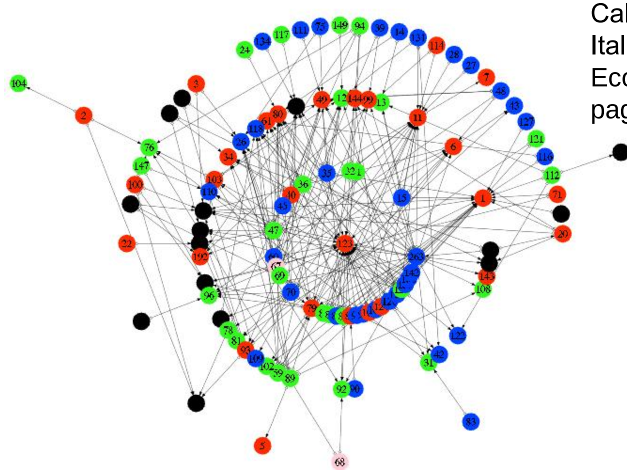
Federal funds



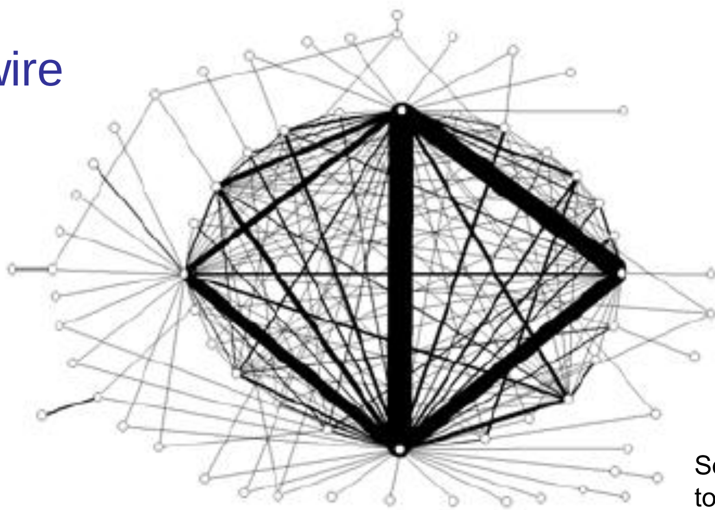
Bech, M.L. and Atalay, E. (2008), “The Topology of the Federal Funds Market”. ECB Working Paper No. 986.

Italian money market

Iori G, G de Masi, O Precup, G Gabbi and G Caldarelli (2008): “A network analysis of the Italian overnight money market”, Journal of Economic Dynamics and Control, vol. 32(1), pages 259-278

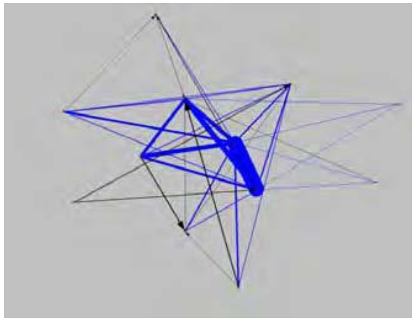


Fedwire



Soramaki, K, M.L. Bech, J. Arnold, R.J. Glass and W.E. Beyeler (2007), “The topology of interbank payment flows”, Physica A, Vol. 379, pp 317-333, 2007.

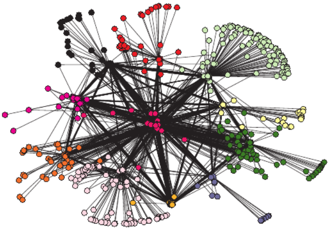
Unsecured sterling money market

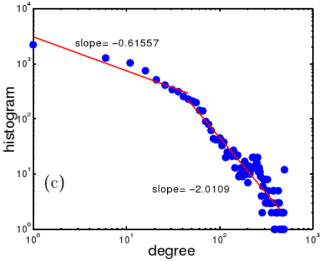


Wetherilt, A. P. Zimmerman, and K. Soramäki (2008), “The sterling unsecured loan market during 2006–2008: insights from network topology“, in Leinonen (ed), BoF Scientific monographs, E 42

Some Empirics: Financial Networks

- **skewed distributions**: few banks interact with many others
- **clusters**: banks with similar investment behavior





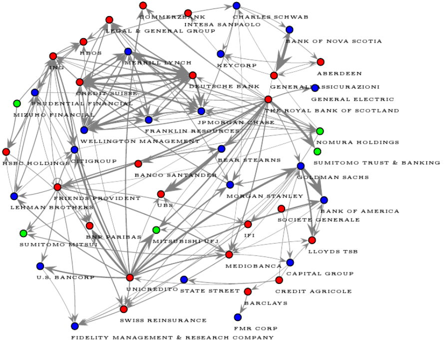
Example: Banking network of Austria (M Boss *et. al*, Quantitative Finance **4** (2004) 677-684)

(left) Clusters are grouped (colored) according to regional and sectorial organization

(right) Degree distribution of the interbank connection network

Some Empirics: Ownership Networks

- **directed network** of ownership ⇒ *control*



Example: Ownership relations in the international financial network
Nodes represent *major* financial institutions, links the strongest existing relations, node colors different geographical areas

Trade networks

- Using aggregate **bilateral imports** from 1950 to 2000, De Benedictis and Tajoli [2011] analyze the structure of the world trade network over time, detecting and interpreting patterns of trade ties among countries.

- World trade tends to be concentrated among a sub-group of countries and a small percentage of the total number of flows accounts for a disproportionately large share of world trade.

- The larger countries account for a generally larger share of world trade and have more partners.

- **Core-periphery structure**, indicating nestedness of their networks.

- All these networks: **dissortativity** arises naturally since “big” agents tend to interact with “small” agents and vice versa.
- Banks seek relationships with each other that are mutually beneficial.
- Small banks interact with large banks for security, lower liquidity risk and lower servicing costs.
- Large banks interact with small banks in part because they can extract a higher premium for services and can accommodate more risk.

- Aim: propose a dynamic network formation model that exhibits not only the standard features of real-world networks (small worlds, high clustering, short path lengths and a power-law degree distributions)
- but also **nestedness** and **dissortativity**.

Structure of real-life networks: characterized by low diameter (small world), high clustering, and “scale-free” degree distributions.

Need to analyze *how* and *why* networks form,

the impact of network structure on agents’ outcomes,

the evolution of networks over time.

This is the aim of this paper

Network formation: Two literatures

Random network formation: mainly dynamic and links are formed randomly

(Erdös and Rényi, PM 1959, Albert and Barabási , RMP 2002; Jackson and Rogers, AER 2004).

Strategic network formation: mainly static (game theory, pairwise stability)

(Myerson, 1991; Jackson and Wolinsky JET 1996; Bala and Goyal, Econometrica 2000)

Random approach: insight into *how* networks form (i.e. matches the characteristics of real-life networks)

Strategic approach: *Why* networks form

Games on networks: take the network as given and study how the network structure impacts on outcomes and individual decisions.

(Ballester, Calvó-Armengol, Zenou, Econometrica 2006, Bramoulle and Kranton, JET 2008; Galeotti, Goyal, Jackson, Vega-Redondo, Yariv, Restud 2009)

Here: introduce strategic interactions in a non-random dynamic network formation game.

Two-stage game at each period of time:

First stage: agents play their equilibrium contributions proportional to their Bonacich centrality (as in Ballester et al, 2006)

Second stage: a randomly chosen agent can update her linking strategy by creating a new link or deleting an existing link as a local best response to the current network.

Result:

At each period of time, the network generated by this dynamic formation process is a *nested split graph*.

Nested split graphs have a very nice and simple structure that make them very tractable to work with.

Here: A complex dynamic network formation model can be characterized by a simple structure in terms of networks it generates.

Result:

There exists a unique stationary network, which is a nested split graph.

We show under which conditions these networks emerge and that there exists a sharp transition between hierarchical and flat network structures.

Result:

Our model is able to match the characteristics of most real-life networks.

We show that the stationary networks emerging in our link formation process are characterized by *short path length* with *high clustering* (small worlds), *exponential degree distributions* with *power law tails* and *negative degree-clustering correlation*.

Without capacity constraints in the number of links an agent can maintain: networks are *dissortative*.

With capacity constraints: networks are *assortative*.

A Network Game with Linear Quadratic Payoffs

- We consider a network game where payoffs are interdependent (Ballester et al., 2006)

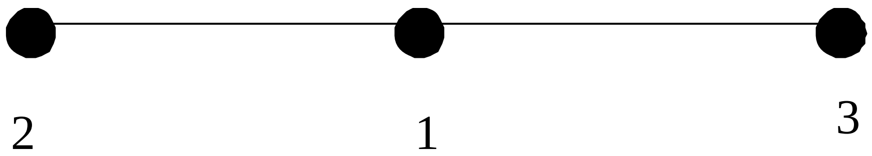
- $\mathbf{A}$ is the symmetric $n \times n$ adjacency matrix of network $\mathbf{G}$.

- Each agent $i = 1, .., n$ selects an effort level $x_i \geq 0$ and obtains a payoff of ($\lambda > 0$):

$$u_i(x_1, ..., x_n) = \underbrace{x_i - \frac{1}{2}x_i^2}_{\text{own concavity}} + \underbrace{\lambda \sum a_{ij} x_i x_j}_{\text{local complementarities}}$$

- $\mathbf{A}$ Adjacency matrix of network $\mathbf{G}$

- Example: Network g with three individuals (star)



- Adjacency matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

For $k \geq 1$

$$\mathbf{A}^{2k} = \begin{bmatrix} 2^k & 0 & 0 \\ 0 & 2^{k-1} & 2^{k-1} \\ 0 & 2^{k-1} & 2^{k-1} \end{bmatrix} \qquad \text{and} \qquad \mathbf{A}^{2k+1} = \begin{bmatrix} 0 & 2^k & 2^k \\ 2^k & 0 & 0 \\ 2^k & 0 & 0 \end{bmatrix}$$

For example

$$\mathbf{A}^3 = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

$\mathbf{A}^3$: two paths of length three between 1 and 2: $12 \to 21 \to 12$ and $12 \to 23 \to 32$.

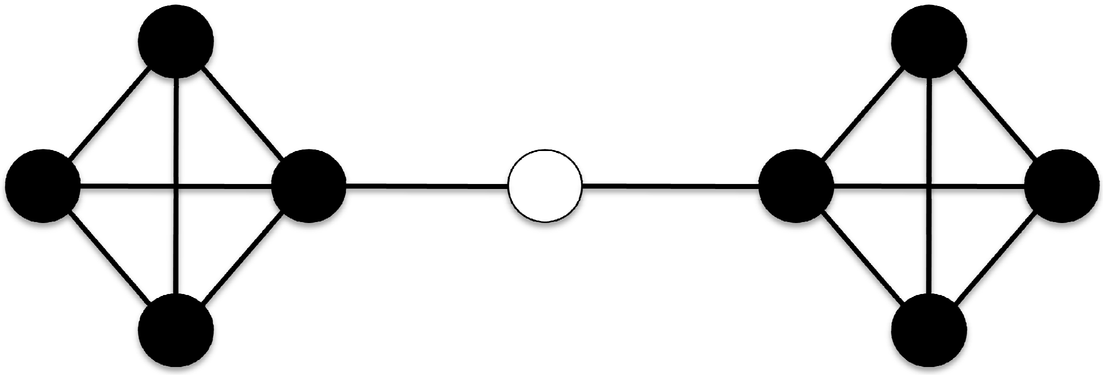
But no path of length three from i to i.

Different *centrality measures* to capture the prominence of actors inside a network.

Degree centrality: counts the number of connections an agent has.

Bonacich centrality: gives to any individual a particular numerical value for each of his/her direct connection. Then, give a smaller value to any connection at distance two and an even smaller value to any connection at distance three; etc. When adding up all these values, we end up with a new numerical value that is now capturing both direct and indirect connections of any order.

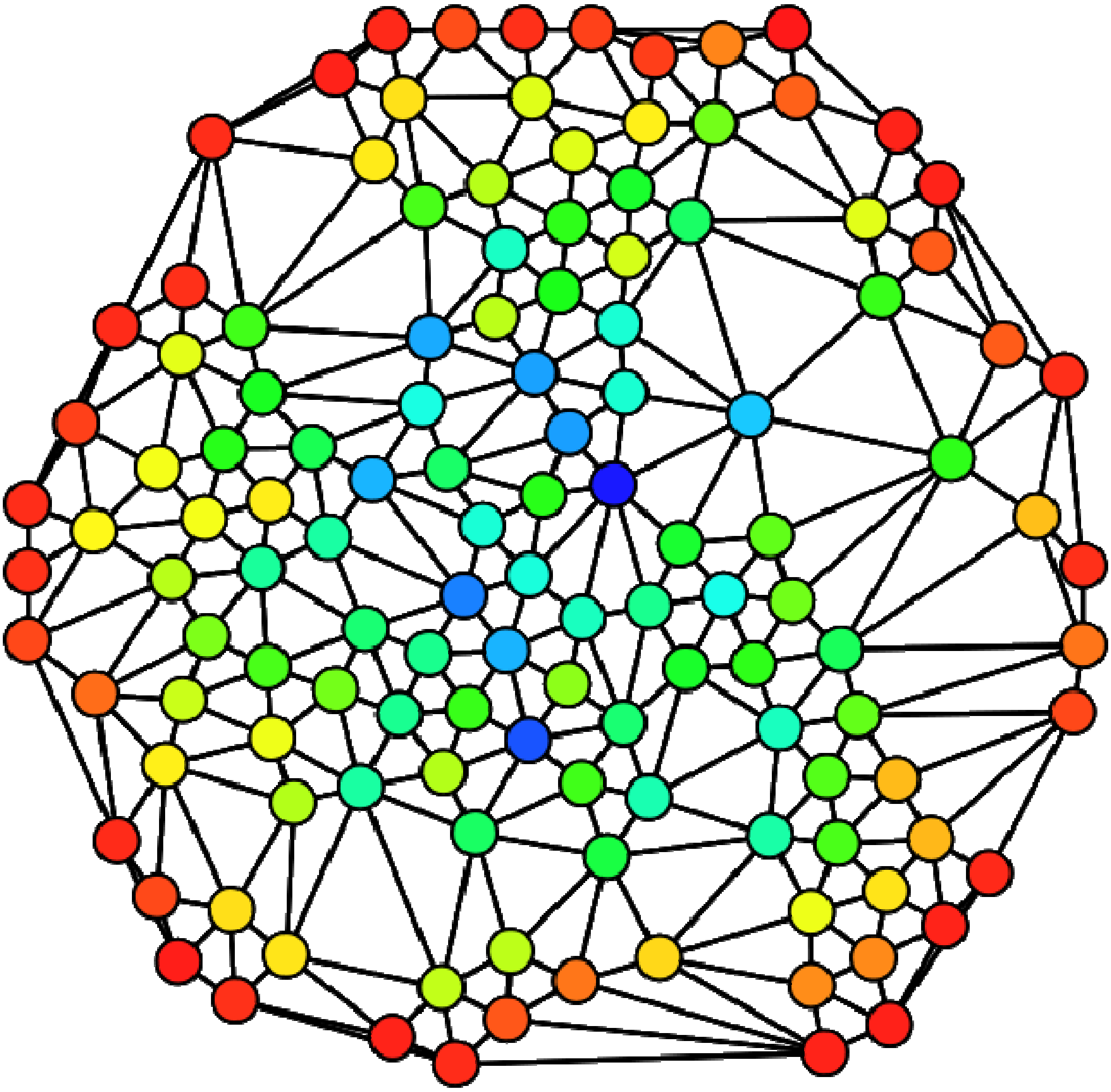
Betweenness centrality: calculates the relative number of indirect connections (or shortest paths) in which the actor into consideration is involved in with respect to the total number of paths in the network.



Agent in the middle: **Lowest degree** centrality, **highest betweenness** centrality.

Bonacich centrality: Depends on the value of the discount factor.

For small discount factors (i.e. indirect links give less benefits), this agent is the less central one while for high levels of discount (i.e. direct links are weighted less), this agent is the most central.



From red=0 to blue=max shows the node betweenness centrality

Bonacich Centrality

- The *Bonacich centrality* vector is

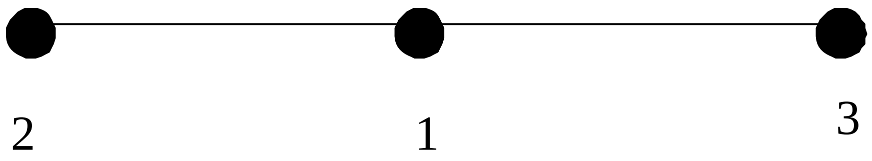
$$\mathbf{b}(G,\lambda) = \sum_{k=0}^{\infty} \lambda^k \mathbf{A}^k \cdot \mathbf{u} = (\mathbf{I} - \lambda \mathbf{A})^{-1} \cdot \mathbf{u}.$$

- For the components of the Bonacich vector we get

$$b_i(G,\lambda) = \sum_{k=0}^{\infty} \lambda^k (\mathbf{A}^k \cdot \mathbf{u})_i = \sum_{k=0}^{\infty} \lambda^k \sum_{j=1}^{n} \left(\mathbf{A}^k\right)_{ij}.$$

- $\sum_{j=1}^{n} \left(\mathbf{A}^k\right)_{ij}$ is the sum of all paths of length k in $\mathbf{A}$ starting from i.
- $b_i(G,\lambda)$ is the sum of all paths in G starting from i, where the paths of length k are weighted by their geometrically decaying factor λ^k.

- Example: Network g with three individuals (star)



When λ is small enough (i.e. $\lambda < 1/2^{0.5}$ since $2^{0.5}$ largest eigenvalue), then the vector of Bonacich network centralities is:

$$\mathbf{b}(G,\lambda) = \begin{bmatrix} b_1(G,\lambda) \\ b_2(G,\lambda) \\ b_2(G,\lambda) \end{bmatrix} = \frac{1}{\left(1-2\lambda^2\right)} \begin{bmatrix} 1+2\lambda \\ 1+\lambda \\ 1+\lambda \end{bmatrix}$$

Nash Equilibrium

Theorem:[3]

- An interior Nash equilibrium in pure strategies $x^* \in \mathbb{R}^n_+$ is such that for all $i = 1, ..., n$ we have that (FOC) $\frac{\partial u_i(\mathbf{x}^*)}{\partial x_i} = 0$. From this one can show that the Nash equilibrium is given by

$$x_i^* = b_i(G, \lambda)$$

where $b_i(G, \lambda)$ is the Bonacich network centrality of parameter $\lambda < 1/\lambda_{\mathrm{PF}}(G)$ and $\lambda_{\mathrm{PF}}(G)$ the largest real eigenvalue in the network G.
- Equilibrium payoffs are given by

$$u_i(\mathbf{x}^*, G) = \frac{1}{2}(x_i^*)^2 = \frac{1}{2} b_i^2(G, \lambda).$$

[3] C. Ballester, A. Calvo-Armengol and Y. Zenou, *Who's Who in Networks. Wanted: The Key Player*, Econometrica 74(5):1403–1417, 2006

Network Formation Game and Optimal Actions

Let time be measured at countable dates $t = 0, 1, 2, 3, ...$ Network formation process: $(G(t))_{t=0}^{\infty}$ with $G(t) = (N, L(t))$ (N: set of agents, $L(t)$: set of links)

Timing: At $t = 0$, we start with the empty network $G(0)$. Then every agent $i \in N$ optimally chooses her effort, which is $x_i^* = 1$, for all $i = 1, ..., n$ since there is no link and $b_i(G(0), \lambda) = 1$ for all $i = 1, .., n$.

Then, an agent i is chosen at random and with probability α forms a link with agent $j \in N \setminus (\mathcal{N}_i \cup \{i\})$ that gives her the highest utility (or equivalently her highest Bonacich centrality). We obtain the network $G(1)$ and all agents provides their optimal effort, which is $x_k^* = 1$, for all $k \in N \setminus (\{i\}, \{j\})$ and $x_i^* = x_j^* = b_i (G(1), \lambda) = b_j (G(1), \lambda)$.

Then, again, a player i is chosen at random and with probability α decides with whom she wants to form a link. For that, she has to calculate all the possible network configurations and chooses the one that gives her the highest utility. An so forth.

How individuals choose among their potential linking partners

If an agent is chosen (i.e. selected uniformly at random from the set N) , then she initiates a link to agent j (*anyone* in the network) which increases her equilibrium payoff the most. Agent j is said to be the *best response* of agent i given the network $G(t)$.

Agent j accepts the link if i also increases j's equilibrium payoff the most. That is, agent i is also a best response of agent j.

The underlying assumption for this is that individuals carefully decide with whom to interact and this decision entails some consent by both parts in a given relationship.

As we will see below, agent i is always a best response of agent j if agent j is a best response of agent i.

Agents create a link in a *myopic* way, that is they only look at any agent in the network that gives them the *current* highest utility.

Formal definition of best responses of an agent in network $G(t)$

Definition 0.1 *Consider the current network* $G(t, L(t))$ *with agents* $N = \{1, ..., n\}$ *and links* $L(t)$*. Let* $G(t) + ij$ *be the graph obtained from* $G(t)$ *by the addition of the edge* $ij \notin G(t)$ *between agents* $i \in N$ *and* $j \in N$*. Further, let* $\boldsymbol{\pi}^*(G(t)) = (\pi_1^*(G(t)), ..., \pi_n^*(G(t)))$ *denote the profile of Nash equilibrium payoffs of the agents in* $G(t)$ *following from the payoff function with parameter* $\lambda < 1/\lambda_{PF}(G(t))$*.*

Then agent j *is a best response of agent* i *if* $\pi_i^*(G(t) + ij) \geq \pi_i^*(G(t) + ik)$ *for all* $j, k \in N \setminus (\mathcal{N}_i \cup \{i\})$*. Agent* j *may not be unique. The set of agent* i*'s best responses is denoted by* $BR_i(G(t))$*.*

Definition 0.2 *We define the network formation process* $(G(t))_{t=0}^{\infty}$, $G(t) = (N, L(t))$, *as a sequence of networks* $G(0), G(1), G(2), ...$ *in which at every step* $t = 0, 1, 2, ...\}$, *an agent* $i \in N$ *is uniformly selected at random. Then one of the following two events occurs:*

a. *With probability* $\alpha \in (0, 1)$ *agent* i *initiates a link to a best response agent* $j \in BR_i(G(t))$. *Then the link* ij *is created if* $i \in BR_j(G(t))$ *is a best response of* j, *given the current network* $G(t)$. *If* $BR_i(G(t))$ *is not unique, then* i *selects randomly one agent in* $BR_i(G(t))$.

b. *With probability* $1-\alpha$ *the link* $ij \in G(t)$ *is removed such that* $\pi_i^*(G(t) - ij) \leq \pi_i^*(G(t) - ik)$ *for all* $j, k \in \mathcal{N}_i$. *If agent* i *does not have any link then nothing happens.*

Network Formation and Nested Split Graphs

- **Proposition:** For any graph G the following are equivalent:
  (i) G has a *stepwise* adjacency matrix $\mathbf{A}$.
  (ii) G is a *nested split graph*.[4]

[4]Nested split graphs are special types of *inter-linked stars*; (see Goyal, Sanjeev, *Connections: an introduction to the economics of networks*, Princeton University Press, 2008.)

Stepwise Adjacency Matrix

Definition (Brualdi '85): A stepwise matrix $\mathbf{A}$ is a matrix with elements a_{ij} satisfying the condition: if $i < j$ and $a_{ij} = 1$ then $a_{hk} = 1$ whenever $h < k \leq j$ and $h \leq i$.

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Definitions

A *dominating set* for a graph $G = (N, L)$ is a subset S of N such that every node not in S is connected to at least one member of S by a link.

An *independent set* is a set of nodes in a graph in which no two nodes are adjacent.

A *clique* in a graph is a subset of its vertices such that every two vertices in the subset are connected by an edge.

For example the central node in a star $K_{1,n-1}$ forms a *dominating set* while the peripheral nodes form an *independent set*.

Definition 0.3 *Let* $G = (N, L)$ *be a graph whose distinct positive degrees are* $d_{(1)} < d_{(2)} < ... < d_{(k)}$*, and let* $d_0 = 0$ *(even if no agent with degree* 0 *exists in* G*). Further, define* $D_i = \{v \in N : d_v = d_{(i)}\}$ *for* $i = 0, ..., k$*. Then the vector* $\mathbf{D} = (D_0, D_1, ..., D_k)$ *is called the degree partition of* G*.*

Let x be a real valued number $x \in \mathbb{R}$. Then, $\lceil x \rceil$ denotes the smallest integer larger or equal than x (the ceiling of x). Similarly, $\lfloor x \rfloor$ denotes the largest integer smaller or equal than x (the floor of x).

Formal definition of a nested split graph.

Definition 0.4 *Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, ..., D_k)$ be its degree partition. Then the nodes N can be partitioned in independent sets D_i, $i = 1, ..., \left\lfloor \frac{k}{2} \right\rfloor$ and a dominating set $\bigcup_{i=\left\lfloor \frac{k}{2} \right\rfloor+1}^{k} D_i$ in the graph $G' = (N \backslash D_0, L)$.*

Moreover, the neighborhoods of the nodes are nested. In particular, for each node $v \in D_i$, $i = 1, ..., k$,

$$N(v) = \begin{cases} \bigcup_{j=1}^{i} D_{k+1-j} & \text{if } i = 1, ..., \left\lfloor \frac{k}{2} \right\rfloor, \\ \bigcup_{j=1}^{i} D_{k+1-j} \setminus \{v\} & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, ..., k. \end{cases}$$

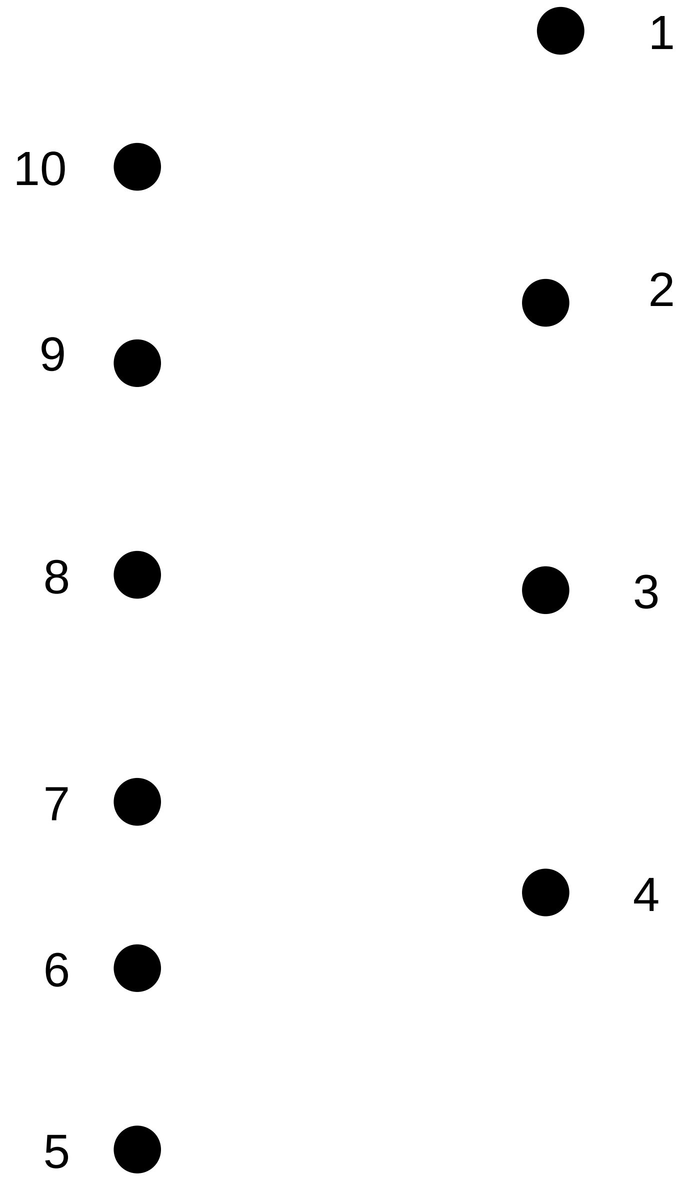
The sets D_i, $i = \left\lfloor \frac{k}{2} \right\rfloor + 1, ..., k$ are called **dominating subsets**, since the set D_i induces a dominating set in the graph obtained by removing the nodes in the set $\bigcup_{j=0}^{j=k-i} D_j$ from G.

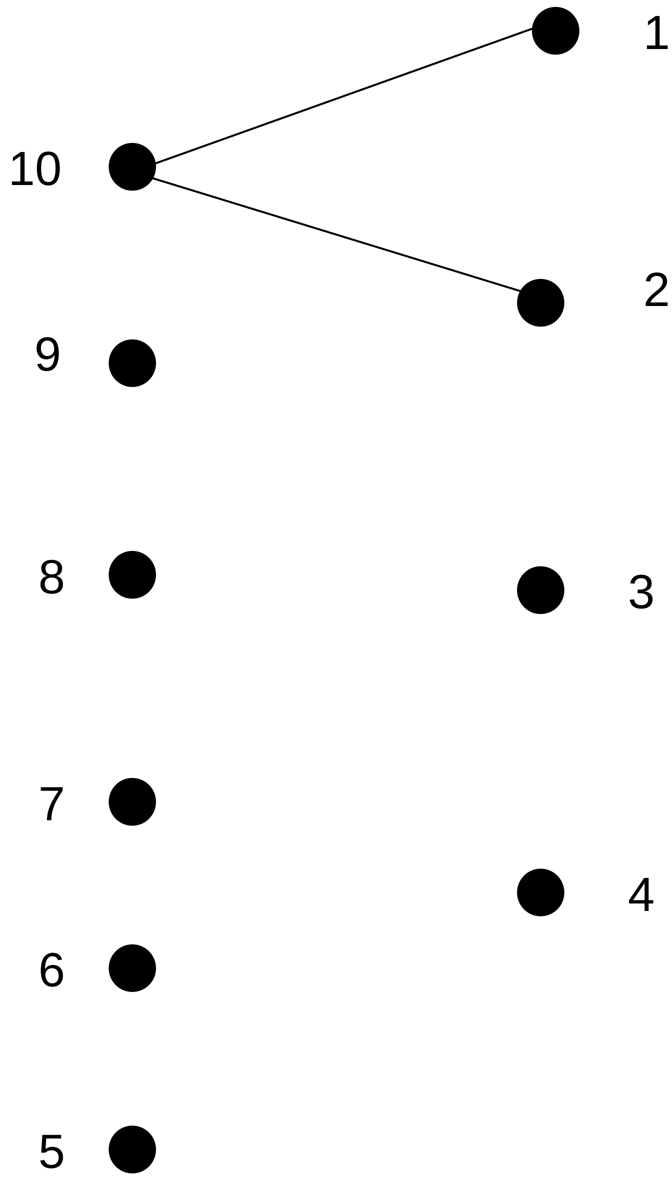
The dominating set (and all the dominating subsets) are cliques while the independent sets are empty subgraphs.

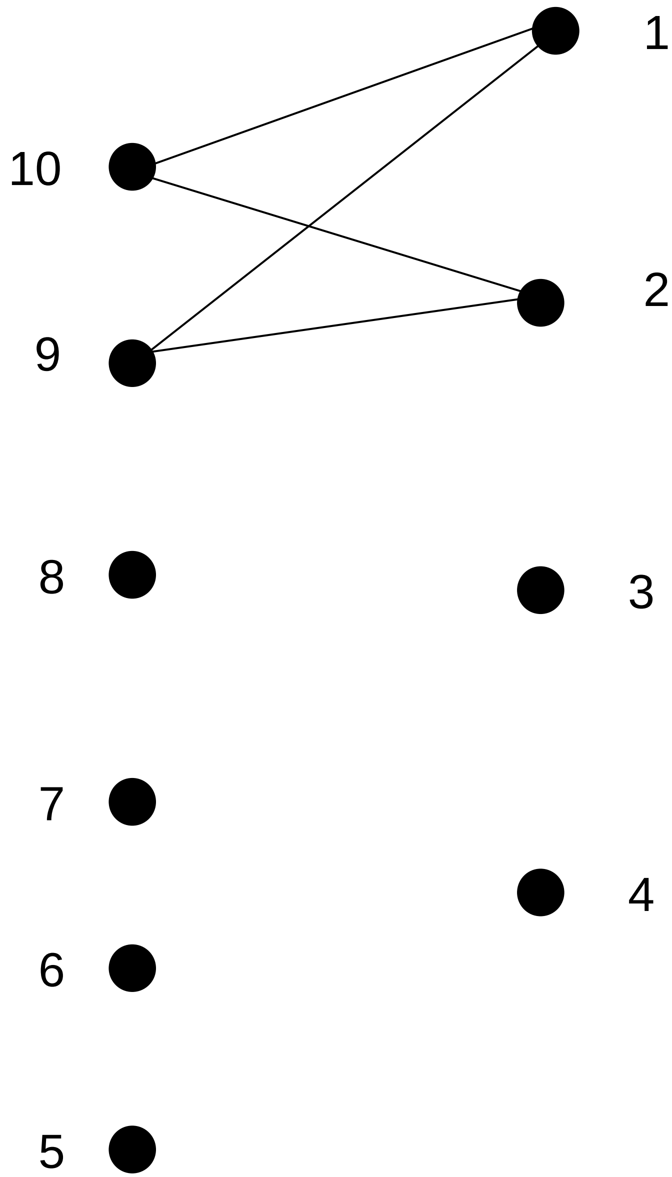
Nested split graphs have a nested neighborhood structure:

The set of neighbors of each node is contained in the set of neighbors of the next higher degree nodes.

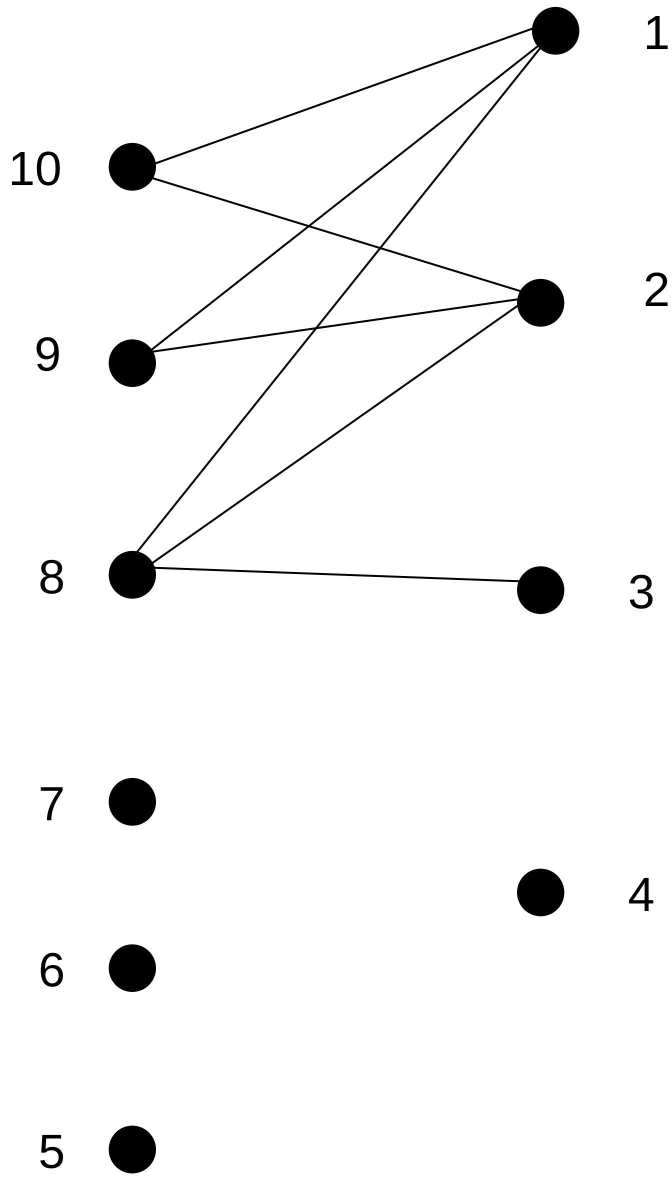
Threshold graphs have many different characterizations, and one of their pleasant features is that they are completely determined up to isomorphism by their vertex **degree sequence** $d = (d_1, ..., d_n)$, listed in weakly decreasing order.


1
10
2
9
8
3
7
4
6
5

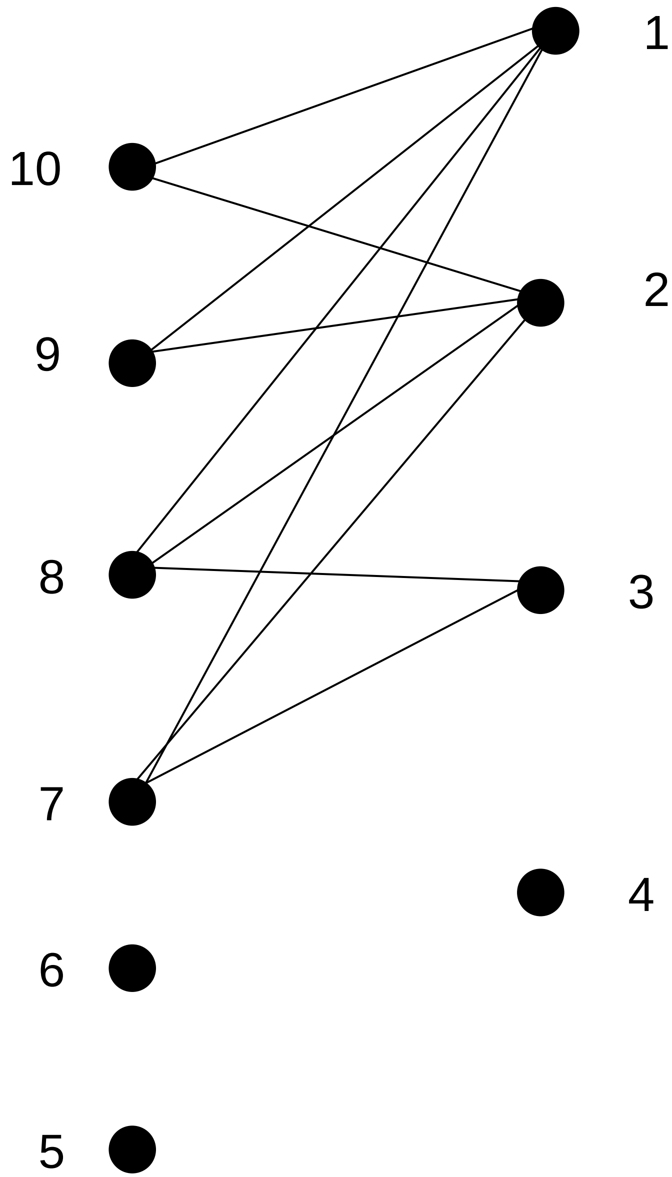

1
10
2
9
8
3
7
4
6
5



$D_1 = \{10,9\}$, $d_{(1)} = 2$

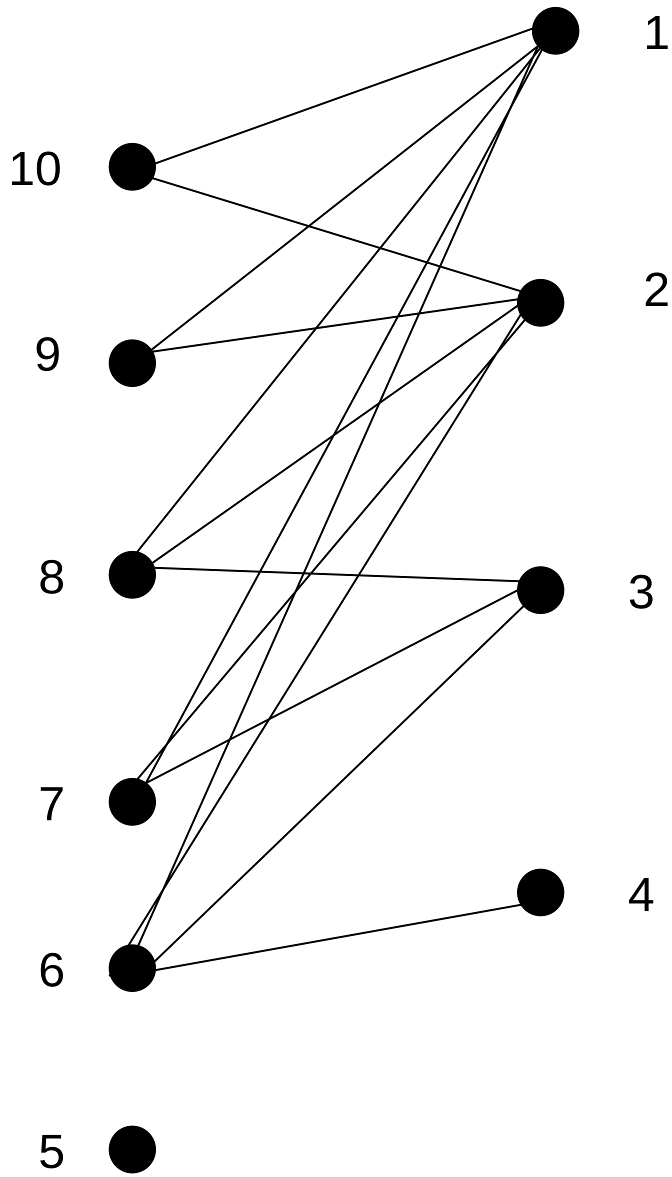


$D_1 = \{10,9\}$, $d_{(1)} = 2$



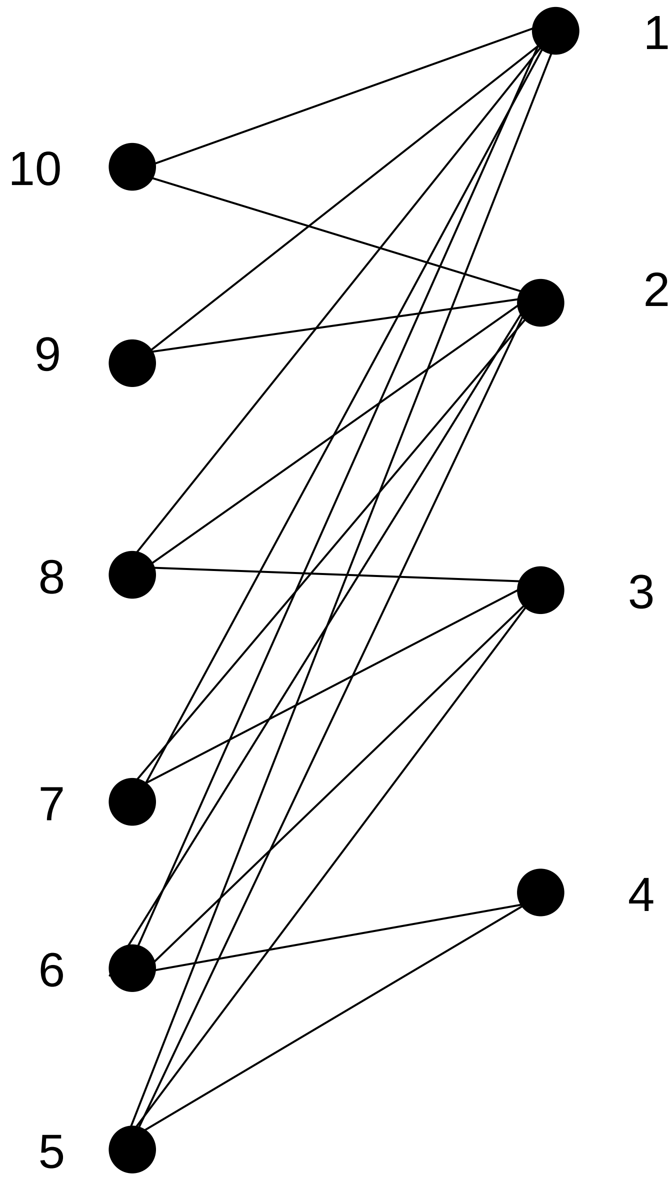
$D_1 = \{10,9\}$, $d_{(1)} = 2$

$D_2 = \{8,7\}$, $d_{(2)} = 3$



$D_1 = \{10,9\}$, $d_{(1)} = 2$

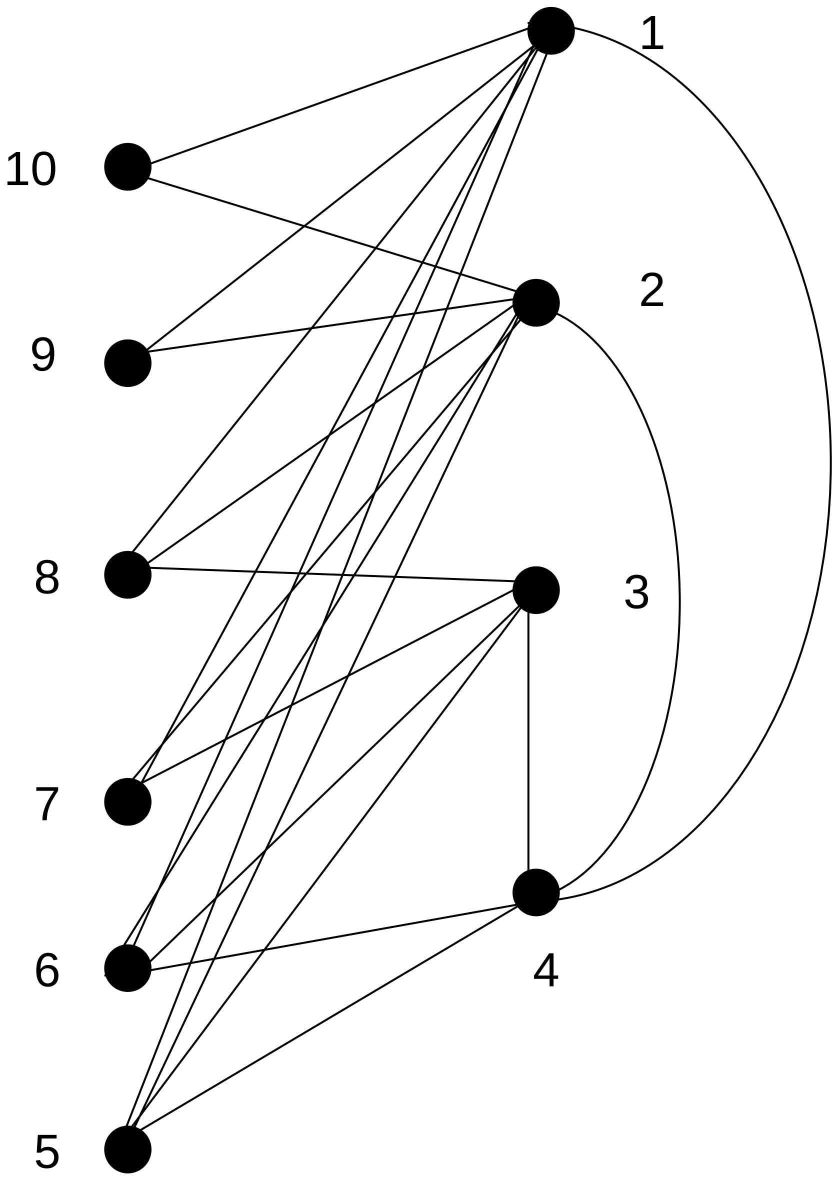
$D_2 = \{8,7\}$, $d_{(2)} = 3$



$D_1 = \{10,9\}$, $d_{(1)} = 2$

$D_2 = \{8,7\}$, $d_{(2)} = 3$

$D_3 = \{6,5\}$, $d_{(3)} = 4$

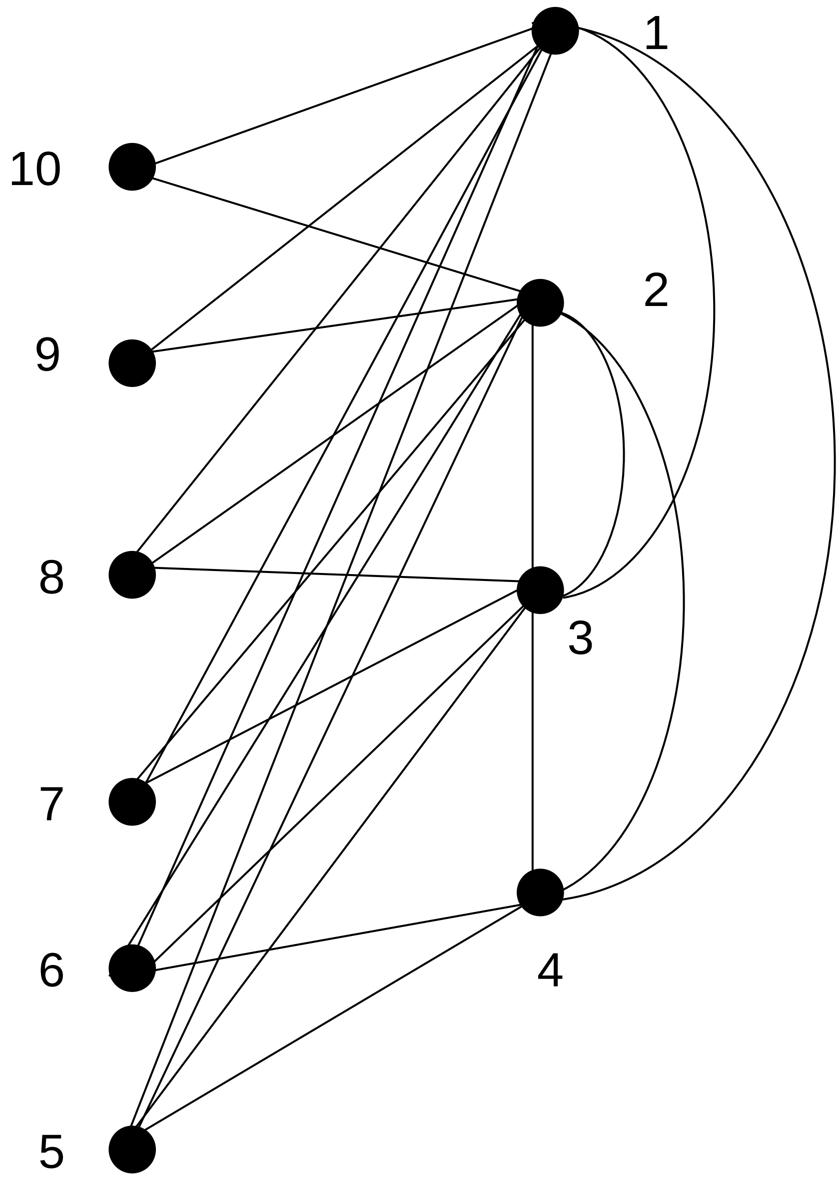


$D_1 = \{10,9\}$, $d_{(1)} = 2$

$D_2 = \{8,7\}$, $d_{(2)} = 3$

$D_3 = \{6,5\}$, $d_{(3)} = 4$

$D_4 = \{4\}$, $d_{(4)} = 5$



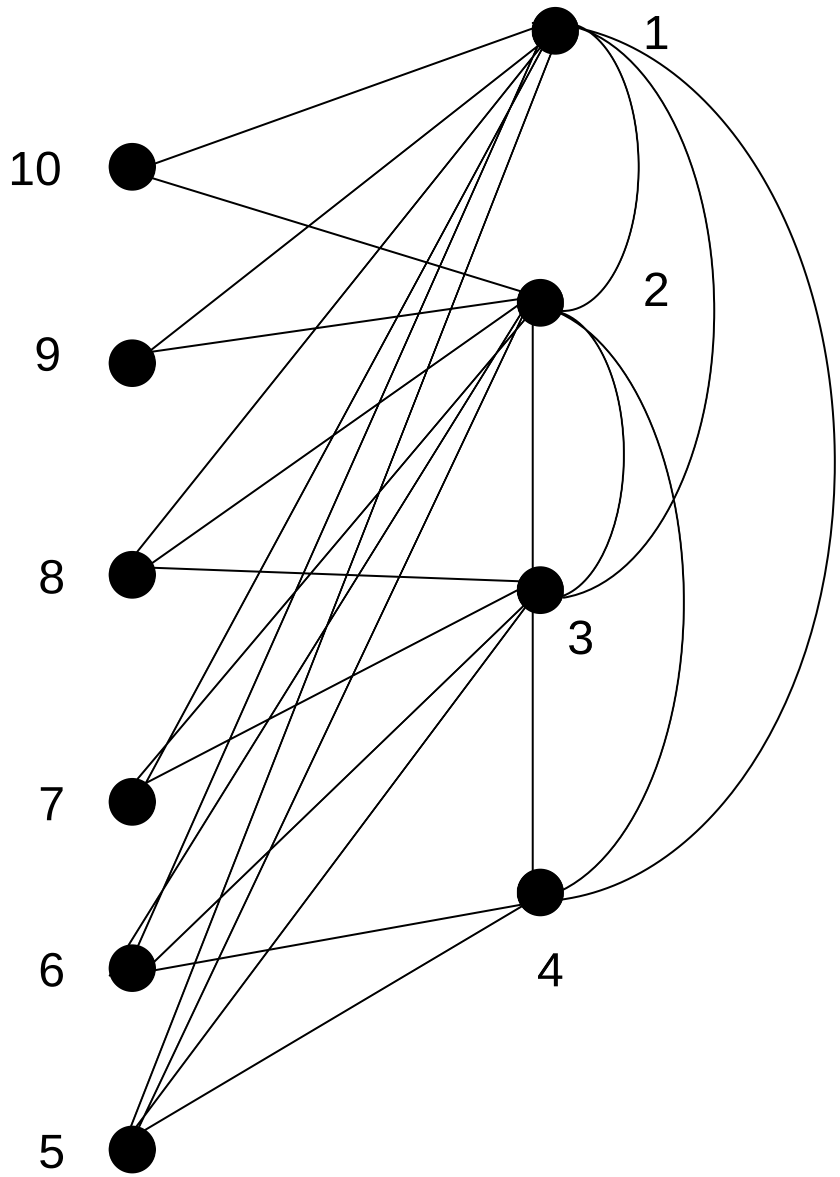
$D_1 = \{10,9\}$, $d_{(1)} = 2$

$D_2 = \{8,7\}$, $d_{(2)} = 3$

$D_3 = \{6,5\}$, $d_{(3)} = 4$

$D_4 = \{4\}$, $d_{(4)} = 5$

$D_5 = \{3\}$, $d_{(5)} = 7$



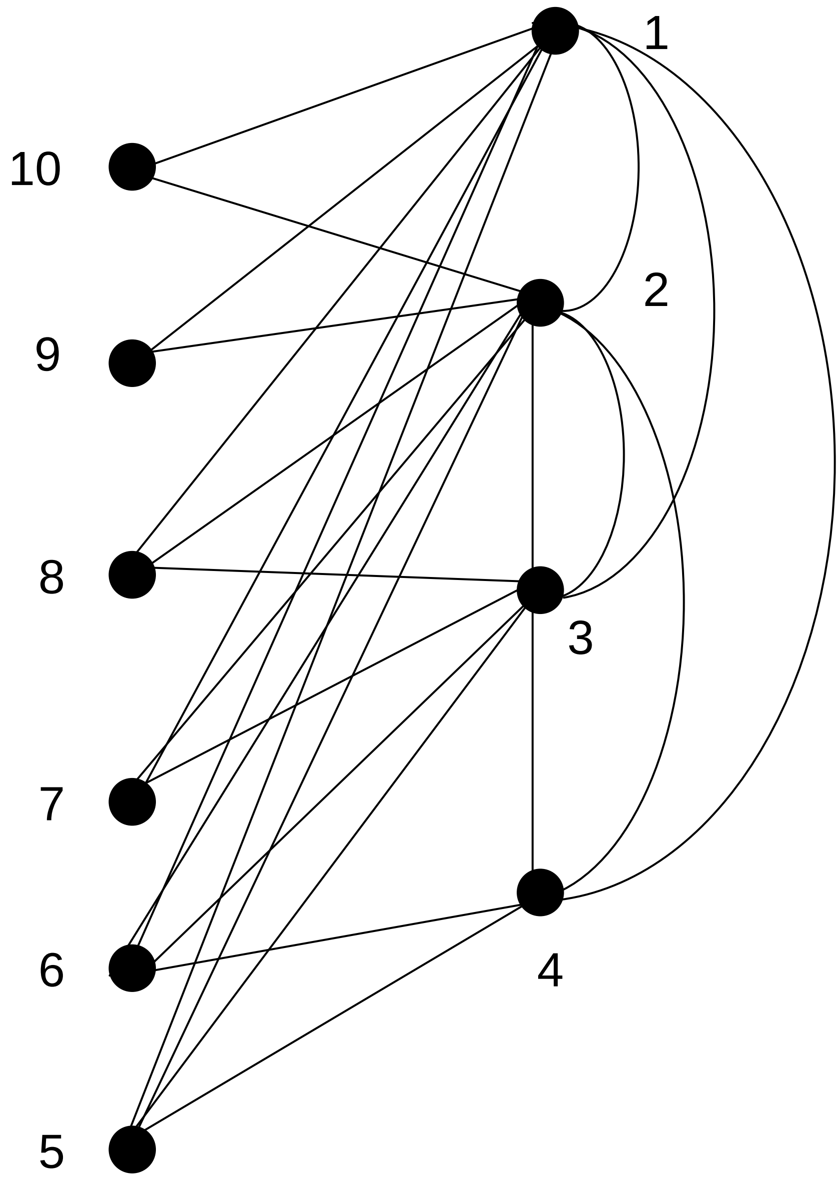
$D_1 = \{10,9\}$, $d_{(1)} = 2$

$D_2 = \{8,7\}$, $d_{(2)} = 3$

$D_3 = \{6,5\}$, $d_{(3)} = 4$

$D_4 = \{4\}$, $d_{(4)} = 5$

$D_5 = \{3\}$, $d_{(5)} = 7$



$D_1 = \{10,9\}$, $d_{(1)} = 2$

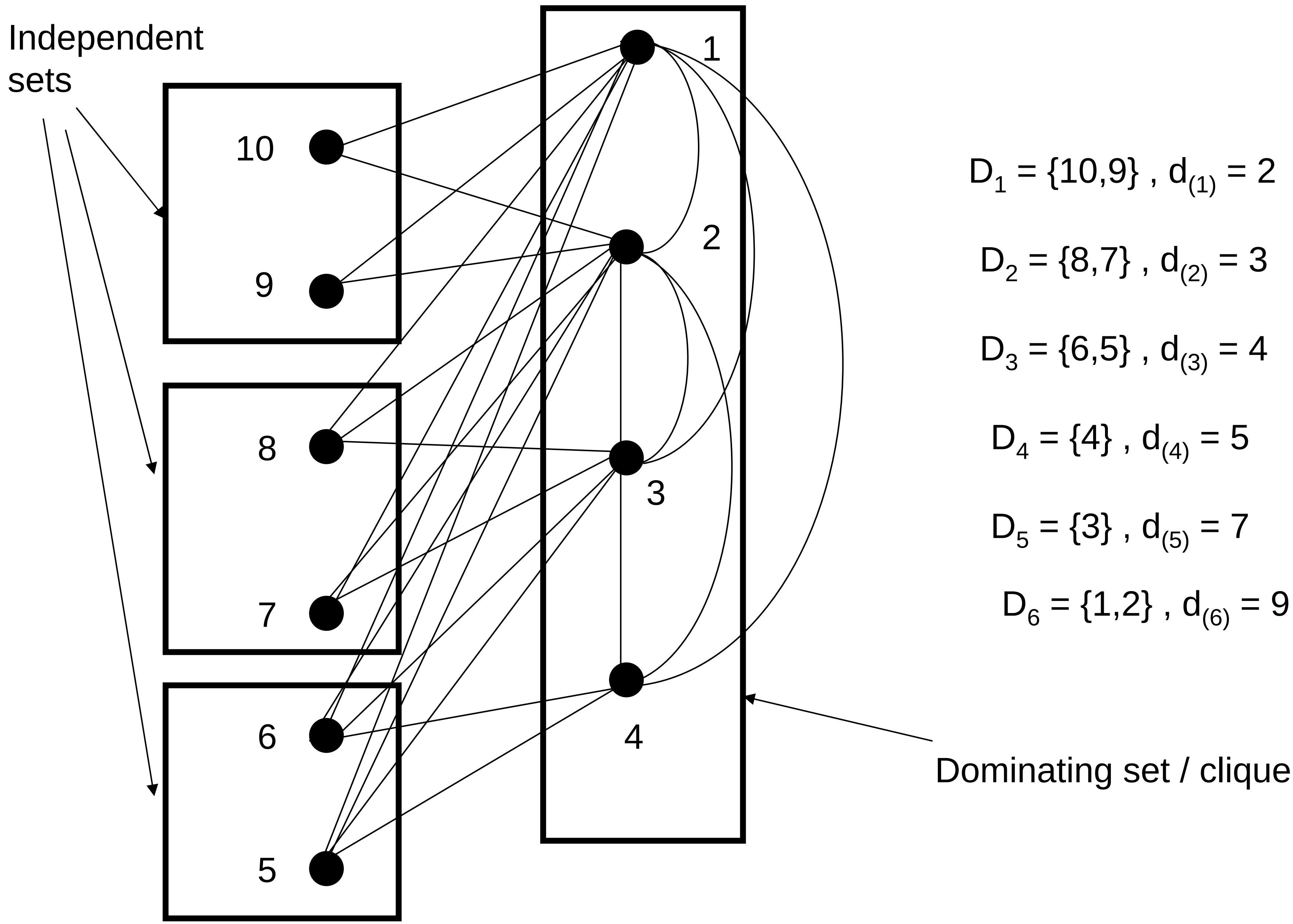
$D_2 = \{8,7\}$, $d_{(2)} = 3$

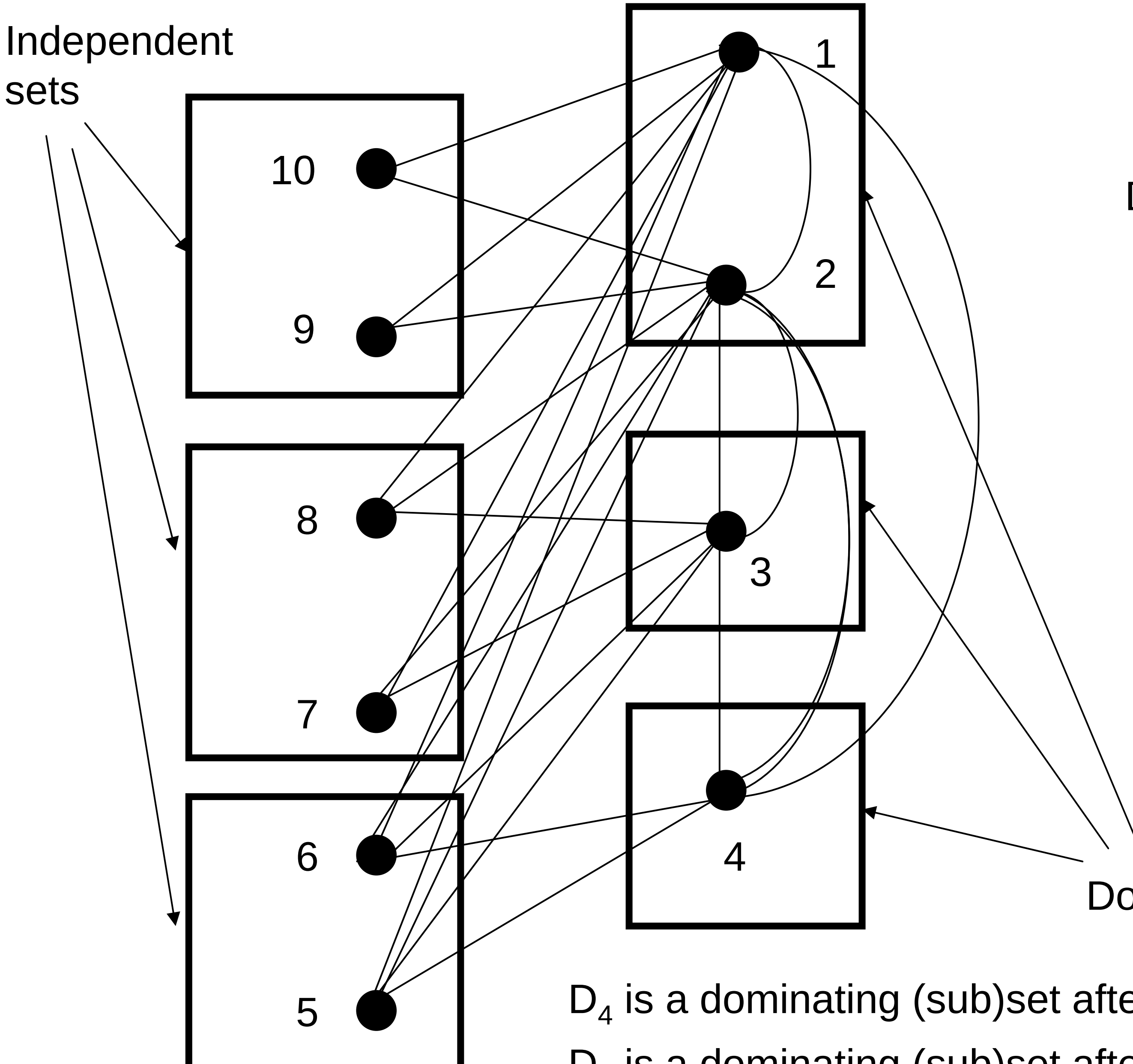
$D_3 = \{6,5\}$, $d_{(3)} = 4$

$D_4 = \{4\}$, $d_{(4)} = 5$

$D_5 = \{3\}$, $d_{(5)} = 7$

$D_6 = \{1,2\}$, $d_{(6)} = 9$


Independent sets
10
9
8
7
6
5
1
2
3
4
Dominating set / clique
$D_1 = \{10,9\}$, $d_{(1)} = 2$
$D_2 = \{8,7\}$, $d_{(2)} = 3$
$D_3 = \{6,5\}$, $d_{(3)} = 4$
$D_4 = \{4\}$, $d_{(4)} = 5$
$D_5 = \{3\}$, $d_{(5)} = 7$
$D_6 = \{1,2\}$, $d_{(6)} = 9$



$D_1 = \{10,9\}$, $d_{(1)} = 2$

$D_2 = \{8,7\}$, $d_{(2)} = 3$

$D_3 = \{6,5\}$, $d_{(3)} = 4$

$D_4 = \{4\}$, $d_{(4)} = 5$

$D_5 = \{3\}$, $d_{(5)} = 7$

$D_6 = \{1,2\}$, $d_{(6)} = 9$

D_4 is a dominating (sub)set after removing D_1 and D_2

D_5 is a dominating (sub)set after removing D_1

D_6 is a dominating (sub)set

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

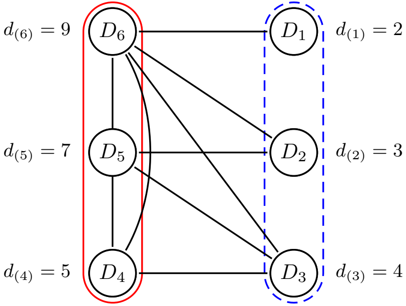


Figure (left) illustrates the degree partition

$$\mathbf{D} = (D_0, D_1, ..., D_6)$$

and the nested neighborhood structure of a nested split graph.

A nested split graph has an associated adjacency matrix which is called *stepwise matrix* and it is defined as follows.

Definition 0.5 *A stepwise matrix* $\mathbf{A}$ *is a matrix with elements* a_{ij} *satisfying the condition: if* $i < j$ *and* $a_{ij} = 1$ *then* $a_{hk} = 1$ *whenever* $h < k \leq j$ *and* $h \leq i$*.*

Figure (right) shows the stepwise adjacency matrix $\mathbf{A}$ corresponding to the nested split graph shown on the left hand side. If we let the nodes by indexed by the order of the rows in the adjacency matrix $\mathbf{A}$ then it is easily seen that for example

$$D_6 = \{1, 2 \in N : d_1 = d_2 = d_{(6)} = 9\}$$

and

$$D_1 = \{9, 10 \in N : d_9 = d_{10} = d_{(1)} = 2\}$$

From the stepwise property of the adjacency matrix it follows that a connected nested split graph contains at least one spanning star, that is, there is at least one agent that is connected to all other agents.

Nested Split Graphs and Interlinked Stars

- **Definition (Goyal '03):** An *interlinked star* has a degree partition $\mathbf{D} = (D_1, D_2, ..., D_m)$ with two features:
  (i) $N_i = n - 1$ for all nodes $i \in D_m(G)$, and
  (ii) $N_j = D_m$ for all nodes $j \in D_1(G)$.
- A nested split graph is an interlinked star but an interlinked star is not necessarily a nested split graph. This is due to the nested neighborhood structure that is also required for the sets h_l for $1 < l < m$ in a nested split graph.

Proposition. A nested split graph consists of at most one connected component and the nodes not in that component are all singletons. The distance between nodes in the connected component is at most two.

The empty network and the complete network are nested split graphs.

On these graphs, two nodes are either friends, friends of friends (i.e. separated by one node) or there exists no path between them.

We also derive the clustering coefficient, the neighbor connectivity and the characteristic path length of a nested split graph.

In particular, we show that connected nested split graphs have small characteristic path length, which is at most two.

We also analyze different measures of centrality in a nested split graph.

One important result is that degree, closeness, and Bonacich centrality induce the same ordering of nodes in a nested split graph. If the ordering is not strict, then this holds also for betweenness centrality.

Bonacich Centrality in Nested Split Graphs

Lemma: In a nested split graph G the following two properties hold:

(i) If and only if agent i has a higher degree than agent j then i has a higher Bonacich centrality than j, i.e.

$$d_i > d_j \Leftrightarrow b_i(G, \lambda) > b_j(G, \lambda).$$

(ii) Assume that neither the links ik nor ij are in G, $ij \notin L$ and $ik \notin L$. Further assume that agent k has a higher degree than agent j, $d_k > d_j$. Then adding the link ik to G increases the Bonacich centrality of agent i more that adding the link ij to G, i.e.

$$d_k > d_j \Leftrightarrow b_i(G + ik, \lambda) > b_i(G + ij, \lambda).$$

Proposition 0.1 *Consider the network formation process $(G(t))_{t \in T}$ introduced in Definition above. Then, at any time t, a network $G(t)$ generated by $(G(t))_{t \in T}$ is a nested split graph.*

This result is due to the fact that agents, when they have the possibility of creating a new link, always connect to the agent who has the highest Bonacich centrality (and the highest degree) among her second-order neighbors.

(Recall that all nodes with the same degree, i.e. belonging to the same partition, are linked to exactly the same nodes. The nodes being identical reveals that centrality ranking will not differ from degree ranking).

This creates a nested neighborhood structure which can always be represented by a stepwise adjacency matrix after a possible relabeling of the agents. The same applies for link removal.

Due to the nested neighborhood structure of nested split graphs, any pair of agents in (the connected component) of a nested split graph is at most two links separated from each other.

From Proposition above it then follows that, in a nested split graph $G(t)$, the best response of an agent i are the agents with the highest degrees in i's second-order neighborhood $\mathcal{N}_i^{(2)} = \bigcup_{j \in \mathcal{N}_i} \mathcal{N}_j \backslash (\mathcal{N}_i \cup \{i\})$.

Moreover, if $G(t)$ is a nested split graph then $i \in BR_j(G(t))$ if and only if $j \in BR_i(G(t))$.

Proof by induction:

Induction basis: $Gt0)$ **to** $G(1)$

Start at $t = 0$ from an empty network $G(0)$, which has a trivial stepwise adjacency matrix.

At $t = 1$, we select an agent and connect her to another one. All isolated agents are best responses of the selected agent. This creates a path of length one whose adjacency matrix is stepwise.

This is true because we can always find a simultaneous columns and rows permutation which makes the adjacency matrix stepwise.

Thus $G(1)$ has a stepwise adjacency matrix.

Induction step: $G(t)$ **to** $G(t+1)$

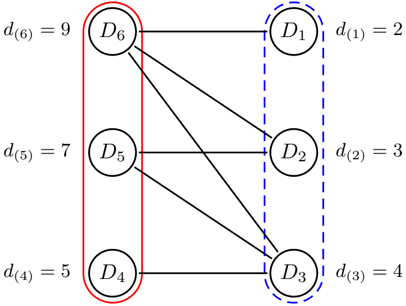
$G(t)$ is a nested split graph with a stepwise adjacency matrix.

Consider the creation of a link ij.

From the stepwise adjacency matrix $A(G(t))$ of $G(t)$, we find that adding the link ij to the network $G(t)$ such that both agents are the agents with the highest degrees results in a matrix $\mathbf{A}(G(t)+ij)$ that is stepwise.

Therefore, the network $G(t)+ij$ is a nested split graph.

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$



Agents are numbered by the rows respectively columns of the adjacency matrix.

0	1	1	1	1	1	1	1	1	1
1	0	1	1	1	1	1	1	1	1
1	1	0	1	1	1	1	1	0	0
1	1	1	0	1	1	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0

Agent 9 is selected to create a link
He will create a link to agent 3 (BR to both 9 and 3)

0	1	1	1	1	1	1	1	1	1
1	0	1	1	1	1	1	1	1	1
1	1	0	1	1	1	1	1	**1**	0
1	1	1	0	1	1	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0
1	1	**1**	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0

Creating the link 39 preserves the stepwise form of the adjacency matrix

0	1	1	1	1	1	1	1	1	1
1	0	1	1	1	1	1	1	1	1
1	1	0	1	1	1	1	1	$\boxed{\mathbf{1}}$	0
1	1	1	0	1	1	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0	0
1	1	$\boxed{\mathbf{1}}$	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0

By creating link 39

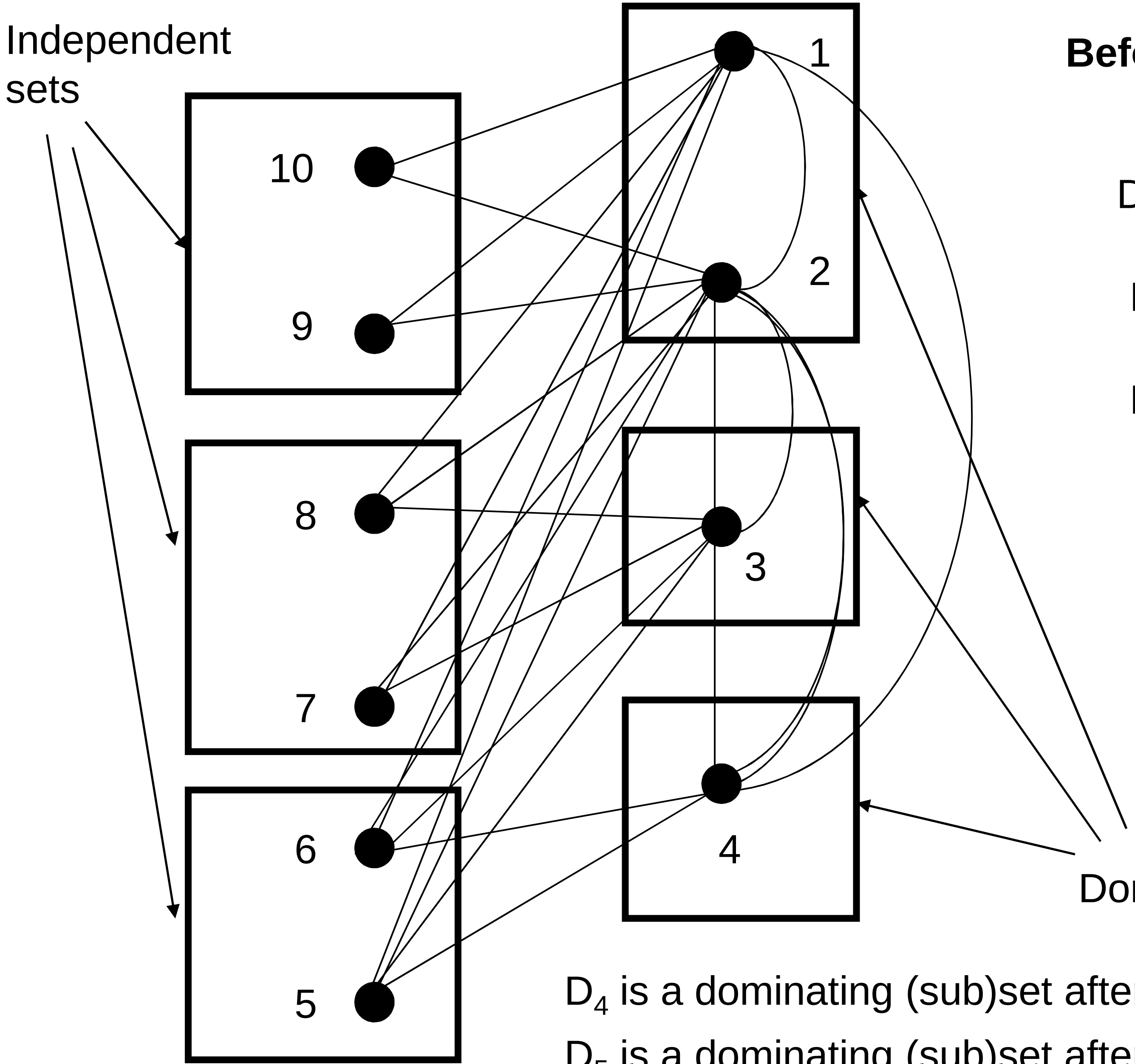
Switch from: $D_1=\{9,10\}$, $D_2=\{7,8\}$, $D_3=\{5,6\}$, $D_4=\{4\}$, $D_5=\{3\}$, $D_6=\{1,2\}$,

with $d_{(1)}=2$, $d_{(2)}=3$, $d_{(3)}=4$, $d_{(4)}=5$, $d_{(5)}=7$, $d_{(6)}=9$.

To: $D_1=\{10\}$, $D_2=\{7,8,9\}$, $D_3=\{5,6\}$, $D_4=\{4\}$, $D_5=\{3\}$, $D_6=\{1,2\}$,

with $d_{(1)}=2$, $d_{(2)}=3$, $d_{(3)}=4$, $d_{(4)}=5$, $d_{(5)}=8$, $d_{(6)}=9$.

Before the creation of 39



$D_1 = \{9,10\}$, $d_{(1)} = 2$

$D_2 = \{7,8\}$, $d_{(2)} = 3$

$D_3 = \{5,6\}$, $d_{(3)} = 4$

$D_4 = \{4\}$, $d_{(4)} = 5$

$D_5 = \{3\}$, $d_{(5)} = 7$

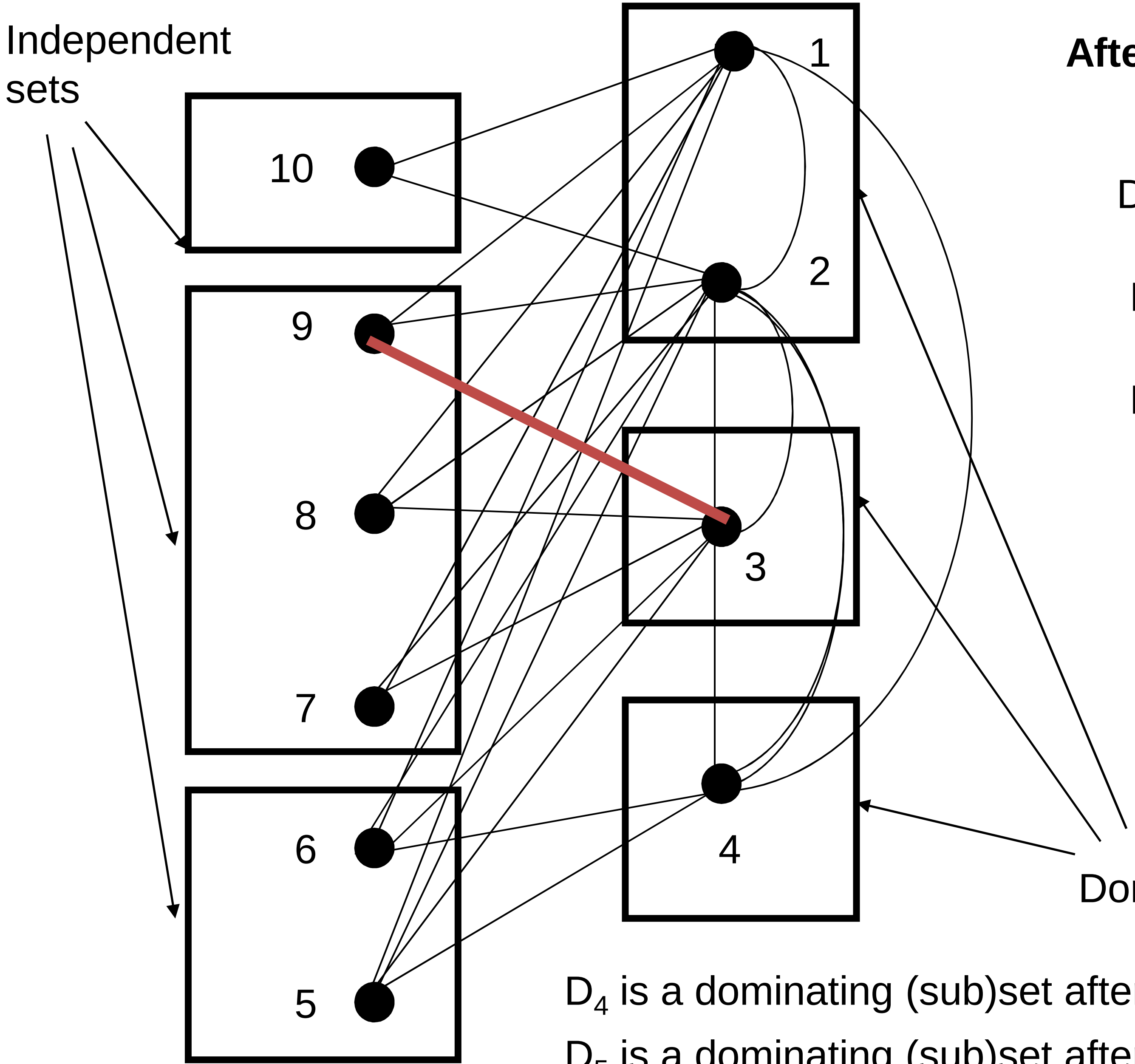
$D_6 = \{1,2\}$, $d_{(6)} = 9$

D_4 is a dominating (sub)set after removing D_1 and D_2

D_5 is a dominating (sub)set after removing D_1

D_6 is a dominating (sub)set

After the creation of 39



$D_1 = \{10\}$, $d_{(1)} = 2$

$D_2 = \{7,8,9\}$, $d_{(2)} = 3$

$D_3 = \{5,6\}$, $d_{(3)} = 4$

$D_4 = \{4\}$, $d_{(4)} = 5$

$D_5 = \{3\}$, $d_{(5)} = 8$

$D_6 = \{1,2\}$, $d_{(6)} = 9$

D_4 is a dominating (sub)set after removing D_1 and D_2

D_5 is a dominating (sub)set after removing D_1

D_6 is a dominating (sub)set

Stationary Networks: Characterization

The network formation process $(G(t))_{t \in T}$ is a Markov process. Denote by Ω the state space of $(G(t))_{t=0}^{\infty}$. Ω consists of all possible unlabeled nested split graphs on n nodes. It can be shown that the number of possible states is $|\Omega| = 2^{n-1}$. Thus the number of states is finite and the transition between states can be represented by a transition matrix $\mathbf{P}$.

$(G(t))_{t=0}^{\infty}$ is a Markov chain. Indeed, the network $G(t+1) \in \Omega$ is obtained from $G(t)$ by removing or adding a link to $G(t)$. Thus, the probability of obtaining $G(t+1)$ depends only on $G(t)$ and not on the previous networks for $t' < t$.

The number of possible networks $G(t)$ is finite for any discrete time $t \geq 0$ and the transition probabilities from a network $G(t)$ to $G(t+1)$ do not depend on t. Therefore, $(G(t))_{t=0}^{\infty}$ is a finite state, discrete, time-homogeneous Markov process.

Transition matrix

$$\mathbf{P}_{ij} = P(G(t+1) = G_j \mid G(t) = G_i) \text{ for any } G_i, G_j \in \Omega$$

Proposition 0.2 *The network formation process* $(G(t))_{t=0}^{\infty}$ *induces an ergodic Markov chain on the state space* Ω *with a unique stationary distribution* $\boldsymbol{\mu}$*. In particular, the state space* Ω *is finite and consists of all possible unlabeled nested split graphs on* n *nodes, where the number of possible states is given by* $|\Omega| = 2^{n-1}$*.*

Uniqueness: This Markov chain is irreducible and aperiodic.

There exists a unique stationary distribution μ satisfying: $\mu\,\mathbf{P} = \mu$

Our network formation process $(G(t))_{t=0}^{\infty}$ is independent of initial conditions $G(0)$.

This means that even when we start from an initial network G which is not a nested split graph then after some finite time the Markov chain will reach a nested split graph.

From then on all consecutive networks visited by the chain are nested split graphs.

A nested split graph is thus an absorbing state.

Proof: Network formation process $(G(t))_{t=0}^{\infty}$ independent of initial conditions $G(0)$

Let Ω be the set of nested split graphs.

For any graph $G \notin \Omega$, there exists a finite probability that in all consecutive steps agents remove their links until the empty network $G(0)$ is obtained.

Let T be the time when this happens starting from some network $G \notin \Omega$. Note that $G(0) \in \Omega$.

A network G is transient if $\sum_{\tau=1}^{\infty} \mathbb{P}(G(t+\tau) = G | G(t) = G) < \infty$.

We have that

$$\begin{aligned}\sum_{\tau=1}^{\infty} \mathbb{P}(G(t+\tau) &= G|G(t)=G) \\ &= \sum_{\tau=1}^{T} \mathbb{P}(G(t+\tau)=G|G(t)=G) < T < \infty\end{aligned}$$

Therefore all networks which are not nested split graphs are transient and they have vanishing probability in the stationary distribution μ

Proposition 0.3 *Consider the network formation process* $(G(t))_{t=0}^{\infty}$ *with link creation probability* α *and the network formation process* $(G'(t))_{t=0}^{\infty}$ *with link creation probability* $1-\alpha$. *Let* $\boldsymbol{\mu}$ *be the stationary distribution of* $(G(t))_{t=0}^{\infty}$ *and* $\boldsymbol{\mu}'$ *the stationary distribution of* $(G'(t))_{t=0}^{\infty}$. *Then for each network* G *in the stationary distribution* $\boldsymbol{\mu}$ *with probability* μ_G *the complement of* G, $\bar{G}$, *has the same probability* μ_G *in* $\boldsymbol{\mu}'$, *i.e.* $\mu'_{\bar{G}} = \mu_G$.

The complement $\bar{G}$ of a nested split graph G is a nested split graph as well.

$\bar{G}$ has an adjacency matrix which is obtained from G by replacing each 1 by 0 and each 0 by 1, except for the elements in the diagonal.

The networks $\bar{G}$ are nested split graphs in which the number of nodes in the dominating subsets corresponds to the number of nodes in the independent sets in G and vice versa.

The structure of nested split graphs implies that if there exist nodes for all degrees between 0 and d^* (in the independent sets), then the dominating subsets with degrees larger than d^* contain each a single node.

Using Proposition 3 (symmetry), for $\alpha > 1/2$, the expected number of nodes in the dominating subsets is given by the expected number of nodes in the independent sets for $1 - \alpha$ from the previous equation (i.e. n_d), while each of the independent sets contains a single node.

For any $G_1, G_2 \in \Omega$,

$$P(G(t+1) = G_2 \mid G(t) = G_1)$$

$$= P(G'(t+1) = \overline{G}_2 \mid G(t) = \overline{G}_1)$$

Notations:

$\{N(t)\}_{t=0}^{\infty}$: degree distribution

$N_d(t)$: Number of nodes with degree d at time t in $G(t)$

$n_d(t) = N_d(t)/n$: Proportion of nodes with degree d at time t in $G(t)$.

$n_d = \lim_{t\to\infty} \mathsf{E}[n_d(t)]$: Asymptotic expected value of $n_d(t)$

The following proposition determines the asymptotic degree distribution (i.e. t large enough) of the nodes in the independent sets for n large enough.

Proposition 0.4 *Let $0 \leq \alpha \leq 1/2$. Then the asymptotic expected proportion n_d of nodes in the independent sets with degrees, $d = 0, 1, ..., d^*$, for large n is given by:*

$$n_d = \frac{(1 - 2\alpha)}{(1 - \alpha)} \left(\frac{\alpha}{1 - \alpha} \right)^d$$

where

$$d^*(n, \alpha) = \frac{\ln \left(\frac{(1-2\alpha)n}{2(1-\alpha)} \right)}{\ln \left(\frac{1-\alpha}{\alpha} \right)}$$

The structure of nested split graphs implies that if there exist nodes for all degrees between 0 and d^* (in the independent sets), then the dominating subsets with degrees larger than d^* contain each a single node.

Using Proposition 3 (symmetry), for $\alpha > 1/2$, the expected number of nodes in the dominating subsets is given by the expected number of nodes in the independent sets for $1 - \alpha$ from the previous equation (i.e. n_d), while each of the independent sets contains a single node.

A nested split graph is uniquely defined by its degree distribution. The degree distribution uniquely determines the corresponding nested split graph up to a permutation of the indices of nodes.

Thus this Proposition delivers us a complete description of a typical network generated by our model in the limit of large t and n.

We call this network the “stationary network”.

We can compute the degree distribution and the corresponding adjacency matrix of the stationary network for different values of α.

Probability Limit of Degree Distribution

Proposition: For any $\epsilon > 0$ we have that

$$\mathrm{Prob}\left(\left|n_d(t) - \mathrm{E}\left(n_d(t)\right)\right| \geq \epsilon\right) \leq 2e^{-\frac{\epsilon^2 n^2}{8t}}.$$

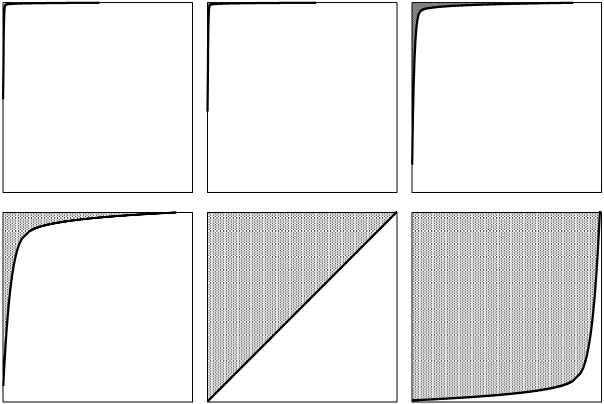
- The empirical degree distribution converges in probability to the expected distribution in the limit of large population sizes n.
- As we will show later, clustering coefficient $\mathcal{C}$, clustering-degree correlation, neighbor connectivity d_{nn}, assortativity γ, characteristic path length $\mathcal{L}$ and various centrality measures are functions of the degree distribution $\{n_d\}_{d=0}^{n-1}$. Since we know the plim of n_d, we also know the plim of these quantities.

Phase Transition

Proposition: In the limit of large n there exists a phase transition in the asymptotic average number of independent sets as n becomes large such that

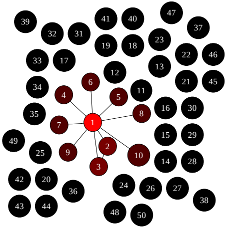
$$\lim_{n\to\infty}\frac{d^*(n,\alpha)}{n}=\begin{cases}0, & \text{if } \alpha<\frac{1}{2},\\ \frac{1}{2}, & \text{if } \alpha=\frac{1}{2},\\ 1, & \text{if } \alpha>\frac{1}{2}.\end{cases}$$

"Expected" Adjacency Matrix

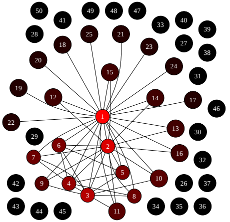


⇒ Low α, hierarchical, centralized network.

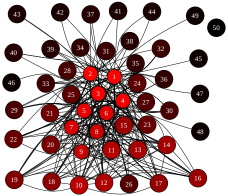
⇒ High α, highly decentralized network.



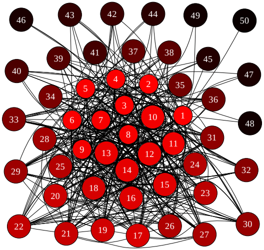
$\alpha = 0.2$



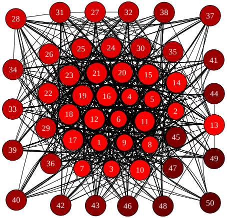
$\alpha = 0.4$



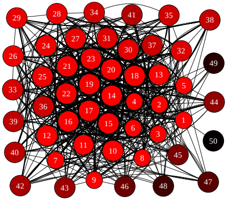
$\alpha = 0.48$



$\alpha = 0.495$



$\alpha = 0.5$



$\alpha = 0.52$

Stationary Network Statistics

Main stylized facts of social and economic networks are

- Sparseness.
- High clustering & short path length (small world).
- Skewed degree distributions (power law, exponential).
- Negative clustering-degree correlation.
- Betweenness centrality degree correlation.
- Degree-degree correlations: dissortativity.
- Core-periphery structure.

Degree Distribution

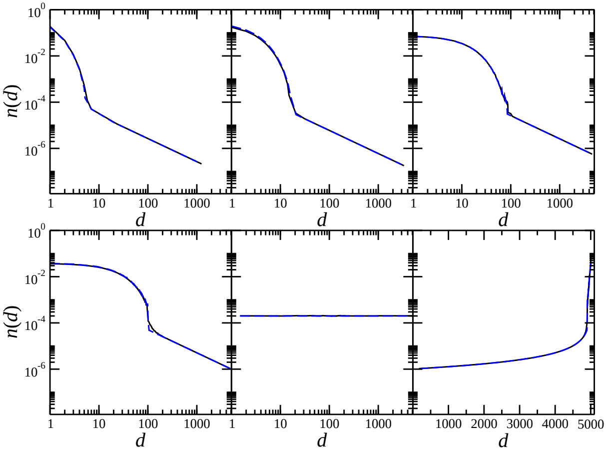
- The expected degree distribution is given by[9]

$$n(d) = \begin{cases} \frac{1-2\alpha}{1-\alpha} e^{-2(1-2\alpha)d}, & \text{if } 0 \leq d \leq d^*, \\ \frac{\alpha}{(1-2\alpha)n} d^{-1}, & \text{if } d^* < d \leq n-1 \end{cases}$$

- Degree distributions with exponential and power-law parts have been found in empirical networks, e.g. in email communication networks.[10]
- For $\alpha = 1/2$ the degree distribution is uniform while for larger values of α most of the agents have a degree close to the maximum degree.

[9] The degree distribution for degrees larger than d^* has been obtained by logarithmic binning.

[10] Guimera, R., Diaz-Guilera, A., Giralt, F., Arenas, A., 2006. *The real communication network behind the formal chart: Community structure in organizations.* Journal of Economic Behavior and Organization 61, 653 667.


n(d)
n(d)
10^0
10^-2
10^-4
10^-6
1
10
100
1000
1000
2000
3000
4000
5000
d

Clustering

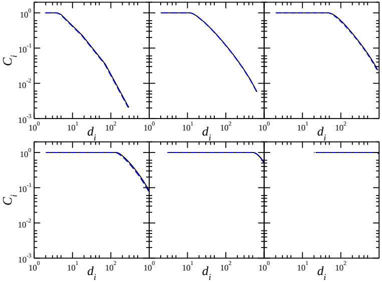
The clustering coefficient $C(v)$ for an agent v is the proportion of links between the agents in her neighborhood $\mathcal{N}_v$ divided by the number of links that could possibly exist between them. That is

$$C(v) = \frac{\{uw : u, w \in \mathcal{N}_v \wedge uw \in L\}}{d_v(d_v - 1)/2}$$

Clustering

- Consider a nested split graph G and let $\mathbf{D} = (D_0, D_1, ..., D_k)$ be the degree partition of G. Denote by $S_D^i = \sum_{j=i}^k |D_j|$. Then for each $v \in D_i$, $i = 0, ..., k$, (k even)

$$
\mathcal{C}(v) = \begin{cases} 0, & \text{if } i = 0, \\ 1, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{1}{d_u(d_u-1)} \left[\left(S_D^{\left\lfloor \frac{k}{2} \right\rfloor + 1} - 1 \right) \left(S_D^{\left\lfloor \frac{k}{2} \right\rfloor + 1} - 2 \right) + \right. & \\ \left. 2|D_{\left\lfloor \frac{k}{2} \right\rfloor}| \left(S_D^i - 1 \right) \right], & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \\ \frac{1}{d_u(d_u-1)} \left[\left(S_D^{\left\lfloor \frac{k}{2} \right\rfloor + 1} - 1 \right) \left(S_D^{\left\lfloor \frac{k}{2} \right\rfloor + 1} - 2 \right) + \right. & \\ \left. 2 \sum_{j=k-i+1}^{\left\lfloor \frac{k}{2} \right\rfloor} |D_j| \left(S_D^{k-j+1} - 1 \right) \right], & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 2 < i \leq k. \end{cases}
$$



- High clustering for all values of α.
- Decrease in clustering coefficient near $\alpha = 1/2$ due to the attachment of isolated agents to the hub.
- Negative degree-clustering correlation.

Characteristic Path Length

The Characteristic Path Length $\mathcal{L}$ of a network is the number of links in the shortest path between two agent, averaged over all pairs of agents.

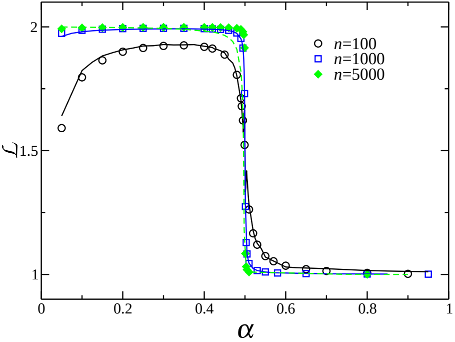
$$\mathcal{L} = \frac{\sum_{u \neq v} d(u, v)}{n(n-1)/2}$$

where $d(u, v)$ is the geodesic distance (shortest path) between u and v.

Characteristic Path Length

- Consider a nested split graph G and let $\mathbf{D} = (D_0, D_1, \ldots, D_k)$ be the degree partition of G. Then the characteristic path length $\mathcal{L}$ of G is given by

$$\mathcal{L} = \frac{1}{n(n-1)/2}\left[\frac{1}{2}\sum_{j=\left\lfloor \frac{k}{2}\right\rfloor+1}^{k}|D_j|\left(\sum_{j=\left\lfloor \frac{k}{2}\right\rfloor+1}^{k}|D_j|-1\right)+\right.$$
$$2\frac{1}{2}\sum_{j=1}^{\left\lfloor \frac{k}{2}\right\rfloor}|D_j|\left(\sum_{j=1}^{\left\lfloor \frac{k}{2}\right\rfloor}|D_j|-1\right)+$$
$$\left.\sum_{l=1}^{\left\lfloor \frac{k}{2}\right\rfloor}|D_l|\left(\sum_{j=k-l+1}^{k}|D_j|+2\left(\sum_{j=1}^{k-l}|D_j|-1\right)\right)\right].$$



⇒ Longer average path length for low values of α.

⇒ Short average path length for high values of α.

Nearest Neighbor Connectivity and Assortativity

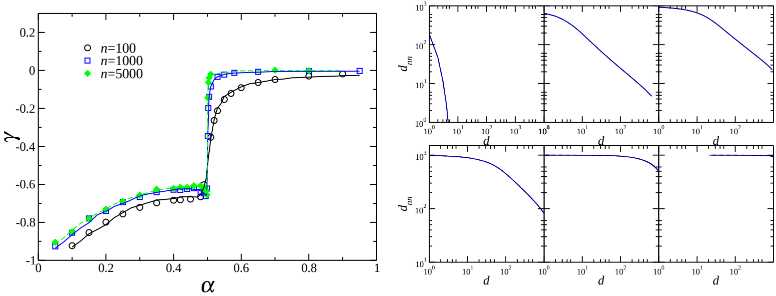
The Average Nearest Neighbor Connectivity $d_{nn}(v)$ is the average degree of the neighbors of an agent with degree d_v, that is

$$d_{nn}(v) = \frac{\sum_{u \in N_v} d_u}{d_v}$$

Nearest Neighbor Connectivity and Assortativity

- Consider a nested split graph G and let $\mathbf{D} = (D_0, D_1, ..., D_k)$ be the degree partition of G. Denote by $S_D^i = \sum_{j=1}^{i} |D_{k+1-j}|$. Then for each $v \in D_i$, $i = 0, ..., k$, (k even)

$$d_{nn}(v) = \begin{cases} \frac{1}{S_D^i} \sum_{j=1}^{i} |D_{k+1-j}| \left(S_D^{k+1-j} - 1\right), & \text{if } i = 1, ..., \left\lfloor \frac{k}{2} \right\rfloor \\ \frac{1}{S_D^{\left\lfloor \frac{k}{2} \right\rfloor + 1} - 1} \left[\sum_{j=\left\lfloor \frac{k}{2} \right\rfloor + 1}^{k} |D_j| \left(S_D^j - 1\right) \right. & \\ \left. + |D_{\left\lfloor \frac{k}{2} \right\rfloor}| S_D^{\left\lfloor \frac{k}{2} \right\rfloor}\right], & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \\ \frac{1}{S_D^i - 1} \left[\sum_{j=\left\lfloor \frac{k}{2} \right\rfloor + 1}^{k} |D_j| \left(S_D^j - 1\right) \right. & \\ \left. + \sum_{j=k-i+1}^{\left\lfloor \frac{k}{2} \right\rfloor} |D_j| S_D^j \right], & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 2, ..., k. \end{cases}$$



⇒ Networks tend to be dissortative (decreasing with α).

⇒ Note that high degree nodes are connected among each other while low degree nodes are not.

Centrality and Centralization

- Consider a particular centrality measure $\mathcal{C}(u)$ for any agent $u \in N$ in the network G.
- The *centralization* $\mathcal{C}$ of G is given by

$$\mathcal{C} = \frac{\sum_{u \in N} (\mathcal{C}(u^*) - \mathcal{C}(u))}{\max_{G'} \sum_{v \in N'} (\mathcal{C}(v^*) - \mathcal{C}(v))},$$

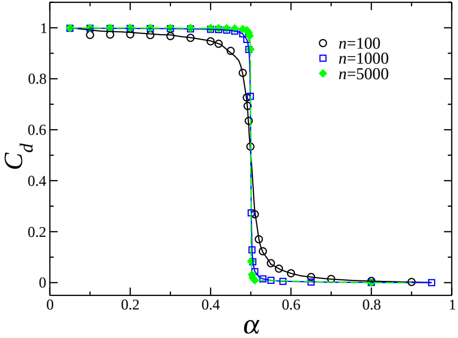
where u^* and v^* are the agents with the highest values of centrality in the network and and the maximum in the denominator is computed over all networks G' with the same number of agents, $|N| = |N'|$

Degree Centrality

- Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, \ldots, D_k)$ be the degree partition of G. Then for each $v \in D_i$, $i = 0, \ldots, k$, the degree centrality is given by

$$\mathcal{C}_d(v) = \begin{cases} \frac{1}{n-1} \sum_{j=1}^{i} |D_{k+1-j}|, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{1}{n-1} \sum_{j=1}^{i} |D_{k+1-j}| - 1, & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases}$$

Degree Centralization



⇒ Highly centralized networks for $\alpha < 1/2$.

⇒ Highly decentralized networks for $\alpha > 1/2$.

Closeness centrality

The Closeness centrality $\mathcal{C}_c(v)$ of agent v is defined as:

$$\mathcal{C}_c(v) = \frac{n-1}{\sum_{u \neq v} d(u,v)}$$

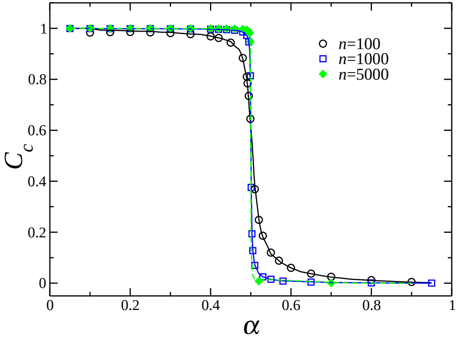
where $d(u,v)$ is the geodesic distance (shortest path) between u and v.

Closeness Centrality

- Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, \ldots, D_k)$ be the degree partition of G. Then for each $v \in D_i$, $i = 0, \ldots, k$, the closeness centrality is given by

$$\mathcal{C}_c(v) = \begin{cases} \frac{n-1}{\sum_{j=k-i+1}^{k} |D_j| + 2\sum_{j=1}^{k-i} |D_j| - 2}, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{n-1}{\sum_{j=k-i+1}^{k} |D_j| + 2\sum_{j=1}^{k-i} |D_j| - 1}, & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases}$$

Closeness Centralization



$\Rightarrow$ Highly centralized networks for $\alpha < 1/2$.

$\Rightarrow$ Highly decentralized networks for $\alpha > 1/2$.

Betweenness centrality

The Closeness centrality $\mathcal{C}_b(v)$ of agent v is defined as:

$$\mathcal{C}_c(v) = \sum_{u \neq v \neq w} \frac{g(u, v, w)}{(u, w)}$$

where $g(u, v, w)$ counts the number of paths from agent u to agent w that pass through agent v and $g(v, w)$ denotes the number of shortest paths from agent u to agent w.

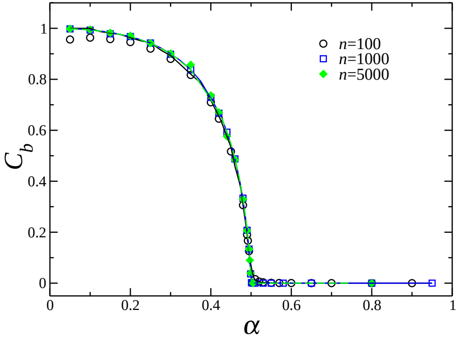
Betweenness Centrality

- Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, \ldots, D_k)$ be the degree partition of G. The betweenness centrality can be computed from $\mathcal{C}_b(u) = 0$ for $u \in D_{\lfloor \frac{k}{2} \rfloor + 1}$ and the following recursive relation for $u \in D_{\lfloor \frac{k}{2} \rfloor + 1 + (l+1)}$ and $v \in D_{\lfloor \frac{k}{2} \rfloor + l + 1}$

$$\begin{aligned}\mathcal{C}_b(u) &= \mathcal{C}_b(v) + \frac{|D_{k-(\lfloor \frac{k}{2} \rfloor + l + 1) + 1}| \left(|D_{k-(\lfloor \frac{k}{2} \rfloor + l + 1) + 1}| - 1 \right)}{\sum_{j=(\lfloor \frac{k}{2} \rfloor + l + 1) + 1}^{k} |D_j|} \\ &+ \frac{2 |D_{k-(\lfloor \frac{k}{2} \rfloor + l + 1) + 1}| \left[\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{(\lfloor \frac{k}{2} \rfloor + l + 1) - 1} (|D_j| + |D_{k-j+1}|) + |D_{\lfloor \frac{k}{2} \rfloor + l + 1}| \right]}{\sum_{j=(\lfloor \frac{k}{2} \rfloor + l + 1) + 1}^{k} |D_j|}\end{aligned}$$

with $l = 0, \ldots, \lfloor \frac{k}{2} \rfloor - 1$ if k is odd and $l = 0, \ldots, \lfloor \frac{k}{2} \rfloor - 2$ if k is even.

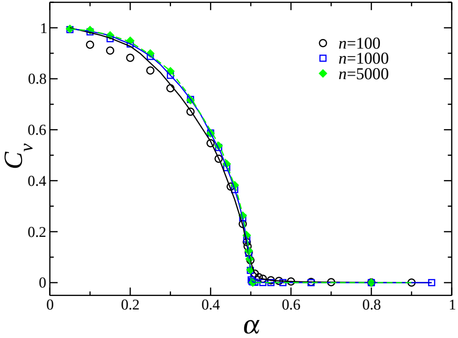
Betweenness Centralization



⇒ Highly centralized networks for $\alpha < 1/2$.

⇒ Highly decentralized networks for $\alpha > 1/2$.

Eigenvector Centralization



⇒ Highly centralized networks for $\alpha < 1/2$.

⇒ Highly decentralized networks for $\alpha > 1/2$.

Efficiency of Stationary Networks

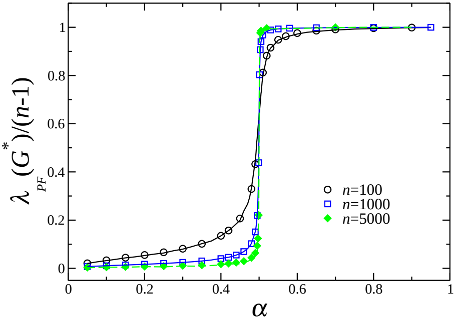
- The efficient network H maximizes aggregate utility.

$$H = \underset{G \in \mathcal{G}(n)}{\operatorname{argmax}} \sum_{i=1}^{n} u_i(\mathbf{x}^*, G).$$

- One can show the efficient network is complete, $H = K_n$.[11]
- Further, as λ approaches $1/\lambda_{\mathrm{PF}}$ maximizing aggregate equilibrium payoffs is equivalent to maximizing the largest real eigenvalue λ_{PF} of the network.[12]

[11] C. Ballester, A. Calvo-Armengol and Y. Zenou, *Who's Who in Networks. Wanted: The Key Player*, Econometrica 74(5):1403–1417, 2006

[12] Corbo, J., Calvo-Armengol, A. and Parkes, D.. *A study of nash equilibrium in contribution games for peer-to-peer networks.* SIGOPS Operation Systems Review 40(3):6166, 2006


$\lambda_{PF}(G^*)/(n\text{-}1)$
0
0.2
0.4
0.6
0.8
1
n=100
n=1000
n=5000
0
0.2
0.4
0.6
0.8
1
α

Rewrite and thus reinterpret the model in terms of noise

The Network Formation Process At every time $t \geq 0$, links can be created or decay with specified rates that depend on the current network $G(t) \in \Omega$.

Consider a continuous time Markov chain $(G(t))_{t \in \mathbb{R}_+}$.

Let $\boldsymbol{\pi}^*(G(t), \lambda) \equiv (\pi_1^*(G(t)), \ldots, \pi_n^*(G(t)))$ denote the vector of Nash equilibrium payoffs of the agents in $G(t)$ with parameter $0 \leq \lambda < 1/\lambda_{\mathsf{PF}}(G(t))$.

(i) At rate $\alpha_i \in (0, 1)$, link creation opportunities arrive to each agent $i \in \mathcal{N}$.

If such an opportunity arrives, then agent i computes the marginal payoff $\pi_i^*(G(t) \oplus (i, j), \lambda)$ for each agent $j \notin \mathcal{N} \backslash (\mathcal{N}_i \cup \{i\})$ she is not already connected to, where this computation **includes an additive, exogenous stochastic** term ε_{ij}, incorporating possible mistakes in the computation of the agent.

We assume that the exogenous random terms ε_{ij} are identically and independently type I extreme value distributed (or Gumbel distributed) with scaling parameter ζ.

Given that agent $i \in \mathcal{N}$ receives a link creation opportunity, she then links to agent $j \in \mathcal{N} \backslash (\mathcal{N}_i \cup \{i\})$ with probability

$$
\begin{aligned}
& b_i^{\zeta}(j|G(t)) \\
& \equiv \mathbb{P}\left(\begin{array}{c} \pi_i^*(G(t) \oplus (i,j), \lambda) + \varepsilon_{ij} \\ = \max_{k \in \mathcal{N} \backslash (\mathcal{N}_i \cup \{i\})} \pi_i^*(G(t) \oplus (i,k), \lambda) + \varepsilon_{ik} \end{array} \right) \\
& = \frac{e^{\pi_i^*(G(t) \oplus (i,j), \lambda)/\zeta}}{\sum_{k \in \mathcal{N} \backslash (\mathcal{N}_i \cup \{i\})} e^{\pi_i^*(G(t) \oplus (i,k), \lambda)/\zeta}}.
\end{aligned}
$$

It follows that the probability that during a small time interval $[t, t + \Delta t)$ a transition takes place from $G(t)$ to $G(t) \oplus (i, j)$ is given by

$$\mathbb{P}(G(t+\Delta t) = G\oplus(i,j)|G(t) = G) = \alpha_i b_i^\zeta(j|G(t))\Delta t+o(\Delta t)$$

(ii) We assume that a link (i, j), once established, has an exponentially distributed life time $\tau_{ij} \in \mathbb{R}_+$ with parameter

$$\nu_{ij}^{\zeta}(G(t)) \equiv \frac{1}{\mathbb{E}(\tau_{ij}|G(t))} = \beta_i f_{ij}(G(t))$$

including an agent specific component $\beta_i \in (0, 1)$, and a link specific component

$$f_{ij}^{\zeta}(G(t)) \equiv \frac{e^{\pi_i^*(G(t)\ominus(i,j),\lambda)/\zeta}}{\sum_{k\in\mathcal{N}_i} e^{\pi_i^*(G(t)\ominus(i,k),\lambda)/\zeta}},$$

for any $i \in \mathcal{N}$ and $j \in \mathcal{N}_i$.

The probability that, during a small time interval $[t, t + \Delta t)$, a transition takes place from $G(t)$ to $G(t) \ominus (i, j)$ is given by

$$\begin{aligned}\mathbb{P}(G(t + \Delta t) &= G \ominus (i, j) | G(t) = G) \\ &= \beta_i f_{ij}^{\zeta}(G(t))\Delta t + o(\Delta t)\end{aligned}$$

If agent i is chosen to form a link (at rate α_i), she will choose the agent that increases the most her utility.

There is, however, a possibility of error, captured by the stochastic term in the profit function.

Furthermore, it is assumed that links do not last forever but have an exponentially distributed life time with an expectation that depends on the relative payoff loss from removing that link.

The specific functional form of the pairwise component $f_{ij}^{\zeta}(\cdot)$ in the expected life time of a link incorporates the fact that links which are more valuable to an agent (i.e. the ones with the highest Bonacich centrality) live longer than the ones which are viewed as less valuable to her.

The value of a link is measured by the perceived loss in payoff incurred by the agent from removing the link.

Focus on networks generated by our model for large times t, and how these depend on the error term parameterized by ζ.

The Markov chain $(G(t))_{t\in\mathbb{R}_+}$ can be described infinitesimally in time by the *generator matrix* $\mathbf{Q}^\zeta$ with elements given by the *transition rates* $q^\zeta : \Omega \times \Omega \to \mathbb{R}$ defined by

$$\lim_{\Delta t \downarrow 0} \mathbb{P}(G(t+\Delta t) = G'^{\zeta}(G, G')$$

Thus

$$q^{\zeta}(G, G \oplus (i,j)) = \alpha_i b_i^{\zeta}(j|G)$$

and

$$q^{\zeta}(G, G \ominus (i,j)) = \beta_i f_{ij}^{\zeta}(G)$$

The transition rates have the property that

$q^{\zeta}(G, G'^{\zeta}(G, G \pm (i,j)) \geq 0$ if G' differs from G by the link (i,j)

and

$q^{\zeta}(G, G') = 0$ if G' differs from G by more than one link.

μ^ζ is the stationary distribution of the Markov chain satisfying $\mu^\zeta(G') = \lim_{t\to\infty} \mathbb{P}(G(t) = G'|G(0) = G)$

The most simple case is the one where ζ diverges, the error term ε_{ij} becomes dominant and the link formation and decay rates are payoff independent.

The link creation and decay rates for any $i \in \mathcal{N}$ are then given by

$$\lambda_i \equiv \lim_{\zeta \to \infty} q^{\zeta}(G, G \oplus (i,j)) = \alpha_i \frac{1}{|\mathcal{N} \backslash (\mathcal{N}_i \cup \{i\})|}, \quad j \in \mathcal{N} \backslash (\mathcal{N}_i \cup \{i\}),$$
$$\mu_i \equiv \lim_{\zeta \to \infty} q^{\zeta}(G, G \ominus (i,j)) = \beta_i \frac{1}{|\mathcal{N}_i|}, \quad j \in \mathcal{N}_i.$$

These transition rates correspond to a birth-death Markov chain with birth rates λ_i and death rates μ_i, and the stationary degree distribution is the one of the corresponding birth-death chain.

A more interesting case is the one where ζ converges to zero and the error term ε_{ij} vanishes.

For each agent $i \in \mathcal{N}$ let the best response be the set-valued map $\mathcal{B}_i : \Omega \rightarrow \mathcal{N}$ defined as

$$\mathcal{B}_i(G) \equiv \text{arg max}_{k \in \mathcal{N} \setminus (\mathcal{N}_i \cup \{i\})} \pi_i^*(G \oplus (i,k), \lambda)$$

We also define the map $\mathcal{M}_i : \Omega \rightarrow \mathcal{N}$ as

$$\mathcal{M}_i(G) \equiv \text{arg max}_{k \in \mathcal{N}_i} \pi_i^*(G \ominus (i,k), \lambda)$$

In the limit $\zeta \longrightarrow 0$, we then have that the link creation and decay rates for any $i \in \mathcal{N}$ are given by

$$q(G, G \oplus (i,j)) \equiv \lim_{\zeta \to 0} q^{\zeta}(G, G \oplus (i,j)) = \alpha_i \frac{1}{|\mathcal{B}_i(G)|}, \quad j \in \mathcal{B}_i(G),$$

$$q(G, G \ominus (i,j)) \equiv \lim_{\zeta \to 0} q^{\zeta}(G, G \ominus (i,j)) = \beta_i \frac{1}{|\mathcal{M}_i(G)|}, \quad j \in \mathcal{M}_i(G).$$

We call a network $G \in \Omega$ *stochastically stable* if $\mu(G) > 0$ where $\mu \equiv \mu^0$ is the stationary distribution of the Markov chain with transition rates given in the equation above (i.e. $\zeta \longrightarrow 0$).

- Some empirical implications of the model

- Simulations of the theoretical model
- Adjacency matrices
- Networks

Estimating the Model's Parameters

- First network: Network of **Austrian banks** in the year 2008 [cf. Boss et al.,2004].
- **Links** in the network represent exposures between Austrian-domiciled banks on a nonconsolidated basis (i.e. no exposures to foreign subsidiaries are included).
- Sample of n = 770 banks with m = 2454 links between them and an **average degree** of

  d = 20.54.

Network of Austrian banks

- The largest connected component comprises 768 banks, which is 99.7% of the total of banks.
- The network is highly clustered with an **average clustering coefficient** of $C = 0.75$ and shows a **low average path length** of $L = 2.27$ in the connected component.
- The network of banks is **dissortative** and an assortativity coefficient of $\gamma = -0.51$

Global banking network

- Second Network: **Global banking network** in the year 2011 obtained from the Bank of
- International Settlements (BIS) locational statistics on exchange-rate adjusted changes in cross-border bank claims.
- Network with n = 239 nodes and m = 2454 links between them.
- **Average degree** d = 20.54.
- Almost all nodes are contained in the largest connected component (93%).
- **High Average clustering coefficient** C = 0.81.
- **Short average path length** (in the connected component) of L = 1.95.
- Network is **dissortative** with a coefficient of $\gamma = -0.76$.

Network of trade relationships

- Third network: Network of trade relationships between countries in the year 2000.
- The trade network is defined as the network of **import-export relationships** between countries in a given year in millions of current-year U.S. dollars.
- **Undirected network**: a link is present between two countries if either one has exported to the other country.

Network of trade relationships

- Trade network contains n = 196 nodes, m = 4138 links, has an **average degree** of d = 42.22.
- The network consists of a giant component with 181 nodes, encompassing 92% of all nodes in the network.
- Highly clustered with C = 0.73,
- a short average path length of L = 2.25,
- Dissortative $\gamma = -0.40$

Arms trade network

- Fourth network: arms trade between countries.
- Data: from the SIPRI Arms Transfers Database holding information on all international transfers between countries of seven categories of major conventional weapons accumulated from 1950 to 2010.
- A link in the network represents a recipient or supply relationship of arms between two countries during this period.

Arms trade network

- Network with n = 246 nodes and m = 2245 links.
- The **average degree** is d = 18.25.
- Network is connected and has an average path length of L = 2.25.
- The average clustering coefficient is C = 0.48
- Dissortativity coefficient of $\gamma = -0.39$.

- All these four real-world networks are of similar size, show **short average path lengths** of around 2, are **dissortative** and have a monotonic decreasing average nearest neighbor connectivity.
- They also show a relatively **high clustering** and the clustering degree distribution is decreasing with the degree.

- An important feature of these networks is that they all show a high degree of **nestedness**.
- This can be witnessed from the adjacency matrices depicted in Figure 11, which resemble the nested matrices we derive from our theoretical model (see Figure 2).

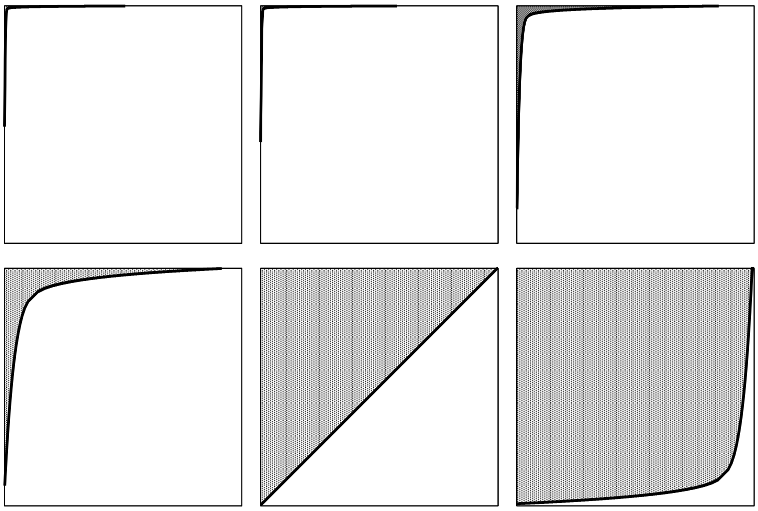


Figure 2: Representation of the adjacency matrices of stationary networks with $n = 1000$ agents for different values of parameter α: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.495$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The solid line illustrates the stepfunction separating the zero from the one entries in the matrix. The matrix top-left for $\alpha = 0.4$ is corresponding to a star-like network while the matrix bottom-right for $\alpha = 0.52$ corresponds to an almost complete network. Thus, there exists a sharp transition from sparse to densely connected stationary networks around $\alpha = 0.5$. Networks of smaller size for the same values of α can be seen in Figure 3.

observe the transition from sparse networks containing a hub and many agents with small degree to a quite homogeneous network with many agents having similar high degrees. Moreover, this transition is sharp around $\alpha = 1/2$. In Figure 3, we show particular networks arising from the network formation process for the same values of α. Again, we can identify the sharp transition from hub-like networks to homogeneous, almost complete networks.

Figure 4 (left) displays the number $\bar{m}$ of links m relative to the total number of possible links $n(n-1)/2$, i.e. $\bar{m} = \frac{2m}{n(n-1)}$, and the number of distinct degrees k as a function of α. We see that there exists a sharp transition from sparse to dense networks around $\alpha = 1/2$ while k reaches a maximum at $\alpha = 1/2$. This follows from the fact that $k = 2d^*$ with d^* given in Equation (9) is monotonic increasing in α for $\alpha < 1/2$ and monotonic decreasing in α for $\alpha > 1/2$.

5. Stochastically Stable Networks: Statistics

We would like now to investigate further the properties of our networks and see how they match real-world networks. There exists a growing number of empirical studies trying to identify the key characteristics of social and economic networks. However, only few theoretical models (a notable exception is Jackson and Rogers [2007]) have tried to reproduce these findings to the full extent. We pursue the same approach. We show that our network formation model leads to properties which are shared with empirical networks. These properties can be summarized as follows:[33]

[33] This list of empirical regularities is far from being extensive and summarizes only the most pervasive patterns found in the literature.

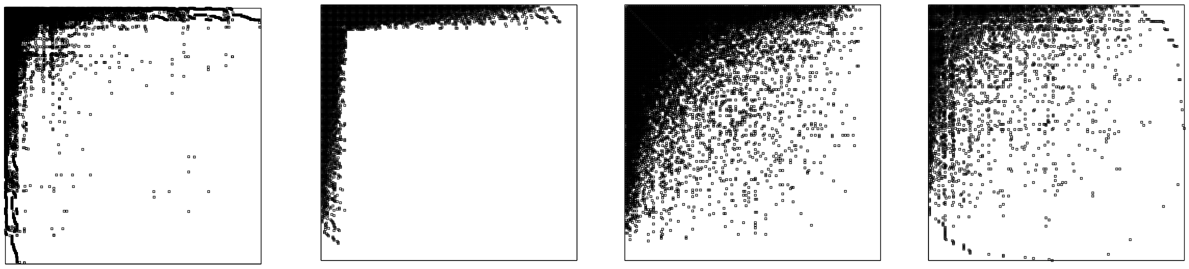


Figure 11: Adjacency matrices (sorted by the eigenvector centralities of the nodes) for the Austrian banking network, the global network of banks obtained from the Bank of International Settlements (BIS) locational statistics, the GDP trade network and the arms trade network (from left to right). All adjacency matrices are significantly nested.

$\boldsymbol{\Theta}$ conditional on the observed statistic $\mathcal{S}^o$.[48] Since this estimation algorithm would require the computation of the Bonacich centrality an extensive number of times, we assume that the complementarity parameter λ is small such that we can approximate the Bonacich centrality by the degree centrality when simulating the network formation process.[49] Note also that the reported estimates of ζ hold only up to a scaling factor, which depends on the choice of λ. Hence, only the relative values of ζ between different samples is meaningful but not its absolute value.

The estimated parameter values are shown in Table 1. We observe that the estimates for ζ are higher for the network of GDP trading countries and the network of arms trade than the corresponding estimates for the networks of banks. This confirms our intuition from Figure 9 (right), where we have seen that with increasing values of ζ stationary networks become less nested (and we obtain a random graph as $\zeta \to \infty$), and the lower values for the matrix temperature T_n for these networks (see also the adjacency matrices in Figure 11). Hence, our estimates support our earlier observation that the networks of banks have a higher degree of nestedness than the networks of trade relationships between countries.

Moreover, Figure 12 shows the empirical distributions (squares) and typical simulated ones (circles) for the bank network, the network of GDP trade and the arms trade network. The comparison of observed and the simulated distributions shown in Figure 12 indicate that the model can relatively well reproduce the observed empirical networks, even though the model is parsimoniously parameterized in relying only on two exogenous variables α and ζ. The fit seems to be best for the networks of banks, which also shows the most distinct nestedness pattern (see

[48]For the implementation of the algorithm we have chosen an initial uniform (prior) parameter distribution. The proposal distribution is a normal distribution. During the "burn-in" phase [Chib, 2001], we consider a monotonic decreasing sequence of thresholds with appropriately chosen values from careful numerical experimentation. For the Austrian banking network we have chosen a burn-in period of 1000 steps, while for the network of GDP trade we have used a period of 3000.

[49]The Bonacich centrality is defined by $b_i(G,\lambda) = \sum_{k=0}^{\infty} \lambda^k \left(\mathbf{A}^k \cdot \mathbf{u}\right)_i = 1 + \lambda d_i + \lambda^2 \sum_{j \in \mathcal{N}_i} d_j + \lambda^3 \sum_{j \in \mathcal{N}_i} \sum_{k \in \mathcal{N}_j} d_k + \ldots = 1 + \lambda d_i + \lambda^2 \sum_{j \in \mathcal{N}_i} d_j + O(\lambda^3)$. Marginal payoff from forming a link ij for agent i can then be written as $\pi_i^*(G \oplus ij, \lambda) - \pi_i^*(G, \lambda) = \frac{\lambda(2+\lambda)}{2} + \frac{\lambda^2}{2} d_i(d_i+1) + \lambda^2 d_j + O(\lambda^3)$. When computing marginal payoffs from forming a link (and the decay rates) we ignore terms of the order $O(\lambda^3)$.

- When we compare the networks simulated from our model (see Figure 3) and the ones
- described in real-world networks (Figure 10), they are relatively similar (in terms of a clear **core-periphery structure** indicating nestedness).

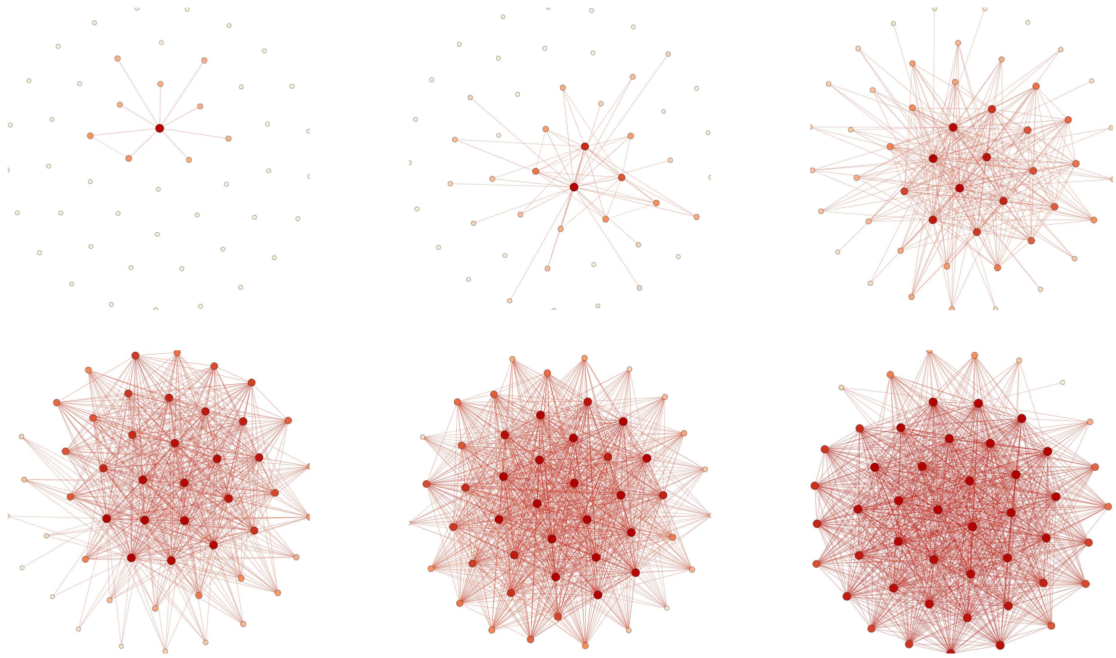


Figure 3: Sample networks with $n = 50$ agents for different values of parameter α: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.495$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The shade and size of the nodes indicate their eigenvector centrality. The networks for small values of α are characterized by the presence of a hub and a growing cluster attached to the hub. With increasing values of α the density of the network increases until the network becomes almost complete.

(i) The average shortest path length between pairs of agents is small [Albert and Barabási, 2002].

(ii) Empirical networks exhibit high clustering [Watts and Strogatz, 1998]. This means that the neighbors of an agent are likely to be connected.

(iii) The distribution of degrees is highly skewed. While some authors [e.g. Barabasi and Albert, 1999] find power-law degree distributions, others find deviations from power-laws in empirical networks, e.g. in Newman [2004], or exponential distributions [Guimera et al., 2006].

(iv) Several authors have found that there exists an inverse relationship between the clustering coefficient of an agent and her degree [Goyal et al., 2006; Pastor-Satorras et al., 2001]. The neighbors of a high degree agent are less likely to be connected among each other than the neighbors of an agent with low degree. This means that empirical networks are characterized by a negative clustering-degree correlation.

(v) Networks in economic and social contexts exhibit degree-degree correlations. Newman [2002, 2003] has shown that many social networks tend to be positively correlated. In this case the network is said to be assortative. On the other hand, technological networks such as the internet [Pastor-Satorras et al., 2001] display negative correlations. In this case the network is said to be dissortative. Others, however, find also negative correlations in social networks such as in the Ham radio network of interactions between amateur radio operators [Killworth and Bernard, 1976] or the affiliation network in a Karate club [Zachary, 1977]. Networks in economic contexts may have features of both technological and social relationships [Jackson, 2008] and so there exist examples with positive degree correlations such

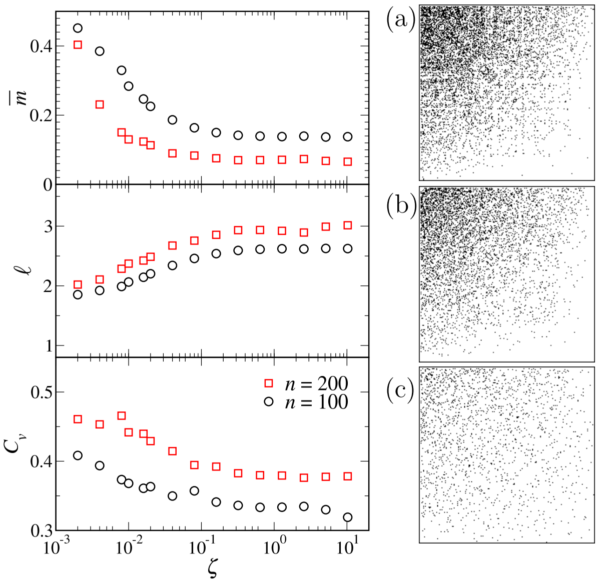


Figure 9: Results of numerical simulations of the model introduced in Definition 2, where the rates of link creation (decay) are Gumbel distributed on the increase (resp. decrease) of payoff of the agents adjacent to the link. In the left panels (from top to bottom) the network density $\bar{m}$, the average path-length (ℓ) and the degree of centralization for the eigenvector centrality C_v are plotted as a function of ζ, the parameter of the Gumbel distribution. The three plots on the right column depict snapshots of the adjacency matrix once the stationary regime has been reached (all the global measures do not change over time, but fluctuate around the stationary state). In these snapshots, the parameters are (a) $\zeta = 0.002$, (b) $\zeta = 0.004$, and (c) $\zeta = 0.08$, respectively.

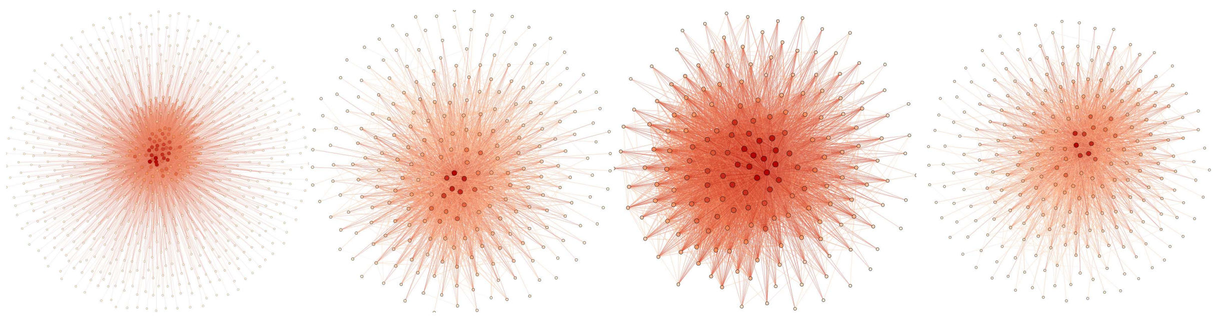


Figure 10: The Austrian banking network, the global network of banks obtained from the Bank of International Settlements (BIS) locational statistics, the GDP trade network and the arms trade network (from left to right). The shade and size of the nodes indicate their eigenvector centrality. The GDP trade network is much more dense than the network of banks and the network of arms trade. All four networks show a core of densely connected nodes.

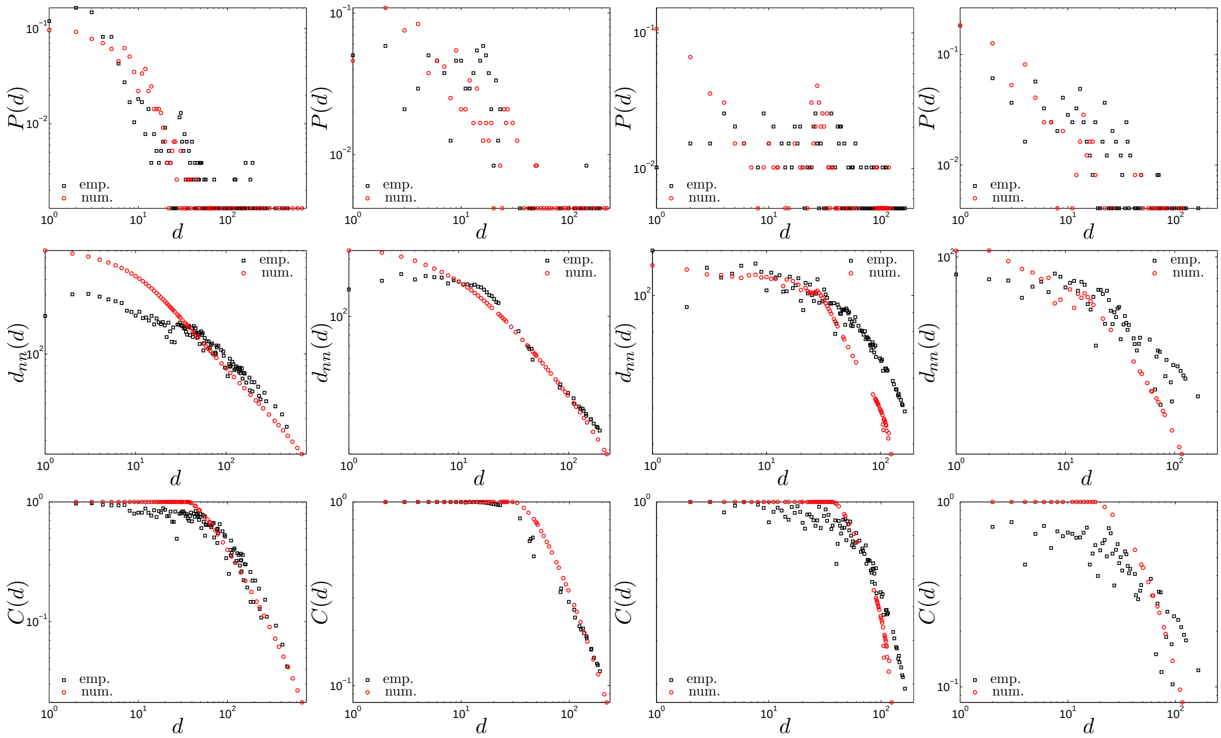


Figure 12: The empirical (□) and an exemplary simulated (○) degree distribution $P(d)$, average nearest neighbor degree $d_{nn}(d)$ and clustering degree distribution $C(d)$ for the Austrian banking network (first column), the network of banks obtained from the Bank of International Settlements (BIS) (second column), the GDP trade network (third column) and the arms trade network (fourth column).

Figure 11).

9. Robustness Analysis

In our model, we describe a dynamic process incorporating both the play of a network game and the endogenous formation of the network. A striking finding is that, starting from any arbitrary graph,[50] the process converges to a nested split graph in the limit of vanishing noise. We like to show that this characterization is *not* an artifact of a very specific protocol of network formation but is quite general.

First, in the model presented in this paper, using Ballester et al. [2006], we give a micro-foundation of why agents choose to create a link with the agent who has the highest Bonacich centrality in the network. By doing so we impose a specific structure of the utility function (i.e. a linear quadratic structure; see Equation (1)) and a condition on the largest real eigenvalue of the graph (i.e. $\lambda < 1/\lambda_{\mathrm{PF}}$; see Theorem 1). In fact, all our results, in particular the fact that the network converges to a nested split graph, hold if we take a general utility function (i.e. no specific structure), that is increasing in their Bonacich centrality. Moreover, imagine that agents do not choose effort $\mathbf{x}$ but just create links following the network formation mechanism described

[50]See Proposition 3 in Section 4.

Robustness

Our model: describe a dynamic process incorporating both the play of a network game and the endogenous formation of the network.

Striking finding: starting from an arbitrary graph, the process converges to a nested split graph.

This characterization is *not* an artifact of a very specific protocol of network formation but is quite general.

Robustness

(1) Linear quadratic utility function and condition on the largest real eigenvalue of the graph (i.e. $\lambda < 1/\lambda_{\mathsf{PF}}$).

All our results hold with a general utility function (i.e. no specific structure), that is increasing in agents' Bonacich centrality.

Assume agents do not choose effort x but just create links following our network formation mechanism and their utility is the sum of the current Bonacich centrality of their neighbors.

Same results (general utility + no condition on largest eigenvalue).

True also with other centrality measures (degree, closeness).

Robustness

(2) Assumption: if selected, an agent must cut a link (the least valuable one).

Justification: Agents cannot have an infinite number of links (for example in the case of friendships, individuals cannot have an infinite number of friends).

Let's develop the same model (same network formation process) but without imposing the link deletion of agents.

Assume that each agent has a *finite maximum number of links* she can maintain and each agent has a different maximum, reflecting their heterogeneity in the cost of maintaining links.

When an agent reaches her maximum number of links, then she will not accept to create any more links.

Proposition 0.5 *Consider the network formation process* $(G(t))_{t=0}^{\infty}$ *in Definition above with link creation probability* $\alpha = 1$*. Further assume that agents have the capacity constraint* $\mathbf{C} = (C_1, ..., C_n)$ *in the number of links they can form. Then* $(G(t))_{t=0}^{\infty}$ *eventually leads to an equilibrium given by the nested split graph* G_C *with degree partition* $\mathbf{D} = \mathbf{C}$*, i.e.* $\mu_G = 1$ *if* $G = G_C$ *and zero otherwise.*

Nested split graphs can also arise in the absence of link removal by assuming that agents have some limitation in the number of interactions they can maintain.

Robustness

Consider the network formation process as before with one modification:

The agent who wants to create a link needs to pay a cost c.

Proposition 0.6 *Consider the network formation process* $(G(t))_{t=0}^{\infty}$ *as above. Assume that there is a cost* c *of creating a link for the agent who initiates that link. Then, if*

$$c < \frac{\lambda}{1-\lambda}$$

agents will always form a link and the emerging network will always be a nested split graph.

The above proposition shows that nested split graphs can also arise even when links are costly to be formed, as long as the costs are not too large.

Robustness

(3) We have assumed *myopic* agents, i.e. maximize *current* utility.

There exists a *farsighted* equilibrium which is exactly like ours.

Agents create and delete links based on *discounted lifetime expected utility*.

They will still create and delete links with the agent who has the highest Bonacich centrality in the network.

because he/she has a higher probability than any other agent in the network to still have the highest Bonacich centrality in the future (any period).

Robustness

(4) We obtain a power-law degree distribution with exponent minus one.

Extend our model to obtain an arbitrary power law degree distribution

For that, make the probability of creating a link for player i, i.e. α, depending on, $|D_i|$, the size of the degree partition she belongs to.

Robustness

(5) Here always obtain negative degree-degree correlation (i.e. dissortative networks).

Extend our game by including capacity constraints.

Obtain assortative networks.

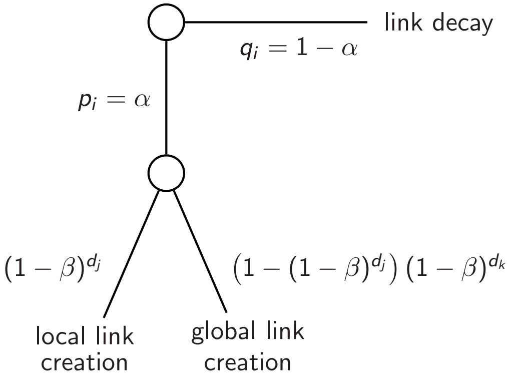
Introducing Capacity Constraints and Global Search

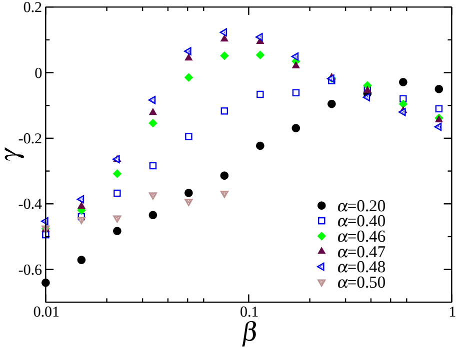
Definition: Network formation process $\Gamma'(G)$

(i) With probability $p_i = \alpha$ player i receives the opportunity to create an additional link. Let j be the player in $N_i^{(2)}$ with the highest degree, that is $d_j \geq d_k$ for all $j, k \in N_i^{(2)}$.
- With probability $(1-\beta)^{d_j}$ the link ij is formed.
- Otherwise, player i connects to a randomly selected player $k \in N \backslash \{N_i \cup i\}$ with probability $\left(1-(1-\beta)^{d_j}\right)(1-\beta)^{d_k}$.

If player i is already connected to all other players then nothing happens.

(ii) With probability $q_i = 1 - p_i = 1 - \alpha$, the link to the player j in N_i with the smallest degree $d_j \leq d_k$ for all $j, k \in N_i$, decays. If player i does not have any links then nothing happens.


link decay
$q_i = 1 - \alpha$
$p_i = \alpha$
$(1-\beta)^{d_j}$
$\left(1-(1-\beta)^{d_j}\right)(1-\beta)^{d_k}$
local link creation
global link creation



⇒ Qualitatively same network statistics, but networks can become assortative.

⇒ Explanation for the distinction between technological and social networks.

Summary and Conclusion

- We introduce a micro-founded network formation game based on Bonacich centrality.[17]
- We show that a particular class of graphs, "nested split graphs" (special case of inter-linked stars), emerge in this network game.
- We can derive any network statistics in the stationary distribution.
- We show that our networks agree with stylized facts of empirical networks.

[17] By endogenizing the network in the model of C. Ballester, A. Calvo-Armengol and Y. Zenou, *Who's Who in Networks. Wanted: The Key Player*, Econometrica 74(5):1403–1417, 2006

- We find a critical dependence of the emerging networks G on the environmental volatility:[18]
  - Strongly volatile environment (low α) $\Rightarrow$ G highly centralized with small core and large periphery.
  - Weakly volatile environment (high α) $\Rightarrow$ G highly decentralized network with large core and small periphery.

[18] Similar to Arenas, A., Cabrales, A., Diaz-Guilera, A., Guimera, R., Vega-Redondo, a F., 2008. *Optimal information transmission in organizations: Search and congestion.* Review of Economic Design.