Dynamics of Degree Distributions of Social Networks

Isabel Fernández,[1] Kevin M. Passino,[1] and Jorge Finke[2]

***Abstract*— Social network models aim to capture the complex structure of social connections. They are a framework for the design of control algorithms that take into account relationships, interactions, and communications between social actors. Based on three formation mechanisms - random attachment, triad formation, and network response - our work characterizes the dynamics of the degree distributions of social networks. In particular, we show that the complementary cumulative in- and out-degree distributions of highly clustered, reciprocal networks can be approximated by infinite dimensional time-varying linear systems. Furthermore, we determine the invariance of both limit distributions and the stability properties of the average degree.**

I. Introduction

Research on social networks has focused on characterizing structural properties such as degree distributions and community structures [**?**], [**?**], [**?**], [**?**]. More recent approaches have tried to close the gap between network analysis and control theory by introducing dynamic models that describe the evolution of social connections [**?**], [**?**]. However, less attention has been paid to understanding the dynamics of two key properties that define a large number of social networks: clustering and reciprocity.

The two properties can be explained based on triad formation and network response mechanisms. Characterizing the impact of clustering and reciprocity on the degree distributions of networks is a first step for the design of control algorithms that leverage the structure of a social network. It is an open challenge to develop models that are rich enough to characterize social structures, but simple enough to be formally analyzed [**?**].

Our work is closely related to the network models introduced in [**?**], [**?**], [**?**], [**?**], which propose various mechanisms for how nodes establish links. In particular, the authors of [**?**], [**?**] study the asymptotic behavior of the degree distributions of highly clustered networks with no reciprocity. The work in [**?**], [**?**] considers the effect of reciprocity on the limit distributions of the in- and out-degrees, the local and global clustering coefficients, and the local and global reciprocity coefficients. The authors of [**?**] also identify conditions under which the dynamics of the global reciprocity and the global clustering coefficients are asymptotically stable.

Here, we use three simple mechanisms to capture the formation of highly clustered, reciprocal networks [**?**], [**?**]. These mechanisms are: *random attachment*, which describes how a new incoming node connects to a network; *triad formation*, which characterizes how the new node establishes transitive relationships; and *network response*, which represents how the network reacts to node attachments.

Unlike the work in [**?**], the focus of this paper is on defining conditions under which the dynamics of the degree distributions can be approximated as an infinite dimensional time-varying linear system. In particular, we characterize how triad formation and network response impact the dynamics of both degree distributions. Finally, we characterize the dynamics of the average degree distribution and derive sufficient conditions that guarantee its stability.

[1] Department of Electrical and Computer Engineering, The Ohio State University, Columbus, Ohio.

[2] Department of Electrical Engineering and Computer Science, Pontificia Universidad Javeriana, Cali, Colombia.

II. The model

Consider a sequence $\mathcal{G} = \{\mathcal{G}(0), \mathcal{G}(1), \ldots\}$, where each directed graph $\mathcal{G}(t) = (\mathcal{H}(t), \mathcal{A}(t))$ describes a network at time index $t \geq 0$. Let N_0 be the number of nodes in the initial network, $\mathcal{H}(t) = \{1, \ldots, N_0 + t\}$ the set of nodes, and $\mathcal{A}(t) = \{(i, j) : i, j \in \mathcal{H}(t)\}$ the set of directed edges. The pair $(j, i) \in \mathcal{A}(t)$ indicates that there exists an edge from node j to node i at time t, and $\mathcal{Q}_i(t) = \{j \in \mathcal{H}(t) : (j, i) \in \mathcal{A}(t)\}$ represents the set of incoming neighbors of node i. Let $k_i(t) = |\mathcal{Q}_i(t)|$ represent the in-degree of node i. Similarly, $\hat{\mathcal{Q}}_i(t) = \{j \in \mathcal{H}(t) : (i, j) \in \mathcal{A}(t)\}$ represents the set of outgoing neighbors of node i and $\hat{k}_i(t) = |\hat{\mathcal{Q}}_i(t)|$ its out-degree. Consider the following mechanisms for the evolution of the network at time t.

- M1 *Random attachment:* A new node attached to the network and links to $m \geq 1$ different nodes, selected uniformly at random over $\mathcal{H}(t-1)$.
- M2 *Triad formation:* For every link established during random attachment, the new node may establish an additional link. Specifically, if node $j \notin \mathcal{H}(t-1)$ connects to some node $j' \in \mathcal{H}(t-1)$, it connects to an outgoing neighbor of node j' with probability $\pi_f > 0$ (selected again uniformly at random over $\hat{\mathcal{Q}}_{j'}(t-1)$).
- M3 *Network response:* There are two ways the network responds to the attachment of a new node. The first approach is based on reciprocity: Each of the m randomly selected nodes establishes a reciprocal link with probability $\pi_r > 0$. The second approach is random, with no preference for establishing reciprocal links: A set of $n \geq 0$ randomly selected nodes of $\mathcal{H}(t-1)$ connect to the new node.

Note that mechanism M1implies that $\mathcal{H}(t) = \{1, \ldots, N_0 + t\}$. Mechanism M2 enables transitive relationships. The first approach of mechanism M3 tends to form reciprocal links. An edge (i, j) is called a reciprocal link involving nodes i and j if $(i, j) \in \mathcal{A}(t)$ and $(j, i) \in \mathcal{A}(t)$. The pair of edges

(i,j) and (j,i) is called a reciprocal cycle and represents the lowest-order cycle in a directed network.

To ensure that the formation mechanisms are properly defined, we require the following assumptions.

A1 *Configuration of the initial network*: The network $\mathcal{G}(0)$ is weakly connected and has more than $2m$ nodes (i.e., $N_0 \geq 2m$), each with at least m outgoing and n incoming neighbors.

A2 *Persistency of the mechanisms*: Mechanisms M1-M3 are triggered every time index $t \geq 0$.

The two assumptions guarantee that mechanisms M1-M3 create on average $m + m\pi_f + m\pi_r + n$ new edges every time $t \geq 0$.

III. Dynamics, invariance, and stability of the average degree

Let $\mu(t)$ represent the average in-degree of the network at time t. Since every node is connected at least to another node, $\mu(t) \geq 1$ for all t. Let μ_0 represent the average in-degree at time $t=0$, i.e., $\mu_0 = \frac{\sum_{i \in \mathcal{H}(0)} k_i(0)}{N_0}$. The following theorem defines the average in-degree of the network as a recurrence function.

Theorem 1 (Dynamics of the average in-degree): Suppose that assumptions A1-A2 hold. For $t \geq 1$, the average in-degree of the network is given by

$$\mu(t) = \frac{(N_0 + t - 1)\mu(t-1) + m(1 + \pi_f + \pi_r) + n}{N_0 + t} \quad (1)$$

Proof: When a node attaches to the network, mechanisms M1-M3 establish on average $m + m\pi_f + m\pi_r + n$ new edges. Since the total number of nodes at time $t-1$ is $N_0 + t - 1$, the total in-degree at time $t \geq 1$ is

$$\sum_{i \in \mathcal{H}(t)} k_i(t) = (N_0 + t - 1)\mu(t-1) + m + m\pi_f + m\pi_r + n$$

So the average in-degree of a randomly selected node is

$$\mu(t) = \frac{(N_0 + t - 1)\mu(t-1) + m + m\pi_f + m\pi_r + n}{N_0 + t}$$

■

To characterize $\mu(t)$ as a non-recursive function, consider the following corollary.

Corollary 1 (Average in-degree): Suppose that assumptions A1-A2 hold. The solution to eq. (1) is given by

$$\mu(t) = \frac{\mu_0 N_0 + t(m(1 + \pi_f + \pi_r) + n)}{N_0 + t} \quad (2)$$

Proof: Using mathematical induction, the proof is immediate from eq. (1). ■

Remark 1: Using similar arguments as in Theorem 1 and Corollary 1, it can be shown that eq.(1) and eq. (2) also characterize the out-degree of the network as a recursive and non-recursive functions. Note that as time approaches infinity, the limit of the average in-degree equals

$$\mu^* = \lim_{t \to \infty} \mu(t) = m(1 + \pi_f + \pi_r) + n$$

We now show that μ^* is invariant.

Lemma 1 (Invariance of the average in-degree): Suppose that assumptions A1-A2 hold. The average in-degree μ^* is invariant.

Proof: Assume that $\mu(t) = \mu^*$ for an arbitrary t. Using eq. (1), we know that at $t+1$

$$\begin{aligned}\mu(t+1) &= \frac{(N_0 + t)\mu^* + m(1 + \pi_f + \pi_r) + n}{N_0 + t + 1} \\ &= \frac{(N_0+t)(m(1+\pi_f+\pi_r)+n)+m(1+\pi_f+\pi_r)+n}{N_0+t+1} \\ &= m(1 + \pi_f + \pi_r) + n\end{aligned}$$

Since $\mu(t+1) = \mu(t) = \mu^*$ for an arbitrary value of t, we conclude that μ^* is invariant. ■

The following theorem defines the stability of the average in-degree.

Theorem 2 (Asymptotic stability of the average in-degree): Suppose that assumptions A1-A2 hold. The average in-degree μ^* has a region of asymptotic stability equal to $\mathbb{R}_{\geq 1}$.

Proof: Consider the candidate Lyapunov function

$$\mathcal{V}(\mu(t)) := |\mu(t) - \mu^*| \quad (3)$$

The following condition is sufficient to guarantee the stability of μ^*. Using eq. (3), we know that for an arbitrary t

$$\begin{aligned}\mathcal{V}(\mu(t+1)) - \mathcal{V}(\mu(t)) &= |\mu(t+1) - \mu^*| - |\mu(t) - \mu^*| \\ &= \left|\frac{\mu_0 N_0 + (t+1)s}{N_0+t+1} - s\right| - \left|\frac{\mu_0 N_0 + ts}{N_0+t} - s\right| \\ &= \left|\frac{N_0(\mu_0 - s)}{N_0+t+1}\right| - \left|\frac{N_0(\mu_0 - s)}{N_0+t}\right| < 0\end{aligned}$$

where $s = m(1 + \pi_f + \pi_r) + n$. So $\mathcal{V}$ is a strictly decreasing function over time. Thus, the average in-degree is asymptotically stable for all $\mu_0 \geq 1$. ■

IV. Dynamics, convergence, and invariance of the in-degree distribution

Let $P_k(t) = P[k_i(t) \leq k]$ be the probability of randomly selecting a node with in-degree less than or equal to k. That is $P_k(t) = \sum_{j=0}^{k} p_j(t)$ where $p_j(t)$ is the probability of selecting a node with in-degree j at time t. The complementary cumulative in-degree distribution is given by $F(t) = \begin{bmatrix} F_0(t) & F_1(t) & \cdots & F_k(t) & \cdots \end{bmatrix}^\intercal$ where $F_k(t) = 1 - P_k(t)$. Note that

$$\mu(t) = \sum_{i=0}^{\infty} F_i(t)$$

The following theorem presents sufficient conditions to model the dynamics of the complementary cumulative in-degree distribution as a linear system.

Theorem 3 (Dynamics of the in-degree distribution): Suppose that assumptions A1-A2 hold. The complementary cumulative in-degree distribution is a time-varying infinite dimensional linear system of the form

$$F(t) = A(t)F(t-1) + B(t) \quad (4)$$

where

$$A_{kj}(t) = \begin{cases} 1 & \text{if } 0 \le k = j < n \\ \frac{v_k}{N_0+t} & \text{if } n \le k-1 = j \\ \frac{w_k}{N_0+t} & \text{if } n \le k = j \\ 0 & \text{otherwise} \end{cases}$$

represents the element of the kth row and the jth column, with

$$v_k = m\left(1 + \frac{k\pi_f}{m(1+\pi_f+\pi_r)+n}\right)$$
$$w_k = N_0 + t - 1 - m\left(1 + \frac{k\pi_f}{m(1+\pi_f+\pi_r)+n}\right)$$

and

$$B_k = \frac{1}{N_0+t-1}\sum_{i=k+1}^{\infty}\binom{m}{i-n}\pi_r^{i-n}(1-\pi_r)^{m-(i-n)}$$

represents the element of the kth row of B(t).

Proof: We know that all mechanisms affect the in-degree distribution. We first analyze the effect of mechanisms M1 and M2 which establish new incoming edges to existing nodes. In particular, we know that these mechanisms increase $F_k(t)$ when new edges are established to nodes that have in-degree k. Consider the effect of mechanism M1. We know that random attachment increases $F_k(t)$ by

$$mp_k(t-1) = m\left(F_{k-1}(t-1) - F_k(t-1)\right)$$

Now, consider the effect of mechanism M2. We know that triad formation increases the number of nodes with in-degree greater than k by

$$m\left(F_{k-1}(t-1) - F_k(t-1)\right)k\frac{\pi_f}{m(1+\pi_f+\pi_r)+n}$$

where $mp_k(t-1)k = m\left(F_{k-1}(t-1) - F_k(t-1)\right)k$ is the probability that the new node selects, during random attachment, an incoming neighbor of node i with in-degree k. The term $\frac{\pi_f}{m(1+\pi_f+\pi_r)+n}$ represents the probability of choosing node i from the out-going neighbors of node j, thereby forming a triad.

Next, consider the effect of mechanism M3 which establishes new incoming edges to the new nodes. It is clear that $n \le k_i(t_{i-N_0}) \le n+m$ where t_{i-N_0} is the time in which node $i > N_0$ attaches to the network. In particular, the probability that the in-degree of the new node at time t equals k is given by

$$b_k = \binom{m}{k-n}\pi_r^{k-n}(1-\pi_r)^{m-(k-n)}$$

Note that $b_k \neq 0$ only for $n \le k \le n+m$. Letting $\beta_k = \sum_{i=k+1}^{\infty} b_i$ represent the probability that the new node has in-degree greater than k, we know that $\beta_k = 0$ for all $k > n+m$.

Since $F_k(t)$ is the probability of finding a node with in-degree greater than k, and N_0+t indicates the number of nodes in the network at time t, the product $(N_0+t)F_k(t)$ denotes the number of nodes with in-degree greater than k. Thus, at time t, we have that for $k \ge m\pi_r + n$

$$\begin{aligned}(N_0+t)F_k(t) &= (N_0+t-1)F_k(t-1) + \beta_k \\ &+ m\left(F_{k-1}(t-1) - F_k(t-1)\right)\left(1 + k\tfrac{\pi_f}{m(1+\pi_f+\pi_r)+n}\right)\end{aligned} \quad (5)$$

Since mechanism M3 affects only the in-degree of the new nodes, according to assumption A1, we know that there are no nodes with in-degree less than n. Thus, eq. (5) is satisfied for all $k \ge n$, and $F_k(t) = 1$ for all $k < n$. The dynamics of $F(t)$ can be represented as a time-varying infinite dimensional linear system defined by eq. (4). ■

The following theorem presents sufficient conditions that guarantee the existence of the limit of $F_k(t)$ as $t \to \infty$ for all $k \ge 0$..

Theorem 4 (The limit of the in-degree distribution): Suppose that assumptions A1-A2 hold. The limit of $F_k(t)$ as t tends to infinity exists for all k.

Proof: Using induction on k, we want to show that as time goes to infinity, the limit of the complementary cumulative in-degree distribution exists. First, note that $\lim_{t\to\infty} F_k(t) = 1$ for all $k < n$. Moreover, according to eq. (4) for nodes with in-degree n, we know that $F_n(t)$ can be written as

$$F_n(t) = \frac{v_n}{N_0+t} + \frac{w_n}{N_0+t}F_n(t-1) + \frac{\beta_n}{N_0+t}$$

with initial condition $F_n(0) = f_0 \in [0,1]$. In particular,

$$F_n(t) = \frac{v_n + \beta_n + \frac{(v_n(f_0-1)+f_0-\beta_n)\Gamma(N_0+1)\Gamma(N_0+t-v_n)}{\Gamma(N_0+t+1)\Gamma(N_0-v_n)}}{1+v_n}$$

where $\Gamma(\cdot)$ represent the gamma function. Hence,

$$\lim_{t\to\infty} F_n(t) = \frac{v_n+\beta_n}{1+v_n}$$

Since $\beta_k < \infty$, $\lim_{t\to\infty} F_n(t)$ is well-defined. Now, consider the case for $n+1$. It can be shown that $\lim_{t\to\infty}\frac{F_k(t)}{F_k(t-1)} = 1$ for all k. Thus, using eq. (4), we know

$$\begin{aligned}F_{n+1}(t) &\simeq \frac{v_{n+1}}{N_0+t}F_n(t) + \frac{w_{n+1}}{N_0+t}F_{n+1}(t) + \frac{\beta_{n+1}}{N_0+t} \\ &\simeq \frac{1}{1+v_{n+1}}\left(v_{n+1}F_n(t) + \beta_{n+1}\right)\end{aligned}$$

and as t goes to infinity $\lim_{t\to\infty} F_k(t-1) = \lim_{t\to\infty} F_k(t)$, so

$$\lim_{t\to\infty} F_{n+1}(t) = \frac{v_{n+1}}{1+v_{n+1}}\lim_{t\to\infty} F_n(t) + \frac{\beta_{n+1}}{1+v_{n+1}}$$

Because we know that $\lim_{t\to\infty} F_n(t)$ exists, we can guarantee that $\lim_{t\to\infty} F_{n+1}(t)$ also exists. Now, assume that the limit of $F_k(t)$ exists for an arbitraty k. We want to show that the limit exists for $k+1$. Again, using eq. (4) we know that

$$\begin{aligned}F_{k+1}(t) &\simeq \frac{v_{k+1}}{N_0+t}F_k(t) + \frac{w_{k+1}}{N_0+t}F_{k+1}(t) + \frac{\beta_{k+1}}{N_0+t} \\ &\simeq \frac{1}{1+v_{k+1}}\left(v_{k+1}F_k(t) + \beta_{k+1}\right)\end{aligned}$$

we know that $\lim_{t\to\infty} F_k(t-1) = \lim_{t\to\infty} F_k(t)$, so

$$\lim_{t\to\infty} F_{k+1}(t) = \frac{v_{k+1}}{1+v_{k+1}} \lim_{t\to\infty} F_k(t) + \frac{\beta_{k+1}}{1+v_{k+1}} \tag{6}$$

Since the limit of $F_k(t)$ exists, we know that the limit of $F_{k+1}(t)$ also exists. Therefore, $\lim_{t\to\infty} F_k(t)$ exists for all k. ■

Remark 2: Note that $F_k^\infty = \lim_{t\to\infty} F_k(t)$ satisfies $F_k^\infty = 1$ for all $k < n$ and

$$F_k^\infty = \Big(\prod_{i=n}^{k} \frac{v_i}{1+v_i}\Big) + \frac{1}{1+v_k} \sum_{j=0}^{k-n} \beta_{k-j} \prod_{l=k-j}^{k-1} \frac{v_{l+1}}{1+v_l}$$

for all $k \geq n$.

Now, let $F(t)$be the in-degree distribution of the network and define

$$\mathcal{F}_{in} := \{F \in [0,1]^\infty : F_k = 1 \forall k < n \wedge F_k \geq F_{k+1} \forall k \geq n\}$$

Moreover, let

$$\mathcal{F}_{in}^* := \{F \in [0,1]^\infty : F_k = F_k^\infty \ \forall k\}$$

represent a particular subset of $\mathcal{F}_{in}$ defined by the in-degree limit distribution. We will show that if $F(t) \in \mathcal{F}_{in}^*$ for some $t \geq 0$, then $F(t') \in \mathcal{F}_{in}^*$ for all $t' \geq t$ (i.e., the set $\mathcal{F}_{in}^*$ is invariant).

Lemma 2 (Invariance of the in-degree distribution): Suppose that assumptions A1-A2 hold. The set $\mathcal{F}_{in}^*$ is invariant.

Proof: Assume that $F(t) \in \mathcal{F}_{in}^*$. In other words, $F_k(t) = F_k^\infty$ for all $k \geq 0$. We want to show that $F_k(t+1) = F_k(t)$ for all $k \geq 0$. First, for all $k < n$, we know that $F_k(t+1) = F_k(t) = 1$. Using eq. (4), since $F(t) \in \mathcal{F}_{in}^*$, we know that for $k = n$

$$\begin{aligned} F_n(t+1) &= \tfrac{w_n}{N_0+t+1} F_n(t) + \tfrac{v_n}{N_0+t+1} F_{n-1}(t) + \tfrac{\beta_n}{N_0+t+1} \\ &= \tfrac{w_n}{N_0+t+1} \tfrac{v_n+\beta_n}{1+v_n} + \tfrac{v_n}{N_0+t+1} + \tfrac{\beta_n}{N_0+t+1} \\ &= \frac{v_n+\beta_n}{1+v_n} \end{aligned}$$

Hence, $F_n(t+1) = F_n(t)$. Now, using eqs. (4) and (6) for a fixed k, since $F(t) \in \mathcal{F}_{in}^*$, we know that

$$\begin{aligned} F_k(t+1) &= \tfrac{w_k}{N_0+t+1} F_k(t) + \tfrac{v_k}{N_0+t+1} F_{k-1}(t) + \tfrac{\beta_k}{N_0+t+1} \\ &= \tfrac{w_k}{N_0+t+1} \tfrac{v_k F_{k-1}^\infty + \beta_k}{1+v_k} + \tfrac{v_k}{N_0+t+1} F_{k-1}^\infty + \tfrac{\beta_{k+1}}{N_0+t+1} \\ &= \frac{v_k F_{k-1}^\infty + \beta_k}{1+v_k} \end{aligned}$$

which implies that $F_k(t+1) = F_k(t)$. This holds for any k, so the set $\mathcal{F}_{in}^*$ is invariant. ■

V. Dynamics, Convergence, and Invariance of the Out-Degree Distribution

Let $\hat{P}_k(t) = P[\hat{k}_i(t) \leq k]$ be the probability of randomly selecting a node with out-degree less than or equal to k. In other words, $\hat{P}_k(t) = \sum_{j=0}^{k} \hat{p}_j(t)$, where $\hat{p}_j(t)$ represents the probability distribution function of the out-degree at time t. The complementary cumulative out-degree distribution is given by $\hat{F}(t) = \begin{bmatrix} \hat{F}_0(t) & \hat{F}_1(t) & \cdots & \hat{F}_k(t) & \cdots \end{bmatrix}^\mathsf{T}$ where $\hat{F}_k(t) = 1 - \hat{P}_k(t)$. The following theorem presents sufficient conditions that guarantee that the dynamics of $\hat{F}(t)$ represent a linear system.

Theorem 5 (Dynamics of the out-degree distribution): Suppose that assumptions A1-A2 hold. The complementary cumulative out-degree distribution can be written as a time-varying infinite dimensional linear system of the form

$$\hat{F}(t) = \hat{A}(t)\hat{F}(t-1) + \hat{B}(t) \tag{7}$$

where

$$\hat{A}_{kj}(t) = \begin{cases} 1 & \text{if } 0 \leq k = j < m \\ \frac{n}{N_0+t} & \text{if } m \leq i-1 = j \\ \frac{N_0+t-1-(m\pi_r+n)}{N_0+t} & \text{if } m \leq i = j \\ 0 & \text{otherwise} \end{cases}$$

represents the element of the kth row and the jth column, and

$$\hat{B}_k = \frac{1}{N_0+t-1} \sum_{i=k+1}^{\infty} \binom{m}{i-m} \pi_f^{i-m} (1-\pi_f)^{2m-i}$$

represents the element of the kth row of $B(t)$.

Proof: Note that the out-degree of a randomly selected node at time index t is affected by the way the network responds to the attachment of new nodes, (mechanism M3). More specifically, $\hat{F}_k(t)$ increases when existing nodes with out-degree k establish new edges to the new node. According to mechanism M1 and assumption A2, when node $j \notin \mathcal{H}(t-1)$ attaches to the network at time $t = t_j$, the probability that it connects to node $i \in \mathcal{H}(t-1)$ with $\hat{k}_i = k$ and that node i establishes a reciprocal edge is

$$m\pi_r \hat{p}_k(t-1) = m\pi_r \left(\hat{F}_{k-1}(t-1) - \hat{F}_k(t-1)\right)$$

In other words, random response increases $\hat{F}_k(t)$ when the new node connects to node $i \in \mathcal{H}(t-1)$ with out-degree $\hat{k}_i = k$ and node i establishes a reciprocal link. The random approach of mechanism M3 increases the number of nodes with out-degree k by

$$n\hat{p}_k(t-1) = n\left(\hat{F}_{k-1}(t-1) - \hat{F}_k(t-1)\right)$$

which represents the probability that a node with out-degree k establishes an edge to the new node.

Next, consider the effect of mechanisms M1-M2 on $\hat{F}_k(t)$ (i.e., on the out-degree of new nodes). In particular, we know that $\hat{F}_k(t)$ increases due to mechanism $M1$ or $M2$ when the new node has out-degree greater than k. Note that $m \leq \hat{k}_i(t_{i-N_0}) \leq 2m$ where t_{i-N_0} is the time in which node $i > N_0$ attached to the network. The probability that the out-degree of the new node at time t equals k is given by

$$\hat{b}_k = \binom{m}{k-m} \pi_f^{k-m} (1-\pi_f)^{2m-k}$$

In particular, $\hat{b}_k \neq 0$ only for $m \leq k \leq 2m$. Since $\hat{\beta}_k = \sum_{i=k+1}^{\infty} \hat{b}_i$ is the probability that the new node has out-degree greater than k, we know that $\hat{\beta}_k = 0$ for all $k > 2m$.

Since $\hat{F}_k(t)$ is the probability of selecting a node with out-degree greater than k and N_0+t indicates the number of existing nodes at time t, the product $(N_0+t)\hat{F}_k(t)$ indicates the number of nodes with out-degree greater than k at time t. So, we have

$$(N_0+t)\hat{F}_k(t) = (N_0+t-1)\hat{F}_k(t-1) + \hat{\beta}_k + (m\pi_r+n)\left(\hat{F}_{k-1}(t-1) - \hat{F}_k(t-1)\right) \tag{8}$$

According to assumption A1 and the fact that mechanisms M1-M2 only affect the out-degree of newly added nodes, we know that there are no nodes with out-degree less than m. So, eq. (8) is satisfied for all $k \geq m$, and $\hat{F}_k(t) = 1$ for all $k < m$. The dynamics of $\hat{F}(t)$ can be written as a time-varying infinite dimensional linear system of the form $\hat{F}(t) = \hat{A}(t)\hat{F}(t-1) + \hat{B}(t)$ as in eq. (7). ■

The following theorem presents sufficient conditions that guarantee that as time goes to infinity, the limit of $\hat{F}(t)$ exists.

Theorem 6 (The limit of the out-degree distribution): Suppose that assumptions A1-A2 hold. The limit of the complementary cumulative out-degree distribution exists.

Proof: Using induction on k, we want to show that $\lim_{t\to\infty} \hat{F}_k(t)$ exists for all k. According to assumption A1, there are no nodes with out-degree less than m. This implies that $\lim_{t\to\infty} \hat{F}(t) = 1$ for all $k < m$. Now, we show that the limit exists for $k = m$. Using eq. (7), we know that

$$\hat{F}_m(t) = \tfrac{m\pi_r+n}{N_0+t} + \tfrac{N_0+t-1-(m\pi_r+n)}{N_0+t}\hat{F}_m(t-1) + \tfrac{\hat{\beta}_k}{N_0+t}$$

with initial condition $\hat{F}_m(0) = \hat{f}_0 \in [0,1]$. In particular, it can be shown that

$$\hat{F}_m(t) = \frac{1}{1+m\pi_r+n}\Big(m\pi_r+n+\hat{\beta}_m + \tfrac{((m\pi_r+n)(\hat{f}_0-1)+\hat{f}_0-\hat{\beta}_m)\Gamma(N_0+1)\Gamma(N_0+t-m\pi_r-n)}{\Gamma(N_0+t+1)\Gamma(N_0-m\pi_r-n)}\Big)$$

where $\Gamma(\cdot)$ represents the gamma function. So

$$\lim_{t\to\infty} \hat{F}_m(t) = \frac{m\pi_r+n+\hat{\beta}_m}{1+m\pi_r+n}$$

Since $\hat{\beta}_m < \infty$, $\lim_{t\to\infty} \hat{F}_m(t)$ is well-defined. Next, we will analyze the limit when $k = m+1$. In particular, it can be shown that for $\lim_{t\to\infty} \frac{\hat{F}_k(t)}{\hat{F}_k(t-1)} = 1$ for all k. Using eq. (7), we know that

$$\begin{aligned}\hat{F}_{m+1}(t) &\simeq \tfrac{m\pi_r+n}{N_0+t}\hat{F}_m(t) + \tfrac{\hat{\beta}_{m+1}}{N_0+t} + \tfrac{N_0+t-1-(m\pi_r+n)}{N_0+t}\hat{F}_{m+1}(t) \\ &\simeq \frac{(m\pi_r+n)\hat{F}_m(t)+\hat{\beta}_{m+1}}{1+m\pi_r+n}\end{aligned}$$

and as t goes to infinity $\lim_{t\to\infty} \hat{F}_k(t-1) = \lim_{t\to\infty} \hat{F}_k(t)$.

$$\lim_{t\to\infty} \hat{F}_{m+1}(t) = \frac{m\pi_r+n}{1+m\pi_r+n}\lim_{t\to\infty} \hat{F}_m(t) + \frac{\hat{\beta}_{m+1}}{1+m\pi_r+n}$$

Since $\lim_{t\to\infty} \hat{F}_m(t)$ exists, we know that $\lim_{t\to\infty} \hat{F}_{m+1}(t)$ exists as well. Now, assume that for an arbitrary value of k, the limit of $\hat{F}_k(t)$ exists. We will show that the limit exists for $k+1$. Using eq. (7), we know that

$$\begin{aligned}\hat{F}_{k+1}(t) &\simeq \tfrac{m\pi_r+n}{N_0+t}\hat{F}_k(t) + \tfrac{N_0+t-1-(m\pi_r+n)}{N_0+t}\hat{F}_{k+1}(t) + \tfrac{\hat{\beta}_{k+1}}{N_0+t} \\ &\simeq \frac{(m\pi_r+n)\hat{F}_k(t)+\hat{\beta}_{k+1}}{1+m\pi_r+n}\end{aligned}$$

We also know that $\lim_{t\to\infty} \hat{F}_k(t-1) = \lim_{t\to\infty} \hat{F}_k(t)$ and

$$\lim_{t\to\infty} \hat{F}_{k+1}(t) = \frac{m\pi_r+n}{1+m\pi_r+n}\lim_{t\to\infty} \hat{F}_k(t) + \frac{\hat{\beta}_{k+1}}{1+m\pi_r+n} \tag{9}$$

Since the limit of $\hat{F}_k(t)$ exists, the limit of $\hat{F}_{k+1}(t)$ exists as well. So $\lim_{t\to\infty} \hat{F}_k(t)$ exists for all k.

Remark 3: Note that $\hat{F}_k^\infty = \lim_{t\to\infty} \hat{F}_k(t)$ satisfies $\hat{F}_k^\infty = 1$ for all $k < m$ and

$$\hat{F}_k^\infty = \left(\tfrac{m\pi_r+n}{1+m\pi_r+n}\right)^{k-m} + \sum_{j=0}^{k-m} \tfrac{\hat{\beta}_{k-j}}{1+m\pi_r+n}\left(\tfrac{m\pi_r+n}{1+m\pi_r+n}\right)^{j-1}$$

for all $k \geq m$.

■

Next, define

$$\mathcal{F}_{out} := \Big\{\hat{F} \in [0,1]^\infty : \hat{F}_k = 1 \forall k < m \tag{10}$$

$$\text{and } \hat{F}_k \geq \hat{F}_{k+1} \forall k \geq m\Big\} \tag{11}$$

Moreover, let

$$\mathcal{F}^*_{out} := \Big\{\hat{F} \in [0,1]^\infty : \hat{F}_k = \hat{F}_k^\infty \;\forall k\Big\}$$

represent a particular subset of $\mathcal{F}_{out}$ defined by the out-degree limit distribution. We will show that if $\hat{F}(t) \in \mathcal{F}^*_{out}$ for some $t \geq 0$, then $\hat{F}(t') \in \mathcal{F}^*_{out}$ for all $t' \geq t$ (i.e., the set $\mathcal{F}^*_{out}$ is invariant).

Lemma 3 (Invariance of the out-degree distribution): Suppose that assumptions A1-A2 hold. The set $\mathcal{F}^*_{out}$ is invariant

Proof: Assume that $\hat{F}(t) \in \mathcal{F}^*_{out}$. In other words, $\hat{F}_k = \hat{F}_k(\infty)$ for all $k \geq 0$. We want to show that for any t, $\hat{F}(t+1) = \hat{F}(t)$. Using eq. (7), we know that for all $k < m$, $\hat{F}_k(t+1) = \hat{F}_k(t) = 1$. Similarly, using eq. (7), since $\hat{F}(t) \in \mathcal{F}^*_{out}$, we know that for $k = m$

$$\begin{aligned}\hat{F}_m(t+1) &= \tfrac{N_0+t-(m\pi_r+n)}{N_0+t+1}\hat{F}_m(t) + \tfrac{m\pi_r+n}{N_0+t+1} + \tfrac{\hat{\beta}_m}{N_0+t+1} \\ &= \tfrac{N_0+t-(m\pi_r+n)}{N_0+t+1}\tfrac{m\pi_r+n+\hat{\beta}_m}{1+m\pi_r+n} + \tfrac{m\pi_r+n}{N_0+t+1} + \tfrac{\hat{\beta}_m}{N_0+t+1} \\ &= \frac{m\pi_r+n+\hat{\beta}_m}{1+m\pi_r+n}\end{aligned}$$

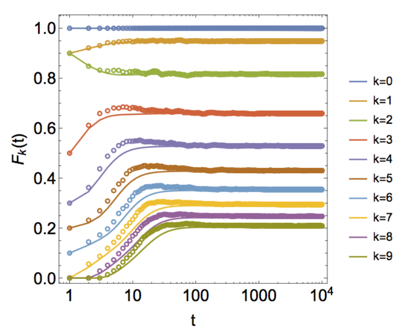
So $\hat{F}_m(t+1) = \hat{F}_m(t)$. Next, using eqs. (7) and (9) for any k, and since $\hat{F}(t) \in \mathcal{F}^*_{out}$, we know that

$$\begin{aligned}\hat{F}_{k+1}(t+1) &= \tfrac{N_0+t-(m\pi_r+n)}{N_0+t+1}\hat{F}_{k+1}(t) + \tfrac{m\pi_r+n}{N_0+t+1}\hat{F}_k(t) \\ &\quad + \tfrac{\hat{\beta}_{k+1}}{N_0+t+1} \\ &= \tfrac{N_0+t-(m\pi_r+n)}{N_0+t+1}\tfrac{(m\pi_r+n)\hat{F}_k(t)+\hat{\beta}_{k+1}}{1+m\pi_r+n} \\ &\quad + \tfrac{m\pi_r+n}{N_0+t+1}\hat{F}_k(t) + \tfrac{\hat{\beta}_{k+1}}{N_0+t+1} \\ &= \frac{(m\pi_r+n)\hat{F}_k(t)+\hat{\beta}_{k+1}}{1+m\pi_r+n}\end{aligned}$$

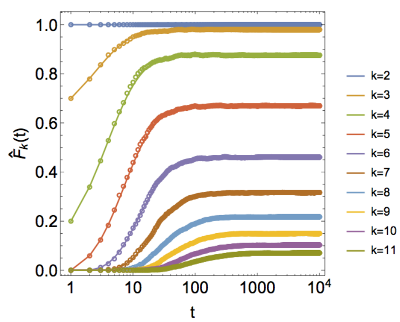
which implies that $\hat{F}_{k+1}(t+1) = \hat{F}_{k+1}(t)$. This holds for any k, so the set $\mathcal{F}^*_{out}$ is invariant. ■

VI. Simulations

To gain further insight into the network model, let $N_0 = 10$, $m = 3$, $\pi_f = 0.6$, $n = 1$, and $\pi_r = 0.4$. Figure 1 illustrates the dynamics of the complementary cumulative degree distributions for nodes with degree $k = 0, \ldots, 11$ for a random initial network satisfying assumptions A1-A2. The top plot indicates the in-degree distribution; the bottom plot the out-degree distribution. The solid lines represent the theoretical value based on eqs. (4) and (7), and the dots indicate the average of 150 simulation runs for $t = 10000$. Since there are no nodes with in-degree lower than $n = 1$, note that the solid lines in Figure 1(a) are horizontal lines with a value of one for $k = 0$. Similarly, for nodes with out-degree lower than $k < m = 3$, the theoretical value is a horizontal line with a value of one for $k = 0, 1, 2$. Note that for a small t, the theoretical predictions of the in-degree of nodes with small degrees are a better fit to the simulations.



(a)



(b)

Fig. 1. Evolution of the complementary cumulative degree distributions; (a) in-degree distribution based on eq (4); (b) out-degree distribution based on eq (7).

Figure 2 shows the evolution of the average degree distribution. The solid line represents the theoretical prediction based on eq. (2) and the dashed line represents its limit value as $t \to \infty$. The dots represent the average degree for 100 simulation runs.

VII. Conclusions and future work

This paper explains how the mechanisms of random attachment, triad formation, and network response impact

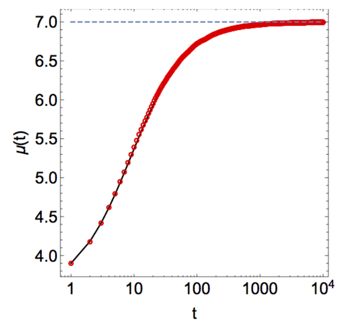


Fig. 2. Evolution of the average degree of the network.

the dynamics of the in- and out-degree distributions of growing networks. In particular, we characterize the dynamics and convergence of the average degree and shows that it is asymptotically stable. Additionally, we use infinite dimensional time-varying linear systems to characterize the evolution of the two degree distributions. Finally, we show that the dynamics of the distributions reach unique invariant sets. Evaluating whether these sets are stable is an important direction for future research.

VIII. Acknowledgements

This research was supported in part by the Center of Excellence and Appropriation in Big Data and Data Analytics (CAOBA) at the Pontificia Universidad Javeriana, the Ministry of Information Technologies and Telecommunications of Colombia (MinTIC), and the Colombian Administrative Department of Science, Technology and Innovation (COLCIENCIAS), under grant no. FP44842-546-2015.

References

[1] M. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

[2] S. L. Johnson, S. Faraj, and S. Kudaravalli, "Emergence of power laws in online communities: The role of social mechanisms and preferential attachment." *MIS Quarterly*, vol. 38, no. 3, pp. 795–808, 2014.

[3] K. Guerrero and J. Finke, "On the formation of community structures from homophilic relationships," in *Proceedings of the American Control Conference*, June 2012, pp. 5318–5323.

[4] S. N. Dorogovtsev, J. F. Mendes, and A. N. Samukhin, "Structure of growing networks with preferential linking," *Phys. Rev. Lett.*, vol. 85, no. 21, pp. 4633–4636, 2000.

[5] M. Newman, A. L. Barabasi, and D. J. Watts, *The Structure and Dynamics of Networks*. Princeton University Press, 2011.

[6] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. part I," *CoRR*, vol. abs/1701.06307, 2017.

[7] O. Hagberg and C. Wiuf, "Convergence properties of the degree distribution of some growing network models," *Bulletin of Mathematical Biology*, vol. 68, no. 6, pp. 1275–1291, 2006.

[8] P. Moriano and J. Finke, "On the formation of structure in growing networks," *Journal of Statistical Mechanisms: Theory and Experiment*, vol. 2013, no. 06, p. 06010, 2013.

[9] I. Fernandez and J. Finke, "Transitivity of growing networks," in *Proceedings of the IEEE Conference on Decision and Control*, Osaka, Japan, July 2015, pp. 1625–1630.

[10] ——, "Stability properties of reciprocal networks," in *Proceedings of the American Control Conference*, Boston, MA, July 2016, pp. 776–781.