A Dynamic Model of Network Formation with Strategic Interactions☆

Michael D. König[a], Claudio J. Tessone[a], Yves Zenou[b]

[a] *Chair of Systems Design, Department of Management, Technology and Economics, ETH Zurich, Kreuzplatz 5, CH-8032 Zurich, Switzerland*
[b] *Stockholm University and Research Institute of Industrial Economics (IFN). Sweden.*

Abstract

We develop a network formation model where links are formed on the basis of agents' centrality while the network is exposed to a volatile environment introducing interruptions in the connections between agents. A remarkable feature of our dynamic network formation process is that, at each period of time, the network is a nested split graph. We show that there exists a unique stationary network whose topological properties completely match features exhibited by real-world networks. We also find that there exists a sharp transition in efficiency and network density from highly centralized to decentralized networks.

Key words: Bonacich centrality, network formation, social interactions, nested split graphs
JEL: A14, C63, D85

☆We would like to thank Phillip Bonacich, Yann Bramoullé, Ulrik Brandes, Guido Caldarelli, Sanjeev Goyal, Patrick Groeber, Matthew Jackson, Matteo Marsili, Fernando Vega-Redondo, Douglas White as well as the participants at the 2009 DIME conference in Paris, the 2009 Trento Summer School on "Innovation and Networks", the 2009 HSC Videoconference at the University of California San Diego, the 2009 workshop on the Economics of Social Networks at Laval University and the Microeconomic Theory seminar at University of California Berkeley, for their helpful comments.

Email addresses: `mkoenig@ethz.ch` (Michael D. König), `tessonec@ethz.ch` (Claudio J. Tessone), `yves.zenou@ne.su.se` (Yves Zenou)

1. Introduction

Social networks are important in several facets of our lives. For example, the decision of an agent of whether or not to buy a new product, attend a meeting, commit a crime or find a job is often influenced by the choices of his or her friends and acquaintances. The emerging empirical evidence on these issues motivates the theoretical study of network effects. For example, job offers can be obtained from direct and indirect, acquaintances through word-of-mouth communication. Also, risk-sharing devices and cooperation usually rely on family and friendship ties. Spread of diseases, such as AIDS infection, also strongly depends on the geometry of social contacts. If the web of connections is dense, we can expect higher infection rates. In terms of structure, real-life networks are characterized by low diameter (the so-called "small world" property), high clustering, and "scale-free" degree distributions.

To fathom these different aspects and to match the observed structure of real-life networks, one needs to analyze how and why networks form, the impact of network structure on agents' outcomes, and the evolution of networks over time. The aim of the present paper is to propose a theoretical model that has all these features.

The literature on network formation is basically divided in two strands that are not communicating very much with each other. In the random network approach (mainly developed by mathematicians and physicists),[1] which is mainly dynamic, the reason why a link is formed is pure chance. Indeed, this literature builds networks either through a purely stochastic process where links appear at random according to some distribution, or else through some algorithm for building links. In the other approach (developed by economists), which is mainly static, the reason for the formation of a link is strategic interactions. Individuals carefully decide with whom to interact and this decision entails some consent by both parts in a given relationship. As Jackson [2007, 2008] pointed out, the random approach gives us a great deal of insight into *how* networks form (i.e. matches the characteristics of real-life networks) while the deterministic approach performs better on *why* networks form.

There is also another strand of the literature (called "games on networks") that takes the network as given and studies how the network structure impacts on outcomes and individual decisions. A prominent paper of this literature is Ballester et al. [2006].[2] They mainly show that if agents'

[1] See Albert and Barabási [2002].

[2] Bramoullé and Kranton [2007], Bramoullé et al. [2010] and Galeotti et al. [2010] are also important papers in this literature. The first paper focuses on strategic substitutabilities while the second one provides a general framework solving any game on networks with perfect information and linear best-reply functions. The last paper investigates the

payoffs are linear-quadratic, then the unique interior Nash equilibrium of an $n-$player game in which agents are embedded in a network is such that each individual effort is proportional to her Bonacich centrality measure. The latter is a well-known centrality measure introduced by Bonacich [1987].[3] The Bonacich centrality of a particular node counts the total number of paths that start from this node in the graph, weighted by a decay factor based on path length. In the model of Ballester et al. [2006], it is mainly the centrality of an agent in a network that explains her outcome.[4]

To the best of our knowledge, there are very few papers that combine the literature on network formation and games on networks. The aim of this paper is to introduce strategic interactions in a non-random dynamic network formation game where agents also choose how much effort they put in their activities. By combining these approaches, we will also be able to match the characteristics of most real-life networks.

To be more precise, we develop a two-stage game where, in the first stage, as in Ballester et al. [2006], agents play their equilibrium contributions proportional to their Bonacich centrality while, in the second stage, a randomly chosen agent can update her linking strategy by creating a new link as a best response to the current network. Furthermore, agents are embedded in a volatile environment which requires them to continually adapt to changing conditions. We assume that a link of a randomly selected agent decays, i.e. can be severed. As a result, the formation of social networks can be regarded as a tension between the search for new linking opportunities and volatility that leads to the decay of existing links.

We first show that, at each period of time, the network generated by this dynamic process is a nested split graph. These graphs, which are relatively well-known in the applied mathematics literature, have a very nice and simple structure that make them very tractable to work with. To the best of our knowledge, this is the first time that a complex dynamic network formation model can be characterized by such a simple structure in terms of networks it generates. By doing so, we are able to bridge the

---

case when agents do not have perfect information about the network. Because of its tractability, in the present paper, we use the model of Ballester et al. [2006] who analyze a network game of local complementarities under perfect information.

[3] Centrality is a fundamental measure of the importance of actors in social networks, dating back to early works such as Bavelas [1948]. See Wasserman and Faust [1994] for an introduction and survey.

[4] In the empirical literature, it has been shown that centrality is important in explaining exchange networks [Cook et al., 1983], peer effects [Calvó-Armengol et al., 2009; Haynie, 2001], creativity of workers [Perry-Smith and Shalley, 2003], workers' performance [Mehra et al., 2001], power in organizations [Brass, 1984], the flow of information [Borgatti, 2005; Stephenson and Zelen, 1989], the formation and performance of R&D collaborating firms and inter-organizational networks [Boje and Whetten, 1981; Powell et al., 1996; Uzzi, 1997] as well as the success of open-source projects [Grewal et al., 2006].

economics literature and the applied mathematics/physics literatures in a simple way. Because of their simple features, we then show that degree, closeness, eigenvector and Bonacich centrality induce the same ordering of nodes in a nested split graph (this is also true for betweenness centrality if the ordering is not strict). This implies, in particular, that if we had a game where agents formed links according to other measures of centrality (such as degree, closeness, or betweenness) than the Bonacich centrality, then all our results would be unchanged. We then show that there exists a unique stationary network, which is a nested split graph. Instead of relying on a mean-field approximation of the degree distribution and related measures as most dynamic network formation models do, because of the nature of nested split graphs, we are able to derive explicit solutions for all network statistics of the stationary network (by computing the adjacency matrix). We also find that, by altering the rate at which linking opportunities arrive and links decay, a sharp transition takes place in the network density. This transition entails a crossover from highly centralized networks when the linking opportunities are rare and the link decay is high to highly decentralized networks when many linking opportunities arrive and only few links are removed. From the efficiency perspective such sharp transition can also be observed in aggregate payoffs in stationary networks.

The intuition of these results is as follows. Agents want to link to others who are more central since this leads to higher actions (as actions are proportional to centrality) and higher actions raise payoffs more. Similarly, agents delete links of those with lower centrality as these agents have lower actions and hence lower payoff effects. Notice moreover that, once someone deletes a link with an agent, she becomes less central and this makes it more likely that the next person who has a choice will also delete a link with her. Thus link gains and losses are self reinforcing. This intuition suggests that if α, the probability of adding links is large then the process should approximate complete network while if it is small then the process should approximate the star network. The key insight of our model is that for intermediate values of α the network is a nested split graph.

We then proceed by showing that our model reproduces the main empirical observations of social networks. Indeed, we show that the stationary networks emerging in our link formation process are characterized by *short path length* with *high clustering* (so called "small worlds", see Watts and Strogatz [1998]), *exponential degree distributions* with *power law tails* and *negative degree-clustering correlation*. These networks also show a clear core-periphery structure. Moreover, we show that, if agents have no "budget constraints"and can form any number of links then stationary networks are *dissortative*. However, if one takes into account capacity constraints in the number of links an agent can maintain, and allows for random global attachment between agents, we keep all the above mentioned network statistics while, at the same time, yielding *assortative* stationary networks.

Finally, it is worth noting that the model we study is rather general and the fact that our network formation generates nested-split graphs is not an artefact of a very specific protocol. First, our results are independent of initial conditions. That is, the network formation process will always converge to a nested split graph starting from any possible network. Second, we would obtain exactly the same results using more general utility functions, e.g. any increasing function of the Bonacich centrality of the agent or using "information centrality" introduced in Stephenson and Zelen [1989]. Third, when creating a link, an agent can choose anyone in the network not only neighbors of neighbors as it is often assumed. In other words, we do not restrict the set of choices of her possible contacts. Fourth, if we relax the fact that agents need to delete a link so that only link creation prevails, and assume that each agent has a finite maximum number of links she can maintain and that they are heterogeneous in this maximum number (reflecting their different costs of maintaining links), then our network formation process still leads to a stable nested split graph. Finally, if it is assumed that link formation is costly, we are able to show that if this cost is not too large, networks will still be nested split graphs. We discuss in more detail all these robustness issues in Section 7.

Relation to the literature

As stated above, there is an important literature on dynamic network formation. The first approach (coming mainly from physics) is to consider a random network formation and to study how emerging networks match real-world networks (see e.g. Ehrhardt et al. [2006b, 2008], Marsili et al. [2004], Vega-Redondo [2006]); for an overview, see Vega-Redondo [2007]). While sharing some common features with this literature, our model is quite different since agents do not create or delete links randomly but in a strategic way, i.e. they maximize their utility function. From the economic literature, there are also dynamic network formation models with strategic interactions. Bala and Goyal [2000], Watts [2001], Jackson and Watts [2002a], Dutta et al. [2005] are prominent papers of this literature. Bala and Goyal [2000] propose a model similar to the connections model (Jackson and Wolinsky [1996]), but with directional flow of communication or information. They focus on the dynamic formation of networks in this directed communications model. The network formation game is played repeatedly, with individuals deciding on link formation in each period. In this setting, for low enough costs to links, the process leads naturally to a limiting network which has the efficient structure of a wheel. Also in the context of the connections model, Watts [2001] considers a framework where pairs of agents meet over time, and decide whether or not to form or sever links with each other. Agents are myopic and so base their decision on how the decision on the given link affects their payoffs, given the current network in place. An interesting result applies to a cost range where a star network is both pairwise stable and efficient. Jackson and Watts [2002b] model network formation as an intertemporal process with myopic individuals breaking and forming links as the network evolves dynamically. Dutta et al. [2005] relax the assumption of myopic agents and assume that agents behave in a farsighted manner by taking into account the intertemporal repercussions of their own decisions. Our model is different than the ones developed in these papers in the sense that we consider both dynamic models of network formation and optimal actions from agents. This allows us to give a microfoundation of the network formation process as equilibrium actions transform into equilibrium utility functions.

There are also some papers that, as in our framework, combine both network formation and endogenous actions. These papers include Bramoullé et al. [2004], Cabrales et al. [2009], Calvó-Armengol and Zenou [2004], Galeotti and Goyal [2010], Goyal and Vega-Redondo [2005], Goyal and Joshi [2003], Jackson and Watts [2002a]. Most of these models are, however, static and the network formation process is different. Goyal and Joshi [2003] is the closest to our model. They consider a standard Cournot model of competition where prior to market stage, firms can form costly links with each other since a link lowers costs of production for the two firms. In this framework, a network defines a cost profile for the competing firms. They show that every pairwise equilibrium connected network is an inter-linked star network. Their model is static but has link formation and choice of action (quantity). Our model is dynamic and has link formation and a choice of action. The details of the payoffs and the methods of analysis, however, differ. In particular, the introduction of stochastic elements in our analysis helps in solving our model, which allows us to obtain more general results, like for example calculating the exact degree distribution of the stationary network.

Finally, our paper is also related to Jackson and Rogers [2007], who also motivate their work by means of statistics of empirical networks. In their model, new nodes are born over time and can attach to existing nodes either by choosing one uniformly at random or through a search process that makes the likelihood of meeting a given node proportional to the number of links the node already has. In their model, m is the average degree while r represents the ratio of how many links are formed uniformly at random compared to how many are formed proportionally to the number of links existing nodes already have. As r approaches 0, the distribution converges to be scale-free, while as m tends to infinity the distribution converges to a negative exponential distribution, which corresponds to the degree distribution of a purely uniform and independent link formation process on a network that grows over time. Our model is quite different since, contrary to Jackson and Rogers [2007], agents choose actions and form links by maximizing their utility. Also we look at the steady-state distribution while they analyze growing networks. As a result, the predictions of the two models are quite different.

To summarize, our main contribution to the literature is that we are able to analyze a dynamic network formation model with endogenous actions and analytically characterize the stationary network distribution. We are also able to match most of the properties of real-world metworks.

Our paper is organized as follows. In Section 2, we introduce the model and discuss the basic properties of the network formation process. In particular, Section 2.1 discusses the first stage of the game. In Section 2.2, we introduce the second stage of the game, where the network formation is explained. Next, Section 3 shows that stationary networks exist and can be computed analytically. After deriving the stationary networks, in Section 5, we analyze their properties in terms of topology and centralization. In Section 6, we study efficiency from the point of view of maximizing total efforts and aggregate payoff in the stationary network. We investigate the efficiency of different stationary networks as a function of the volatility of the environment. Section 7 discusses our results and their robustness, especially when we consider general utility functions. Appendix A gives all the necessary definitions and characterizations of general networks. In Appendix B, we focus on a class of networks (nested split graphs) that are important in our analysis and provide general results in terms of their topology properties and centralization measures. We extend our analysis in Appendices C, D and E by including capacity constraints in the number of links an agent can maintain. Finally, all proofs can be found in Appendix F.

2. The model

In this section, we develop a two-stage game. In the first stage, following Ballester et al. [2006], all agents simultaneously choose their effort level in a fixed network structure. It is a game with local complementarities where players have linear-quadratic payoff functions. In the second stage, a randomly chosen agent decides with whom she wants to form a link while a volatile environment forces the least frequently used link of a randomly selected agent to decay. This introduces two different time scales, one in which agents are choosing their efforts in a simultaneous move game and the second in which an agent forms a link as a best response to the current network

2.1. Nash Equilibrium and Bonacich Centrality

Consider a static network G in which the nodes represent a set of agents/players $N = \{1, 2, ..., n\}$. Following Ballester et al. [2006], each agent $i \in N$ in the network G selects an effort level $x_i \geq 0$, $\mathbf{x} \in \mathbb{R}_+^n$, and receives a payoff $\pi_i(x_1, ..., x_n)$ of the following form

$$\pi_i(x_1, ..., x_n) = x_i - \frac{1}{2}x_i^2 + \lambda \sum_{j=1}^{n} a_{ij}x_ix_j, \tag{1}$$

where $\lambda \geq 0$ and $a_{ij} \in \{0,1\}$, $i,j = 1,...,n$ are the elements of the symmetric $n \times n$ *adjacency matrix* $\mathbf{A}$ of G. This utility function is additively separable in the idiosyncratic effort component $(x_i - 1/2x_i^2)$ and the peer effect contribution $(\lambda \sum_{j=1}^{n} a_{ij}x_i x_j)$. Payoffs display strategic complementarities in effort levels, i.e., $\partial^2 \pi_i(x_1,...,x_n)/\partial x_i \partial x_j = \lambda a_{ij} \geq 0$. In order to find the Nash equilibrium solution associated with the above payoff function, we define a network centrality measure introduced by Bonacich [1987]. Let $\lambda_{\text{PF}}(G)$ be the largest real eigenvalue of the adjacency matrix $\mathbf{A}$ of network G. The adjacency matrix is a matrix that lists the direct connections in the network. If $\mathbf{I}$ denotes the $n \times n$ identity matrix and $\mathbf{u} = (1,...,1)^T$ the n-dimensional vector of ones then we can define Bonacich centrality as follows:

Definition 1. *The matrix* $\mathbf{B}(G,\lambda) = (\mathbf{I} - \lambda\mathbf{A})^{-1}$ *exists and is non-negative if and only if* $\lambda < 1/\lambda_{PF}(G)$.[5] *Then*

$$\mathbf{B}(G,\lambda) = \sum_{k=0}^{\infty} \lambda^k \mathbf{A}^k.$$

The Bonacich centrality vector is given by

$$\mathbf{b}(G,\lambda) = \mathbf{B}(G,\lambda) \cdot \mathbf{u}. \tag{2}$$

We can write the Bonacich centrality vector as

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

For the components $b_i(G,\lambda)$, $i = 1,...,n$, we get

$$b_i(G,\lambda) = \sum_{k=0}^{\infty} \lambda^k (\mathbf{A}^k \cdot \mathbf{u})_i = \sum_{k=0}^{\infty} \lambda^k \sum_{j=1}^{n} \left(\mathbf{A}^k\right)_{ij}, \tag{3}$$

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

Now we can turn to the equilibrium analysis of the game.

Theorem 1 (Ballester et al. [2006]). *Let* $\mathbf{b}(G,\lambda)$ *be the Bonacich network centrality of parameter* λ. *For* $\lambda < 1/\lambda_{PF}(G)$ *the unique interior Nash equilibrium solution of the simultaneous* n*–player move game with payoffs*

[5] The proof can be found e.g. in Debreu and Herstein [1953].

given by Equation (1) and strategy space $\mathbb{R}^n_+$ is given by

$$x_i^* = b_i(G, \lambda), \tag{4}$$

for all $i = 1, ..., n$.

Moreover, the payoff of agent i in the equilibrium is given by[6]

$$\pi_i(\mathbf{x}^*, G) = \frac{1}{2}(x_i^*)^2 = \frac{1}{2}b_i^2(G, \lambda). \tag{5}$$

The parameter λ measures the effect on agent i of agent j's contribution, if they are connected. If we assume that we have strong network externalities so that λ approaches its highest possible value $1/\lambda_{\mathrm{PF}}(G)$ then the Bonacich centrality becomes proportional to the standard eigenvector measure of centrality [Wasserman and Faust, 1994]. The latter result has been shown by Bonacich [1987] and Bonacich and Lloyd [2001].

Furthermore, Ballester et al. [2006] have shown that the equilibrium outcome and the payoff for each player increases with the number of links in G (because the number of network walks increases in this way).[7] This implies that, if an agent is given the opportunity to change her links, she will add as many links as possible. On the other hand, if she is only allowed to form one link at a time, she will form the link to the agent that increases her payoff the most. In both cases, eventually, the network will then become complete, i.e. each agent is connected to every other agent. However, to avoid this latter unrealistic situation, we assume that the agents are living in a volatile environment that causes links to decay such that the complete network can never be reached.[8] Instead the architecture of the network adapts to the volatile environment. We will treat these issues more formally in the next section.

2.2. The Network Formation Game

We now introduce a network formation process that incorporates the idea that agents with high Bonacich centrality (their equilibrium effort levels) are more likely to connect to each other.

Let time be measured at countable dates $t = 0, 1, 2, ...$ and consider the network formation process $(G(t))_{t=0}^{\infty}$ with $G(t) = (N, L(t))$ comprising

[6]As discussed in more detail in Section 7, all our results would be unchanged in a setup where agents do not choose efforts but where their payoff is given by any monotonic increasing function of their Bonacich centrality.

[7]See Theorem 2 in Ballester et al. [2006].

[8]As discussed in more detail in Section 7, all our results would remain unchanged if there were no link deletion but each agent would have a finite maximum number of links she could maintain. In particular, in that case, the emerging network would be a nested split graph.

the set of agents $N = \{1, ..., n\}$ together with the set of links $L(t)$ at time t.[9] The timing is as follows: At $t = 0$, we start with the empty network $G(0) = \bar{K}_n$.[10] Then every agent $i \in N$ optimally chooses her effort $x_i \in \mathbb{R}_+$, which is $x_i^* = 1$, since $b_i(\bar{K}_n, \lambda) = 1$ for all $i = 1, ..., n$.[11] Then, an agent $i \in N$ is chosen at random and with probability $p_i \in (0, 1)$ forms a link with agent $j \in N \backslash (\mathcal{N}_i \cup \{i\})$ that gives her the highest payoff (or equivalently her highest Bonacich centrality as given in Equation (5)). We obtain the network $G(1)$. Then, again, a player i is chosen at random and with probability p_i 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. And so forth.

The key question is how individuals choose among their potential linking partners. Let us explain the way someone is selected to form a link. At every t, an agent i, selected uniformly at random from the set N, enjoys an updating opportunity of her current links at a rate p_i. If an agent receives such an opportunity, then she initiates a link to agent j 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. Note that, 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.

Observe that when agents decide to create a link, they do it in a *myopic* way, that is they only look at the agents that gives them the *current* highest payoff. There is literature on farsighted networks where agents calculate their lifetime-expected utility when they want to create a link (see, e.g. Konishi and Ray [2003]). We adopt here a myopic approach because of its tractability and because our model also incorporates effort decision.[12]

Let us give a formal definition of the best responses of an agent given the prevailing network $G(t)$.[13]

[9] We give a formal characterization of this stochastic process in Section 3.

[10] For simplicity we start from the empty network. However, all results of Section 3 and thereafter hold if we start from any initial network (see footnote 21 in Section 2.3 and Section 7).

[11] The adjacency matrix $\mathbf{A}$ of the empty network $\bar{K}_n$ contains only zero entries and therefore $\mathbf{b}(\bar{K}_n, \lambda) = (\mathbf{I} - \lambda\mathbf{A})^{-1}\mathbf{u} = \mathbf{u}$.

[12] Jackson and Watts [2002b] argue that this form of myopic behavior makes sense if players discount heavily the future.

[13] In order to guarantee an interior solution of the Nash equilibrium efforts corresponding to the payoff functions in equation (1), we assume that the parameter $\lambda \geq 0$ is smaller than the inverse of the largest real eigenvalue of $G(t)$ for any t. Testing the impact of the Bonacich centrality measure on educational outcomes in the United States, Calvó-Armengol et al. [2009] found that only 18 out of 199 networks (i.e. 9 percent) do

Definition 2. *Consider the current network $G(t) = (N, L(t))$ 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 $ij \notin L(t)$ between agents $i, j \in N$. Further, let $\pi^*(G(t)) = (\pi_1^*(G(t)), ..., \pi_n^*(G(t)))$ denote the profile of Nash equilibrium payoffs of the agents in $G(t)$ following from the payoff function (1) with parameter $0 \leq \lambda < 1/\lambda_{PF}(G(t))$. Then agent j is a best response of agent i if $\pi_i^*(G(t) + ij) \geq \pi_i^*(G(t) + ik)$ for all $j, k \in N \backslash (\mathcal{N}_i \cup \{i\})$. Agent j may not be unique. The set of agent i's best responses is denoted by $BR_i(G(t))$.*

Note that the best response strategies for the network games introduced in Bala and Goyal [2000]; Haller et al. [2007]; Haller and Sarangi [2005] allow an agent to remove or create an arbitrary number of links while we restrict the link formation (strategy space) of an agent to one additional link only. We omit the removal of links since agents payoffs are monotonic increasing in the number of links in the network. Since the removal of a link would always decrease an agent's payoffs, link removal is strictly dominated by link creation.

We assume that during the time interval from t to $t + 1$ an agent i is selected and either has the possibility to create a link (with probability $p_i \in (0, 1)$) or to severe a link (with probability $q_i \in (0, 1)$). Note that taking into account the possibility of an agent remaining quiescent only modifies the time-scale of the process discussed, thus yielding identical results to the model proposed. This implies that, without any loss of generality, it is possible to assume $p_i + q_i = 1$. For simplicity, we also assume that these probabilities are the same across agents. Accordingly, we will use one parameter α and $1 - \alpha$ to denote the probabilities at which links are formed and removed respectively, that is, $p_i = \alpha$ and $q_i = 1 - \alpha$, with $\alpha \in (0, 1)$, for all $i \in N$.

Definition 3. *We define the network formation process $(G(t))_{t=0}^{\infty}$, $G(t) = (N, L(t))$, as a sequence of networks $G(0) = \bar{K}_n, G(1), G(2), ...$ 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:*

(i) *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 $BR_i(G(t)) = \emptyset$ or $BR_j(G(t)) = \emptyset$ nothing happens. If $BR_i(G(t))$ is not unique, then i selects randomly one agent in $BR_i(G(t))$.*

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

---

not satisfy this eigenvalue condition. In Section 7, we show that our results still hold in a more general framework (i.e. more general utility function) where this condition on eigenvalue is not needed.

In words, with probability α, the selected agent will create a link with the agent who increases the most her utility, while with probability $1-\alpha$, the selected agent will delete a link with her direct neighbor who reduces the least her utility. This link is for the selected agent the least important and thus the least frequently used. Note that the newly established link also affects the overall network structure and therewith the centralities and payoffs of all other agents (in the same connected component). The formation of links thus can introduce large, unintended and uncompensated externalities.

2.3. Network Formation and Nested Split Graphs

An essential property of the link formation process $(G(t))_{t=0}^{\infty}$ introduced in Definition 3 is that it produces networks in a well defined class of graphs denoted by "nested split graphs" [Aouchiche et al., 2008].[14] We will give a formal definition of these graphs and discuss an example in this section. Nested split graphs include many common networks such as the star or the complete network. Moreover, as their name already indicates, they have a *nested neighborhood structure*. This means that the set of neighbors of each agent is contained in the set of neighbors of each higher degree agent. Nested split graphs have particular topological properties and an associated adjacency matrix with a well defined structure.

In order to characterize nested split graphs, it will be necessary to consider the degree partition of a graph, which is defined as follows:

Definition 4 (Mahadev and Peled [1995]). *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.*

With the definition of a degree partition, we can now give a more formal definition of a nested split graph.[15,16]

Definition 5 (Mahadev and Peled [1995]). *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, $i = 1, ..., \lfloor \frac{k}{2} \rfloor$ and a dominating set $\bigcup_{i=\lfloor \frac{k}{2} \rfloor +1}^{k} D_i$ in the graph $G' = (N \backslash D_0, L)$. Moreover, the*

[14] Nested split graphs are also called "threshold networks" [Hagberg et al., 2006; Mahadev and Peled, 1995].

[15] 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).

[16] In general, split graphs are graphs whose nodes can be partitioned in a set of nodes which are all connected among each other and sets of nodes which are disconnected. A nested split graph is a special case of a split graph.

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

$$\mathcal{N}_v = \begin{cases} \bigcup_{j=1}^{i} D_{k+1-j} & \text{if } i = 1, ..., \left\lfloor \frac{k}{2} \right\rfloor, \\ \bigcup_{j=1}^{i} D_{k+1-j} \setminus \{v\} & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, ..., k. \end{cases} \tag{6}$$

Figure 1 (left) illustrates the degree partition $\mathbf{D} = (D_0, D_1, ..., D_6)$ and the nested neighborhood structure of a nested split graph. A line between D_i and D_j indicates that every node in D_i is linked to every node in D_j for any $i, j = 1, ..., 6$. The nodes in the dominating set included in the solid frame induce a clique while the nodes in the independent sets that are included in the dashed frame induce an empty subgraph. In the following we will call the sets D_i, $i = \left\lfloor \frac{k}{2} \right\rfloor + 1, ..., k$, *dominating subsets*, since the set D_i induces a dominating set in the graph obtained by removing the nodes in the set $\bigcup_{j=0}^{k-i} D_j$ from G.

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

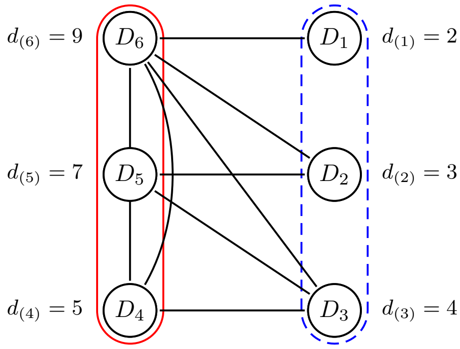
Definition 6 (Brualdi and Hoffman [1985]). *A stepwise matrix* $\mathbf{A}$ *is a matrix with elements a_{ij} satisfying the condition: if $i < j$ and $a_{ij} = 1$ then $a_{hk} = 1$ whenever $h < k \leq j$ and $h \leq i$.*

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

If a nested split graph is connected we call it a connected nested split graph. The representation and the adjacency matrix depicted in Figure 1 actually shows a connected nested split graph. 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. In Appendix B, 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 (see Wasserman and Faust [1994]) 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 (see Section B.2.5 in the Appendix).

In the next proposition, we identify the relationship between the Bonacich centrality of an agent and her degree in a nested split graph.

Proposition 1. *Consider a pair of agents $i, j \in N$ of a nested split graph $G = (N, L)$.*



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

Figure 1: Representation of a connected nested split graph (left) and the associated adjacency matrix (right) with $n = 10$ agents and $k = 6$ distinct positive degrees. A line between D_i and D_j indicates that every node in D_i is linked to every node in D_j. The solid frame indicates the dominating set and the nodes in the independent sets are included in the dashed frame. Next to the set D_i the degree of the nodes in the set is indicated. The neighborhoods are nested such that the degrees are given by $d_{(i+1)} = d_{(i)} + |D_{k-i+1}|$ for $i \neq \lfloor \frac{k}{2} \rfloor$ and $d_{(i+1)} = d_{(i)} + |D_{k-i+1}| - 1$ for $i = \lfloor \frac{k}{2} \rfloor$. In the corresponding adjacency matrix $\mathbf{A}$ to the right the zero-entries are separated from the one-entries by a stepfunction.

(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 than adding the link ij to G, i.e.*

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

From part (ii) of Proposition 1 we find that when agent i has to decide to create a link either to agents k or j, with $d_k > d_j$, in the link formation process $(G(t))_{t=0}^{\infty}$ then i will always connect to agent k because this link gives i a higher Bonacich centrality than the other link to agent j. We can make use of this property in order to show that the networks emerging from the link formation process defined in the previous section actually are nested split graphs. This result is stated in the next proposition.

Proposition 2. *Consider the network formation process $(G(t))_{t=0}^{\infty}$ introduced in Definition 3. Then, at any time $t \geq 0$, a network $G(t)$ 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 by Proposition 1 the highest degree). This creates a nested neighborhood structure which can always be represented by a stepwise adjacency matrix after a possible relabeling of the agents.[17] The same applies for link removal.

Let us give some more intuition of this crucial result. Agents want to link to others who are more central since this leads to higher actions (as actions are proportional to centrality) and higher actions raise payoffs more. Similarly, individuals delete links of those with lower centrality as these agents have lower actions and hence lower payoff effects. Notice moreover that, once someone deletes a link with an agent, she becomes less central and this makes it more likely that the next person who has a choice will also delete a link with her. Thus link gains and losses are self reinforcing. This intuition suggests that if α, the probability of adding links is large then the process should approximate complete network while if it is small then the process should approximate the star network. The key insight of our model is that for intermediate values of α the network is a nested split graph.[18]

Observe that it is assumed that there is no cost to forming links. If links represent a social tie, then there is a cost to maintaining a link since agents must spend time with the person they are linked to. Because of the assumption of no link cost, each agent wants to connect to every other agent, which leads to the formation of nested split graphs. In Section 7 and Appendices C, D and E, we extend the model to see what would happen to the results if links were costly to maintain. We consider two types of extensions. First, we look at a model where either agents do not delete links or face capacity constraints in the number of links they can maintain. Second, we examine the case where there is a cost to create a link. In all these cases, we show that networks always converges to nested split graphs and that all our results hold.

Due to the nested neighborhood structure of nested split graphs, any pair of agents in (the connected component) a nested split graph is at most two links separated from each other. From Proposition 1 it then follows that in a nested split graph $G(t)$ the best response of an agent i are the

[17] Two graphs $G = (N, L)$ and $G' = (N', L')$ are the same unlabeled graph when they are isomorphic, i.e., when there exists a permutation $\pi\colon N \to N'$ such that $ij \in L$ if and only if $\pi(i)\pi(j) \in L'$. Further, we will show in Proposition 3 that $(G(t))_{t=0}^{\infty}$ induces a finite state Markov chain with state space Ω consisting of all unlabeled nested split graphs. Consequently, two states $x, y \in \Omega$ of the Markov chain $(G(t))_{t=0}^{\infty}$ are identical, $x = y$, if they correspond to the same unlabeled graph.

[18] In Appendix C, we show that, if agents face some capacity constraint in the number of links they can maintain, then, even when $\alpha = 1$ (i.e. no link removal), the network is still a nested split graph.

agents with the highest degrees in i's second-order neighborhood $\mathcal{N}_i^{(2)}$.[19] Moreover, if $G(t)$ is a nested split graph then $i \in BR_j(G(t))$ if and only if $j \in BR_i(G(t))$.

From the fact that $G(t)$ is a nested split graph with an associated stepwise adjacency matrix it further follows that at any time t in the network evolution, $G(t)$ consists of a single connected component and possibly isolated nodes.

Corollary 1. *Consider the network formation process $(G(t))_{t=0}^{\infty}$ introduced in Definition 3. Then, at any time $t \geq 0$, a network $G(t)$ consists of a connected component and possibly isolated nodes.*

Nested split graphs are not only prominent in the literature on spectral graph theory [Cvetkovic et al., 1997] but they have also appeared in the recent literature on economic networks. Nested split graphs are so called "inter-linked stars" found in Goyal and Joshi [2003].[20] Subsequently, Goyal et al. [2006] identified inter-linked stars in the network of scientific collaborations among economists. It is important to note that nested split graphs are characterized by a distinctive core-periphery structure. Core-periphery structures have been found in several empirical studies of interfirm collaborations networks [Baker et al., 2008]. The wider applicability of nested split graphs suggests that a network formation process that generates these graphs as it is defined in Definition 3 are of general relevance for understanding economic and social networks.

Finally, note that the network formation process $(G(t))_{t=0}^{\infty}$ introduced in Definition 3 is independent of initial conditions $G(0)$.[21] This means that even when we start from an initial network $G(0)$ which is not a nested split graph then after some finite time the Markov chain will reach a nested split

[19] Let $\mathcal{N}_i = \{k \in N : ik \in L(t)\}$ be the set of neighbors of agent $i \in N$ and $\mathcal{N}_i^{(2)} = \bigcup_{j \in \mathcal{N}_i} \mathcal{N}_j \backslash (\mathcal{N}_i \cup \{i\})$ denote the second-order neighbors of agent i in the current network $G(t)$. Note that the connectivity relation is symmetric such that j is a second-order neighbor of i if i is a second-order neighbor of j, i.e. $i \in \mathcal{N}_j^{(2)}$ if and only if $j \in \mathcal{N}_i^{(2)}$ for all $i, j \in N$.

[20] Nested split graphs are inter-linked stars but an inter-linked star is not necessarily a nested split graph. Nested split graphs have a nested neighborhood structure for all degrees while in an inter-linked star this holds only for the nodes with the lowest and highest degrees.

[21] 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 $\bar{K}_n$ is obtained. Let T be the time when this happens starting from some network $G \notin \Omega$. Note that $\bar{K}_n \in \Omega$, and therefore Proposition 2 implies that all networks $G(t)$, $t > T$, visited by the chain will be in Ω. A network G is transient if $\sum_{\tau=1}^{\infty} \mathbb{P}(G(t+\tau) = G|G(t) = G) < \infty$ [e.g. Grimmett and Stirzaker, 2001, p. 221]. We have that $\sum_{\tau=1}^{\infty} \mathbb{P}(G(t+\tau) = G|G(t) = G) = \sum_{\tau=1}^{T} \mathbb{P}(G(t+\tau) = G|G(t) = G) < T < \infty$. Therefore all networks which are not nested split graphs are transient and they have vanishing probability in the stationary distribution μ (see Section 3).

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

3. Stationary Networks: Characterization

In this section we analyze in more detail the network formation process $(G(t))_{t=0}^{\infty}$ defined in the previous section, where $G(t)$ is the random variable realized at time $t \geq 0$. Let $\mathcal{F}$ define the σ-algebra $\sigma\left(G(t) : t \in \mathbb{N}_0\right)$ generated by $G(0), G(1), ...$ and let Ω denote the state space of $(G(t))_{t=0}^{\infty}$. Then the probability space is given by the triple $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}\colon \mathcal{F} \mapsto [0,1]$ is the probability measure satisfying $\sum_{G \in \Omega} \mathbb{P}(G) = 1$. In the following, we show that the network formation process $(G(t))_{t=0}^{\infty}$ induces an ergodic Markov chain and we analyze the asymptotic states of this process as the number n of agents becomes large.

Proposition 3. *The network formation process $(G(t))_{t=0}^{\infty}$ introduced in Definition 3 induces an ergodic Markov chain on the state space Ω with a unique stationary distribution μ. 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 $|\Omega| = 2^{n-1}$.*

The symmetry of the network formation process $(G(t))_{t=0}^{\infty}$ with respect to the link formation probability α and the link removal probability $1-\alpha$ allows us to state the following proposition.

Proposition 4. *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 μ be the stationary distribution of $(G(t))_{t=0}^{\infty}$ and μ' the stationary distribution of $(G'(t))_{t=0}^{\infty}$. Then for each network G in the stationary distribution μ with probability μ_G the complement of G, $\bar{G}$, has the same probability μ_G in μ', i.e. $\mu'_{\bar{G}} = \mu_G$.*

Proposition 4 allows us to derive the stationary distribution μ for any value of $1/2 < \alpha < 1$ if we know the corresponding distribution for $1-\alpha$. This follows from the fact that the complement $\bar{G}$ of a nested split graph G is a nested split graph as well [Mahadev and Peled, 1995]. In particular, 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, conversely, the number of nodes in the independent sets in $\bar{G}$ corresponds to the number of nodes in the dominating subsets in G.

With this symmetry in mind we restrict our analysis in the following to the case of $0 < \alpha \leq 1/2$. Let $\{N(t)\}_{t=0}^{\infty}$ be the degree distribution with the d-th element $N_d(t)$, giving the number of nodes with degree d in $G(t)$, in the t-th sequence $N(t) = \{N_d(t)\}_{d=0}^{n-1}$. Further, let $n_d(t) = N_d(t)/n$ denote the proportion of nodes with degree d and let $n_d = \lim_{t\to\infty} \mathbb{E}(n_d(t))$

be its asymptotic expected value (as given by μ). In the following proposition we determine the asymptotic degree distribution of the nodes in the independent sets for n large enough.

Proposition 5. *Let $0 < \alpha \leq 1/2$. 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*

$$n_d = \frac{1-2\alpha}{1-\alpha}\left(\frac{\alpha}{1-\alpha}\right)^d, \tag{7}$$

where[22]

$$d^*(n,\alpha) = \frac{\ln\left(\frac{(1-2\alpha)n}{2(1-\alpha)}\right)}{\ln\left(\frac{1-\alpha}{\alpha}\right)}. \tag{8}$$

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 only a single node. Further, using Proposition 4, we know that 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 in Equation (7) for $1-\alpha$, while each of the independent sets contains a single node. This determines the asymptotic degree distribution for the independent or dominating subsets, respectively, for all values of α in the limit of large n.

Moreover, we can show that the empirical degree distribution converges in probability to the expected distribution in the limit of large network sizes n.

Proposition 6. *For any $\epsilon > 0$ we have that*

$$\mathbb{P}\left(|n_d(t) - \mathbb{E}\left(n_d(t)\right)| \geq \epsilon\right) \leq 2e^{-\frac{\epsilon^2 n^2}{8t}}. \tag{9}$$

Furthermore, from Equation (8) we can directly derive the following corollary.

Corollary 2. *There exists a phase transition in the asymptotic average number of independent sets, $d^*(n,\alpha)$, as n becomes large such that*

$$\lim_{n\to\infty}\frac{d^*(n,\alpha)}{n} = \begin{cases} 0, & \text{if } \alpha < \frac{1}{2}, \\ \frac{1}{2}, & \text{if } \alpha = \frac{1}{2}, \\ 1, & \text{if } \alpha > \frac{1}{2}. \end{cases} \tag{10}$$

[22] Note that $d^*(n,\alpha)$ from Equation (8) might in general not be an integer. In this case we take the closest integer value to Equation (8), that is, we take $[d^*(n,\alpha)] = \lfloor d^*(n,\alpha) + \frac{1}{2} \rfloor$. The error we make in this approximation is negligible for large n.

Corollary 2 implies that as n grows without bound the networks in the stationary distribution μ are either sparse or dense, depending on the value of the link creation probability α. Moreover, from the functional form of $d(n, \alpha)$ in Equation (8) we find that there exists a sharp transition from sparse to dense networks as α crosses $1/2$ and the transition becomes sharper the larger is n.

Observe that, because a nested split graph is uniquely defined by its degree distribution,[23] Proposition 5 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.[24] The latter is shown in Figure 2. From the structure of these matrices we 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 displays the number of links m 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 (8) is monotonic increasing in α for $\alpha < 1/2$ and monotonic decreasing in α for $\alpha > 1/2$.

4. How realistic are nested-split networks?

In this section, we would like to discuss two main characteristics of nested-split networks and see how "realistic" they are. First, are real-world networks nested? Nestedness appears in various social contexts, including the organization of the New York garment industry (Uzzi [1996]), and as disassortativity in the topology of the Fedwire network (May et al. [2008]; Soramäki et al. [2007]). If we consider the latter (i.e. bank networks), then as reported in May et al. [2008], the topology of interbank payment flows within the US Fedwire service (see Figures 1 and 2 in Soramäki et al. [2007]) is clearly nested. The sample from this network amounted to around 700,000 transfer funds, with just over 5,000 banks involved on an average day. The

[23] The degree distribution uniquely determines the corresponding nested split graph up to a permutation of the indices of nodes.

[24] Non-integer values for the partition sizes can be approximated with the closest integer while preserving the nested structure of the degree partitions.

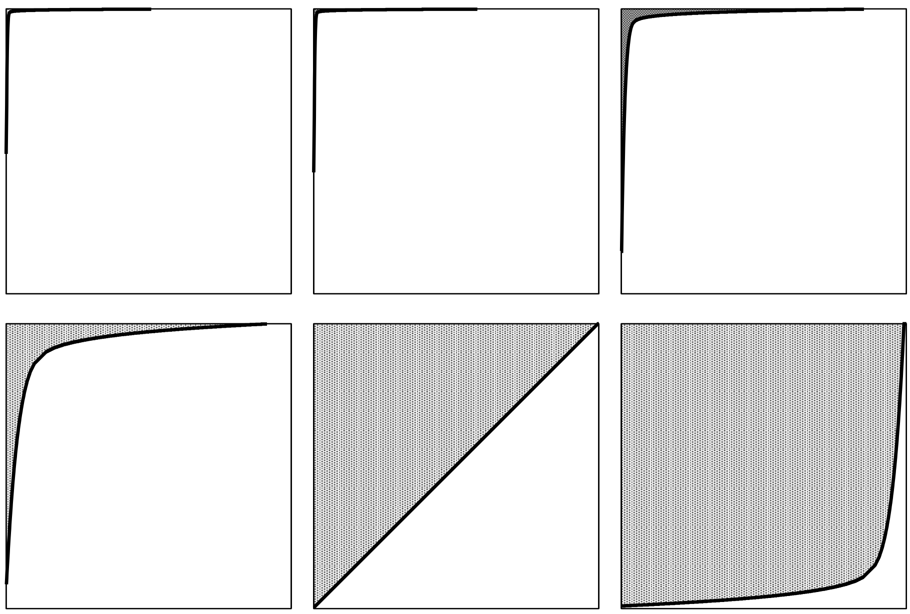


Figure 2: Representation of the adjacency matrices of stationary networks with $n = 1000$ agents for different values of parameter α: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.495$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The matrix top-left for $\alpha = 0.4$ is corresponding to an inter-linked star 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.

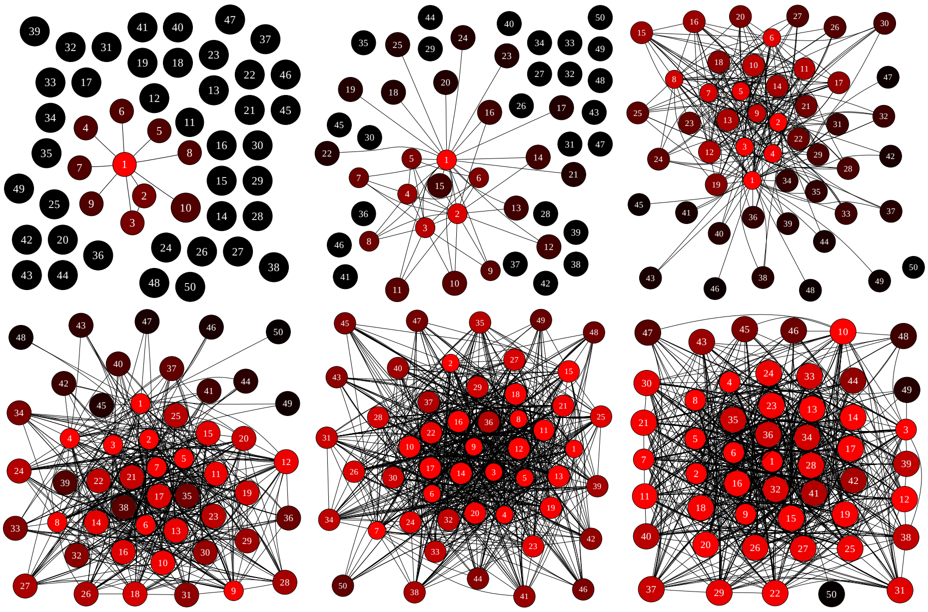


Figure 3: Sample networks with $n = 50$ agents for different values of parameter α: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.495$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). Nodes with brighter shapes correspond to agents with a higher 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. The network plots have been generated using a Fruchterman-Reingold algorithm [Fruchterman and Reingold, 1991].

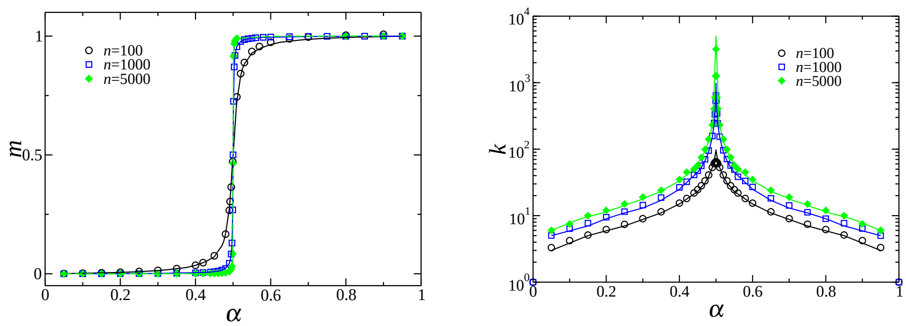


Figure 4: In the left panel we show the number of links m of the stationary network. The number of distinct degrees $k = 2d^*$,with d^* from Equation (8), found in the stationary network for different values of α are shown in the right panel. The figures display both, the results obtained by recourse of numerical simulations (symbols) and respecting theoretical predictions (lines) of the model.

authors (Soramäki 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 zero (few to no flows). On a daily basis, 75 percent of the payment flows involve fewer than 0.1 percent of the nodes, and only 0.3 percent of the observed linkages between nodes (which are already extremely sparse). Overall, the topology of this Fedwire network is highly disassortative: large banks were disproportionately connected to small banks, and vice versa; the average bank was connected to 15 others. In other words, most banks have only a few connections while a small number of "hubs" have thousands. Here, disassortativity may arise naturally when banks seek relationships with each other that are the most mutually beneficial: for example, small banks may interact with large banks for security, lower liquidity risk and lower servicing costs, and large banks may interact preferentially with small banks in part because they can extract a higher premium for services and can in principle accommodate more risk.

Åkerman and Larsson [2010] study the evolution of the global arms trade network using a unique dataset on all international transfers of major conventional weapons over the period 1950-2007. They find that networks are nested and dissortative in the sense that big countries mainly trade arms with small countries but small countries do not trade with each other. Using aggregate bilateral imports from 1950 to 2000, De Benedictis and Tajoli [2010] analyze the structure of the World Trade Network over time, detecting and interpreting patterns of trade ties among countries. They also find that 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 disproportionally large share of world trade. The larger countries account for a generally larger share of world trade and have more partners. Figure 3 in their paper shows a clear core-periphery structure similar to our Figure 3. Indeed, in the 1950s, the trade network was relatively close to that of our Figure 3 when α is small while, in the 1990s, it resembles more to the case when α is high.

Second, nested-split graphs have a diameter equals to 2. Most real-world networks have, in fact, very low diameter. For example, if we consider the Fedwire bank network discussed above (Soramäki et al. [2007]), then the average path length (which is the average distance from a node to any other node) is 2.6 while the diameter is around 6 (see their Table 3 on page 324). This means that this interbank payment network exhibits the small world phenomenom common to many networks. Similarly, Åkerman and Larsson [2010] who study arm-trade networks find also small diameters: between 4 and 6. Finally, Banerjee et al. [2010] have detailed network data on 75 villages in rural Karnataka (2-3 hours outside Bangalore in India) before a particular microfinance institution starts working in the villages. They study how the idea of joining microfinance spread through the network. In this dataset, social network consists of friends and relatives and from whom they borrow or lend small amount of money. The average path length is 1.97 and the diameter is 3. All these networks have small diameters but, of course, not equal exactly to 2. We can extend our model to obtain networks with diameters closer to 6 than to 2 (see, for example, Appendix E) but adding these features would reduce the mathematical elegance of our formulation in terms of nested-split graphs. In this respect, nested-split graphs are a good approximation of real-world networks, even though they are not exactly what we observed in reality.

5. Stationary 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:[25]

(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. Barabási 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

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

> 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 as in the network between venture capitalists [Mas et al., 2007] as well as negative degree correlations as it can be found in the world trade web [Serrano and Boguñá, 2003], online social communities [Hu and Wang, 2009] and in networks of banks [De Masi and Gallegati, 2007; May et al., 2008].

In the following sections, we analyze some of the topological properties of the stationary networks in our model. With the asymptotic expected degree distribution derived in Proposition 5, we can calculate the expected clustering coefficient, the clustering-degree correlation, the neighbor connectivity, the assortativity, and the characteristic path length by using the expressions derived for these quantities in Appendix B, where we show that these statistics are all functions of the degree distribution.[26] These network measures are interesting because they can be compared to key empirical findings of social and economic networks. In fact, we show that the stationary networks exhibit all the well-known stylized facts of real-world networks. Moreover, we show in Appendix E that, by introducing capacity constraints in the number of links an agent can maintain, we are able to produce both, assortative as well as dissortative networks.

Note that since the stationary distribution μ is unique, we can recover the expected value of any statistic by averaging over a large enough sample of empirical networks generated by numerical simulations. We then superimpose the analytical predictions of the statistic derived from Proposition 5 with the sample averages in order to compare the validity of our theoretical results, also for small network sizes n. As we will show, there is a good agreement of the theory with the empirical results for all network sizes.

5.1. Degree Distribution

From Proposition 5, we find that the degree distribution follows an exponential decay with a power-law tail.[27] The power-law tail has an exponent of

[26] By virtue of Proposition 6, we know that the probability limit of the degree distribution is its expected value. Therefore we can compute the probability limit of these statistics in good approximation, by evaluating them at the expected degree distribution [Kim and Motter, 2007]. In particular, we compute the integer valued sizes of the degree partition from the real-valued asymptotic degree distribution by taking closest integers.

[27] For $0 < \alpha \leq 1/2$ and n large enough the asymptotic expected degree distribution for the degrees d smaller or equal than d^* is given by an exponential function

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

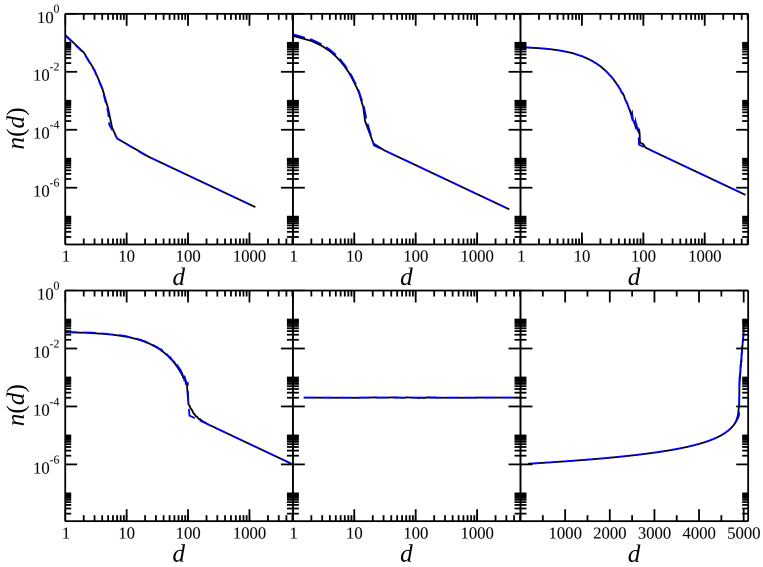


Figure 5: Degree distribution n_d for different values of parameter α and a network size $n = 10\,000$: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.49$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The solid line corresponds to the average of simulations while the dashed line indicates the theoretical degree distribution from Proposition 5. The degrees have been binned to smoothen the degree distribution.

minus one.[28] Degree distributions with power-law tails have been found in empirical networks, e.g. in scientific collaboration networks Newman [2004]. 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.

On the other hand, if we assume (i) that the degree of a node in a dominating subset is symmetrically distributed around its expected value, (ii) we compute the integral over the probability density function by a rectangle approximation and (iii) further assume that the degree distribution obtained in this way has the same functional form for all degrees d larger than d^* then one can show that for $0 < \alpha \leq 1/2$ and n large enough the asymptotic expected degree distribution $n(d)$ is given by

$$n(d) = \frac{\alpha}{(1-2\alpha)n} d^{-1}.$$

The power-law tail of the degree distribution can be confirmed by the empirical distribution from a logarithmic binning of numerical simulations, as can be seen in Figure 5.

[28] We can extend our model to obtain a degree distribution with an arbitrary power law tail by making the probability of selecting an agent depending on the number of links she already has, while preserving the nested structure of the network she is embedded in.

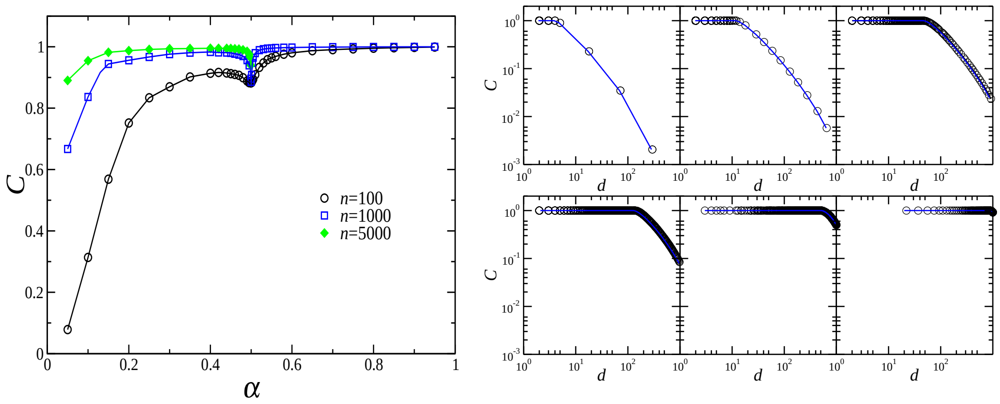


Figure 6: The left panel shows the clustering coefficient $\mathcal{C}$ and the right panel the clustering-degree correlation of stationary networks. The symbols correspond to the results obtained by recourse of numerical simulations. The solid lines correspond to the analytical results. We show that the clustering-degree correlation is negative for different values of α and a network size of $n = 1000$. The different plots show different values of α: $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.49$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot).

5.2. Clustering

The clustering coefficient is shown in Figure 6 (left). We find that for practically all values of α, the clustering in the stationary networks is high. This finding is in agreement with the vast literature on social networks that have reported high clustering being a distinctive feature of social networks. Moreover, Goyal et al. [2006] have shown that there exists a negative correlation between the clustering coefficient of an agent and her degree. We find this property in the stationary networks as well, as it is shown in Figure 6 (right).

5.3. Assortativity and Nearest Neighbor Connectivity

We now turn to the study of correlations between the degrees of the agents and their neighbors. This property is usually measured by the network assortativity γ [Newman, 2002, 2003] and nearest neighbor connectivity $d_{nn}(d)$ [Pastor-Satorras et al., 2001]. Dissortative networks are characterized by negative degree correlations between a node and its neighbors and assortative networks show positive degree correlations. In dissortative networks γ is negative and $d_{nn}(d)$ monotonic decreasing while in assortative networks γ is positive and $d_{nn}(d)$ monotonic increasing. We find that in our basic model without capacity constraints (see Appendix E for an extension including capacity constraints in the number of links an agent can maintain) we observe dissortative networks.

Assortativity and average nearest neighbor connectivity for different values of the link creation probability α are shown in Figure 7. Clearly, sta-

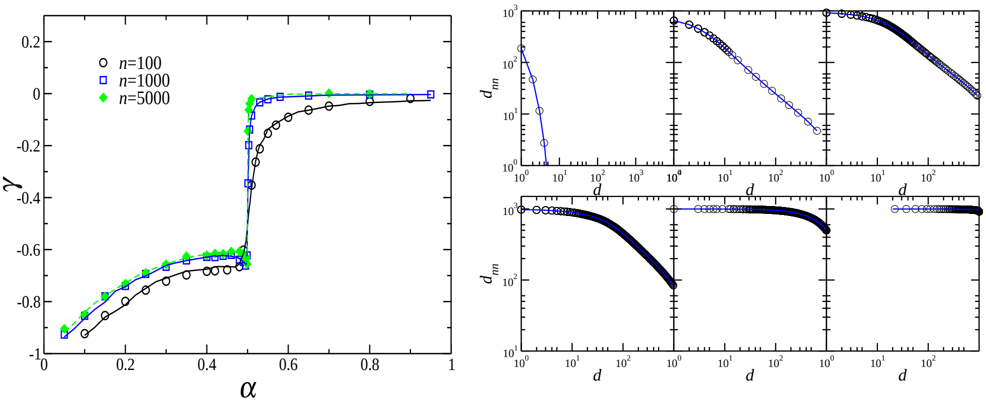


Figure 7: In the left panel we show the assortativity γ of stationary networks. In the right panel we show the average nearest neighbor connectivity d_{nn} for $\alpha = 0.2$ (top-left plot), $\alpha = 0.4$ (top-center plot), $\alpha = 0.48$ (top-right plot), $\alpha = 0.49$ (bottom-left plot), $\alpha = 0.5$ (bottom-center plot), and $\alpha = 0.52$ (bottom-right plot). The symbols correspond to the results obtained by recourse of numerical simulations. The solid lines correspond to the analytical results.

tionary networks are dissortative while the degree of dissortativity decreases with increasing α. However, if we recall the structure of the nested split graphs in Definition 5, to the class the stationary networks belong to, we can see that high degree agents are connected among each other while it is only the low degree agents that are not connected among each other. In this sense agents with high degrees tend to be connected to other agents with high degree. Considering only the agents with high degrees, we can call the network assortative. However, the agents with low degrees, that are only connected to agents with high degrees but are disconnected to agents with low degree, are so numerous in the stationary network (for low values of α) that we obtain an overall negative value for the assortativity of the network.

The dissortativity of stationary networks simply reflects the fact that stationary networks are strongly centralized for values of α below 1/2. As an example consider a star $K_{1,n-1}$. $K_{1,n-1}$ is completely dissortative with $\gamma = -1$. Peripheral agents all have minimum degree one and are only connected to the central agent with maximum degree while the central agent is only connected to the agents with minimum degree. In this sense dissortativity is simply a measure of centralization in the network.

5.4. Characteristic Path Length

Figure 8 shows the characteristic path length $\mathcal{L}$ and the network efficiency $\mathcal{E}$ (defined in Section B.1.4 in Appendix B). From these figures one can see that the characteristic path length $\mathcal{L}$ never exceeds a distance of two. This means that for all parameter values of α stationary networks

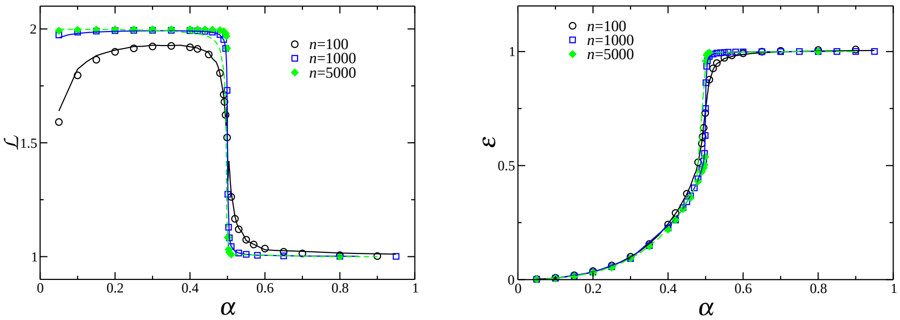


Figure 8: The left panel shows the characteristic path length $\mathcal{L}$ of stationary networks and the right panel shows the results for the network efficiency $\mathcal{E}$, obtained by recourse of numerical simulations (symbols) and respecting theoretical predictions (lines) of the model.

are characterized by short distances between agents. Together with the high clustering shown in this section the stationary networks can be seen as "small worlds"[Watts and Strogatz, 1998]. Stationary networks are efficient for values of α larger than $1/2$, in terms of short average distance between agents, while for values of α smaller than $1/2$ they are not. However, this short average distance is attained at the expense of a large number of links.

5.5. Centrality and Centralization in Stationary Networks

In the following section, we analyze the degree of centralization in stationary networks. As we will show, there exists a sharp transition in the centralization as a function of the link creation probability α. This means that stationary networks are either strongly centralized and hierarchical or decentralized and homogeneous, depending on α. In Section 6, we will also find such a transition in the aggregate payoffs and effort levels of the agents.

We use the centralization index introduced by Freeman [1978]. The centralization of a network $G = (N, L)$ is given by

$$\mathcal{C} = \frac{\sum_{u \in N} (\mathcal{C}(u^*) - \mathcal{C}(u))}{\max_{G'} \sum_{v \in N'} (\mathcal{C}(v^*) - \mathcal{C}(v))}, \tag{11}$$

where u^* and v^* are the agents with the highest values of centrality in the current network and and the maximum in the denominator is computed over all networks $G' = (N, L')$ with the same number of agents. For the degree, closeness, betweenness and eigenvector centrality measures one obtains the

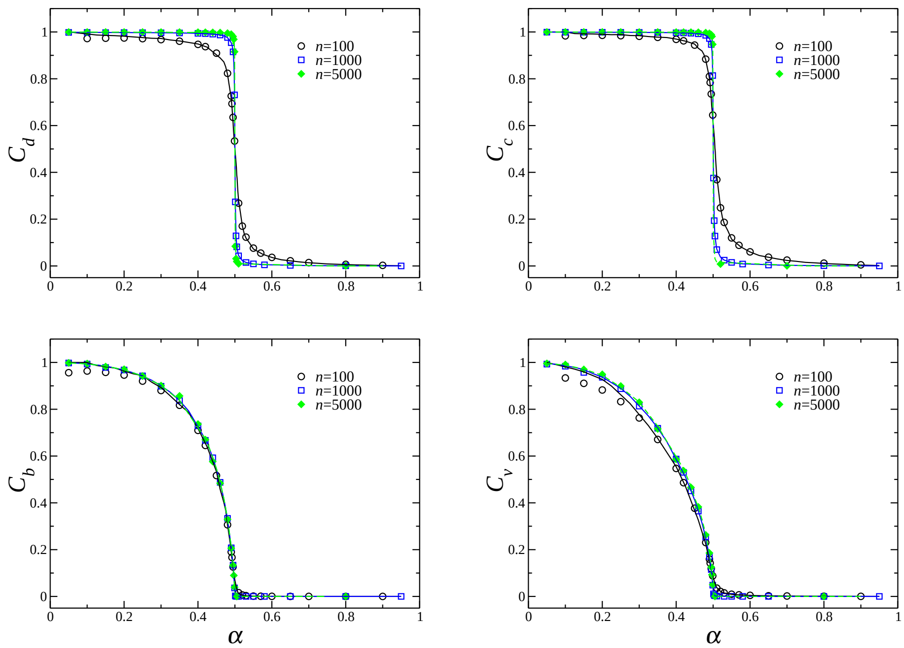


Figure 9: Degree, closeness, betweenness and eigenvector centralization in the stationary networks for different values of α. For all centralization measures we obtain a sharp transition between strongly centralized networks for lower values of α and decentralized networks for higher values of α. Note that we have only considered the connected component for the computation of the different centralization measures.

following indices[29]

$$\begin{aligned}
\mathcal{C}_d &= \frac{\sum_{u \in V} \left(\mathcal{C}_d(u^*) - \mathcal{C}_d(u) \right)}{n^2 - 3n + 2}, \\
\mathcal{C}_c &= \frac{\sum_{u \in V} \left(\mathcal{C}_c(u^*) - \mathcal{C}_c(u) \right)}{(n^2 - 3n + 2)/(2n - 3)}, \\
\mathcal{C}_b &= \frac{\sum_{u \in V} \left(\mathcal{C}_b(u^*) - \mathcal{C}_b(u) \right)}{n^3 - 4n^2 + 5n - 2}, \\
\mathcal{C}_v &= \frac{\sum_{u \in V} \left(\mathcal{C}_v(u^*) - \mathcal{C}_v(u) \right)}{\sqrt{(n-1)/2}(\sqrt{n-1} - 1)}.
\end{aligned}$$

From Figure 9, showing degree, closeness, betweenness and eigenvector centralization, we clearly see that there exists a phase transition at $\alpha = 1/2$

[29] For the normalization of all the centralization indices we have used the star $K_{1,n-1}$. For degree, closeness, betweenness and eigenvector centralization it can be shown that $K_{1,n-1}$ is the network that maximizes the sum of differences in centrality [Bolland, 1988; Freeman, 1978].

from highly centralized to highly decentralized networks. This means that for low arrival rates of linking opportunities α (and a strong link decay) the stationary network is strongly polarized, composed mainly of a star (or an inter-linked star as in Goyal and Joshi [2003]), while for high arrival rates of linking opportunities (and a weak link decay) stationary networks are largely homogeneous. We can also see that the transition between these states is sharp.

Our findings are in line with previous works studying the optimal internal communication structure of organizations [Guimerà et al., 2002]. Other works [Calvó-Armengol and De Martí, 2009; Dodds et al., 2003; Dupouet and Yildizoglu, 2006; Huberman and Hogg, 1995] have discussed the conditions under which informal organizational networks outperform centralized structures in complex, changing environments and under which conditions hierarchies are more efficient. Similar to Arenas et al. [2010] and Ehrhardt et al. [2006a], we find sharp transitions between largely homogeneous and centralized networks. Moreover, the stationary networks in our model are polarized and strongly centralized for a low volatility in the environment associated with many linking opportunities whereas they are homogeneous and largely decentralized for a highly volatile environment with few linking opportunities and a strong link decay. The hierarchical structure of stationary networks and its dependency on the volatility is similar to the findings for optimal networks in Arenas et al. [2010].

6. Stationary Networks: Efficiency

We now turn to the investigation of the optimality and efficiency of stationary networks. Following Jackson and Wolinsky [1996] and Jackson [2008], we define the social welfare as the sum of the agents' individual payoffs

$$\Pi(\mathbf{x}^*, G) = \sum_{i=1}^{n} \pi_i(\mathbf{x}^*, G). \tag{12}$$

We are interested in the solution of the following social planner's problem. Let $\mathcal{G}(n)$ denote the set of connected graphs having n agents in total. The social planner's solution is given by

$$G^* = \underset{G \in \mathcal{G}(n)}{\text{argmax}} \quad \Pi(\mathbf{x}^*, G). \tag{13}$$

A graph G^* solving the maximization problem in equation (13) will be denoted as "efficient". The efficient network has been derived in Ballester et al. [2006] and we state their result in the following proposition.

Proposition 7 (Ballester et al. [2006]). *Let $\mathcal{G}(n)$ denote the set of connected graphs having n agents and consider $G \in \mathcal{G}(n)$. Then the efficient* *network G^* in Equation (13) maximizing aggregate equilibrium contribution and payoff is the complete graph K_n.*

This proposition is a direct consequence of Theorem 2 in Ballester et al. [2006] where more links is always better. Moreover, Corbo et al. [2006] have shown that, in the case of strong complementarities, when λ approaches $1/\lambda_{\mathrm{PF}}(G)$, maximizing aggregate equilibrium payoffs is equivalent to maximizing the largest real eigenvalue $\lambda_{\mathrm{PF}}(G)$ of the network G.[30]

Proposition 8 (Corbo et al. [2006]). *Let $\mathcal{G}(n,m)$ denote the set of connected graphs having n agents and m links and consider $G \in \mathcal{G}(n,m)$. As $\lambda \uparrow 1/\lambda_{PF}(G)$, maximizing aggregate equilibrium contribution and payoff reduces to*

$$\max\{\lambda_{PF}(G) : G \in \mathcal{G}(n,m)\}.$$

Proposition 8 tells us that, if we want to compare aggregate payoffs of any two networks G_1 and G_2, we can compare their largest real eigenvalues, $\lambda_{\mathrm{PF}}(G_1)$ and $\lambda_{\mathrm{PF}}(G_2)$, in the case of strong complementarities λ. Moreover, from Proposition 7 we know that aggregate payoff is highest in the complete network K_n. K_n also has the highest possible largest real eigenvalue, namely $\lambda_{\mathrm{PF}}(K_n) = n-1$ [Cvetkovic and Rowlinson, 1990]. We assume that λ is close to $1/(n-1)$. The closer is the largest real eigenvalue $\lambda_{\mathrm{PF}}(G(t))$ of a network $G(t)$ to the one of the complete network $(n-1)$, the closer it comes to being efficient. Following these observations we show the ratio of the largest real eigenvalue of stationary networks to $n-1$ for different values of α. We find that for values of α below $1/2$, stationary networks are highly inefficient and a sharp transition occurs for increasing values of α above $1/2$. It is also seen that the transition becomes sharper the larger the network is. This implies that a highly volatile environment and the strong competition of the agents for becoming a hub induces highly inefficient network structures.[31]

7. 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 an arbitrary graph,[32] the process converges to

[30] Stationary networks might not always be connected. Instead, they can have a single connected component and isolated agents. If there are k isolated agents then their contribution to the aggregate payoff is k (with an equilibrium effort equal to one for each agent). We neglect the contribution of these isolated agents because it is negligible for large complementarities, and consider only the connected component of the network.

[31] It can be shown that the largest real eigenvalue can be increased by concentrating all the links in a densely connected core (clique) for fixed values of the number of links m and nodes n [Cvetkovic and Rowlinson, 1990].

[32] See also Footnote 21 in Section 2.3.

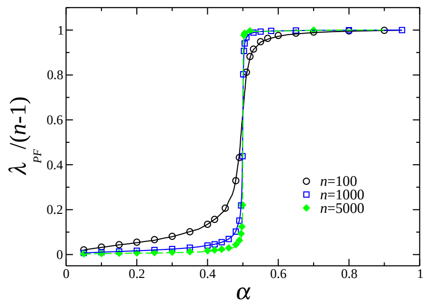


Figure 10: We show the largest eigenvalue of the adjacency matrix of the stationary network relative to the eigenvalue to the complete graph, which is the efficient network, obtained by recourse of numerical simulations (symbols) and respecting theoretical predictions (lines) of the model for different values of α and $n = 200$ agents. For higher values of α the stationary network comes close to the efficient graph which has a largest real eigenvalue of $n - 1$.

a nested split graph. 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 in Definition 3, where their utility is given by the sum of the current Bonacich centrality of their neighbors, then all our results still hold. As a consequence of Corollary 11 (see Appendix B.2.5), this is also true for any other centrality measure we have considered. Moreover, we could also define as utility "information centrality" introduced by Stephenson and Zelen [1989], where the payoff of agents is given by the information they receive along different paths in the network. If we let agents choose the importance (or weights, see Stephenson and Zelen [1989]) of the paths such that they receive maximum information (similar to choosing efforts $\mathbf{x}$) then all our results hold without making any assumption on the eigenvalue of the network.[33]

[33] The same reasoning as in the proof of Proposition 1 that applies for walks also applies for paths. This fact can be used to show that the agent with the highest degree in a nested split graph has the highest information centrality.

Second, we assume that, if selected, an agent must cut a link. In our model, the agent is assumed to cut the least valuable link. One could question the fact that we impose agents to severe a link that they would maybe want to keep. For example, an agent may choose to keep all her links (as it is the case if the payoff is given by Bonacich centrality). Our justification for this assumption is that agents cannot have an infinite number of links (for example in the case of friendships, individuals cannot have an infinite number of friends), because it would be too costly. In order to investigate this assumption further and to show that it is not crucial to the fact that the network always converges to a nested split graph, in Appendix C, we develop the same model but without imposing the link deletion of agents. In other words, consider exactly the same network formation process as in our original model (see Definition 3) without link removal (i.e. setting $\alpha = 1$). We assume, however, that each agent has a *finite maximum number of links* she can maintain. Furthermore, 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. If $\mathbf{C}$ denotes the vector of these constraints, then we show (see Proposition 9 in Appendix C) that our network formation process leads to a stable nested split graph with degree partition $\mathbf{D} = \mathbf{C}$.

Third, to follow on the previous point, we consider in Appendix D the same model as in Section 2 but we assume that the agent who wants to form a link must pay a cost c. Proposition 10 in Appendix D shows that if c is less than $\lambda/(1-\lambda)$, then agents will always want to form a link when selected and, as a result, the emerging network will always be a nested split graph. In other words, even with costly link formation, all our results hold as long as c is not too large.

Fourth, in Section 5.1, for the nodes in the dominating subsets, we obtain a power-law degree distribution with an exponent of minus one. We can extend our model to obtain a degree distribution with an arbitrary power law tail by making the probability of selecting an agent depending on the number of links she already has, while preserving the nested structure of the network she is embedded in.

Fifth, with our network formation game, we always obtain negative degree-degree correlations (i.e. our networks are dissortative). In Appendix E, we extend our game by including capacity constraints in the number of links an agent can maintain and a random search mechanism for new linking partners when an agent refuses to accept a new link due to her capacity constraints. We find that by introducing capacity constraints and random search, stationary networks can become assortative. Thus, we are able to reproduce all topological properties of empirically observed social and economic networks. The emergence of assortativity and positive degree-correlations, respectively, can be explained by considering limitations in the number of links an agent can maintain. This may be of particular relevance for social networks and give an explanation for the distinction between assortative social networks and dissortative technological networks suggested by Newman [2002].

Sixth, we have assumed myopic agents, i.e. agents maximize their current utility. Assume, instead, that agents create and delete links based on their discounted lifetime expected utilitie. They will still create and delete links with the agent who has the highest Bonacich centrality in the network since this agent has a higher probability than any other agent in the network to have the highest Bonacich centrality in the future (at any period of time). In other words, there exists a farsighted equilibrium which has the same properties as ours, i.e., the network at any period of time is a nested-split graph.

Finally, the value added of our approach as compared to more strategic based network formation models (such as Jackson and Wolinsky [1996]) is that we can match most features of real-world networks (e.g. the small world property, i.e. a low diameter and high clustering with a power-law degree distribution) while strategic based network formation models cannot. On the other hand, the value added of our approach as compared to random based network formation models (such as Barabási and Albert [1999]) is that, as in strategic network formation models, agents create and delete links based on economic incentives (i.e. maximize their utility function). Of course, our network formation process is not as rich as strategic based network formation models (such as Jackson and Wolinsky [1996]) because we look at the dynamics of network formation and only one agent at a time can create or delete a link. This is the trade off we face when combining these two approaches (random and strategic network formation).

References

Åkerman, A. and Larsson, A. (2010). The global arms trade network 1950-2007. Stockholm University Working Paper No. 2010:2.

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics*, 74(1):47–97.

Aouchiche, M., Bell, F., Cvetkovic, D., Hansen, P., Rowlinson, P., Simic, S., and Stevanovic, D. (2008). Variable neighborhood search for extremal graphs, 16. Some conjectures related to the largest eigenvalue of a graph. *European Journal of Operation Research*, 191:661–676.

Arenas, A., Cabrales, A., Díaz-Guilera, A., Guimerá, R., and Vega-Redondo, F. (2010). Optimal information transmission in organizations: Search and congestion. *Review of Economic Design*, 14:75–93.

Baker, G., Gibbons, R., and Murphy, K. (2008). Strategic Alliances: Bridges Between "Islands of Conscious Power"? *Journal of the Japanese and International Economies*, 22:146-163.

Bala, V. and Goyal, S. (2000). A noncooperative model of network formation. *Econometrica*, 68(5):1181–1230.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417.

Banerjee, A., Chandrasekhar, A.G., Duflo, E. and Jackson, M.O. (2010). Social networks and microfinance. Unpublished manuscript, MIT.
Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286:509–512.
Bavelas, A. (1948). A mathematical model for group structures. *Human Organization*, 7:16–30.
Beauchamp, M. A. (1965). An improved index of centrality. *Behavioral Science*, 10:161–163.
Boje, D. M. and Whetten, D. A. (1981). Effects of organizational strategies and contextual constraints on centrality and attributions of influence in interorganizational networks. *Administrative Science Quarterly*, 26:378–395.
Bolland, J. (1988). Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks. *Social Networks*, 10(3):233–253.
Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5):1170.
Bonacich, P. and Lloyd, P. (2001). Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, 23(3):191–201.
Borgatti, S. (2005). Centrality and network flow. *Social Networks*, 27(1):55–71.
Bramoullé, Y. and Kranton, R. (2007). Public goods in networks. *Journal of Economic Theory*, 135(1):478–494.
Bramoullé, Y., Kranton, R. and D'Amours, M.(2010). Strategic interaction and networks. Unpublished manuscript, Duke University.
Bramoullé, Y., Lopez, D., Goyal, S., and Vega-Redondo, F. (2004). Social interaction in anti-coordination games. *International Journal of Game Theory*, 33(1):1–19.
Brass, D. J. (1984). Being in the right place: A structural analysis of individual influence in an organization. *Administrative Science Quarterly*, 29(4):518–539.
Brualdi, R. A. and Hoffman, A. J. (1985). On the spectral radius of 0-1 matrices. *Linear Algebra and Application*, 65:133–146.
Cabrales, A., Calvó-Armengol, A., and Zenou, Y. (2009). Social interactions and spillovers: Incentives, segregation and topology. Unpublished manuscript, Stockholm University.
Calvó-Armengol, A. and De Martí, J. (2009). Information gathering in organizations: Equilibrium, welfare and optimal network structure. *Journal of the European Economic Association*, 7:116–161.
Calvó-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer effects and social networks in education. *Review of Economic Studies*, 76:1239-1267.
Calvó-Armengol, A. and Zenou, Y. (2004). Social Networks and Crime Decisions: The Role of Social Structure in Facilitating Delinquent Behavior. *International Economic Review*, 45(3): 939-958.
Cook, K. S., Emerson, R. M., Gillmore, M. R., and Yamagishi, T. (1983). The distribution of power in exchange networks: Theory and experimental results. *The American Journal of Sociology*, 89(2):275–305.
Corbo, J., Calvó-Armengol, A., and Parkes, D. (2006). A study of Nash equilibrium in contribution games for peer-to-peer networks. *SIGOPS Operation Systems Review*, 40(3):61–66.
Cvetkovic, D. and Rowlinson, P. (1990). The largest eigenvalue of a graph: A survey. *Linear and Multinilear Algebra*, 28:3–33.

Cvetkovic, D., Rowlinson, P., and Simic, S. (1997). *Eigenspaces of Graphs*, volume 66. Cambridge University Press, Encyclopedia of Mathematics.

De Benedictis,L. and Tajoli, L. (2010). The world trade network. Unpublished manuscript, University of Macerata, Italy.

De Masi, G. and Gallegati, M. (2007). *Econophysics of Markets and Business Networks*, chapter Debt-credit Economic Networks of Banks and Firms: the Italian Case, pages 159–171. Springer Milan.

Debreu, G. and Herstein, I. N. (1953). Nonnegative square matrices. *Econometrica*, 21(4):597–607.

Dodds, P. S., Watts, D. J., and Sabel, C. F. (2003). Information exchange and robustness of organizational networks. *Proceedings of the National Academy of Sciences*, 100(21):12516–12521.

Dupouet, O. and Yildizoglu, M. (2006). Organizational performance in hierarchies and communities of practice. *Journal of Economic Behavior and Organization*, 61:668–690.

Dutta, B., Ghosal, S., and Ray, D. (2005). Farsighted network formation. *Journal of Economic Theory*, 122(2):143–164.

Ehrhardt, G., Marsili, M., and Vega-Redondo, F. (2006a). Diffusion and growth in an evolving network. *International Journal of Game Theory*, 34(3):383-397.

Ehrhardt, G. C. M. A., Marsili, M., and Vega-Redondo, F. (2006b). Phenomenological models of socio-economic network dynamics. *Physical Review Letters E*, 74.

Ehrhardt, G., Marsili, M., and Vega-Redondo, F. (2008). Networks emerging in a volatile world. Unpubllished manuscript, European University Institute.

Ellison, G. (2000). Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *Review of Economic Studies*,67:17–45.

Erdös, P. and Gallai, T. (1960). Graphs with prescribed degrees of vertices. *Lapok*, 11:264–274.

Fagiolo, G. (2005). Endogenous neighborhood formation in a local coordination model with negative network externalities. *Journal of Economic Dynamics and Control*, 29(1-2):297–319.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40:35–41.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1:215–239.

Fruchterman, T. and Reingold, E. (1991). Graph drawing by force-directed placement. *Software- Practice and Experience*.

Galeotti, A. and Goyal, S. (2010). The law of the few. *American Economic Review*, forthcoming.

Galeotti, A., Goyal, S., Jackson, M.O., Vega-Redondo, F. and L. Yariv (2010). Network games. *Review of Economic Studies*, 77:218–244.

Goyal, S. and Joshi, S. (2003). Networks of collaboration in oligopoly. *Games and Economic Behavior*, 43:57–85.

Goyal, S., van der Leij, M., and Moraga-Gonzalez, J. L. (2006). Economics: An emerging small world. *Journal of Political Economy*, 114(2):403–412.

Goyal, S. and Vega-Redondo, F. (2005). Network formation and social coordination. *Games and Economic Behavior*, 50:178–207.

Grassi, R., Stefani, S., and Torriero, A. (2007). Some new results on the eigenvector centrality. *The Journal of Mathematical Sociology*, 31(3):237 – 248.

Grewal, R., Lilien, G. L., and Mallapragada, G. (2006). Location, location, location: How network embeddedness affects project success in open source systems. *Management Science*, 52(7):1043–1056.

Grimmett, G. and Stirzaker, D. (2001). *Probability and random processes.* Oxford University Press.

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

Guimerà, R., Díaz-Guilera, A., Vega-Redondo, F., Cabrales, A., and Arenas, A. (2002). Optimal network topologies for local search with congestion. *Physical Review Letters*, 89(24):248701.

Hagberg, A., Swart, P., and Schult, D. (2006). Designing threshold networks with given structural and dynamical properties. *Physical Review E*, 74(5):56116.

Haller, H., Kamphorst, J., and Sarangi, S. (2007). (Non-)existence and scope of nash networks. *Economic Theory*, 31:597–604.

Haller, H. and Sarangi, S. (2005). Nash networks with heterogeneous links. *Mathematical Social Sciences*, 50:181–201.

Haynie, D. (2001). Delinquent peers revisited: Does network structure matter? *American Journal of Sociology*, 106:1013–1057.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, pages 13–30.

Hu, H. and Wang, X. (2009). Disassortative mixing in online social networks. *Europhysics Letters*, 86:18003.

Huberman, Bernardo, A. and Hogg, T. (1995). Communities of practice: Performance and evolution. *Computational and Mathematical Organization Theory*, 1(1):73–92.

Jackson, M. (2007). *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress*, volume I, chapter The economics of social networks, pages 1–56. Cambridge University Press.

Jackson, M. (2008). *Social and Economic Networks.* Princeton University Press.

Jackson, M. O. and Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, 97(3):890–915.

Jackson, M. and Watts, A. (2002a). On the formation of interaction networks in social coordination games. *Games and Economic Behavior*, 41(2):265–291.

Jackson, M. O. and Watts, A. (2002b). The evolution of social and economic networks. *Journal of Economic Theory*, 106(2):265–295.

Jackson, M. O. and Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74.

Kemeny, J. and Snell, J. (1960). *Finite Markov chains.* Princeton: Van Nostrand Reinhold.

Killworth, P. and Bernard, H. (1976). Informant accuracy in social network data. *Human Organization*, 35(3):269–286.

Kim, D.-H. and Motter, A. E. (2007). Ensemble averageability in network spectra. *Physical Review Letters*, 98(24):248701.

Konishi, H. and Ray, D. (2003). Coalition formation as a dynamic process. *Journal of Economic Theory*, 110(1):1–41.

Latora, V. and Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters*, 87(19):198701.

Mahadev, N. and Peled, U. (1995). *Threshold Graphs and Related Topics*. North Holland.

Marsili, M., Vega-Redondo, F., and Slanina, F. (2004). The rise and fall of a networked society: a formal model. *Proceedings of the National Academy of Sciences*, 101:1439–1442.

Mas, D., Vignes, A., and Weisbuch, G. (2007). Networks and syndication strategies: Does a venture capitalist need to be in the center? ERMES-CNRS Universite Pantheon Assas Paris II, Working Paper No. 07-14.

May, R., Levin, S., and Sugihara, G. (2008). Ecology for bankers. *Nature*, 451(21):893–895.

Mehra, A., Kilduff, M., and Brass, D. J. (2001). The social networks of high and low self-monitors: Implications for workplace performance. *Administrative Science Quarterly*, 46(1):121–146.

Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters*, 89(20):208701.

Newman, M. E. J. (2003). Mixing patterns in networks. *Physical Review E*, 67(2):026126.

Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, 101(Suppl 1):5200–5205.

Pastor-Satorras, R., Vázquez, A., and Vespignani, A. (2001). Dynamical and correlation properties of the internet. *Physical Review Letters*, 87:258701.

Perry-Smith, J. and Shalley, C. (2003). The social side of creativity: A static and dynamic social network perspective. *Academy of Management Review*, 28(1):89–106.

Powell, W. W., Koput, K. W., and Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 41(1):116–145.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4):581–603.

Seneta, E. (1973). *Markov Chains*. George Allen & Unwin Ltd.

Seneta, E. (2006). *Non-Negative Matrices and Markov Chains*. Springer.

Serrano, M. A. and Boguñá, M. (2003). Topology of the world trade web. *Phys. Rev. E*, 68(1):015101.

Soramäki, K., Bech, M.L., Arnold, J., Glass, R.J., and Beyeler, W.E. (1960). The topology of interbank payment flows. *Physica A*, 379:317–333.

Stephenson, K. and Zelen, M. (1989). Rethinking centrality: Methods and examples. *Social Networks*, 11:1–37.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, 61:674–698..

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, 42(1):35–67.

Vega-Redondo, F. (2006). Building up social capital in a changing world. *Journal of Economic Dynamics and Control*, 30(11):2305–2338.

Vega-Redondo, F. (2007). *Complex Social Networks*. Cambridge: Cambridge University Press.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Watts, A. (2001). A dynamic model of network formation. *Games and*

Economic Behavior, 34:331–341.
Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393:440–442.
Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33(4):452–473.

Appendix

A. Network Definitions and Characterizations (NOT FOR PUBLICATION)

A *network* (*graph*) G is the pair (N, L) consisting of a set of agents (*vertices* or *nodes*) $N = \{1, ..., n\}$ and a set of links L (*edges*) between them. A link ij is *incident* with the vertex $v \in N$ in the network G whenever $i = v$ or $j = v$. There exists a link between vertices i and j such that $a_{ij} = 1$ if $ij \in L$ and $a_{ij} = 0$ if $ij \notin L$. The *neighborhood* of an agent $i \in N$ is the set $\mathcal{N}_i = \{j \in N : ij \in L\}$. The *degree* d_i of an agent $i \in N$ gives the number of links incident to agent i. Clearly, $d_i = |\mathcal{N}_i|$. Let $\mathcal{N}_i^{(2)} = \bigcup_{j \in \mathcal{N}_i} \mathcal{N}_j \backslash (\mathcal{N}_i \cup \{i\})$ denote the second-order neighbors of agent i. Similarly, the k-th order neighborhood of agent i is defined recursively from $\mathcal{N}_i^{(0)} = i$, $\mathcal{N}_i^{(1)} = \mathcal{N}_i$ and

$$\mathcal{N}_i^{(k)} = \bigcup_{j \in \mathcal{N}_i^{(k-1)}} \mathcal{N}_j \backslash \left(\bigcup_{l=0}^{k-1} \mathcal{N}_i^{(l)} \right).$$

A *walk* in G of length k from i to j is a sequence $p = \langle i_0, i_1, ..., i_k \rangle$ of agents such that $i_0 = i$, $i_k = j$, $i_p \neq i_{p+1}$, and i_p and i_{p+1} are directly linked, for all $0 \leq p \leq k-1$. Agents i and j are said to be *indirectly linked* in G if there exists a walk from i to j in G. An agent $i \in N$ is *isolated* in G if $a_{ij} = 0$ for all j. The network G is said to be *empty* when all its agents are isolated.

A *subgraph*, G', of G is the graph of subsets of the agents, $N(G') \subseteq N(G)$, and links, $L(G') \subseteq L(G)$. A graph G is *connected*, if there is a path connecting every pair of agents. Otherwise G is disconnected. The *components* of a graph G are the maximally connected subgraphs. A component is said to be *minimally connected* if the removal of any link makes the component disconnected.

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

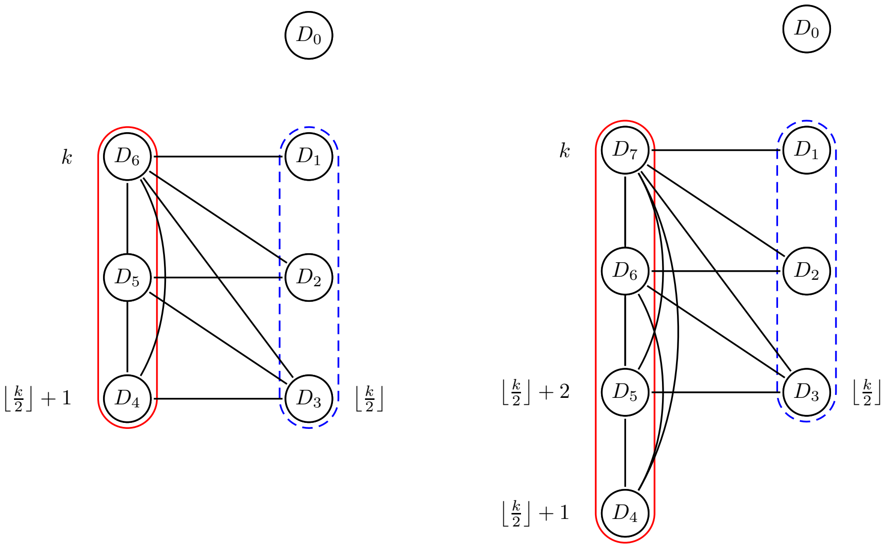
In a *complete* graph K_n, every agent is adjacent to every other agent. The graph in which no pair of agents is adjacent is the empty graph $\overline{K}_n$. A *clique* $K_{n'}$, $n' \leq n$, is a complete subgraph of the network G. A graph is *k-regular* if every agent i has the same number of links $d_i = k$ for all $i \in N$. The complete graph K_n is $(n-1)$-regular. The cycle C_n is 2-regular. In a *bipartite graph* there exists a partition of the agents in two disjoint sets V_1 and V_2 such that each link connects an agent in V_1 to an agent in V_2. V_1 and V_2 are independent sets with cardinalities n_1 and n_2, respectively. In a complete bipartite graph K_{n_1,n_2} each agent in V_1 is connected to each other agent in V_2. The *star* $K_{1,n-1}$ is a complete bipartite graph in which $n_1 = 1$ and $n_2 = n - 1$.

The *complement* of a graph G is a graph $\bar{G}$ with the same nodes as G such that any two nodes of $\bar{G}$ are adjacent if and only if they are not adjacent in G. For example the complement of the complete graph K_n is the empty graph $\bar{K}_n$.

Let $\mathbf{A}$ be the symmetric $n \times n$ *adjacency matrix* of the network G. The element $a_{ij} \in \{0, 1\}$ indicates if there exists a link between agents i and j such that $a_{ij} = 1$ if $ij \in L$ and $a_{ij} = 0$ if $ij \notin L$. The k-th power of the adjacency matrix is related to walks of length k in the graph. In particular, $\left(\mathbf{A}^k\right)_{ij}$ gives the number of walks of length k from agent i to agent j. The *eigenvalues* of the adjacency matrix $\mathbf{A}$ are the numbers $\lambda_1, \lambda_2, ..., \lambda_n$ such that $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ has a nonzero solution vector $\mathbf{v}_i$, which is an *eigenvector* associated with λ_i for $i = 1, ..., n$. Since the adjacency matrix $\mathbf{A}$ of an undirected graph G is real and symmetric, the eigenvalues of $\mathbf{A}$ are real, $\lambda_i \in \mathbb{R}$ for all $i = 1, ..., n$. Moreover, if $\mathbf{v}_i$ and $\mathbf{v}_j$ are eigenvectors for different eigenvalues, $\lambda_i \neq \lambda_j$, then $\mathbf{v_i}$ and $\mathbf{v}_j$ are orthogonal, i.e. $\mathbf{v}_i^T \mathbf{v}_j = 0$ if $i \neq j$. In particular, $\mathbb{R}^n$ has an orthonormal basis consisting of eigenvectors of $\mathbf{A}$. Since $\mathbf{A}$ is a real symmetric matrix, there exists an orthogonal matrix $\mathbf{S}$ such that $\mathbf{S}^T\mathbf{S} = \mathbf{S}\mathbf{S}^T = \mathbf{I}$ (that is $\mathbf{S}^T = \mathbf{S}^{-1}$) and $\mathbf{S}^T\mathbf{A}\mathbf{S} = \mathbf{D}$, where $\mathbf{D}$ is the diagonal matrix of eigenvalues of $\mathbf{A}$ and the columns of $\mathbf{S}$ are the corresponding eigenvectors. The *Perron-Frobenius eigenvalue* $\lambda_{\mathrm{PF}}(G)$ is the *largest real eigenvalue* of $\mathbf{A}$ associated with G, i.e. all eigenvalues λ_i of $\mathbf{A}$ satisfy $|\lambda_i| \leq \lambda_{\mathrm{PF}}(G)$ for $i = 1, ..., n$ and there exists an associated nonnegative eigenvector $\mathbf{v}_{\mathrm{PF}} \geq 0$ such that $\mathbf{A}\mathbf{v}_{\mathrm{PF}} = \lambda_{\mathrm{PF}}(G)\mathbf{v}_{\mathrm{PF}}$. For a connected graph G the adjacency matrix $\mathbf{A}$ has a unique largest real eigenvalue $\lambda_{\mathrm{PF}}(G)$ and a positive associated eigenvector $\mathbf{v}_{\mathrm{PF}} > 0$. There exists a relation between the number of walks in a graph and its eigenvalues. The number of closed walks of length k from a agent i in G to herself is given by $\left(\mathbf{A}^k\right)_{ii}$ and the total number of closed walks of length k in G is $\mathrm{tr}\left(\mathbf{A}^k\right) = \sum_{i=1}^n \left(\mathbf{A}^k\right)_{ii} = \sum_{i=1}^n \lambda_i^k$. We further have that $\mathrm{tr}\left(\mathbf{A}\right) = 0$, $\mathrm{tr}\left(\mathbf{A}^2\right)$ gives twice the number of links in G and $\mathrm{tr}\left(\mathbf{A}^3\right)$ gives six times the number of triangles in G.

B. Topological Properties of Nested Split Graphs (NOT FOR PUBLICATION)

In this Appendix we discuss in more detail the topological properties of nested split graphs that arise from our network formation process. We first derive several network statistics for nested split graphs. We compute the degree distribution, the clustering coefficient, average nearest neighbor neighbor connectivity and the characteristic path length in a nested split graph. In particular, we show that connected nested split graphs have small characteristic path length, which is at most two. We then analyze different measures of centrality in a nested split graph.[34] From the expressions of these centrality measures we then can show that degree, closeness, eigenvector 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. As we elaborate in more detail in Section 7 this has important implications for the generality of our model.



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

Figure 11: Representation of nested split graphs and their degree partitions $\mathbf{D}$ (top) with corresponding adjacency matrices $\mathbf{A}$ (bottom). A line between D_i and D_j indicates that every node in D_i is adjacent to every node in D_j. The partitions included in the solid frame (D_i with $\lfloor \frac{k}{2} \rfloor + 1 \leq i \leq k$) are the dominating subsets while the partitions in the dashed frame (D_i with $1 \leq i \leq \lfloor \frac{k}{2} \rfloor$) are the independent sets. The figure at the top left considers the case of $k = 6$ (even) and the figure at the top right the case of $k = 7$ (odd). The illustration follows Mahadev and Peled [1995, p. 11].

[34]See Wasserman and Faust [1994] for an overview of different measures of centrality.

B.1. Network Statistics

In the following sections we will compute the degree connectivity, the clustering coefficient, assortativity and average nearest nearest neighbor connectivity and the characteristic path length in a nested split graph G as a function of the degree partition $\mathbf{D}$ (introduced in Definition 4).

B.1.1. Degree Connectivity

The nested neighborhood structure of a nested split graph allows us to compute the degrees of the nodes according to a recursive equation that is stated in the next corollary.

Corollary 3. *Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, ..., D_k)$ be the degree partition of G. Then $d_u = 0$ if $u \in D_0$ and for each $u \in D_i$, $v \in D_{i-1}$, $i = 1, ..., k$, we get*

$$d_u = \begin{cases} d_v + |D_{k-i+1}|, & \text{if } i \neq \left\lfloor \frac{k}{2} \right\rfloor + 1, \\ d_v + |D_{k-i+1}| - 1, & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \end{cases} \tag{14}$$

or equivalently

$$d_u = \begin{cases} \sum_{j=1}^{i} |D_{k+1-j}|, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \sum_{j=1}^{i} |D_{k+1-j}| - 1, & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases} \tag{15}$$

Equation (14) shows that the neighborhoods of the agents in a nested split graph are nested (see also Definition 5). The degrees of the agents in ascending order of the graph in Figure 11, top left, are $2, 3, 4, 5, 7, 9$ while in the graph in Figure 11, top right, they are $1, 2, 3, 4, 7, 8, 9$.

B.1.2. Clustering Coefficient

The clustering coefficient $\mathcal{C}(u)$ for an agent u is the proportion of links between the agents within her neighborhood $\mathcal{N}_u$ divided by the number of links that could possibly exist between them [Watts and Strogatz, 1998]. It is given by

$$\mathcal{C}(u) = \frac{|\{vw : v, w \in \mathcal{N}_u \wedge vw \in L\}|}{d_u(d_u - 1)/2}. \tag{16}$$

In a nested split graph the clustering coefficient can be derived from the degree partition, as the following corollary shows.

Corollary 4. *Consider a nested split graph $G = (N, L)$ 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*

$u \in D_i$, $i = 0, ..., k$, and $d_u \geq 2$, the clustering coefficient is given by

$$\mathcal{C}(u) = \begin{cases} 0, & \text{if } i = 0, \\ 1, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{1}{d_u(d_u-1)} \left(S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 1 \right) \left[\left(S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 2 \right) + \right. & \\ \left. 2|D_{\lfloor \frac{k}{2} \rfloor}| \right], & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \text{ even}, \\ \frac{1}{d_u(d_u-1)} \left(S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 1 \right) \left(S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 2 \right), & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \text{ odd}, \\ \frac{1}{d_u(d_u-1)} \left[\left(S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 1 \right) \left(S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 2 \right) + \right. & \\ \left. 2 \sum_{j=k-i+1}^{\lfloor \frac{k}{2} \rfloor} |D_j| \left(S_D^{k-j+1} - 1 \right) \right], & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 2 < i \leq k, \end{cases} \tag{17}$$

where d_u is given by Equation (15).

Proof of Corollary 4. Note that for all agents in the independent sets, $u \in D_i$ with $1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor$, the clustering coefficient is one, since their neighbors are all connected among each other. Next, we consider the agents $u \in D_i$ with $\left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k$ and degree $d_u = \sum_{j=1}^{i} |D_{k+1-j}| - 1$. The neighbors of agent u in the dominating subsets are all connected among each other with a total of $\frac{1}{2} \left(\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{k} |D_j| - 1 \right) \left(\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{k} |D_j| - 2 \right)$ links, excluding agent u from the dominating subset. The neighbors of u in the independent sets are not connected. Finally, we consider the links between neighbors for which one neighbor is in a dominating subset and one neighbor is in an independent set. If k is even we get $\sum_{j=k-i+1}^{\lfloor \frac{k}{2} \rfloor} |D_j| \left(\sum_{l=k-j+1}^{k} |D_l| - 1 \right)$ links, excluding agent u in the dominating subset (see Figure 11 (left)). If k is odd there is no such contribution for the agents in the set $D_{\lfloor \frac{k}{2} \rfloor + 1}$ (see Figure 11 (right)). Putting these contributions together we obtain the clustering coefficient of an agent $u \in D_i$ for all $i = 1, ..., k$, as given by Equation (17). □

The total clustering coefficient is the average of the clustering coefficients over all agents,

$$\mathcal{C} = \frac{1}{n} \sum_{u \in N} \mathcal{C}(u). \tag{18}$$

The clustering coefficients of the agents in ascending order of the graph in Figure 11, top left, are $5/12, 5/12, 13/21, 9/10, 1, 1, 1, 1, 1, 1$, with a total clustering coefficient of $\mathcal{C} = 0.84$. In the graph in Figure 11, top right, it is $13/36, 13/28, 4/7, 1, 1, 1, 1, 1, 1$, with a total clustering of $\mathcal{C} = 0.74$.

B.1.3. Assortativity and Nearest Neighbor Connectivity

There exists a measure of degree correlation called "average nearest neighbor connectivity" [Pastor-Satorras et al., 2001]. More precisely, the average nearest neighbor connectivity $d_{nn}(u)$ is the average degree of the neighbors of an agent with degree d_u. It is defined by

$$d_{nn}(u) = \frac{1}{d_u} \sum_{v \in \mathcal{N}_u} d_v. \tag{19}$$

In a nested split graph the average nearest neighbor connectivity is determined by its degree partition.

Corollary 5. *Consider a nested split graph $G = (N, L)$ 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 $u \in D_i$, $i = 0, ..., k$,*

$$d_{nn}(u) = \begin{cases} 0, & \text{if } i = 0, \\ \frac{1}{S_D^i} \sum_{j=1}^{i} |D_{k+1-j}| \left(S_D^{k+1-j} - 1 \right), & \text{if } i = 1, ..., \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{1}{S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 1} \left[\sum_{j=\lfloor \frac{k}{2} \rfloor + 2}^{k} |D_j| \left(S_D^j - 1 \right) \right. & \\ \left. + \left(|D_{\lfloor \frac{k}{2} \rfloor + 1}| - 1 \right) \left(S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 1 \right) + |D_{\lfloor \frac{k}{2} \rfloor}| S_D^{\lfloor \frac{k}{2} \rfloor} \right], & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \text{ even}, \\ \frac{1}{S_D^{\lfloor \frac{k}{2} \rfloor + 1} - 1} \left[\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{k} |D_j| \left(S_D^j - 1 \right) \right] - 1, & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \text{ odd}, \\ \frac{1}{S_D^i - 1} \left[\sum_{j=\lfloor \frac{k}{2} \rfloor + 1}^{k} |D_j| \left(S_D^j - 1 \right) \right. & \\ \left. + \sum_{j=k-i+1}^{\lfloor \frac{k}{2} \rfloor} |D_j| S_D^j \right] - 1, & \text{if } i = \left\lfloor \frac{k}{2} \right\rfloor + 2, ..., k \end{cases} \tag{20}$$

Proof of Corollary 5. First, consider an agent $u \in D_i$ with $i = 1, ..., \left\lfloor \frac{k}{2} \right\rfloor$ corresponding to the independent sets. We know that the number of neighbors (degree) of agent u is given by $\sum_{j=1}^{i} |D_{k+1-j}|$. The neighbors of agent u are the agents in the dominating subsets with degrees given in Equation (15). Thus, the number of neighbors of the neighbors of u in the sets D_{k+1-j} is $\sum_{l=1}^{k+1-j} |D_{k+1-l}| - 1$. Putting the above results together, we obtain for the average nearest neighbor connectivity of agent $u \in D_i$, $i = 1, ..., \left\lfloor \frac{k}{2} \right\rfloor$, the following expression.

$$d_{nn}(u) = \frac{1}{\sum_{j=1}^{i} |D_{k+1-j}|} \sum_{j=1}^{i} |D_{k+1-j}| \left(\sum_{l=1}^{k+1-j} |D_{k+1-l}| - 1 \right). \tag{21}$$

Next, we consider an agent u in the set D_i with $\left\lfloor \frac{k}{2} \right\rfloor + 2 \leq i \leq k$ corresponding to the dominating subsets. The number of neighbors of agent u is given by

$\sum_{j=1}^{i} |D_{k+1-j}| - 1$. The number of neighbors of an agent $v \in D_j$, $\lfloor \frac{k}{2} \rfloor + 1 \leq j \leq k$ in the dominating subsets is given by $\sum_{l=1}^{j} |D_{k+1-l}| - 1$. Since agent u is connected to all other agents in the dominating subsets, we can sum over all their neighborhoods with a total of $\sum_{j=\lfloor \frac{k}{2} \rfloor+1}^{k} |D_j| \left(\sum_{l=1}^{j} |D_{k+1-l}| - 1 \right)$ neighbors. Note however, that we have to subtract agent u herself from this sum. Morover, the number of neighbors of an agent $w \in D_j$, $1 \leq j \leq \lfloor \frac{k}{2} \rfloor$ in the independent sets is given by $\sum_{l=1}^{j} |D_{k+1-l}|$. Thus, the average nearest neighbor connectivity of agent $u \in D_i$, $\lfloor \frac{k}{2} \rfloor + 2 \leq i \leq k$, is given by

$$\begin{aligned} d_{nn}(u) &= \tfrac{1}{\sum_{j=1}^{i} |D_{k+1-j}| - 1} \left[\sum_{j=\lfloor \frac{k}{2} \rfloor+1}^{k} |D_j| \left(\sum_{l=1}^{j} |D_{k+1-l}| - 1 \right) \right. \\ &\quad \left. + \sum_{j=k-i+1}^{\lfloor \frac{k}{2} \rfloor} |D_j| \sum_{l=1}^{j} |D_{k+1-l}| \right] - 1. \end{aligned} \tag{22}$$

In a similar way we can consider the cases $i = \lfloor \frac{k}{2} \rfloor + 1$ for both k even and k odd. □

When the average nearest neighbor connectivity is a monotonic increasing function of the degree d, then the network is assortative, while, if it is monotonic decreasing with d, it is dissortative [Newman, 2002; Pastor-Satorras et al., 2001]. Nested split graphs are dissortative, since for $i < j$ and $d_u \in D_i < d_v \in D_j$ it follows that $d_{nn}(u) > d_{nn}(v)$. This is because the higher is the degree of an agent in a dominating subset, the more neighbors she has from the independent sets with low degrees, which decreases her average nearest neighbor connectivity. For example, the average nearest neighbor connectivities of the agents in the graph in Figure 11, top left, in ascending order are $13/3, 13/3, 37/7, 33/5, 15/2, 15/2, 25/3, 25/3, 9, 9$ while in the graph in Figure 11, top right, they are $35/9, 35/8, 34/7, 7, 7, 8, 8, 8, 17/2, 9$.

B.1.4. Characteristic Path Length

The characteristic path length is defined as the number of links in the shortest path between two agents, averaged over all pairs of agents [Watts and Strogatz, 1998]. This can be written as

$$\mathcal{L} = \frac{1}{n(n-1)/2} \sum_{u \neq v} d(u, v), \tag{23}$$

where $d(u, v)$ is the geodesic (shortest path) between agent u and agent v in $N \backslash D_0$.[35] Then the characteristic path length in a nested split graph is given by the following corollary.

[35]Note that we do not consider the isolated agents in the set D_0 because the characteristic path length $\mathcal{L}$ is not defined for disconnected networks.

Corollary 6. *Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, ..., D_k)$ be the degree partition of G. Then the characteristic path length of G is given by*

$$\begin{aligned}\mathcal{L} &= \tfrac{1}{n(n-1)/2}\left[\tfrac{1}{2}\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k}|D_j|\left(\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k}|D_j|-1\right)+\right.\\ &\sum_{j=1}^{\lfloor\frac{k}{2}\rfloor}|D_j|\left(\sum_{j=1}^{\lfloor\frac{k}{2}\rfloor}|D_j|-1\right)+\\ &\left.\sum_{l=1}^{\lfloor\frac{k}{2}\rfloor}|D_l|\left(\sum_{j=k-l+1}^{k}|D_j|+2\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k-l}|D_j|\right)\right].\end{aligned} \tag{24}$$

Proof of Corollary 6. We first consider all pairs of agents in the dominating subsets. All theses agents are adjacent to each other and thus the shortest path between them has length one. Moreover, there are $\frac{1}{2}\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k}|D_j|\left(\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k}|D_j|-1\right)$ pairs of agents in the dominating subsets.

Next, we consider all pairs of agents in the independent sets. From Equation (30) we know that all of them are at a distance of two links separated from each other. Moreover, there are $\frac{1}{2}\sum_{j=1}^{\lfloor\frac{k}{2}\rfloor}|D_j|\left(\sum_{j=1}^{\lfloor\frac{k}{2}\rfloor}|D_j|-1\right)$ pairs of agents in which both agents stem from an independent set.

Finally, we consider the pairs of agents in which one agent is in the independent set D_1 and the other in a dominating subset. Then there are $|D_1||D_k|$ pairs of agents with shortest path 1 and $|D_1|\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k-1}|D_j|$ pairs of agents with shortest path 2. Similarly, we can consider the pairs in which one agent is in the set D_2. Then we have $|D_2|(|D_k|+|D_{k-1}|)$ pairs of agents with shortest path 1 and $|D_2|\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k-2}|D_j|$ pairs of agents with shortest path 2. Finally, if one agent is in the set $D_{\lfloor\frac{k}{2}\rfloor}$ then we have $|D_{\lfloor\frac{k}{2}\rfloor}|\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k}|D_j|$ pairs of agents with distance 1 and none with distance 2, if k is even (see Figure 11, top left). If k is odd (see Figure 11, top right), and we have one agent is in the set $D_{\lfloor\frac{k}{2}\rfloor}$ then we have $|D_{\lfloor\frac{k}{2}\rfloor}|\sum_{j=\lfloor\frac{k}{2}\rfloor+2}^{k}|D_j|$ pairs of agents with distance 1 and $|D_{\lfloor\frac{k}{2}\rfloor}||D_{\lfloor\frac{k}{2}\rfloor+1}|$ pairs with distance 2.

Therefore, the average path length $\mathcal{L}$ defined in Equation (23) is given by the following equation

$$\begin{aligned}\tfrac{n(n-1)}{2}\mathcal{L} &= \tfrac{1}{2}\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k}|D_j|\left(\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k}|D_j|-1\right)+\\ &2\tfrac{1}{2}\sum_{j=1}^{\lfloor\frac{k}{2}\rfloor}|D_j|\left(\sum_{j=1}^{\lfloor\frac{k}{2}\rfloor}|D_j|-1\right)+\\ &\sum_{l=1}^{\lfloor\frac{k}{2}\rfloor}|D_l|\left[\sum_{j=k-l+1}^{k}|D_j|+2\sum_{j=\lfloor\frac{k}{2}\rfloor+1}^{k-l}|D_j|\right].\end{aligned} \tag{25}$$

□

Considering the graph in Figure 11, top left, the characteristic path length is $\mathcal{L} = 22/15$ while in the graph in Figure 11, top right, we get

$\mathcal{L} = 68/45$.

Taking the inverse of the shortest path length one can introduce a related measurement, the network efficiency[36] $\mathcal{E}$, that is also applicable to disconnected networks [Latora and Marchiori, 2001]

$$\mathcal{E} = \frac{1}{n(n-1)} \sum_{u \neq v} \frac{1}{d(u,v)}. \tag{26}$$

Finally, we find that in a connected nested split graph agents are at most two links separated from each other and thus these graphs are characterized by a short characteristic path length.

B.2. Centrality

In the next sections we analyze different measures of centrality in a nested split graph G. We derive the expressions for degree, closeness and betweenness centrality as a function of the degree partition of G. Finally, we show that these measures are similar in the sense that they induce the same ordering of the nodes in G based on their centrality values.

B.2.1. Degree Centrality

The degree centrality of an agent $u \in N$ is given by the proportion of agents that are adjacent to u [Wasserman and Faust, 1994]. We obtain the normalized degree centrality simply by dividing the degree of agent u with the maximum degree $n-1$. This yields the following corollary.

Corollary 7. *Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, ..., D_k)$ be the degree partition of G. Then for each $u \in D_i$, $i = 0, ..., k$, the degree centrality is given by*

$$\mathcal{C}_d(u) = \begin{cases} \frac{1}{n-1} \sum_{j=1}^{i} |D_{k+1-j}|, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{1}{n-1} \left(\sum_{j=1}^{i} |D_{k+1-j}| - 1 \right), & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases} \tag{27}$$

Proof of Corollary 7. The result follows directly from Corollary 3. □

We observe that degree centrality as well as the degree are increasing with increasing index i of the set D_i to which agent u belongs. Degree centralities for the graphs shown in Figure 11 can be derived from the degrees given in Section B.1.1 by dividing the degrees with $n-1$.

[36] The network efficiency must not be confused with the efficiency of a network. The first is related to short paths in the network while the latter measures social welfare, that is, the efficient network maximizes aggregate payoff.

B.2.2. Closeness Centrality

Excluding the isolated nodes in G, closeness centrality of agent $u \in N \setminus D_0$ is defined as [Beauchamp, 1965; Sabidussi, 1966]:

$$\mathcal{C}_c(u) = \frac{n-1}{\sum_{v \neq u} d(u,v)}. \tag{28}$$

where $d(u,v)$ measures the shortest path between agent u and agent v in $N \setminus D_0$. For a nested split graph we obtain the following corollary.

Corollary 8. *Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, ..., D_k)$ be the degree partition of G. Then for each $u \in D_i$, $i = 0, ..., k$, the closeness centrality is given by*

$$\mathcal{C}_c(u) = \begin{cases} \frac{n-1}{\sum_{j=k-i+1}^{k} |D_j| + 2\sum_{j=1}^{k-i} |D_j| - 2}, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{n-1}{\sum_{j=k-i+1}^{k} |D_j| + 2\sum_{j=1}^{k-i} |D_j| - 1}, & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases} \tag{29}$$

Proof of Corollary 8. For both agents in the independent sets, $u \in D_i$ with $1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor$, and in the dominating subsets, $u \in D_i$ with $\left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k$, we can compute the length of the shortest paths as follows:

$$d(u,v) = \begin{cases} 1 & \text{for all } v \in \bigcup_{j=k-i+1}^{k} D_j, \\ 2 & \text{for all } v \in \bigcup_{j=1}^{k-i} D_j. \end{cases} \tag{30}$$

In order to compute the closeness centrality we have to consider all pairs of agents in the graph and compute the length of the shortest path between them, which is given in Equation (30). We obtain for any agent $u \in D_i$, $i = 1, ..., k$, the following expression

$$\mathcal{C}_c(u) = \begin{cases} \frac{n-1}{\sum_{j=k-i+1}^{k} |D_j| + 2\sum_{j=1}^{k-i} |D_j| - 2}, & \text{if } 1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor \\ \frac{n-1}{\sum_{j=k-i+1}^{k} |D_j| + 2\sum_{j=1}^{k-i} |D_j| - 1}, & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 1 \leq i \leq k. \end{cases} \tag{31}$$

Note that we have subtracted 1 and 2 in the denominator, respectively, since the sums would otherwise include the contribution of agent u herself. □

We have that closeness centrality is identical for all agents in the same set. Also note that $\mathcal{C}_c(u) = 1$ for $u \in D_k$. Moreover, closeness centrality is increasing with increasing degree. The closeness centralities of the agents in descending order for the graph in Figure 11, top left, are $1, 1, 9/11, 9/13, 9/14, 9/14, 9/15, 9/15, 9/16, 9/16$ while in the graph in Figure 11, top right, they are $1, 9/10, 9/11, 9/14, 9/14, 9/15, 9/15, 9/15, 9/16, 9/17$.

B.2.3. Betweenness Centrality

Betweenness centrality is defined as [Freeman, 1977]

$$\mathcal{C}_b(u) = \sum_{u \neq v \neq w} \frac{g(v,u,w)}{g(v,w)}, \tag{32}$$

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

The betweenness centrality for a nested split graph can be derived from its degree partition as follows.[37]

Corollary 9. *Consider a nested split graph $G = (N, L)$ and let $\mathbf{D} = (D_0, D_1, ..., D_k)$ be the degree partition of G. Then $\mathcal{C}_b(u) = 0$ if $u \in D_i$, $i = 0, ..., \left\lfloor \frac{k}{2} \right\rfloor$ and for each $u \in D_i$, $v \in D_{i-1}$, $i = \left\lfloor \frac{k}{2} \right\rfloor + 1, ..., k$, the betweenness centrality is given by*

$$\mathcal{C}_b(u) = \begin{cases} 0 & \text{if } , i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \quad \text{odd} \\ \frac{|D_{\left\lfloor \frac{k}{2} \right\rfloor}|\left(|D_{\left\lfloor \frac{k}{2} \right\rfloor}|-1\right)}{\sum_{j=\left\lfloor \frac{k}{2} \right\rfloor+1}^{k} |D_j|}, & \text{if } , i = \left\lfloor \frac{k}{2} \right\rfloor + 1, \quad k \quad \text{even} \\ \mathcal{C}_b(v) + \frac{|D_{k-i+1}|(|D_{k-i+1}|-1)}{\sum_{j=i}^{k} |D_j|} & \\ + \frac{2|D_{k-i+1}|\sum_{j=k-i+2}^{i-1} |D_j|}{\sum_{j=i}^{k} |D_j|}, & \text{if } \left\lfloor \frac{k}{2} \right\rfloor + 2 \leq i \leq k. \end{cases} \tag{33}$$

Proof of Corollary 9. In this proof, we follow closely Hagberg et al. [2006]. The agents in the independent sets D_i, $0 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor$ do not lie on any shortest path between two other agents in the network and thus their betweenness centrality vanishes. For the agents in the dominating subsets we have that the betweenness centrality of the agent $u \in D_{\left\lfloor \frac{k}{2} \right\rfloor+1}$ vanishes if k is odd and is given by $|D_{\left\lfloor \frac{k}{2} \right\rfloor}|\left(|D_{\left\lfloor \frac{k}{2} \right\rfloor}|-1\right) / \sum_{j=\left\lfloor \frac{k}{2} \right\rfloor+1}^{k} |D_j|$ if k is even. The latter result is due the shortest path between agents that are both in $D_{\left\lfloor \frac{k}{2} \right\rfloor}$. Next, consider an agent $u \in D_i$ and $v \in D_{i-1}$, with $\left\lfloor \frac{k}{2} \right\rfloor + 2 \leq i \leq k$. Then the betweenness centrality of agent u is given by the following recursive relationship

$$\mathcal{C}_b(v) + \frac{|D_{k-i+1}|\left(|D_{k-i+1}|-1\right)}{\sum_{j=i}^{k} |D_j|} + \frac{2|D_{k-i+1}|\sum_{j=k-i+2}^{i-1} |D_j|}{\sum_{j=i}^{k} |D_j|}. \tag{34}$$

The first term in Equation (34) is due to the fact that all shortest paths through lower dominating nodes $v \in D_{i-1}$ have the same length as through

[37] A similar result can be found in Hagberg et al. [2006].

$u \in D_i$. The second term in Equation (34) represents the contribution of paths between nodes in D_{k-i+1}, divided by the number of shortest path passing through the agents in the dominating subsets D_j, $i \leq j \leq k$. The third term in Equation (34) represents all path between an agent in D_{k-i+1} and the other being in D_j, $k-i+2 \leq j \leq i-1$, divided by the number of shortest path passing through the agents in the dominating subsets D_j, $i \leq j \leq k$. □

From Corollary 9, we find that the agents in the independent sets D_i with $1 \leq i \leq \left\lfloor \frac{k}{2} \right\rfloor$ have vanishing betweenness centrality. From the above equation we also observe that the betweenness centrality is increasing with degree such that the agents in D_k have the highest betweenness centrality, the agents in D_{k-1} the second highest betweenness centrality and so on. Thus, the ordering of betweenness centralities follows the degree ordering for all agents in the dominating subsets while the agents in the independent sets have vanishing betweenness centrality. For the betweenness centralities of the agents in the graph in Figure 11, top left, we obtain in descending order $109/6, 109/6, 31/6, 1/2, 0, 0, 0, 0, 0, 0$ while in the graph in Figure 11, top right, they are $28, 12, 6, 0, 0, 0, 0, 0, 0, 0$.

B.2.4. Eigenvector Centrality

There is a central property that holds for nested split graphs in relation to Bonacich centrality, namely that the agents with higher degree also have higher Bonacich centrality. Similar to part (i) of Proposition 1 we can give the following corollary.[38]

Corollary 10. *Let* $\mathbf{v}$ *be the eigenvector associated with the largest real eigenvalue* $\lambda_{PF}(G)$ *of the adjacency matrix* $\mathbf{A}$ *of a nested split graph* $G = (N, L)$. *For each* $i = 1, ..., n$, v_i *is the eigenvector centrality of agent* i. *Consider a pair of agents* $i, j \in N$. *If and only if agent* i *has a higher degree than agent* j *then* i *has a higher eigenvector centrality than* j, *i.e.*

$$d_i > d_j \Leftrightarrow v_i > v_j.$$

PROOF OF COROLLARY 10. The proof is identical to the proof of part (i) of Proposition 1. □

B.2.5. Centrality Rankings

Putting together the results for different centrality measures derived in the previous sections, we can make the following observation of the rankings of agents for different centrality measures in a nested split graph.

[38] A similar result can be found in [Grassi et al., 2007].

Corollary 11. *Consider a nested split graph $G = (N, L)$. Let $\mathcal{C}_d$, $\mathcal{C}_c$, $\mathcal{C}_b$, $\mathcal{C}_v$ denote the degree, closeness, betweenness and eigenvector centrality in G. Then for any $l, m \in \{d, c, v\}$, $l \neq m$ and $i, j \in N$ we have that*

$$\mathcal{C}_l(i) \geq \mathcal{C}_l(j) \Leftrightarrow \mathcal{C}_m(i) \geq \mathcal{C}_m(j), \tag{35}$$

and

$$\mathcal{C}_l(i) \geq \mathcal{C}_l(j) \Rightarrow \mathcal{C}_b(i) \geq \mathcal{C}_b(j). \tag{36}$$

Proof of Corollary 11. The proof is a direct application of Corollaries 7, 8 9 and Proposition 1. □

If and only if an agent i has the k-th highest degree centrality then i is the agent with the k-th highest closeness and eigenvector centrality. This result also holds for Bonacich centrality (see Proposition 1). Moreover, if an agent i has the k-th highest degree centrality then she also has the k-th highest betweenness centrality and this also holds for closeness, eigenvector and Bonacich centrality, respectively. The ordering induced by degree, closeness eigenvector and Bonacich centrality coincide and these orderings also apply in a weak sense for betweenness centrality. We discuss in Section 7 that this allows us to generalize our model to various other centrality measures beyond Bonacich centrality.

C. Capacity Constraints in the Absence of Link Removal

Consider the network formation process of Definition 3 with link creation probability $\alpha = 1$. This means that there is no link decay. However, agents face constraints in the number of links they can maintain. Such constraints can arise from a possible information overload and congestion [Arenas et al., 2010; Dodds et al., 2003; Fagiolo, 2005; Guimerà et al., 2002; Huberman and Hogg, 1995]. Let $\mathbf{C} = (C_1, ..., C_n)$ be the vector of capacity constraints of the agents.[39] If an agent i, $1 \leq i \leq n$ has degree $d_i = C_i$, i.e. she has reached her maximum number of links (capacity constraint), then she stops accepting any additional links. We can then state the following proposition.

Proposition 9. *Consider the network formation process $(G(t))_{t=0}^{\infty}$ in Definition 3 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*

[39] We assume that the sequence $\mathbf{C}$ is a proper degree sequence of a simple graph. This is satisfied if $\sum_{i=1}^{n} C_i$ is even and $\sum_{i=1}^{k} d_i \leq k(k-1) + \sum_{i=k+1}^{n} \min(d_i, k)$ for $k \in \{1, \ldots, n\}$ [Erdös and Gallai, 1960].

split graph G_C with degree partition $\mathbf{D} = \mathbf{C}$, i.e. $\mu_G = 1$ if $G = G_C$ and zero otherwise.

Proof of Proposition 9. Note that Proposition 2 holds also for the case of $\alpha = 1$, when no link is removed. We therefore know that $G(t)$ is a nested split graph for any $t \geq 0$. Consider the nested split graph G_C with degree partition $\mathbf{D} = \mathbf{C}$. Under the link formation process from Definition 3 for $\alpha = 1$ and capacity constraint $\mathbf{C}$, G_C is stationary. For contradiction assume that there is another nested split graph G' which is stationary. By definition of the capacity constraint, G' cannot have more links than G_C. On the other hand, if G' has less links than G_C then there exist at least two agents who can create a link between each other so that their payoff increases. Thus, there exists a positive probability that this link is created and we call the resulting network G''. If G'' is not identical to G_C then we can again find a pair of agents that can create a link, and so on. If G'' is identical to G_C then the network formation process stays in G_C forever. Therefore, G' cannot be stationary, but G_C is. For a given number of links G_c is unique (up to a permutation of nodes) and the claim follows. □

The above proposition shows that nested split graphs can also arise in the absence of the particular link removal mechanism introduced in Definition 3 by assuming that agents have some limitation in the number of interactions they can maintain.

D. Introducing the cost of forming a link

Consider the network formation process of Definition 3 with one modification: The agent who wants to create a link needs to pay a cost c. We have the following result.

Proposition 10. *Consider the network formation process $(G(t))_{t=0}^{\infty}$ in Definition 3. Assume that there is a cost c of creating a link for the agent who initiates that link. Then, if c is smaller than $\lambda/(1-\lambda)$, agents will always form a link and the emerging network will always be a nested split graph.*

Proof of Proposition 10. Note that the Bonacich centrality of agent $i \in G$ can be written as

$$b_i(G, \lambda) = 1 + \sum_{j \in N_i} b_j(G, \lambda). \tag{37}$$

We consider the network $G + ij$ obtained by adding the link $ij \notin G$. The corresponding change in the Bonacich centrality is given by

$$b_i(G+ij,\lambda) - b_i(G,\lambda) = \underbrace{\sum_{k \in N_i \setminus \{j\}} (b_k(G+ij,\lambda) - b_k(G,\lambda))}_{>0} + \lambda b_j(G+ij,\lambda) \quad (38)$$

$$\geq \lambda b_j(G+ij,\lambda) \quad (39)$$

$$\geq \lambda \min_{k \in G} b_k(G+ij,\lambda). \quad (40)$$

In the first line we have used the fact that in a connected graph (as it is the case for a nested split graph) the number of walks are increasing with the addition of a link and so is the Bonacich centrality. However, the inequality in the second line stems from the fact that a node might be disconnected initially and thus the increase in Bonaich centrality is exactly $\lambda b_j(G+ij,\lambda)$.

The smallest Bonacich centrality in a non-empty graph G (after the creation of a link the graph is always non-empty) is obtained in a dyad, for which

$$b_i(G,\lambda) = \frac{1}{1-\lambda}. \quad (41)$$

Hence, the marginal payoff from forming a link is bounded from below by

$$b_i(G+ij,\lambda) - b_i(G,\lambda) > \frac{\lambda}{1-\lambda}. \quad (42)$$

This bound might seem crude but in our network formation process we always start from an empty graph. Therefore, if the cost of a link is lower than $\lambda/(1-\lambda)$, a link will always be formed and the network will 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.

E. Capacity Constraints and Assortative Networks (NOT FOR PUBLICATION)

A natural generalization of the model presented in the paper is to allow for the possibility that agents are not accepting to establish a link from another agent that wants to connect to them. The underlying assumption is that agents face capacity constraints in the number of links they can maintain. Agents preferably connect to those that give them the highest payoff, but here we assume that if this fails they search for new contacts at random. This means that, if capacity constraints prevent an agent from forming a link to her best response, we assume that she tries to link to an agent out of the whole population of agents at random. This mechanism introduces a global random search mechanism in the link formation process (see Marsili et al. [2004]; Vega-Redondo [2006] for a similar approach). We find that by introducing capacity constraints and global random search, differently to the model presented in the main text, stationary networks can become assortative. Thus, we are able to reproduce all topological properties of empirically observed social and economic networks.

We assume that capacity constraints arise from the fact that an agent can only interact with one other agent out of her neighborhood at a time. Each neighbor of an agent requests information with probability β.[40] Assuming that information requests are independent, the probability that an agent $j \in N$ with d_j links does not receive any information requests from her neighbors is given by $(1-\beta)^{d_j}$. If an agent does not receive such an information request, she can accept an additional link, otherwise not.

In the following, we make a technical assumption. In the model exposed in the text, the Bonacich centrality of an agent increases the most if she forms a link to the agent with the highest degree. For the current model we will assume that this property is still approximately true. In most cases this approximation can be made albeit there exist exceptions in which the degree and Bonacich centrality ranking do not coincide [Grassi et al., 2007].

We have seen in Section 2.3 that the the best response j of an agent i is in her second-order neighborhood, $j \in \mathcal{N}_i^{(2)}$. In contrast, here we allow for the formation of links between agents that are not connected through a common neighbor. This means that agents search globally for new contacts if they cannot connect to the agent with highest degree among their neighbors' neighbors. When an agent i is selected, she tries to connect to the agent j with the highest degree in her second-order neighborhood $\mathcal{N}_i^{(2)}$. However, agent $j \in \mathcal{N}_i^{(2)}$ only accepts the link ij with probability $(1-\beta)^{d_j}$, otherwise agent i selects another agent $k \in N \backslash (\mathcal{N}_i \cup \{i, j\})$ out of the whole population of agents (excluding agents i and j) uniformly at random, and this link also has the same acceptance probability $(1-\beta)^{d_k}$ based on the degree of agent k.

Taking into account the above mentioned capacity constraints in the number of links an agent can form and the possibility to form links outside the second-order neighborhood, we generalize the link formation process $(G(t))_{t=0}^{\infty}$ introduced in Section 2.2 as follows:

Definition 7. *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:*

(i) *With probability $\alpha \in (0, 1)$ agent i receives the opportunity to create an*

[40] If an agent has to process such a request, she cannot accept an additional link.

additional link. Let j be the agent in $\mathcal{N}_i^{(2)}$ with the highest degree, that is $d_j \geq d_k$ for all $j,k \in \mathcal{N}_i^{(2)}$. Then with probability $(1-\beta)^{d_j}$ the link ij is formed. Otherwise agent i connects to a randomly selected agent $k \in N\backslash(\mathcal{N}_i \cup \{i,j\})$ with probability $\left(1-(1-\beta)^{d_j}\right)(1-\beta)^{d_k}$. If agent i is already connected to all other agents then nothing happens.

(ii) *With probability $1-\alpha$, the link to the agent j in $\mathcal{N}_i$ with the smallest degree $d_j \leq d_k$ for all $j,k \in \mathcal{N}_i$, decays. If agent i does not have any links then nothing happens.*

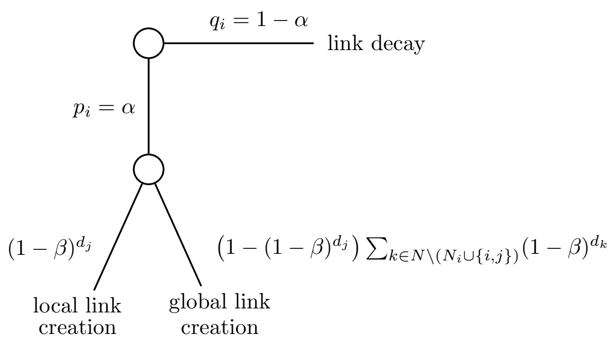


Figure 12: Probabilities with which a randomly selected agent i creates a link and a link of agent i decays, respectively, when capacity constraints are taken into account (assuming that the agent is neither isolated nor fully connected).

An illustration of the above link formation process $(G'(t))_{t=0}^{\infty}$ is shown in Figure 12. An agent i is selected at random either creates a link or the link to the neighbor with lowest degree decays with probability $q_i = 1-\alpha$. However, with probability $p_i = \alpha$ agent i is selected to create a link. In this case, agent i forms the link to agent j with highest degree among her second-order neighbors with probability $(1-\beta)^{d_j}$ and to another agent out of the whole population of agents at random with probability $\left(1-(1-\beta)^{d_j}\right)\sum_{k\in N\backslash(\mathcal{N}_i\cup\{i,j\})}(1-\beta)^{d_k}$.

Having introduced the extended network formation process $(G'(t))_{t=0}^{\infty}$ we now investigate its properties by means of computer simulations for values of $\alpha \in [0.2, 0.5]$ and $\beta \in [0.01, 1]$.[41] We consider a set of $n = 200$ agents and use a sample of 30 to 40 simulation runs from which we compute the average as an approximation to the stationary network.

[41] In the simulation results shown in this section, we assume that if an agent is not free to accept an additional link, (or the agent that is the target of the link can not form an additional link) another agent is selected at random, until a link is formed. In this way, the values of α in the generalized model are comparable with the basic model without capacity constraints in which α is a measure of the network density.

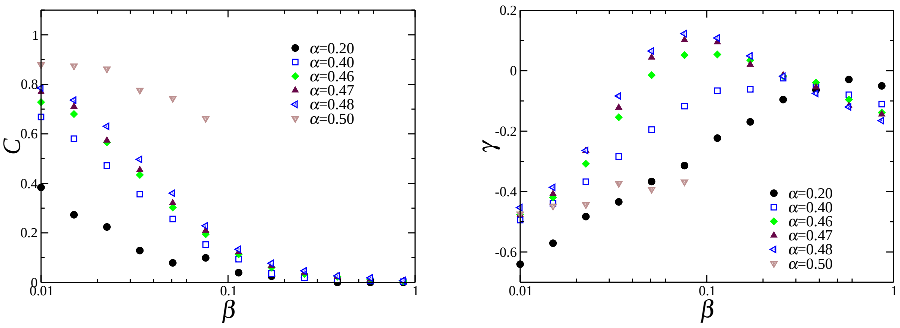


Figure 13: In the left panel we show the clustering coefficient obtained by recourse of numerical simulations of the extended model with capacity constrains for different values of α and β in a network with $n = 200$ agents. In the right panel we show the corresponding network assortativity. Each different curve corresponds to a different value of α. Only agents that are not isolated are considered.

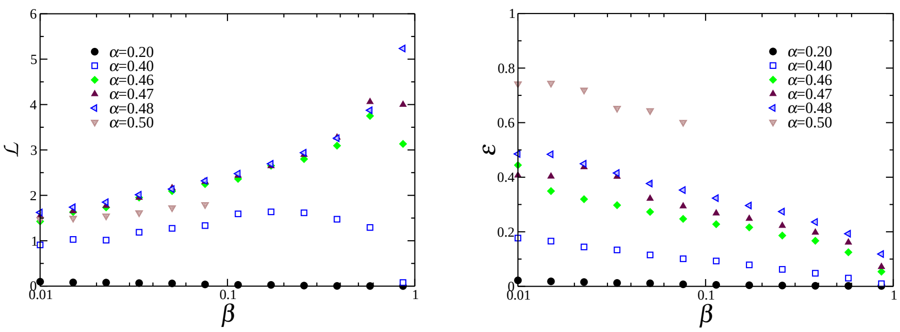


Figure 14: The left panel shows the characteristic path length $\mathcal{L}$ of the stationary network and the right panel shows the results for the network efficiency $\mathcal{E}$ by recourse of numerical simulations of the model with capacity constrains for different values of α and β in a system comprised of $n = 200$ agents.

Figure 13 shows the clustering and assortativity of stationary networks for different values of α and β. We find that for values of β around 0.1 and in $\alpha \in [0.45, 0.5]$ stationary networks are assortative while displaying a high clustering (albeit lower than in the basic model without capacity constraints). In Figure 14 we show the characteristic path length $\mathcal{L}$ and the efficiency $\mathcal{E}$ in terms of short connections in the network. The plots indicate that stationary networks in the extended model exhibit short path lengths between the agents. However, we find that the stationary network may not just consist of one connected component and possibly isolated agents but it may have multiple components. However, there exists a giant component encompassing at least 90% of the agents in all the simulations we studied. We can further analyze the degree distribution of stationary networks and we find that it is highly skewed following an exponential function.

Moreover, we find that the results for different centralization measures show a similar behavior as we have seen already in Section 5.5. There exists a sharp, albeit less pronounced, transition from highly centralized networks to homogeneous networks by increasing α above $1/2$.

In Figure 15 we show the fraction of the largest real eigenvalue (as a measure of efficiency) of the stationary network compared to the corresponding value of the complete network. The figure resembles the findings in Section 6. For values of $\alpha < 1/2$ stationary networks are highly inefficient with respect to the complete network.

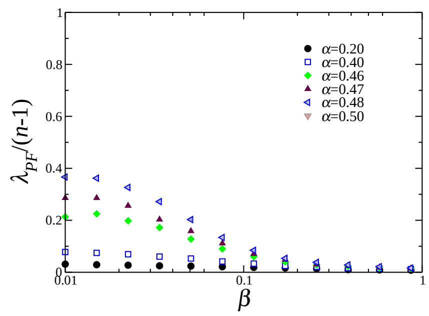


Figure 15: We show the largest eigenvalue of the adjacency matrix normalized to the largest one in a complete graph (which is the efficient network), obtained by recourse of numerical simulations for the model with capacity constrains for different values of α and β in a system comprised of $n = 200$ agents.

In this section we have studied different network statistics for different values of α and β. We find that, by introducing capacity constraints and global random search, stationary networks become assortative while exhibiting an exponential degree distribution, high clustering, short average path length and negative clustering-degree correlation. These characteristics can be found in many social and economic networks. Thus, our model is able to reproduce characteristics of real world networks to the whole extent, ranging from assortative to dissortative networks.

Our findings have an implication for the distinction between assortative and dissortative networks in the literature. As we have discussed in the main text, our network formation process generates stationary networks that are characterized by negative degree-degree correlation and dissortativity. On the other hand, capacity constraints transform stationary networks to exhibiting positive degree-degree correlations and assortativity. This effect may shed some light on the origin of the distinction between technological and social networks suggested in Newman [2002, 2003] where technological networks are characterized by dissortativity and social networks by assortativity. Following our findings, technological networks are facing capacity constraints to a much lower extent than social networks. Consider for example the internet as a technological network and the email network in an organization as a prototype of a social network. The number of hyper-links a website can contain may not be limited as much as the number of social contacts (measured e.g. by mutual email exchange) an individual in an organization may keep. Thus, the distinction between technological and social networks and the degree of assortativity and degree-degree correlations can be derived from the severity of capacity constraints imposed on the number of links an agent can maintain.

F. Proofs of Propositions, Corollaries and Lemmas

In this section we give the proofs of the propositions, corollaries and lemmas stated earlier in the paper.

PROOF OF PROPOSITION 1.

(i) A graph having a stepwise adjacency matrix is a nested split graph G. A nested split graph has a nested neighborhood structure. The neighborhood $\mathcal{N}_j$ of an agent j is contained in the neighborhood $\mathcal{N}_i$ of the next higher degree agent i with $|\mathcal{N}_i| = d_i > |\mathcal{N}_j| = d_j$ with $\mathcal{N}_j \subset \mathcal{N}_i$. For a symmetric adjacency matrix the vector of Bonacich centralities is given by $\mathbf{b}(G, \lambda) = \lambda \mathbf{A}\mathbf{b} + \mathbf{u}$, $\mathbf{u} = (1, ..., 1)^T$. For agent i we get

$$b_i(G, \lambda) = \lambda \sum_{k=1}^{n} a_{ik} b_k(G, \lambda) + 1 = \lambda \sum_{k \in \mathcal{N}_i} b_k(G, \lambda) + 1, \tag{43}$$

and similarly for agent j

$$b_j(G, \lambda) = \lambda \sum_{k \in \mathcal{N}_j} b_k(G, \lambda) + 1. \tag{44}$$

Since $\mathcal{N}_j \subset \mathcal{N}_i$ and $d_j = |\mathcal{N}_j| < |\mathcal{N}_i| = d_i$ we get

$$\frac{b_i(G, \lambda)}{b_j(G, \lambda)} = \frac{\lambda \sum_{k \in \mathcal{N}_i} b_k(G, \lambda) + 1}{\lambda \sum_{k \in \mathcal{N}_j} b_k(G, \lambda) + 1} > 1. \tag{45}$$

The inequality follows from the fact that the Bonacich centrality is nonnegative and the numerator contains the sum over the same positive numbers as the denominator plus some additional values.

Conversely, in a nested split graph we must either have $\mathcal{N}_i \subset \mathcal{N}_j$ or $\mathcal{N}_j \subset \mathcal{N}_i$. Assuming that $b_i(G, \lambda) > b_j(G, \lambda)$ we can conclude from the above equation that $\mathcal{N}_j \subset \mathcal{N}_i$ and therefore $|\mathcal{N}_i| = d_i > |\mathcal{N}_j| = d_j$. If there are l distinct degrees in G then the ordering of degrees $d_1 >$

$d_2 > ... > d_l$ is equivalent to the ordering of the Bonacich centralities $b_1(G,\lambda) > b_2(G,\lambda) > ... > b_l(G,\lambda)$.

(ii) Consider the agents i, j and k in the nested split graph $G(t)$, such that $d_j \leq d_k$. Let G' be the graph obtained from $G(t)$ by adding the link ij and G'' be the graph obtained from $G(t)$ by adding the link ik. We want to show that the Bonacich centrality of agent i in G'' is higher than in G', that is, $b_i(G',\lambda) < b_i(G'',\lambda)$. For this purpose we count the number of walks emanating at agent i when connecting either to agent j or agent k. Since G is a nested split graph, we have

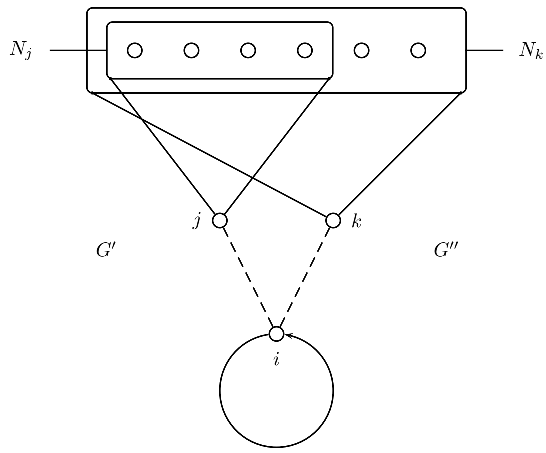


Figure 16: An illustration of the two networks G' and G'', which differ in the links ij and ik. The neighborhood $\mathcal{N}_j$ of agent j and the neighborhood $\mathcal{N}_k$ of agent k are indicated by corresponding boxes. Note that the neighborhood of agent j is contained in the neighborhood of agent k. The loop at agent i indicates a walk starting at i and coming back to i before proceeding to either agent j or k.

that $\mathcal{N}_j \subset \mathcal{N}_k$. An illustration is given in Figure 16. We consider a walk W_l of length $l \geq 2$ starting at agent i in G'. We want to know how many such walks there are in G' and G'', respectively. For this purpose we distinguish the following cases:

(a) Assume that W_l does not contain the link ij nor the link ik. Then each such walk W_l in G' is also contained in G'', since G' and G'' differ only in the links ij and ik.

(b) Consider the graph G' and a walk W_l starting at agent i and proceeding to agent j. For each walk W_l in G' there exists a walk $\tilde{W}_l$ in G'' being identical to W_l except of proceeding from i to j it proceeds from i to k and then to the neighbor of j that is visited after j in W_l. This is always possible since the neighbors of j are also neighbors of k.

(c) Consider a walk W_l in G' that starts at i but first takes a detour returning to i before proceeding from i to j. Using the same argument as in (ii) it follows that for each such walk W_l in G' there exists a walk of the same length in G''.

(d) Consider a walk W_l in G' that starts at agent i and at some point in its sequence of agents and links proceeds from agent j to agent i. For each such walk W_l in G' there exists a walk $\tilde{W}_l$ in G'' that is identical to W_l except that it does not proceed from a neighbor of j to j and then to i it proceeds from a neighbor of j to k and then to i.

The above cases take into account all possible walks in G' and G'' of an arbitrary length l and show that in G'' there are at least as many walks of length l starting from agent i as there are in G'.

Now consider the walks of length two, W_2, in G' starting at agent i and proceeding to agent j. Then there are $|\mathcal{N}_j|$ such walks in G'. However, there are $|\mathcal{N}_k| > |\mathcal{N}_j|$ such walks in G'' of length two that start at agent i.

The Bonacich centrality $b_i(G(t), \lambda)$ is computed by the number of all walks in $G(t)$ starting from i, where the walks of length l are weighted by their geometrically decaying factor λ^l. We have shown that for each l the number of walks in G'' is larger or equal than the number of walks in G' and for $l = 2$ it is strictly larger. Thus, the Bonacich centrality of agent i in G'' is higher than in G'. Note that all agents in a nested split graph are at most two links separated from each other (if there exists any walk between them). Thus, the agent with the highest degree is also the agent with the highest degree among the neighbors' neighbors. From this discussion we see 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. □

Proof of Proposition 2. We give a proof by induction. Let $G(t)$ be a network generated by $(G(t))_{t=0}^{\infty}$. The induction basis is trivial. We start at $t = 0$ from an empty network $G(0) = \bar{K}_n$, which has a trivial stepwise adjacency matrix (see also the Definition 6). Since there are no link present in $\bar{K}_n$ we can omit the removal of a link. 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.

Next we consider the induction step $G(t)$ to $G(t + 1)$. By the induction hypothesis, $G(t)$ is a nested split graph with a stepwise adjacency matrix.

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

Figure 17: Two possible positions for the creation of a link from agent 4, either to agent 7 or to agent 10, are indicated with boxes. Agent 7 has degree 3 while agent 10 has degree 1. Creating a link to an agent with higher degree results in higher equilibrium payoffs. Thus, the best response of agent 4 is agent 7 and not agent 10.

First, we consider the creation of a link ij. Now let agent j be a best response of agent i, that is $j \in BR_i(G(t))$. Now, a link is created only if agent i is also a best response of agent j, that is $i \in BR_j(G(t))$. Using Proposition 1, this means that agent i must be the agent with the highest degree not already connected to j. From the stepwise adjacency matrix $\mathbf{A}(G(t))$ of $G(t)$ (see Definition 6) 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.

We give an example in Figure 17. Let the agents be numbered by the rows respectively columns of the adjacency matrix. We assume that agent 4 is selected to create a link. Two possible positions for the creation of a link from agent 4, either to agent 7 or to agent 10 are indicated with boxes. Since, in a stepwise matrix, the best response agent has the highest degree, agent 7 is a best response of agent 4 while agent 10 is not. We now can turn to the best response of agent 7. The agents not connected to agent 7 are indicated by zero entries in the seventh column of the adjacency matrix. There we find that agent 4 is also a best response of agent 7, since agent 4 is the agent with the highest degree not already connected to agent 7. Finally, we observe that creating the link 47 preserves the stepwise form of the adjacency matrix (see also Definition 6).[42]

For the removal of a link a similar argument can be applied as in the preceding discussion. Disconnecting from the agent with the smallest degree decreases the Bonacich centrality and equilibrium payoffs the least. From

[42]The adjacency matrix is uniquely defined up to a permutation of its rows and columns. Applying such a permutation, we can always find an adjacency matrix which is stepwise.

the properties of the stepwise matrix $\mathbf{A}(G(t))$ it then follows that the matrix $\mathbf{A}(G(t) - ij)$ is stepwise.

Thus, in any step t in the network formation process $(G(t))_{t=0}^{\infty}$, $G(t)$ is a nested split graph with an associated stepwise adjacency matrix $\mathbf{A}(G(t))$. □

Proof of Corollary 1. In Proposition 2 we have shown that $G(t)$ generated by $(G(t))_{t=0}^{\infty}$ is a nested split graph for all times t. In a nested split graph, any node in the connected component is directly connected to the node(s) with maximum degree. Thus, there exists a path of at most length two from any node to any other node in the connected component. It follows that $G(t)$ consists of a connected component and possible isolated nodes. □

Proof of Proposition 3. We will show that the network formation process $(G(t))_{t=0}^{\infty}$ introduced in Definition 3 induces a Markov chain on a finite state space Ω. Ω contains all unlabeled nested split graphs with n nodes. It can be shown that $|\Omega| = 2^{n-1}$ [Mahadev and Peled, 1995]. Therefore, the number of states is finite and the transition between states can be represented with a transition matrix $\mathbf{P}$. In the following we show that this Markov chain is irreducible and aperiodic. We then say that $(G(t))_{t=0}^{\infty}$ is ergodic.

First, we show that $(G(t))_{t=0}^{\infty}$ is a Markov chain. The network $G(t+1)$ 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 $G(t')$ for $t' < t$, that is

$$\mathbb{P}\left(G(t+1) = G_j | G(0) = G_{i_0}, G(1) = G_{i_1}, ..., G(t) = G_{i_t}\right) = \mathbb{P}\left(G(t+1) = G_j | G(t) = G_{i_t}\right). \quad (46)$$

The number of possible networks $G(t)$ is finite for any time t and the transition probabilities from a network $G(t)$ to $G(t+1)$ do not depend on t but only on α and the current number of agents in the independent sets and dominating subsets, respectively. Therefore, $(G(t))_{t=0}^{\infty}$ is a finite state, discrete time, homogeneous Markov chain. Moreover, the transition matrix $\mathbf{P}$ is defined by $(\mathbf{P})_{ij} = \mathbb{P}\left(G(t+1) = G_j | G(t) = G_i\right)$ for any $G_i, G_j \in \Omega$.

Next, we show that the Markov chain is irreducible. Consider two networks $G, G' \in \Omega$. $(G(t))_{t=0}^{\infty}$ is irreducible if there exists a positive probability to pass from any G to any other G' in Ω. We say that G' is accessible from G. For any G there exists a positive probability that in all consecutive steps in the Markov chain links are removed and no links are created until the empty network $\bar{K}_n \in \Omega$ is reached. Then there exists a positive probability that from $\bar{K}_n$ only those links are created that generate exactly the network G'. Therefore, there exists a positive probability to pass from any network G to any other network G' with positive probability. Similarly, one can show that

G is accessible from G'. States G and G' are accessible from one-another. We say that they communicate and Ω is a communicating class.

Moreover, the Markov chain is aperiodic. Observe that with positive probability the empty network $\bar{K}_n$ can stay empty in the next time step. This happens when an agent in $\bar{K}_n$ is selected for removing a link with probability $1-\alpha$. Since this agent has no links, nothing happens. Thus, the state $\bar{K}_n$ is aperiodic. The existence of an aperiodic state in the communicating class Ω implies that the Markov chain induced by $(G(t))_{t=0}^{\infty}$ is aperiodic.

Since we have shown that $(G(t))_{t=0}^{\infty}$ induces a finite Markov chain that is irreducible and aperiodic, we say that the Markov chain is ergodic. Further, this means that there exists a unique stationary distribution μ satisfying $\mu\mathbf{P} = \mu$ [see e.g. Seneta, 1973, 2006]. □

PROOF OF PROPOSITION 4. We consider the network formation process $(G(t))_{t=0}^{\infty}$ on Ω introduced in Definition 3. At every step $t = 0, 1, 2, ...$ a link is created with probability α and a link is removed with probability $1 - \alpha$. Further, we consider the complementary network formation process $(G'(t))_{t=0}^{\infty}$ on Ω where in every period t a link is created with probability $\alpha' = 1 - \alpha$ and a link is removed with probability $1 - \alpha' = \alpha$.[43] This means that a link is removed in $(G'(t))_{t=0}^{\infty}$ whenever a link is created in $(G(t))_{t=0}^{\infty}$ and a link is created whenever a link is removed in $(G(t))_{t=0}^{\infty}$.

As an example, consider the network G represented by the adjacency matrix $\mathbf{A}$ in Figure 17. The complement $\bar{G}$ has an adjacency matrix $\bar{\mathbf{A}}$ obtained from $\mathbf{A}$ by replacing each one element in $\mathbf{A}$ by zero and each zero element by one, except for the elements on the diagonal. Let H be the network obtained from G by adding the link 47 (setting $a_{47} = a_{74} = 1$ in $\mathbf{A}$). The probability of this link being created and thus the probability of reaching H after the process was in G is $3\alpha/n$, either by selecting one of the two nodes with degrees three or the node with degree five to create a link. Observe that this is identical to the probability of reaching the network $\bar{H}$ from $\bar{G}$ if either the two nodes with degrees seven or the node with degree four in $\bar{G}$ are selected to remove a link (with probability $\alpha' = 1 - \alpha$).

In general we can say that, for any $G_1, G_2 \in \Omega$ we have that

$$\mathbb{P}\left(G(t+1) = G_2 | G(t) = G_1\right) = \mathbb{P}\left(G'(t+1) = \bar{G}_2 | G'(t) = \bar{G}_1\right). \qquad (47)$$

[43] Two nodes of $G'(t)$ are adjacent if and only if they are not adjacent in $G(t)$. Note that the complement of a nested split graph is a nested split graph as well [Mahadev and Peled, 1995]. In particular, the networks $G'(t)$ 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 $G(t)$ and the number of nodes in the independent sets in $G'(t)$ corresponds to the number of nodes in the dominating subsets in $G(t)$. Thus, $(G'(t))_{t=0}^{\infty}$ has the same state space Ω as $(G(t))_{t=0}^{\infty}$, namely the space consisting all unlabeled nested split graphs on n nodes.

Next consider the stationary distribution μ of $(G(t))_{t=0}^{\infty}$ and the corresponding transition matrix $\mathbf{P}$. Similarly, consider the stationary distribution μ' of $(G'(t))_{t=0}^{\infty}$ and the corresponding transition matrix $\mathbf{P}'$. Further, consider an ordering of states $G_1, G_2, \ldots$ in Ω and the transition matrix $\mathbf{P}$ with elements $(\mathbf{P})_{ij}$ giving the probability of observing G_j after the Markov chain $(G(t))_{t=0}^{\infty}$ was in G_i. Similarly, consider an ordering of states $\bar{G}_1, \bar{G}_2, \ldots$ in Ω and the transition matrix $\mathbf{P}'$ with elements $(\mathbf{P}')_{ij}$ giving the probability of observing $\bar{G}_j$ after the Markov chain $(G'(t))_{t=0}^{\infty}$ was in $\bar{G}_i$. Equation (47) implies that $\mathbf{P} = \mathbf{P}'$. Moreover, for the stationary distributions it must hold that $\mu\mathbf{P} = \mu$ and $\mu'\mathbf{P}' = \mu'$. Since $\mathbf{P}$ is irreducible and aperiodic, $\mathbf{P}$ has a unique positive eigenvector and therefore $\mu' = \mu$. It follows that for any network $G \in \Omega$ with probability μ_G we can take the complement $\bar{G} = G'$ and assign it the probability μ_G to get the corresponding probability in μ', i.e. $\mu_G = \mu'_{G'}$. □

Proof of Proposition 5. Before we proceed with the proof of Proposition 5, we state two useful lemmas.

Lemma 1. *Consider the ergodic Markov chain $(G(t))_{t=0}^{\infty}$ with the parameter $0 < \alpha \leq 1/2$ and state space Ω consisting of all nested split graphs. Let X denote the set of states in Ω in which there is exactly one node with degree $d+1$ and Y the set of states where there is no node with degree $d+1$. Denote by μ_X the probability of the states in X in the stationary distribution μ of $(G(t))_{t=0}^{\infty}$ and by μ_Y the probability of states in Y. If the number of nodes with degree N_d in Y is $\Theta(n)$ such that $\lim_{n\to\infty} N_d/n > 0$ then $\lim_{n\to\infty} \mu_Y = 0$.*[44]

Proof of Lemma 1. Let $N(X, Y, y)$ be the expected number of times states in X occur before the process reaches Y (not counting the process as having immediately reached Y if $y \in Y$) when the process starts in y. Then the following relation holds (see Theorem 6.2.3 in Kemeny and Snell [1960] and also Ellison [2000])

$$\frac{\mu_X}{\mu_Y} = N(X, Y, y). \tag{48}$$

Let p_{YX} denote a lower bound on the probability that a state in X occurs after the process is in a state in Y and, conversely, let p_{XY} denote the probability that a state in Y occurs after the process is in a state in X. This probability is the same for all states in X, since from the properties

[44] By $f = \Theta(g)$ we mean that

$$0 < \liminf_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \leq \limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| < \infty.$$

In particular, $f = \Theta(1)$ implies that $0 < \lim_{n\to\infty} f(n) < \infty$.

of the Markov chain $(G(t))_{t=0}^{\infty}$, it follows that $p_{XY} = 2(1-\alpha)/n$, because there exist two possibilities to remove the link of the node with degree $d+1$ and the probability to select a node for link removal is $(1-\alpha)/n$. Observe that this probability vanishes for large n, $\lim_{n\to\infty} p_{XY} = 0$. Moreover, we have that

$$\begin{aligned} N(X,Y,y) \geq\ & p_{YX}p_{XY} \\ & + 2p_{YX}(1-p_{XY})p_{XY} \\ & + 3p_{YX}(1-p_{XY})^2 p_{XY} \\ & + \ldots \\ & = p_{YX}p_{XY}\sum_{i=1}^{\infty} i(1-p_{XY})^{i-1} \\ & = \frac{p_{YX}}{p_{XY}}. \end{aligned}$$

The right hand side of the above inequality takes into account the fact that states in X can be reached once, twice, etc., before a state in Y is reached and assigns the corresponding probabilities to compute the expected value.

By assuming that there exists a number N_d of nodes with degree d which is $\Theta(n)$, we have that $p_{YX} \geq \alpha N_d/n$ and $\lim_{n\to\infty} p_{YX} > 0$. It then follows that

$$\frac{\mu_X}{\mu_Y} = N(X,Y,y) \geq \frac{p_{YX}}{p_{XY}} = \frac{\alpha}{2(1-\alpha)} N_d \xrightarrow[n\to\infty]{} \infty. \tag{49}$$

Since μ_X is a probability with $\mu_X \leq 1$, Equation (49) implies that $\lim_{n\to\infty}\mu_Y = 0$. □

Lemma 2. *For $0 < \alpha \leq 1/2$ the asymptotic expected proportion of isolated nodes in the limit of large n is given by*

$$n_0 = \frac{1-2\alpha}{1-\alpha}. \tag{50}$$

Proof of Lemma 2. We consider the expected change in the number of links $m(t)$ in $G(t)$ from t to $t+1$.[45] The number of links increases by one if any node which does not have the maximum degree $n-1$ is selected for creating a link. This happens with probability $\alpha\left(n - N_{n-1}(t)\right)/n$. The number of links decreases whenever a node with degree higher than zero is selected for removing a link. This happens with probability $(1-\alpha)\left(n - N_0(t)\right)/n$. Putting the above contributions together we can write for the expected

[45] We have that $2m(t) = \sum_{d=0}^{n-1} N_d(t) d$.

change in the total number of links from t to $t+1$

$$\mathbb{E}\left(m(t+1)|N(t)\right) - m(t) = \frac{\alpha}{n}\left(n - N_{n-1}(t)\right) - \frac{1-\alpha}{n}\left(n - N_0(t)\right). \quad (51)$$

Taking expectations on both sides of the above equation and denoting by $n_d(t) = \mathbb{E}\left(N_d(t)/n\right)$ we obtain

$$\mathbb{E}\left(m(t+1)\right) - \mathbb{E}\left(m(t)\right) = \alpha\left(1 - n_{n-1}(t)\right) - (1-\alpha)\left(1 - n_0(t)\right). \quad (52)$$

Let ρ denote the initial distribution of states, with $\rho_i = 1$ if $G_i = \bar{K}_n$ and zero otherwise. Further, let m be the column vector whose j-th coordinate, m_j, is the value of m at state $G_j \in \Omega$. Let $G_i = \bar{K}_n$ then we can write

$$\begin{aligned}\mathbb{E}\left(m(t)\right) &= \mathbb{E}\left(m(t)|G(0) = G_i\right) \\ &= \sum_{G_j \in \Omega} m_j \mathbb{P}\left(G(t) = G_j | G(0) = G_i\right) \\ &= \sum_{G_j \in \Omega} \left(\mathbf{P}^t\right)_{ij} (m)_j = \left(\mathbf{P}^t m\right)_i = \rho \mathbf{P}^t m.\end{aligned}$$

For large times t the expectation is computed over the invariant distribution μ. In particular, $\lim_{t\to\infty} \rho\mathbf{P}^t = \mu$ and therefore $\lim_{t\to\infty} \mathbb{E}\left(m(t)\right) = \lim_{t\to\infty} \rho\mathbf{P}^t m = \mu m$. Similarly, we have that $\lim_{t\to\infty} \mathbb{E}\left(m(t+1)\right) = \lim_{t\to\infty} \rho\mathbf{P}^{t+1} m = \mu\mathbf{P}m = \mu m$. Therefore, $\lim_{t\to\infty} \mathbb{E}\left(m(t+1)\right) = \lim_{t\to\infty} \mathbb{E}\left(m(t)\right)$. Thus, we can set the left hand side of Equation (51) to zero, in the limit of large t, and obtain a relationship between the asymptotic expected proportion of nodes of degree zero and one, respectively,

$$1 - 2\alpha = (1-\alpha)n_0 - \alpha n_{n-1}, \quad (53)$$

where we have denoted by $n_d = \lim_{t\to\infty} n_d(t)$. Next, we consider the chain $(G'(t))_{t=0}^{\infty}$ which is constructed from $(G(t))_{t=0}^{\infty}$ by taking the complement of each network $G(t)$ in every period t (see also the proof of Proposition 4). In the following, denote the asymptotic expected number of links $\lim_{t\to\infty} \mathbb{E}\left(m(t)\right)$ of $(G(t))_{t=0}^{\infty}$ by m and of $(G'(t))_{t=0}^{\infty}$ by m'. By construction, we must have that $m = n(n-1)/2 - m'$. From Proposition 4 we know that the Markov chain $(G'(t))_{t=0}^{\infty}$ has the same stationary distribution μ' as a process from Definition 3 for a link creation probability of $\alpha' = 1-\alpha$. For $\alpha = 1/2$ the two processes are identical and we must have that also their expected number of links are the same. This implies that for $\alpha = 1/2$, $m = m' = n(n-1)/4$. The only nested split graph with this number of links, for which the complement has the same number of links as the original graph, is the one in which each independent set is of size one and also each dominating subset has size one (except possibly for the set corresponding to the $\left(\left\lfloor \frac{k}{2} \right\rfloor + 1\right)$-th partition). Thus, for $\alpha = 1/2$ it must hold that

$n_0 = n_{n-1} = 1/n$.

Moreover, we know that for $\alpha < 1/2$ the expected number of maximally connected nodes (with degree $n-1$) is at most as large as the expected number for $\alpha = 1/2$, since the probability of links being created strictly decreases while the probability of links being removed increases for values of α below $1/2$ (and the probability of a maximally connected node losing a link strictly increases). Thus $n_{n-1} \leq 1/n$ for $\alpha \leq 1/2$, and for large n we can write Equation (53) as follows

$$1 - 2\alpha = (1-\alpha)n_0, \tag{54}$$

or equivalently

$$n_0 = \frac{1-2\alpha}{1-\alpha}. \tag{55}$$

For $\alpha = 0$ no links are created and all nodes are isolated, that is $n_0 = 1$, while for $\alpha = 1/2$ the asymptotic expected number of isolated nodes vanishes in the limit of large n. $\square$

With these two lemmas in hand, let us now prove Proposition 5. Note that $G(t)$ is completely determined by $N(t)$ and vice versa. Thus it follows that $\{N(t)\}_{t=0}^{\infty}$ is a Markov chain. Denote by $n_d(t) = \mathbb{E}\left(N_d(t)/n\right)$ the expected proportion of nodes with degree d at time t and let us denote by $n_d = \lim_{n\to\infty} n_d(t)$; n_d is determined by the invariant distribution μ in the limit of large times t. Lemma 2 shows that Equation (7) holds for $d = 0$. In the following we show by induction that, given that Equation (7) holds for n_{d-1} and n_d, as n becomes large, also n_{d+1} satisfies Equation (7) for all $0 \leq d < d^*$, in the limit of large n. For this purpose we consider (a) the expected number of isolated nodes $\mathbb{E}\left(N_0(t+1)|N(t)\right)$ and (b) the expected number of nodes with degree $d = 1, ..., d^*$, $\mathbb{E}\left(N_d(t+1)|N(t)\right)$ at time $t+1$, conditional on the current degree distribution $N(t)$.

(a) Consider a particular network $G(t)$ in period t generated by $(G(t))_{t=0}^{\infty}$ and its associated degree distribution $N(t)$. Figure 18 shows an illustration of the corresponding stepwise matrix. In the following we compute the expected change of the number $N_0(t)$ of isolated nodes in $G(t)$.
The expected change of $N_0(t)$ due to the creation of a link has the following contributions. An agent with the highest degree k in $N_k(t)$ can create a link to an isolated agent and thus decreases the number of isolated agents by one. The expected change from this link is $-\alpha N_k(t)/n$. On the other hand, if an isolated agent creates a link then the expected change in the number of isolated agents is $-\alpha N_0(t)/n$.
Moreover, the removal of links can affect $N_0(t)$ if there is only one agent with maximal degree, i.e. $N_k(t) = 1$. In this case, if the agent with the highest degree removes a link, then an additional isolated agent is created yielding an expected increase in $N_0(t)$ of $(1-\alpha)N_k(t)/n$. Next,

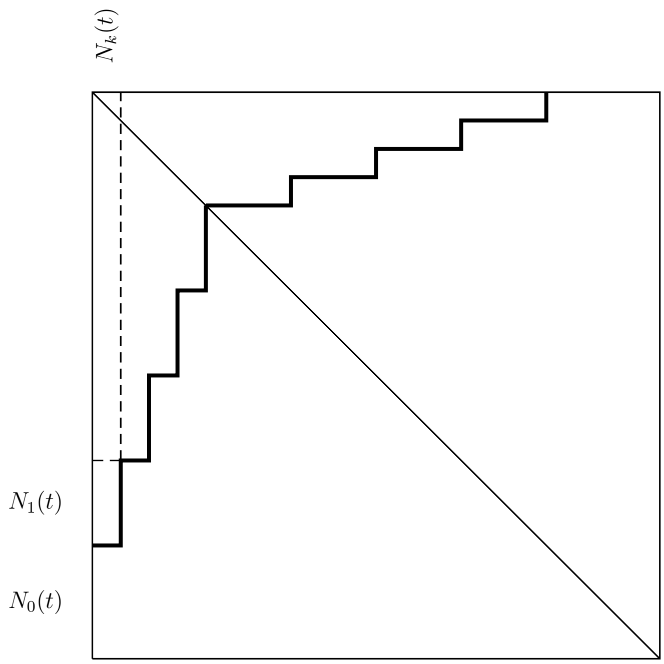


Figure 18: Representation of the stepwise matrix $\mathbf{A}$ of a nested split graph G and some selected degree partitions. The stepfunction separating the zero entries in the matrix from the one entries is shown with a thick line.

if an agent with degree one in $N_1(t)$ removes a link, then the number of isolated agents increases. Note that in a nested split graph $N_1(t) > 0$ implies that $N_k(t) = 1$ and vice versa. This gives an expected change of $N_0(t)$ given by $(1-\alpha)N_1(t)/n$.

Putting the above contributions together, the expected change in the number of isolated nodes at time $t+1$, conditional on $N(t)$, is given by the following expression[46]

$$\mathbb{E}\left(N_0(t+1)|N(t)\right) - N_0(t) = \\ -\frac{\alpha}{n}\left(N_0(t) + N_k(t)\right) + \frac{1-\alpha}{n}\left(N_1(t)+1\right)\delta_{N_k(t),1}. \quad (56)$$

We can take expectations on both sides of Equation (56). For large times t the expectation is computed on the basis of the invariant distribution μ and similarly to the proof of Lemma 2, after taking expectations, we can set the left hand side of Equation (56) to zero for large times t. Note that from Lemma 2 we know that the asymptotic expected proportion n_0 of isolated nodes is $\Theta(1)$, for n large. Thus we can apply the result of Lemma 1 which tells us that the networks in which there does not exist a node with degree one have vanishing probability in μ for large n. Since the existence of a node with degree one implies that $N_k(t) = 1$, in the limit of large n we can set $\delta_{N_k(t),1} = 1$. We then obtain from Equation (56)

$$n_1 = \frac{\alpha}{1-\alpha} n_0. \quad (57)$$

This shows that also n_1 satisfies Equation (7). Together with Lemma 2 this proves the induction basis.

(b) We give a proof by induction on the number $N_d(t)$ of nodes with degree $0 < d < d^*$ in a network $G(t)$ in the support of the stationary distribution μ. In the following, we compute the expected change in $N_d(t)$ due to the creation or the removal of a link. An illustration can be found in Figure 19.

Let us investigate the creation of a link. With probability α/n a link is created from an agent in $N_{k-d}(t)$ to an agent in $N_d(t)$. This yields a contribution to the expected change of $N_d(t)$ of $-\alpha N_{k-d}(t)/n$. If a link is created from an agent in $N_{k-d+1}(t)$ to an agent in $N_d(t)$ then the expected change is α/n, if $N_{k-d+1}(t)$ contains only a single agent. Similarly, if a link is created from an agent in $N_{d-1}(t)$ to an agent in $N_d(t)$ then the expected change of $N_d(t)$ is $\alpha N_{d-1}(t)/n$, if $N_{k-d+1}(t) = 1$. Moreover, if an agent in $N_d(t)$ is selected for link creation then we get an expected decrease of $-\alpha N_d(t)/n$.

Now we consider the removal of a link. If a link is removed from the

[46] $\delta_{i,j}$ denotes the usual Kronecker delta which is 1 if $i = j$ and 0 otherwise.

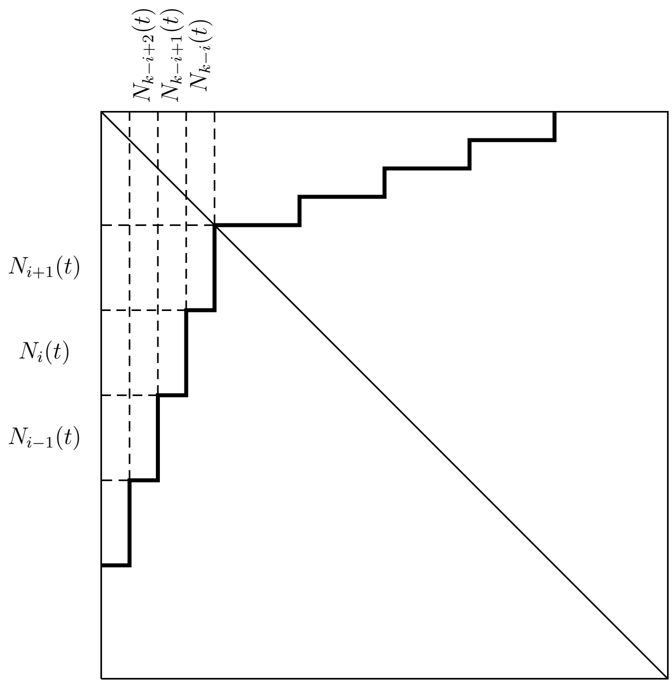


Figure 19: Representation of the stepwise matrix $\mathbf{A}$ of a nested split graph G. The stepfunction separating the zero entries in the matrix from the one entries is shown with a thick line.

agent in $N_{k-d+1}(t)$ to an agent in $N_d(t)$ then the expected change of $N_d(t)$ is $-(1-\alpha)N_{k-d+1}(t)/n$. If a link is removed from an agent in $N_{k-d}(t)$ to an agent in $N_{d+1}(t)$ then the expected increase of $N_d(t)$ is $(1-\alpha)/n$, if $N_{k-d}(t) = 1$. Moreover, if an agent in $N_{d+1}(t)$ is selected for removing a link, then we get an expected increase of $(1-\alpha)N_{i+d}(t)/n$, if $N_{k-d}(t) = 1$. Finally, if an agent in $N_d(t)$ is selected for removing a link, then we get an expected change of $-(1-\alpha)N_d(t)/n$.
Putting the above contributions together, the expected change in $N_d(t)$ is given by

$$\begin{aligned}
&\mathbb{E}\left(N_d(t+1)|N(t)\right) - N_d(t) = \\
&\qquad \frac{\alpha}{n}\left(-N_d(t) + \left(N_{d-1}(t)+1\right)\delta_{N_{k-d+1}(t),1} - N_{k-d}(t)\right) \\
&\quad + \frac{1-\alpha}{n}\left(-N_d(t) + \left(N_{d+1}(t)+1\right)\delta_{N_{k-d}(t),1} - N_{k-d+1}(t)\right). \quad (58)
\end{aligned}$$

We can take expectations on both sides of Equation (58) and similarly to part (a) of this proof we can set the left-hand-side of Equation (58) as t becomes large. For large times t the above expectation is computed on the basis of the invariant distribution μ. By the induction assumption, the asymptotic expected proportion n_{d-1} of nodes with degree $d-1$ is $\Theta(1)$ in the limit of large n (as follows from Equation (7)). Thus we can apply Lemma 1 and neglect the networks in which there does not exist a node with degree d since they have vanishing probability in μ for large n. Similarly, we know from the induction assumption that the asymptotic proportion n_d of nodes with degree d is $\Theta(1)$ and, by virtue of Lemma 1, we know that the networks in which there does not exist a node with degree $d+1$ have vanishing probability in μ for large n. Thus, in the limit of large n we can set $\delta_{N_{k-d+1}(t),1} = \delta_{N_{k-d}(t),1} = 1$, since the existence of nodes with degrees d and $d+1$ implies that $N_{k-d+1}(t) = N_{k-d}(t) = 1$ in the limit of large t and n. Therefore, we get from Equation (58) the following relationship

$$n_{d+1} = \frac{1}{1-\alpha}n_d - \frac{\alpha}{1-\alpha}n_{d-1}. \quad (59)$$

Inserting the expressions for n_{d-1} and n_d from Equation (7) into Equation (59) yields

$$\begin{aligned}
n_{d+1} &= \frac{1}{1-\alpha}\frac{1-2\alpha}{1-\alpha}\left(\frac{\alpha}{1-\alpha}\right)^d - \frac{\alpha}{1-\alpha}\frac{1-2\alpha}{1-\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{d-1} \\
&= \frac{1-2\alpha}{1-\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{d+1}
\end{aligned}$$

Thus, Equation (7) also holds for n_{d+1}. This proves the induction step.

Finally, we have that the degree distribution must be normalized to one, i.e. $\sum_{d=0}^{n-1} n_d = 1$. We know that the number of agents in the dominating subsets with degrees larger than d^* is d^* (since each set contains only one node and there are d^* such sets).[47] Adding this to the number of agents in the independent sets with degree $d = 0, ..., d^*$ yields

$$n \sum_{d=0}^{d^*} n_d + d^* = n. \tag{60}$$

Further, inserting Equation (7) we can derive the number d^* of independent sets as a function of n and α

$$d^*(n, \alpha) = \frac{\ln\left(\frac{2(1-\alpha)}{(1-2\alpha)n}\right)}{\ln\left(\frac{\alpha}{1-\alpha}\right)}. \tag{61}$$

d^* is a monotonic decreasing function of n for a fixed value of α. Conversely, for a fixed value of n we get the limits $\lim_{\alpha \to 0} d^* = 0$ and $\lim_{\alpha \to 1/2} d^* = n/2$. This completes the proof. $\square$

Proof of Corollary 2. The results follows directly from the functional form of $d(n, \alpha)$ in Proposition 5. $\square$

Proof of Proposition 6. Let us define the following random variable

$$Y_s = \mathbb{E}\left(N_d(t)|N(s)\right). \tag{62}$$

Since $\{N(t)\}_{t=0}^{\infty}$ is a Markov chain, the sequence $\{Y_s\}_{s=0}^{t}$ is a Martingale with respect to $\{N(t)\}_{t=0}^{\infty}$.[48] Moreover, the change in the number of nodes with degree d per period t is bounded by two, i.e. $|N_d(t) - N_d(t-1)| \leq 2$, since at most one link is added or removed in every period t and this can change the degrees of at most two nodes. Therefore, we can apply Hoeffding's inequality [Hoeffding, 1963], which states that for any $0 < s \leq t$ with $|Y_s - Y_{s-1}| \leq c$ and any $\epsilon > 0$

$$\mathbb{P}\left(|Y_t - Y_0| > \epsilon\right) \leq 2e^{-\frac{\epsilon^2}{2tc^2}}. \tag{63}$$

With $Y_t = \mathbb{E}\left(N_d(t)|N(t)\right) = N_d(t)$ and $Y_0 = \mathbb{E}\left(N_d(t)|N(0)\right) = \mathbb{E}\left(N_d(t)\right)$ it

[47] Note that since those networks in which there does not exist a node with degree $0 \leq d \leq d^*$ in the corresponding independent set can be neglected, the structure of nested split graphs implies that all dominating subsets have size one.

[48] We have that $\mathbb{E}\left(Y_s|N(s-1)\right) = \mathbb{E}\left(\mathbb{E}\left(N_d(t)|N(s)\right)|N(s-1)\right) = \mathbb{E}\left(N_d(t)|N(s-1)\right) = Y_{s-1}$. Further, one can show that the first and second moments of $\{Y_s\}_{s=0}^{t}$ are bounded. Thus, $\{Y_s\}_{s=0}^{t}$ is a Martingale with respect to $\{N(t)\}_{t=0}^{\infty}$ [see also Grimmett and Stirzaker, 2001].

follows from Equation (63) that

$$\mathbb{P}\left(\left|\frac{N_d(t)}{n} - \mathbb{E}\left(\frac{N_d(t)}{n}\right)\right| \geq \epsilon\right) = \mathbb{P}\left(|N_d(t) - \mathbb{E}\left(N_d(t)\right)| \geq n\epsilon\right) \leq 2e^{-\frac{\epsilon^2 n^2}{8t}}. \tag{64}$$

This implies that the empirical proportion $N_d(t)/n$ of nodes with degree d converges in probability to its expected value $\mathbb{E}\left(N_d(t)/n\right)$, as n becomes large.

Proof of Corollary 3. See Theorem 1.2.4 in Mahadev and Peled [1995]. □