

## Preferential Attachment as a Unique Equilibrium

Chen Avin, Avi Cohen, Pierre Fraigniaud, Zvi Lotker, David Peleg

► **To cite this version:**

Chen Avin, Avi Cohen, Pierre Fraigniaud, Zvi Lotker, David Peleg. Preferential Attachment as a Unique Equilibrium. World Wide Web Conference (WWW), 2018, Lyon, France. pp.559-568, 10.1145/3178876.3186122 . hal-01964759

**HAL Id: hal-01964759**

**<https://hal.inria.fr/hal-01964759>**

Submitted on 3 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Preferential Attachment as a Unique Equilibrium

## Extended Abstract

Chen Avin

Ben Gurion University, Beer Sheva  
avin@cse.bgu.ac.il

Avi Cohen

Weizmann Institute, Rehovot  
avi.cohen@weizmann.ac.il

Pierre Fraigniaud

CNRS and University Paris Diderot  
pierre.fraigniaud@irif.fr

Zvi Lotker

Ben Gurion University, Beer Sheva  
zvilo@bgu.ac.il

David Peleg

Weizmann Institute, Rehovot  
david.peleg@weizmann.ac.il

### ABSTRACT

This paper demonstrates that the Preferential Attachment rule naturally emerges in the context of evolutionary network formation, as the *unique* Nash equilibrium of a simple social network game. In this game, each node aims at maximizing its degree in the future, representing its social capital in the “society” formed by the nodes and their connections. This result provides additional formal support to the commonly used Preferential Attachment model, initially designed to capture the “rich get richer” aphorism. In the process of establishing our result, we expose new connections between Preferential Attachment, random walks, and Young’s Lattice.

### CCS CONCEPTS

• **Theory of computation** → **Social networks; Network formation;**

### KEYWORDS

preferential attachment; network formation games; social networks; Young’s lattice; random walks

### ACM Reference Format:

Chen Avin, Avi Cohen, Pierre Fraigniaud, Zvi Lotker, and David Peleg. 2018. Preferential Attachment as a Unique Equilibrium: Extended Abstract. In *Proceedings of The Web Conference 2018 (WWW 2018)*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3178876.3186122>

## 1 INTRODUCTION

The Preferential Attachment [1] model is definitely one of the most commonly used models for real life networks in general, and social networks in particular. The model is based on the Preferential Attachment rule, which dictates that an existing node in the network with degree  $k$  will receive a new link with probability  $k/Z$  where  $Z$  is a normalization constant. This model generates synthetic graphs with characteristics intriguingly similar to those of many existing networks. These characteristics include small diameter [4], a power-law degree sequence [5, 24], and correlation between neighbor degrees [14]. Yet, the reason why Preferential Attachment fits so well with observations is still widely open. Indeed, while

This paper is published under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW 2018, April 23–27, 2018, Lyons, France

© 2018 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC BY 4.0 License.

ACM ISBN 978-1-4503-5639-8/18/04.

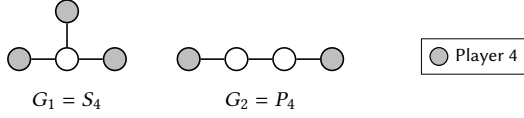
<https://doi.org/10.1145/3178876.3186122>

there have been many empirical observations [3, 21] that justify the use of Preferential Attachment in the social networks framework (including the “rich get richer” aphorism), there are little known formal analytical justifications for that usage, to our knowledge. In this paper, we make a significant step towards filling the gap between empirical explanations and formal justifications. We show that Preferential Attachment is the *only* rational choice for players involved in a simple natural network formation game.

**The wealth-based recommendation game.** We model the network evolution as a network formation game, called the *wealth-based recommendation (WBR)* game, in which the goal of each node is to maximize its expected degree. The only choice each node makes occurs when it joins the network: it consists in choosing whom to ask for a recommendation. Specifically, we consider an evolutionary process in which nodes arrive one after the other. Initially, the network consists of a single node  $v_1$ . The node  $v_t$ , arriving at discrete time  $t > 1$ , proposes itself for connection to some existing node chosen at random. The choice is made on the basis of the actual social capital of each node present in the network, modeled as its current degree. That is, node  $v_t$  chooses an index  $i \in \{1, \dots, t-1\}$  according to some probability distribution  $\pi_t$  over the current degree sequence  $D = (d_1, \dots, d_{t-1})$ ,  $d_1 \geq d_2 \geq \dots \geq d_{t-1}$ , corresponding to the nodes of the actual network. Then  $v_t$  contacts some node  $u$  of degree  $d$ , and asks for connection to  $u$ . The chosen node  $u$ , named the *host*, may either accept the connection, or delegate it to one of its neighbors. The motivation for this feature of the game is that, depending on some form of *wealth* of the current society formed by the actual network, hosts may be willing or reluctant to accept new connections (i.e., new contracts, new social, emotional or economical dependencies, etc.), since such connections induce risks. The wealth of the society at time  $t$  is captured by a parameter  $\alpha_t \in [0, 1]$ , where the higher  $\alpha_t$ , the higher the wealth. The contacted host  $u$  accepts the connection proposed by  $v_t$  with probability  $\alpha_t$ , or delegates it to one of its neighbors, with probability  $1 - \alpha_t$ . This neighbor  $w$  is chosen uniformly at random among  $u$ ’s neighbors, and  $w$  accepts the connection.

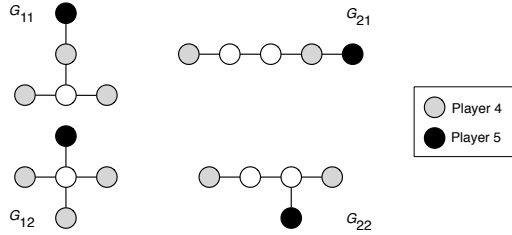
An alternative perspective of this model is the following. Each new node  $v_t$  *meets* its host  $u$  randomly according to  $\pi_t$  when it enters the network. Then it starts to perform a *simple random walk* [18] with stopping probability  $\alpha_t > 0$ . The new node connects to the node where the walk stops. For the sake of simplifying the presentation, we do not enforce recursion and only consider a one-step recommendation or a one-step random walk.

Note that the network state at time 3 is always a chain of 3 nodes, since node  $v_2$  has only a single strategy and the strategy of  $v_3$  does not change the outcome. Uncertainty only starts at time  $t \geq 4$ . Figures 1 and 2 depict the possible scenarios of the network formation after 4 and 5 steps, respectively.



**Figure 1:** The two possible states after the first four players have played, with the possible positions of the fourth node.

Note that, in Figure 2, the two graphs  $G_{11}$  and  $G_{22}$  are isomorphic. We have distinguished the two cases on the figure because the two graphs arise from two different histories: in  $G_{11}$ , the fifth node has connected to a leaf of the graph  $G_1$  depicted in Figure 1, while in  $G_{22}$ , the fifth node has connected to a leaf of the graph  $G_2$ . Note that the connection of each node may be a direct connection to the contacted node (i.e., the host), or a delegated connection to a neighbor of the contacted nodes.



**Figure 2:** The four possible states after the first five players have played, with the possible positions of the fourth and fifth nodes.

The *utility* of a node is defined as its expected degree, measuring its *social capital*. Note that the choice to whom to propose connection when a node  $v_t$  joins has no impact on the social capital of  $v_t$  in the short term, because every new node has degree 1 regardless of the chosen node with whom the connection is established. The main consideration of players is the long term affect of their choice, in light of the current state of the network (captured by its degree sequence), and future players' strategies. Thus, the game captures some important psychological features of decision-making, where individuals must balance between the past, present, and future, and where the short and long term implications of a decision must be weighed. In particular, it is not known when the game will end, and nodes should plan their strategies to fit with long term scenarios. This is consistent with concepts like "best friends forever" (BFF) in social networks, and the terminology used in traditional wedding ceremonies: one does not know when the game will end, but one plans for the long term. In other words, unknown stopping time is the natural assumption, which is consistent with the fact that, sadly or luckily, one does not know when our life will end.

To sum up, our network formation game is characterized by the sequence  $\bar{\alpha} = (\alpha_t)_{t \geq 1}$  measuring the evolution of the wealth of the

society along with time, and the ending time  $\tau$  where the utilities are measured. Both  $\bar{\alpha}$  and  $\tau$  are unknown to the players. They represent the natural uncertainty about the future, that decision making mechanisms have to cope with. The probability distribution  $\pi_t$  over the degree sequences  $(d_1, \dots, d_{t-1})$ ,  $d_1 \geq d_2 \geq \dots \geq d_{t-1}$ , defines the strategy of the  $t$ -th node joining the network. A strategy profile  $\Pi$  is defined as the collection of all strategies of the players, i.e.,  $\Pi = (\pi_t)_{t \geq 1}$ .

**Our Results.** We show that the Preferential Attachment strategy profile (defined formally later) is the *unique* Nash equilibrium of the game that is *universal*, i.e., it holds for all sequences  $\bar{\alpha} = (\alpha_t)_{t \geq 1}$  and all stopping times  $\tau$ . The uniqueness property is the striking point: the uncertainty about the future behaviors of the players makes Preferential Attachment the *only* viable choice for establishing connections when joining the network. Any other strategy profile  $\Pi$  might be a Nash equilibrium for some pairs  $(\bar{\alpha}, \tau)$ , but there must exist some pair  $(\bar{\alpha}', \tau')$  for which  $\Pi$  is not a Nash equilibrium. That is, in the game defined by the pair  $(\bar{\alpha}', \tau')$ , if there exists some player  $v_t$  at time  $t$ ,  $4 < t \leq \tau$ , that decides to deviate from the Preferential Attachment strategy, then there exists at least one player  $v_{t'}$  at time  $t' < t$  that gains (in expected degree) at time  $\tau$  by deviating from the Preferential Attachment strategy.

It is worth discussing the impact of the amount of uncertainty about the future on the uniqueness of the Nash equilibrium. First, the unnatural assumption of a finite game, i.e., with ending time  $\tau$  fixed a priori, and known to the players, would obviously result in many Nash equilibria, distinct from Preferential Attachment. In particular, the last player,  $v_\tau$ , could use any strategy, as this would have no impact on its utility, necessarily equal to 1 at the end of the game. On the other hand, the assumption that the wealth varies with time is not crucial to the existence of a unique universal Nash equilibrium. Indeed, even if the wealth of the society is constant over time, i.e.,  $\alpha_t = \alpha_{t'}$  for all  $t, t' \geq 1$ , Preferential Attachment would still be essentially the unique universal Nash equilibrium. More precisely, any universal Nash equilibrium under this assumption must satisfy that every player plays Preferential Attachment on all networks that are different from a star (i.e., an  $n$ -node tree with  $n - 1$  leaves). Hence, assuming  $\alpha$  varies over time is not central for demonstrating the relevance of Preferential Attachment. What is crucial is that the players do not know the value of  $\alpha$  (even if it is fixed over time). Nevertheless, our proofs do not require using the whole range  $[0, 1]$  of values for  $\alpha$ , and having  $\alpha$  taken from a finite set of at least two distinct values is sufficient for establishing Preferential Attachment as the unique universal Nash equilibrium. Assuming at least two values for  $\alpha$  seems, however, necessary. Indeed, we show that if  $\alpha$  is fixed a priori, say  $\alpha = 1$ , and is known to the players, then there are many universal Nash equilibria. In fact, we characterize a sub-class of such universal Nash equilibria for  $\alpha = 1$ , as an intermediate result toward demonstrating our main result.

**Related Work.** The Preferential Attachment (PA) model is a stochastic process used to randomly generate power-law degree distribution (a.k.a., scale-free) networks. It was originally proposed by Price [22] to study citation networks, and was popularized in Barabasi and Albert [1], who explained the structure of the World

Wide Web using the PA rule. A few other examples of networks well captured by the PA rule include collaboration networks, the internet, interbank payment networks, airline networks, and protein-protein interaction networks [20], among others. Hence, the PA model is a powerful tool for studying real-world networks in general, and social networks in particular, but it does not help fully explain the existence of the very property it aims at simulating, that is, the prevalence of social networks with a power law degree distribution.

Some authors approach the problem of formally justifying PA. In particular, D’souza et al. [7] showed the emergence of preferential attachment from underlying optimization mechanisms. Our objective is to approach the problem from a simpler, and perhaps more natural game theoretical perspective. The same approach was suggested for future research by Jackson [10]. Several network formation games have been defined. The one introduced by Fabrikant et al. [8] (see also [19]) reaches an equilibrium on either the clique or the star graph, both of which rarely occur in large social networks. The PageRank game (see [2, 6, 9, 13]) better fits the social network setting. In this game, the players represent web pages, and each player forms directed links to other players, with the goal of maximizing his PageRank. In particular, equilibria of the PageRank game that are insensitive to the “jumping probability” of PageRank, and therefore are universal (in a similar sense to that defined in the current paper) were presented in [9]. Nevertheless, the PageRank game assumes that the number of players is fixed, and connections between individuals may evolve with time. Instead, we consider the formation of a possibly infinitely growing society, in which connections between individuals involve a permanent commitment.

We complete this section by noticing that another family of random models that generates scale scale-free networks is based on the *copying* mechanism [11, 12, 15, 16, 25]. In these models the process by which a new node selects its connections is similar to ours, but these models (as well as the original PA) are not game-based.

## 2 FORMAL STATEMENT OF THE RESULTS

In this section we fully specify of the game presented in the introduction, and then formally state our main results.

### 2.1 The games $\text{WBR}(\bar{\alpha}, \tau)$

To formalize our game, we define the network evolution as a network formation game called the *wealth-based recommendation* (WBR) game. The evolution of the network occurs at discrete times, with one event occurring at each time step. We start at time 1 with a single node denoted  $v_1$ , and no edges. This graph is denoted by  $T^{(1)}$ . At each time step  $t > 1$ , a new player  $v_t$  arrives and adds a single edge from itself to an existing node  $v_j$  in  $T^{(t-1)}$ , and thus creates the graph  $T^{(t)}$ . Hence, at time  $t$ , the network  $T^{(t)}$  consists of a tree with  $t$  nodes (and  $t - 1$  edges). When the player  $v_t$  arrives, it contacts an existing player  $v_i$  and requests a *host*. When a host  $v_i$  is contacted, it must recommend a neighbor for  $v_t$ , which is determined as follows: with probability  $\alpha_t$ , the host recommends itself, and, with probability  $1 - \alpha_t$ , it recommends one of its own neighbors uniformly at random. Subsequently,  $v_t$  connects to the recommended node.

We now clarify the process for choosing the host. Let  $\pi_t$  denote the *strategy* of  $v_t$ , which dictates the distribution by which  $v_t$  chooses its host, for every instance of the graph  $T^{(t-1)}$ . For a node  $v_i$  in a graph  $G$ , we denote by  $\deg_G(v_i)$  its degree (i.e., number of neighbors). The *degree sequence*  $\text{DS}(G)$  of an  $n$ -node graph  $G$  is the non-increasing sequence  $(d_1, d_2, \dots, d_n)$  of its node degrees. Let  $\mathcal{D}^{(t)}$  be the set of all degree sequences of trees with  $t$  nodes. For every  $t \geq 1$ , the strategy  $\pi_t$  is a function

$$\pi_t : \mathbb{N} \times \mathcal{D}^{(t-1)} \rightarrow \mathbb{R}.$$

That is, for every degree sequence  $D \in \mathcal{D}^{(t-1)}$ , and for every node  $v_i$  of degree  $k$  in a tree  $T^{(t-1)}$  with degree sequence  $D$ , the probability that  $v_i$  is contacted at time  $t$  is  $\pi_t(k, D)$ . Note that, while the probability that a node is selected as a host at time  $t$  depends on the strategy  $\pi_t$ , the probability that a node is recommended at time  $t$  (i.e., adopts a new neighbor), depends additionally on the edge set of the tree  $T^{(t-1)}$ , and on the parameter  $\alpha_t$ . A *strategy profile*  $\Pi$  is the set of strategies for all players,  $\Pi = (\pi_t)_{t \geq 1}$ .

*Definition 2.1.* For a sequence  $\bar{\alpha} = (\alpha_t)_{t \geq 1}$  with  $\alpha_t \in [0, 1]$  for every  $t \geq 1$ , and a time  $\tau > 1$ , let  $\text{WBR}(\bar{\alpha}, \tau)$  denote the wealth-based recommendation (WBR) game that is played with  $\alpha_t$  at round  $t = 1, \dots, \tau$ , and which stops right after time  $\tau$ .

For a given game  $\text{WBR}(\bar{\alpha}, \tau)$ , let  $T^{(t)}(\Pi)$  be the random tree obtained at time  $t$  when strategy profile  $\Pi$  is used. Note that  $T^{(t)}(\Pi)$  is actually a node-labeled tree, where each node is labeled by its arrival time in the network. In the following, an  $n$ -node labeled tree is a tree with its node labeled by distinct integers in  $\{1, \dots, n\}$ . For every  $t$ -node labeled tree  $T$ , let

$$\varphi(T) = \Pr[T^{(t)}(\Pi) = T].$$

*Definition 2.2.* In the WBR game, the *utility* of  $v_i$  at time  $t$  is defined as  $\mathbf{u}_i^{(t)}(\Pi) = \mathbb{E}[\deg_{T^{(t)}(\Pi)}(v_i)] = \sum_{T \in \mathcal{T}^{(t)}} \deg_T(v_i) \cdot \varphi(T)$ , where  $\mathcal{T}^{(t)}$  is the set of all  $t$ -node labeled trees.

A strategy profile  $\Pi$  is a Nash equilibrium for a game  $\text{WBR}(\bar{\alpha}, \tau)$  if no player  $v_t$  can increase its utility by unilaterally changing its own strategy  $\pi_t$ , and a strategy profile  $\Pi$  is a *universal Nash equilibrium* if it is a Nash equilibrium for  $\text{WBR}(\bar{\alpha}, t)$  for every finite time  $\tau$  and every sequence  $\bar{\alpha}$ . More formally, for a game  $\text{WBR}(\bar{\alpha}, \tau)$ , and  $t \in \{1, \dots, \tau\}$ , let  $\Pi_{-t}$  be the  $(\tau - 1)$ -dimensional vector of strategies played by all players different from  $t$ , and let  $(\pi'_t, \Pi_{-t})$  denote the strategy profile in which  $v_t$  plays the strategy  $\pi'_t$  and all other players play according to the vector  $\Pi_{-t}$ .

*Definition 2.3.* A strategy profile  $\Pi = (\pi_t)_{t \geq 1}$  is a *universal Nash equilibrium* for the WBR game if, for every game  $\text{WBR}(\bar{\alpha}, \tau)$ , and every strategy profile  $\Pi' = (\pi'_t)_{t \geq 1}$ , we have  $\mathbf{u}_i^{(\tau)}(\pi_t, \Pi_{-t}) \geq \mathbf{u}_i^{(\tau)}(\pi'_t, \Pi_{-t})$  for every player  $t = 1, \dots, \tau$ .

### 2.2 Formal statements

Let us denote by PA the *preferential attachment* strategy, that is, PA is the individual node strategy by which, given a degree sequence  $D = (d_1, \dots, d_n)$ , a node of degree  $k \in D$  is contacted with probability

$$\text{PA}(k, D) = \frac{k}{\sum_{j=1}^n d_j}.$$

Note that, in the WBR game, the graph at time  $t$  after the insertion of the  $t$ -th node is a tree  $T^{(t)}$  with  $t$  nodes. So, regardless of the history of the game, for any  $t > 1$ ,

$$\text{PA}(k, D) = \text{PA}(k) = \frac{k}{2(t-1)}.$$

*Definition 2.4.* The *preferential attachment profile*, denoted  $\Pi_{\text{PA}}$ , is the strategy profile where players  $v_t$  play arbitrarily for  $t \leq 4$ , and players  $v_t$  play according to PA for  $t \geq 5$ .

Note that, for every strategy profile, the second player has no choice, and the choice of the third player is irrelevant (as any choice yields the same network). The assumption that the fourth player plays arbitrarily in  $\Pi_{\text{PA}}$  is a minor artifact due to technical reasons which are popping up in the proofs. The following result shows that PA is a natural strategy to play.

**THEOREM 2.5.** *The preferential attachment profile  $\Pi_{\text{PA}}$  is a universal Nash equilibrium for the WBR Game.*

The following theorem is the main result in this paper. It shows that, in fact, for rational players, there is no alternative to play PA.

**THEOREM 2.6.** *For any strategy profile  $\Pi \neq \Pi_{\text{PA}}$ ,  $\Pi$  is not a universal Nash equilibrium for the WBR Game.*

The *time-invariant* WBR game is the variant of the WBR game where the wealth parameter  $\alpha$  does not evolve with time. The following result shows that, apart from pathological scenarios, Preferential Attachment remains essentially the only strategy to be played by rational players.

**THEOREM 2.7.** *If a strategy profile  $\Pi$  is a universal Nash equilibrium for the time-invariant WBR game, then each player plays PA on every graph that is not a star, and if player  $t$  plays PA on the star  $S_{t-1}$  then all subsequent players  $t' > t$  play PA on all graphs. Moreover,*

- if  $\pi_t(t-2, \text{DS}(S_{t-1})) > \frac{1}{2}$  then, with probability 1, the players take alternate choices: player  $t-1$  chooses a leaf in the star, player  $t-2$  chooses the center, player  $t-3$  chooses a leaf, etc.;
- if  $\pi_t(t-2, \text{DS}(S_{t-1})) < \frac{1}{2}$  then, with probability 1, the players take alternate choices, where player  $t-1$  starts with choosing the star center.

The *static* WBR game is the variant of the time-invariant WBR game where the wealth parameter equals 1. To characterize universal Nash equilibria for the static WBR game, we define the following concept. A strategy  $\pi_t$  is *degree- $k$  consistent* if, for every degree- $k$  node, the probability of selecting that node is independent of the degree sequence. A strategy  $\pi_t$  is *degree consistent* if it is degree- $k$  consistent for every  $k \in \mathbb{N}$ . Note that PA is a degree consistent strategy. A strategy profile  $\Pi = (\pi_t)_{t \geq 1}$  is degree consistent if  $\pi_t$  is degree consistent for every  $t \geq 1$ . The following result shows that the universal Nash equilibria for the static WBR game are essentially degree consistent.

**THEOREM 2.8.** *Let  $\Pi$  be a universal Nash equilibrium for the static WBR game. If the strategy  $\pi_{t'}$  is degree consistent for every  $t' \in \{1, 2, \dots, t-1\}$ , and  $\pi_{t'}(k) > 0$  for every  $k \in \{1, \dots, t-1\}$ , then  $\pi_t$  is a degree consistent strategy. In particular, if every player  $t' \in \{1, 2, \dots, t-1\}$  played PA, then  $\pi_t$  is a degree consistent strategy.*

It is important to note that PA is not the only degree consistent strategy. Actually, Lemma 4.15 describes an interesting family of degree consistent strategies, which consists of PA as well as the uniform distribution (i.e., connecting each host with probability  $1/(t-1)$ ), among other distributions.

The rest of the paper is dedicated to proving these theorems. Due to lack of space, some proofs are deferred to the full version.

### 3 PREFERENTIAL ATTACHMENT IS A UNIVERSAL NASH EQUILIBRIUM

We now show that the preferential attachment (PA) profile is universal. We do this by showing an intimate connection between the WBR( $\bar{\alpha}, \tau$ ) game and the simple random walk on a graphs. First recall a well known result about the stationary distribution of the simple random walk with self-loops (e.g., see Levin et al. [17]§1.5). For any undirected graph  $G = (V, E)$  with  $n$  nodes and  $m$  edges the stationary distribution of the simple random walk with self-loop probability  $\alpha$ , for any  $0 < \alpha < 1$ , is  $\pi(v) = \frac{\deg_G(v)}{2m}$ . The interesting observation here is that the stationary distribution for the actual tree  $T$  at time  $t$  is the same as the PA strategy for  $T$ , namely, for every node  $v \in T$ ,  $\pi(v) = \text{PA}(\deg_T(v), \text{DS}(T))$ . We can now use this observation to show that *for any wealth*, the probability that a new arriving node *connects* to any given existing node is equal to the probability of choosing that node as a host if the node plays according to PA.

**LEMMA 3.1.** *Suppose that the preferential attachment profile is being used in WBR( $\bar{\alpha}, \tau$ ), and let us consider a node  $u$  joining the network  $T$  at time  $t \geq 4$ . For every sequence  $\bar{\alpha} = (\alpha_t)_{t \geq 1}$ , we have*

$$\Pr[u \text{ connects to } v \mid T] = \text{PA}(\deg_T(v), \text{DS}(T)) = \frac{\deg_T(v)}{2(t-2)}.$$

**PROOF.** The proof follows from the fact that the process by which  $u$  decides to which node it *connects* in our WBR( $\bar{\alpha}, \tau$ ) game, is identical to a one step simple random walk with self-loop probability  $\alpha_t$  and starting the walk from the distribution  $\pi_t$ . By definition, if  $\pi_t$  is the stationary distribution of  $T$  then the probability to *choose*  $v$  as a host is equal to the probability to *connect* to  $v$  (after one step of the random walk). The proof concludes by recalling that PA is the stationary distribution for  $T$ .  $\square$

We can now prove Theorem 2.5.

**PROOF OF THEOREM 2.5.** Suppose that the preferential attachment profile  $\Pi_{\text{PA}}$  is being used, and assume towards contradiction that there exists a sequence  $\bar{\alpha}$  and some player  $v_t$  for  $t \geq 4$  who could increase his utility by deviating from the strategy PA to a strategy  $\pi'_t \neq \text{PA}$ . Let  $d_t^{(s)}$  denote the random variable indicating the degree of player  $v_t$  at time  $s$ . Then  $d_t^{(t)} = 1$  and, by Lemma 3.1, for every  $s > t$  the following recurrence holds:

$$d_t^{(s)} = \begin{cases} d_t^{(s-1)} + 1 & \text{with probability } \frac{d_t^{(s-1)}}{2(s-2)}; \\ d_t^{(s-1)} & \text{otherwise.} \end{cases}$$

This recurrence is independent of the strategy  $\pi'_t$  of player  $v_t$  and of the sequence  $\bar{\alpha}$ , in contradiction to the assumption. It follows that  $\Pi_{\text{PA}}$  is a universal Nash equilibrium.  $\square$

## 4 PREFERENTIAL ATTACHMENT IS UNIQUE

We now turn to prove the main technical result of the paper, namely, that Preferential Attachment is the unique Nash Equilibrium. We first explain and prove the case of five players, and then extend the result to the general case using what we call *game event operators* and relying on an interesting connection to Young's Lattice [23].

### 4.1 Preliminaries

We make use of the following three simple lemmas.

LEMMA 4.1. *Every tree  $T$  of size  $n$  has some positive probability of occurring after  $n$  steps of the WBR game with the preferential attachment profile.*

LEMMA 4.2. *A tree  $T$  of size  $n$  with degree sequence  $(d_1, \dots, d_n)$  exists if and only if  $\sum_{i=1}^n d_i = 2(n-1)$ .*

LEMMA 4.3. *Let  $T$  be a tree with the degree sequence  $(d_1, \dots, d_n)$ , and let  $v_i$  be a node in  $T$  such that  $d_i > 1$ . Then a tree  $T'$  may be constructed with the same degree sequence as  $T$  where  $v_i$  has a neighbor of degree 1.*

### 4.2 The First Six Steps

Let  $\Pi = (\pi_i)_{i \geq 1}$  be a strategy profile, where  $\pi_i$  is the individual strategy of the  $i^{\text{th}}$  player, denoted by  $v_i$ . The strategies  $\pi_1, \pi_2$ , and  $\pi_3$  are trivial since  $\pi_1$  consists, by definition, in creating one node, while the second and third players invariably face degree distributions (0) and (1, 1), respectively. That is,  $\pi_2$  is the unique strategy of contacting the unique degree-0 node, and  $\pi_3$  is the strategy of contacting a degree-1 node (both with probability 1). After the third player has played, the graph consists of the 3-node path  $P_3$ , so player  $v_4$  can be facing only one kind of degree sequence, namely 211 (we omit parenthesis and commas for clarity). Still, player  $v_4$  has infinitely many strategies. More specifically, the strategy of  $v_4$  can be described by just one parameter  $p \in [0, 1]$ , namely, the probability with which it contacts a degree-1 node. After player  $v_4$  has played, the resulting tree can be of two forms only, denoted hereafter by  $G_1$  and  $G_2$ , where  $G_1$  is the 4-node star  $S_4$  and  $G_2$  is the 4-node path  $P_4$ , with respective degree sequences  $\text{DS}(G_1) = 3111$  and  $\text{DS}(G_2) = 2211$ . These graphs are depicted in Figure 1.

It follows that the strategy  $\pi_5$  may be described by just two parameters: the probabilities of connecting to a leaf in  $S_4$  and in  $P_4$ . After player 5 has played, the resulting tree can be of three forms only, the 5-node star  $S_5$ , the 5-node path  $P_5$ , and the unique tree  $T_5$  with degree sequence 32111 (see Figure 2). For the sake of the analysis, we distinguish two different states corresponding to  $T_5$ , depicted as  $G_{11}$  and  $G_{22}$  in Figure 2. Indeed,  $G_{11}$  and  $G_{12}$  in Figure 2 result from  $v_5$  playing on  $G_1$  in Figure 1, while  $G_{21}$  and  $G_{22}$  result from player  $v_5$  playing on  $G_2$ .

The strategy  $\pi_6$  is played on one of three graphs:  $S_5 (= G_{12})$ ,  $P_5 (= G_{21})$  and  $T_5 (= G_{11} = G_{22})$ , and may thus be described by four parameters: one parameter for each of the graphs  $S_5$  and  $P_5$ , and two parameters for the tree  $T_5$  (since it contains three distinct degrees). After player  $v_6$  has played, the resulting tree can take several forms, which we consider in the full version of the paper.

In Section 4.5 we collapse the evolution of the game further to consider only degree distributions. This new structure discards

knowledge of the topology of the graph, and the history of the game, but is easier to generalize and has several advantages, to be discussed later.

### 4.3 Proof of Theorem 2.6 for Five Players

The strategies  $\pi_1, \pi_2$ , and  $\pi_3$  are unique, as discussed earlier. The strategy of  $v_4$  can be described by just one parameter  $p \in [0, 1]$ , where  $\pi_4(1) = p$  and  $\pi_4(2) = 1 - 2p$  (note that  $\text{PA}$  corresponds to  $p = \frac{1}{4}$ ). According to our definition of the preferential attachment profile, player  $v_4$  is free to use any of the strategies available to it. Theorem 2.6 claims that, in particular,  $v_5$ , must play the  $\text{PA}$  strategy in order to ensure a universal Nash equilibrium. This is shown by the next lemma.

LEMMA 4.4. *Let  $\Pi = (\pi_i)_{i \geq 1}$  be a strategy profile. If  $\pi_5 \neq \text{PA}$ , then  $\Pi$  is not a universal Nash equilibrium. Moreover, if  $\pi_5 \neq \text{PA}$ , then  $\Pi$  is not even a Nash equilibrium for all WBR( $\bar{\alpha}, 5$ ) game (i.e., even if the games stop after time 5).*

PROOF. To prove the lemma we show that if  $\pi_5 \neq \text{PA}$ , then for every  $\pi_4$ , there exists a wealth parameter  $\alpha_5$  such that  $\pi_4$  is not the optimal strategy for  $v_4$ . After  $v_4$  has played, the resulting tree can be of two forms only, denoted earlier by  $G_1$  and  $G_2$ , corresponding to the 4-node star  $S_4$  and the 4-node path  $P_4$  (depicted in Fig. 1), with respective degree sequences  $\text{DS}(G_1) = 3111$  and  $\text{DS}(G_2) = 2211$ . By the way  $S_4$  and  $P_4$  emerge from  $S_3$  we have

$$\begin{cases} \varphi(S_4) &= 2p(1 - \alpha_4) + (1 - 2p)\alpha_4; \\ \varphi(P_4) &= 2p\alpha_4 + (1 - 2p)(1 - \alpha_4). \end{cases}$$

The strategy of  $v_5$  can be described by just two parameters  $r, s \in [0, 1]$ , where  $\pi_5(1, S_4) = r$  and  $\pi_5(1, P_4) = s$ . Note that  $\text{PA}$  corresponds to  $r = \frac{1}{6}$  and  $s = \frac{1}{6}$ . After  $v_5$  has played, the resulting tree can be of three forms only, the 5-node star  $S_5$ , the 5-node path  $P_5$ , and the unique tree  $T_5$  with degree sequence  $\text{DS}(T_5) = 32111$  (see Figure 2). Note that in accordance with our general notation, and for the sake of the analysis, we distinguish two different states corresponding to  $T_5$ , depicted as  $G_{11}$  and  $G_{22}$  in Figure 2. We have

$$\begin{cases} \varphi(G_{11}) &= \varphi(S_4) \cdot [3r\alpha_5 + (1 - 3r)(1 - \alpha_5)]; \\ \varphi(G_{12}) &= \varphi(S_4) \cdot [3r(1 - \alpha_5) + (1 - 3r)\alpha_5]; \\ \varphi(G_{21}) &= \varphi(P_4) \cdot [2s\alpha_5 + \frac{1}{2}(1 - 2s)(1 - \alpha_5)]; \\ \varphi(G_{22}) &= \varphi(P_4) \cdot [2s(1 - \alpha_5) + \frac{1}{2}(1 - 2s)(1 + \alpha_5)]. \end{cases}$$

By Definition 2.2, the utility  $u_4^{(5)}(\Pi)$  of player  $v_4$  at time 5 in strategy  $\Pi$  satisfies:

$$\begin{aligned} u_4^{(5)}(\Pi) &= \sum_{i,j \in \{1,2\}} \varphi(G_{ij}) \mathbb{E}[\text{deg}_{G_{ij}}(v_4)] \\ &= \frac{4}{3}\varphi(G_{11}) + \varphi(G_{12}) + \frac{3}{2}\varphi(G_{21}) + \varphi(G_{22}) \\ &= 1 + \frac{1}{3}\varphi(G_{11}) + \frac{1}{2}\varphi(G_{21}). \end{aligned}$$

This utility is a linear function of  $p$ , that we write as  $u_4^{(5)}(\Pi) = A + Bp$ , where

$$A = 1 + \frac{1}{3}\alpha_4 \left( 3r\alpha_5 + (1 - 3r)(1 - \alpha_5) \right) + \frac{1}{2}(1 - \alpha_4) \left( 2s\alpha_5 + \frac{1}{2}(1 - 2s)(1 - \alpha_5) \right)$$

and

$$B = \frac{2}{3}(1 - 2\alpha_4) \left( 3r\alpha_5 + (1 - 3r)(1 - \alpha_5) \right) + (2\alpha_4 - 1) \left( 2s\alpha_5 + \frac{1}{2}(1 - 2s)(1 - \alpha_5) \right).$$

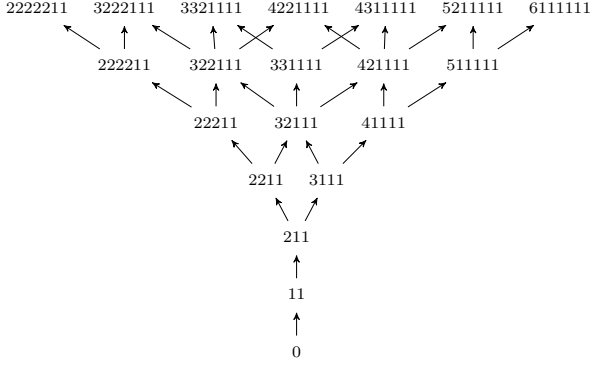


Figure 3: The lattice of degree sequences connected by  $\oplus$  events.

Assume that  $s \neq \frac{1}{6}$ , and write  $s = \frac{1}{6} + \delta$  with  $\delta \in [-\frac{1}{6}, \frac{5}{6}]$ . Let  $\alpha = \alpha_4 = \alpha_5 = \frac{1}{2} + \epsilon$  where  $|\epsilon| > 0$  is arbitrarily small. Then we have  $B = \delta\epsilon + O(\epsilon^2)$ . Therefore, if  $\delta \neq 0$ , choosing  $\alpha > \frac{1}{2}$  or  $\alpha < \frac{1}{2}$  yields the best response of player  $v_4$  to be  $p = 0$  or  $p = \frac{1}{2}$  depending on the sign of  $B$ , which is essentially governed by the sign of  $\delta\epsilon$ . Therefore, to be a universal Nash equilibrium, the strategy profile  $\Pi$  must be such that  $s = \frac{1}{6}$ . Hence, we now assume that  $s = \frac{1}{6}$ . Under the assumption that  $s = \frac{1}{6}$ , writing  $r = \frac{1}{6} + \delta$  with  $\delta \in [-\frac{1}{6}, \frac{5}{6}]$  yields  $B = -2\delta(1-2\alpha_4)(1-2\alpha_5)$ . Therefore, if  $\delta \neq 0$ , by choosing  $\alpha_4$  and  $\alpha_5$  both larger than  $\frac{1}{2}$ , or one larger than  $\frac{1}{2}$  and the other small than  $\frac{1}{2}$ , one can change the sign of  $B$ , yielding the best response of player  $v_4$  to be  $p = 0$  or  $p = \frac{1}{2}$  depending on whether  $B < 0$  or  $B > 0$ , respectively. Therefore, to be a universal Nash equilibrium, the strategy  $\Pi$  must be such that  $s = \frac{1}{6}$  and  $r = \frac{1}{6}$ . That is, player  $v_5$  must play PA.  $\square$

**Remark.** It would be tempting to use Lemma 4.4 as the basis for an induction on the arrival time  $t$  in order to establish Theorem 2.6. However, this is not as simple as it may seem at a first glance. Indeed, it turns out that the fact that player  $v_t$  does not play PA may benefit not to player  $v_{t-1}$  but to some other earlier player  $v_{t'}$  for  $t' < t - 1$ . In particular, there is an example in which the fact that  $v_6$  does not play PA benefits to  $v_4$  but not to  $v_5$ . Establishing Theorem 2.6 requires more sophisticated tools, which are presented in the next two sub-sections.

#### 4.4 Game Event Operators

The aim of this section and the next one is to concisely represent a full sequence of the game. Each player has a single turn in which it takes a simple action: choosing the degree of its host in the current graph. For any tree  $T$  of size  $n$ , let  $T \oplus k$  be a random tree of size  $n + 1$  uniformly distributed over all trees obtained from  $T$  by adding a leaf to some node of degree  $k$  in  $T$ . Let  $D = (d_1, d_2, \dots, d_n)$  be a degree sequence of size  $n$ . Let  $k$  be a degree in  $D$ , and let  $i$  be the minimal index such that  $d_i = k$ . We denote by  $D \oplus k$  the degree sequence of size  $n + 1$  obtained by increasing  $d_i$  to  $d_i + 1$ , and adding 1 at the end of the sequence  $D$ . That is,

$$D \oplus k = (d_1, d_2, \dots, d_{i-1}, k + 1, d_{i+1}, \dots, d_n, 1).$$

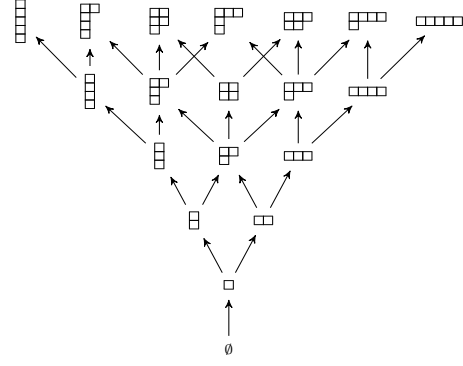


Figure 4: Young's Lattice.

We have  $DS(T \oplus k) = DS(T) \oplus k$ . Additionally, by Lemma 4.2, if  $D$  is a degree sequence of a tree then  $D \oplus k$  is a degree sequence of a tree, for every  $k$  occurring in  $D$ . Similarly,  $T \ominus k$  is a random tree of size  $n - 1$  obtained by selecting a random node  $u$  of degree  $k$ , selecting a random neighbor  $v$  of  $u$ . Then delete the edge  $(u, v)$ , select a random leaf  $w$  from the tree of  $u$ , delete it and connect  $v$  to the parent of  $w$ . Let  $D = (d_1, d_2, \dots, d_n)$  be a degree sequence of size  $n$ , let  $k > 1$  be a value occurring in  $D$ , and let  $i$  be the maximal index such that  $d_i = k$ .  $D \ominus k$  denotes the degree sequence of size  $n - 1$  obtained by decreasing  $d_i$  to  $d_i - 1$ , and by removing  $d_n = 1$  from  $D$ . That is,

$$D \ominus k = (d_1, d_2, \dots, d_{i-1}, k - 1, d_{i+1}, \dots, d_{n-1}).$$

By Lemma 4.3, if  $D$  is a degree sequence of a tree then  $D \ominus k$  is well defined for every degree  $k$  in  $D$ .

**LEMMA 4.5.** *The operations  $\oplus k$  and  $\ominus(k + 1)$  are inverse operations on degree sequences. Namely, let  $D$  and  $D'$  be degree sequences of size  $n$  and  $n + 1$ , respectively.  $D' = D \oplus k \iff D = D' \ominus (k + 1)$ .*

We say that  $D \in \mathcal{D}^{(n+1)}$  covers  $D' \in \mathcal{D}^{(n)}$ , denoted by  $D > D'$ , if  $D = D' \oplus k$  for some degree  $k$  in  $D'$ . We say  $D \in \mathcal{D}^{(n)}$  dominates  $D' \in \mathcal{D}^{(k)}$ , denoted by  $D > D'$ , if  $n > k$  and for every  $1 \leq i \leq k$ ,  $d_i \geq d'_i$ . Let  $D_1, D_2 \in \mathcal{D}^{(n)}$  be degree sequences of the same size, and let  $S = \{D' \mid D' < D_1, D_2\}$ . We say  $D$  is the meet of  $D_1$  and  $D_2$ , denoted by  $D = D_1 \wedge D_2$ , if  $D \in S$ , and, for every  $D' \neq D$  in  $S$ , it holds that  $D' < D$ . Similarly, let  $S' = \{D' \mid D' > D_1, D_2\}$ . We say  $D$  is the join of  $D_1$  and  $D_2$ , denoted by  $D = D_1 \vee D_2$ , if  $D \in S'$ , and, for every  $D' \neq D$  in  $S'$  it holds that  $D' > D$ . Two degree sequences  $D_1 \neq D_2$  of size  $n$  are called a covering pair if there exists a degree sequence that is covered by both. Similarly,  $D_1$  and  $D_2$  are called a covered pair if there exists a degree sequence that covers both.

#### 4.5 The Degree Sequence Lattice

The game events defined in Section 4.4 describe a lattice structure of degree sequences as in Figure 3. Note that this lattice is a less detailed version of the evolution described in Section 4.2, disregarding the topology of the graph, as well as the history of the game. This collapsed version allows us to visualize the game evolution in terms of the operators  $\ominus$  and  $\oplus$ . Moreover, if  $\alpha_i = 1$  for all  $i > 0$ , then a

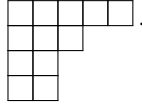
path in the degree sequence lattice uniquely defines the evolution of the game, including the chosen hosts.

The lattice of degree sequences is equivalent to a poset called *Young's lattice* (see Figure 4). We will start with a few definitions (taken from Sagan [23]) necessary to explain this equivalence.

A *partition* of  $n$ , denoted  $\lambda \vdash n$ , is a non-increasing sequence of non-negative integers  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_\ell)$  where  $\sum_{i=1}^{\ell} \lambda_i = n$ .

*Definition 4.6* ([23], p. 54). Suppose  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_\ell) \vdash n$ . The *Young diagram* of  $\lambda$  is an array of  $n$  boxes having  $\ell$  left-justified rows with row  $i$  containing  $\lambda_i$  boxes for  $1 \leq i \leq \ell$ .

As an example, the Young diagram of  $\lambda = (5, 3, 2, 2)$  is



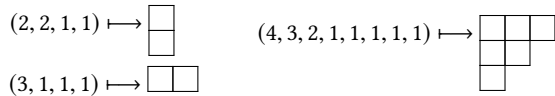
*Definition 4.7* ([23], p. 192). Let  $(A, \leq)$  be a partially ordered set. Then  $a, b \in A$  have a *greatest lower bound*, or *meet*, if there is  $c \in A$  such that

$$c \leq a, c \leq b \text{ and if } d \leq a, d \leq b, \text{ then } d \leq c. \quad (1)$$

The meet of  $a$  and  $b$ , if it exists, is denoted by  $a \wedge b$ . The *least upper bound*, or *join*, of  $a$  and  $b$  is defined by reversing all the inequalities in Eq. (1) and is written  $a \vee b$ . A poset where every pair of elements has a meet and a join is called a *lattice*.

*Definition 4.8* ([23], p. 192). The *Young lattice* is the poset of all integer partitions ordered by inclusion of their Young diagrams.

Consider a degree sequence  $D = (d_1, d_2, \dots, d_n) \in \mathcal{D}^{(n)}$ . Let us map  $D$  to the Young diagram with  $d_1 - 1$  boxes in the first row,  $d_2 - 1$  boxes in the second row, etc. For example:



By Lemma 4.2,  $D$  is a partition of  $2n - 2$ , and thus the sequence  $d_1 - 1, d_2 - 1, \dots, d_n - 1$  sums up to  $n - 2$ . Moreover, by the same lemma, every possible partition of  $n - 2$  can be similarly mapped to a degree sequence of a tree of size  $n$ . Now let  $D'$  be a degree sequence of size  $n + 1$  such that  $D' > D$ . The Young diagram that matches  $D'$  covers the Young diagram that matches  $D$ . Therefore, the lattice of degree sequences illustrated in Figure 3 is equivalent to Young's lattice.

The equivalence between the two lattices implies the following lemma, since a corresponding lemma holds for Young's lattice (see Prop. 5.1.3 in [23]). Moreover, the following subsection was mostly inspired by analogous results on Young's lattice.

**LEMMA 4.9.** *Let  $D_1 \neq D_2$  be degree sequences of size  $n$ . Then  $D_1$  and  $D_2$  are a covering pair if and only if they are a covered pair. Explicitly, there exists a degree sequence that is covered by both  $D_1$  and  $D_2$  if and only if there exists a degree sequence that covers both  $D_1$  and  $D_2$ .*

## 4.6 Nash Equilibria of the Static WBR game

In this section, we assume that the wealth  $\alpha_t$  is 1 at all steps of the game (i.e., the game is static: each node connects to the host it contacts). The result obtained in this section will provide us with the tools we need for the proof of Theorem 2.6. Moreover, characterizing the Nash equilibria of this simpler variant of the game is interesting in itself, and establish Theorem 2.8.

For a given strategy profile  $\Pi$ , a degree sequence  $D$  is called *reachable* if there exists a sequence of player actions with nonzero probability of occurring such that

$$D = D^{(1)} \oplus k_1 \oplus k_2 \oplus \dots \oplus k_n$$

where  $D^{(1)} = (0)$  is the degree sequence of the initial graph  $T^{(1)}$ , containing one isolated node. (Recall that Lemma 4.1 states that every tree is reachable when the PA strategy profile is used).

There may be Nash equilibrium for the static WBR game that are not degree consistent. For example, consider a strategy profile  $\Pi$  such that  $\pi_i(1, D) = 0$  for every degree sequence  $D$  and every  $i \geq 4$ . Namely, the players never choose to connect to a leaf. Then  $\Pi$  is not degree consistent, yet it is a Nash equilibrium for the static WBR game. As another example, let  $P_n$  denote the path of length  $n$ . The path  $P_n$  is the only graph with the degree sequence  $(2, 2, \dots, 2, 1, 1)$ , so with a slight abuse of notation, we denote the degree sequence of  $P_n$  as  $P_n$  as well. If  $\pi_i(1, P_{i-1}) = 1/2$  and  $\pi_i(1, D) = 0$  for all  $D \neq P_{i-1}$  then  $\Pi$  is a Nash equilibrium for the static WBR game.

Our next goal is to establish a sufficient condition for a strategy profile that is a Nash equilibrium for the static WBR game to be degree consistent. The following lemma will be used later, and contributes to establish part of the statement on Theorem 2.8.

**LEMMA 4.10.** *Let  $\Pi$  be a Nash equilibrium for the static WBR game, where  $\pi_i = \text{PA}$  for all  $i \in \{1, \dots, t-1\}$ . Then the strategy  $\pi_t$  is degree consistent.*

In order to prove Lemma 4.10 we will need several helping lemmas. Let  $\Gamma_n$  be a graph with the set of vertices  $V(\Gamma_n) = \mathcal{D}^{(n)}$ , the set of degree sequences of size  $n$  trees, and the set of edges

$$E(\Gamma_n) = \{(D_1, D_2) \mid D_1 \text{ and } D_2 \text{ are a covering pair}\}.$$

**LEMMA 4.11.** *The graph  $\Gamma_n$  is connected for all  $n$ .*

**PROOF.** By induction on  $n$ . For  $n = 4$ ,  $V(\Gamma_4) = \{(2211), (3111)\}$ , and both degree sequences cover (211), so there is an edge between them. Suppose the statement holds for  $n = k$ , and consider  $n = k + 1$ .

Let  $\lambda, \mu \in \mathcal{D}^{(k+1)}$ . Let  $\lambda' \in \mathcal{D}^{(k)}$  be a degree sequence such that  $\lambda' < \lambda$ , and let  $\mu' \in \mathcal{D}^{(k)}$  be a degree sequence such that  $\mu' < \mu$ . By the induction hypothesis there exists a path between  $\lambda'$  and  $\mu'$  in  $\Gamma_k$ . Let that path be  $\lambda' = \lambda'_0, \lambda'_1, \dots, \lambda'_\ell = \mu'$ .

For every  $i = 0, \dots, \ell - 1$ , the degree sequences  $\lambda'_i$  and  $\lambda'_{i+1}$  have an edge between them in  $\Gamma_k$ , hence they are a covering pair. Therefore, by Lemma 4.9 they are also a covered pair, so they are covered by a degree sequence in  $\mathcal{D}^{(k+1)}$ , which we call  $\lambda_i$ , for  $i = 0, \dots, \ell - 1$ .

Now for every  $i = 0, \dots, \ell - 1$ , the elements  $\lambda_i$  and  $\lambda_{i+1}$  both cover the element  $\lambda'_{i+1}$ , hence they are a covering pair, and thus have an edge between them. Moreover,  $\lambda$  and  $\lambda_0$  both cover  $\lambda'$  (and possibly  $\lambda = \lambda_0$ ), and  $\lambda_{\ell-1}$  and  $\mu$  both cover  $\mu'$ . Therefore there exists a path between  $\lambda$  and  $\mu$  in  $\Gamma_n$ .  $\square$



LEMMA 4.12. *Let  $\Pi$  be a Nash equilibrium for the static WBR game where  $\pi_i = \text{PA}$  for all  $i \in \{1, \dots, t-1\}$ . Then the strategy  $\pi_t$  is a degree-1 consistent strategy.*

PROOF. Consider a covering pair  $D_1, D_2 \in \mathcal{D}^{(t-1)}$ . They both cover some degree sequence  $D' \in \mathcal{D}^{(t-2)}$ . By definition, there exist some  $k_1 \in D_1$  and  $k_2 \in D_2$  such that  $D_1 \ominus k_1 = D_2 \ominus k_2 = D'$ . The player  $v_{t-1}$  may deterministically choose between creating  $D_1$  and  $D_2$  when it sees the degree sequence  $D'$ . Hence, if  $\pi_t(1, D_1) > \pi_t(1, D_2)$  then the player  $v_{t-1}$  would gain by deviating from PA to take the action  $D' \oplus (k_1 - 1)$  with probability one, contradicting the assumption that  $\Pi$  is a Nash equilibrium for the static WBR game. Moreover, because the players  $v_1, \dots, v_{t-2}$  all play PA, by Lemma 4.1, every  $D' \in \mathcal{D}^{(t-2)}$  is reachable at time  $t-2$ . So it must hold that for every  $D_1, D_2 \in E(\Gamma_{t-1})$ ,

$$\pi_t(1, D_1) = \pi_t(1, D_2). \quad (2)$$

Now let us consider any two degree sequences  $\lambda, \mu \in \mathcal{D}^{(t-1)}$ , which are not necessarily neighbors in  $\Gamma_{t-1}$ . The graph  $\Gamma_{t-1}$  is connected by Lemma 4.11, so there exists a path in  $\Gamma_{t-1}$  from  $\lambda$  to  $\mu$ . Let the path be  $\lambda = \lambda_0, \lambda_1, \dots, \lambda_\ell = \mu$ . By Equation (2),  $\pi_t(1, \lambda_i) = \pi_t(1, \lambda_{i+1})$  for every  $i = 0, \dots, \ell-1$ . Therefore,  $\pi_t(1, \lambda) = \pi_t(1, \mu)$  for every two degree sequences  $\lambda$  and  $\mu$  in  $\mathcal{D}^{(t-1)}$ . In other words,  $\pi_t$  is a degree-1 consistent strategy.  $\square$

LEMMA 4.13. *Let  $\Pi$  be a strategy profile in which  $\pi_i = \text{PA}$  for all  $i \in \{1, \dots, t-1\}$  and  $\pi_t$  is a degree- $d$  consistent strategy, for  $d = 1, \dots, k$ . Then none of the players  $v_{t-1}, v_{t-2}, \dots, v_{t-k}$  can change their utility by unilaterally deviating from  $\Pi$ .*

**Remark.** Conversely, provided the assumptions of the lemma above hold, and provided that  $\pi_t$  is not degree- $(k+1)$  consistent, then the player  $v_{t-k-1}$  can profit by deviating from the PA strategy as we show in the next lemma.

LEMMA 4.14. *Let  $\Pi$  be a Nash equilibrium for the static WBR game where  $\pi_i = \text{PA}$  for all  $i \in \{1, \dots, t-1\}$ . Then the strategy  $\pi_t$  is a degree- $k$  consistent strategy, for every  $k \in \{1, 2, \dots, t-4\}$ .*

PROOF. We prove the lemma by induction on  $k$ . The statement is true for  $k = 1$  by Lemma 4.12, so assume it holds for  $k \leq \ell$ , and consider  $k = \ell + 1$ . By the induction hypothesis, the strategy  $\pi_t$  is degree  $1, \dots, \ell$  consistent. Thus, by Lemma 4.13, the players  $v_{t-1}, v_{t-2}, \dots, v_{t-\ell}$  have no reason to deviate from PA. Let  $I = t - \ell - 1$ , and define the event

$$\mathcal{E} = \{\text{players } v_{t-1}, v_{t-2}, \dots, v_{t-\ell} \text{ all connected to the node } v_I\}.$$

If  $\mathcal{E}$  does not occur, then player  $v_I$  has degree at most  $\ell$  in  $T^{(t-1)}$  and thus  $v_I$  cannot affect its own utility (since  $\pi_t$  is degree  $1, \dots, \ell$  consistent). Furthermore, the probability of event  $\mathcal{E}$  is constant. Namely,

$$\Pr[\mathcal{E}] = \prod_{j=1}^{\ell} \text{PA}_{t-\ell-1+j}(j) = \prod_{j=1}^{\ell} \frac{j}{2(t-\ell-3+j)}.$$

It is left for player  $v_I$  to consider only the cases in which its degree is  $\ell+1$  in the graph  $T^{(t-1)}$ . That is, for every degree sequence  $D \in \mathcal{D}^{(t-1)}$  that may occur at the beginning of player  $v_I$ 's turn, and for every degree  $k \in D$  that  $v_I$  selects in its turn,  $v_I$  should

consider the possibility that all subsequent players connect to it. This scenario would generate the following degree sequence:

$$D \oplus k \oplus 1 \oplus 2 \oplus \dots \oplus \ell.$$

Let  $f : \mathcal{D}^{(t)} \rightarrow \{D \in \mathcal{D}^{(t-1)} \mid (\ell+1) \in D\}$  be the following function:

$$f(D) = D \oplus 1 \oplus 2 \oplus \dots \oplus \ell.$$

Hence in the above scenario, for degree  $k \in D$  chosen by player  $v_I$  in its turn, the resulting degree sequence is  $f(D) \oplus k = f(D \oplus k)$ .

Now for any two distinct degrees  $k_1, k_2 \in D$  that  $I$  considers to select in its turn,  $I$  should consider the following degree sequences in  $\mathcal{D}^{(t-1)}$ :

$$D_1 = f(D \oplus k_1) \text{ and } D_2 = f(D \oplus k_2).$$

If  $\pi_t(\ell+1, D_1) > \pi_t(\ell+1, D_2)$ , then the player  $v_I$  can improve its utility by deviating from PA to select the action  $D \oplus k_1$  whenever the PA strategy prescribes the action  $D \oplus k_2$ . Thus,  $\Pi$  is a Nash equilibrium for the static WBR game only if

$$\pi_t(\ell+1, D_1) = \pi_t(\ell+1, D_2). \quad (3)$$

Accordingly, Eq. (3) translates to:

$$\pi_t(\ell+1, f(D \oplus k_1)) = \pi_t(\ell+1, f(D \oplus k_2))$$

for every  $D \in \mathcal{D}^{(t-1)}$ , and every  $k_1, k_2 \in D$ . Equivalently, we may say that for every edge  $(D_1, D_2)$  in the graph  $\Gamma_t$ ,

$$\pi_t(\ell+1, f(D_1)) = \pi_t(\ell+1, f(D_2)) \quad (4)$$

By Lemma 4.11, the graph  $\Gamma_t$  is connected. Therefore, as in the proof of Lemma 4.12, Eq. (4) must hold for every two degree sequences  $D_1$  and  $D_2$  in the graph  $\Gamma_t$ . In other words,  $\pi_t$  is a degree- $(\ell+1)$  consistent strategy over all degree sequences in the range of the function  $f$ . So if we show that  $f$  is onto  $\{D \in \mathcal{D}^{(t-1)} \mid (\ell+1) \in D\}$ , we are done. Thus, let  $D \in \mathcal{D}^{(t-1)}$  be a degree sequence that contains a node of degree  $\ell+1$ . We define the function  $g : \{D \in \mathcal{D}^{(t-1)} \mid (\ell+1) \in D\} \rightarrow \mathcal{D}^{(t)}$  as follows:

$$g(D) = D \ominus (\ell+1) \ominus \ell \ominus \dots \ominus 2.$$

We observe that the function  $g$  is well defined. For every degree sequence  $D$  in  $\{D \in \mathcal{D}^{(t-1)} \mid (\ell+1) \in D\}$ : the degree sequence  $D \ominus (\ell+1)$  contains a node of degree  $\ell$ , the degree sequence  $D \ominus (\ell+1) \ominus \ell$  contains a node of degree  $\ell-1$ , the degree sequence  $D \ominus (\ell+1) \ominus \ell \ominus (\ell-1)$  contains a node of degree  $\ell-2$ , etc. By applying Lemma 4.5 repeatedly, we get that for every  $D$  in  $\{D \in \mathcal{D}^{(t-1)} \mid (\ell+1) \in D\}$

$$\begin{aligned} f(g(D)) &= \left[ (D \ominus (\ell+1) \ominus \ell \ominus \dots \ominus 3) \ominus 2 \right] \oplus 1 \oplus 2 \oplus \dots \oplus \ell \\ &= \left[ (D \ominus (\ell+1) \ominus \ell \ominus \dots \ominus 4) \ominus 3 \right] \oplus 2 \oplus 3 \oplus \dots \oplus \ell \\ &= \dots = D \ominus (\ell+1) \oplus \ell = D. \end{aligned}$$

It follows that the function  $f$  is onto  $\{D \in \mathcal{D}^{(t-1)} \mid (\ell+1) \in D\}$ , and thus Eq. (4) implies that  $\pi_t$  is a degree  $\ell+1$  consistent strategy.  $\square$

PROOF OF LEMMA 4.10. By Lemma 4.14 the strategy  $\pi_t$  is a degree- $k$  consistent strategy, for every  $k \in \{1, 2, \dots, t-4\}$ . Degrees larger than  $t-2$  are impossible because, at the beginning of step  $t$ , the graph contains exactly  $t-2$  edges, so it is left to consider the degrees  $t-2$  and  $t-3$ . The issue is that the only nodes that can be of degree  $t-2$  or  $t-3$  at time  $t-1$  are nodes 1, 2 and 3, which

have no meaningful choice in the game and thus have no effect on Nash equilibria. However, every possible strategy  $\pi'_t$  at time  $t$  is degree- $(t-2)$  and degree- $(t-3)$  consistent. Because for  $D \in \mathcal{D}^{(t-1)}$ , it holds that  $t-2 \in D$  only if  $D$  is the degree sequence of the star graph, and  $t-3 \in D$  only if  $D = (t-3, 2, 1, \dots, 1)$ . Since both of those degrees occur in only one degree sequence, there is no way for the strategy  $\pi'_t$  to select nodes of those degrees with probabilities that are inconsistent over different degree sequences.  $\square$

To establish the first part of the statement of Theorem 2.8, notice that the proof for Lemma 4.10 applies in this case as well. The assumption that every player up to  $t$  plays PA implied two properties of the strategy profile that were needed in the proof of the theorem. The first property is that  $\pi_i$  is degree consistent for every  $i \in \{1, \dots, t-1\}$ , and the second property is that every degree sequence in  $\bigcup_{i=1}^{t-1} \mathcal{D}^{(i)}$  is reachable. The first property applies here by definition. The second property applies since every possible action at steps  $1, \dots, t-1$  has a positive probability of occurring.

#### 4.7 Proof of Theorem 2.6

We need one last preliminary result.

LEMMA 4.15. *If  $\pi_n$  is a degree consistent strategy, then  $\pi_n(1)$ , the probability of contacting a leaf, determines the value of  $\pi_n(k)$  for all  $k > 1$ . Moreover, if  $\pi_n(1) = \frac{1}{2(n-2)}$  then  $\pi_n = \text{PA}$ .*

PROOF. Let  $\pi_n(1) = x$  (note that necessarily  $x \leq 1/(n-2)$ ). We may calculate the probability of contacting a node of degree 2 by considering the path graph  $P_{n-1}$ . Contact probabilities must sum up to 1 in each graph, so  $\pi_n(2) = (1-2x)/(n-3)$ . For every  $k = 3, \dots, n-2$  we may calculate the probability of contacting a node of degree  $k$  by considering a graph with the degree sequence  $(k, 2^{n-k-2}, 1^k)$  to get:

$$\pi_n(k) = 1 - (n-k-2) \cdot \frac{1-2x}{n-3} - k \cdot x. \quad (5)$$

One can verify that for every degree sequence  $D = (d_1, \dots, d_{n-1})$  in  $\mathcal{D}^{(n-1)}$ ,  $\pi_n$  defines a probability distribution over the nodes.

The second part of the lemma follows by direct calculation of  $\pi_n(k)$  for every  $k$  when  $x = \frac{1}{2(n-2)}$ .  $\square$

PROOF OF THEOREM 2.6. Let  $\Pi$  be a strategy profile. We show that if  $\Pi$  is not the preferential attachment profile, then  $\Pi$  is not a universal Nash equilibrium. Let  $n$  be the first player that does not play according to PA. From Lemma 4.4 if  $n = 5$  we are done so assume  $n > 5$  and so player  $v_{n-1}$  must play PA.

Let the graph at player  $v_{n-1}$ 's turn be  $P_{n-2}$ , the path of length  $n-2$ . Note that the graph  $P_{n-2}$  is reachable by Lemma 4.1. Let  $\pi'_{n-1}$  be the strategy in which  $v_{n-1}$  selects a node of degree 1 with probability one, and let  $\pi''_{n-1}$  be the strategy in which  $v_{n-1}$  selects a node of degree 2 with probability one. Let us denote by  $T_{n-1}$  the graph of  $n-1$  nodes with degree sequence  $(3, 2, \dots, 2, 1, 1, 1)$ , which results from  $v_{n-1}$  connecting to a node of degree 2 in  $P_{n-2}$ . Let  $\alpha_{n-1} = 1$ , then the strategies  $\pi'_{n-1}$  and  $\pi''_{n-1}$  result in the graphs  $P_{n-1}$  and  $T_{n-1}$ , respectively.

Since we assumed player  $v_{n-1}$  plays a mixed strategy and  $\Pi$  is a Nash equilibrium for every  $\alpha_{n-1}$ , it follows that

$$\mathbf{u}_{n-1}^{(n)}(\pi'_{n-1}, \Pi_{-(n-1)}) = \mathbf{u}_{n-1}^{(n)}(\pi''_{n-1}, \Pi_{-(n-1)}),$$

By Lemma 4.10,  $\pi_n$  is a degree consistent strategy, so we may rewrite this equation as

$$1 + \alpha_n \pi_n(1) + \frac{1}{2}(1 - \alpha_n) \pi_n(2) = 1 + \alpha_n \pi_n(1) + \frac{1}{3}(1 - \alpha_n) \pi_n(3),$$

In turn this yields (for  $\alpha_n \neq 1$ )

$$\frac{1}{2} \pi_n(2) = \frac{1}{3} \pi_n(3). \quad (6)$$

Plugging Equation (5) into Equation (6) yields

$$\pi_n(1) = \frac{1}{2(n-2)},$$

and thus the strategy  $\pi_n$  must equal PA by Lemma 4.15, so we reach a contradiction. This completes the proof of Theorem 2.6.  $\square$

## 5 CONCLUSION

This paper is, up to our knowledge, the first paper providing formal game-theoretic analytical arguments giving hints of why preferential attachment inherently pops up in the context of social networks. In short, there is no alternative: one must play preferential attachment for maximizing our social capital. This holds for a natural game capturing essential aspects of how connections between individuals are created in some realistic contexts.

This work can be extended in many ways. First, one may ask what if the recommendation proceeds recursively, that is, the node recommended by the host may in turn recommend one of its neighbors with certain probability, and so on. Actually, the same argument as the one developed in this paper shows that PA remains a universal Nash equilibrium in this case too. A more challenging extension of this work is when each new node connects to  $m > 1$  existing nodes. The resulting networks are not trees anymore, and extending our results to this framework appears to be non trivial. Finally, extensions where, in addition to node-events, edge-events could also occur, creating new connections between already existing nodes, or where the players have more knowledge about the actual structure than just its degree sequence, are challenging research directions which are definitely worth being investigated.

## ACKNOWLEDGMENTS

The authors would like to thank the anonymous referees for their valuable comments and helpful suggestions.

Pierre Fraigniaud received additional support from ANR under project DESCARTES - Grant No: ANR-16-CE40-0023 (<http://www.agence-nationale-recherche.fr/Projet-ANR-16-CE40-0023>), and from Inria under Project GANG (<https://www.inria.fr/equipes/gang>).

Zvi Lotker received additional support from LabEx Sciences Mathématiques de Paris (SMP). This work was partially done when he was visiting IRIF, CNRS and University Paris Diderot, under the framework of CNRS LIA FILOFOCS.

## REFERENCES

- [1] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.
- [2] David Avis, Kazuo Iwama, and Daichi Paku. Verifying nash equilibria in pagerank games on undirected web graphs. In *International Symposium on Algorithms and Computation*, pages 415–424. Springer, 2011.
- [3] Albert-László Barabási. *Network science*. Cambridge university press, 2016.
- [4] Béla Bollobás and Oliver Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.
- [5] Béla Bollobás, Oliver Riordan, Joel Spencer, Gábor Tusnády, et al. The degree sequence of a scale-free random graph process. *Random Structures & Algorithms*, 18(3):279–290, 2001.
- [6] Wei Chen, Shang-Hua Teng, Yajun Wang, and Yuan Zhou. On the alpha-sensitivity of nash equilibria in pagerank-based network reputation games. In *FAW*, pages 63–73. Springer, 2009.
- [7] Raissa M D’souza, Christian Borgs, Jennifer T Chayes, Noam Berger, and Robert D Kleinberg. Emergence of tempered preferential attachment from optimization. *Proceedings of the National Academy of Sciences*, 104(15):6112–6117, 2007.
- [8] Alex Fabrikant, Ankur Luthra, Elitza Maneva, Christos H Papadimitriou, and Scott Shenker. On a network creation game. In *Proceedings of the twenty-second annual symposium on Principles of distributed computing*, pages 347–351. ACM, 2003.
- [9] John Hopcroft and Daniel Sheldon. Network reputation games. Technical report, 2008.
- [10] Matthew O Jackson. The economics of social networks. In *Proceedings of the 9th World Congress of the Econometric Society*, vol. 1, pages 1–56. Cambridge University Press, 2006.
- [11] Matthew O Jackson and Brian W Rogers. Meeting strangers and friends of friends: How random are social networks? *The American economic review*, 97(3):890–915, 2007.
- [12] Jon Kleinberg, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew Tomkins. The web as a graph: measurements, models, and methods. *Computing and combinatorics*, pages 1–17, 1999.
- [13] Georgios Kouroupas, Evangelos Markakis, Christos Papadimitriou, Vasileios Rigas, and Martha Sideri. The web graph as an equilibrium. In *International Symposium on Algorithmic Game Theory*, pages 203–215. Springer, 2015.
- [14] Paul L Krapivsky and Sidney Redner. Organization of growing random networks. *Physical Review E*, 63(6):066123, 2001.
- [15] Pavel L Krapivsky and Sidney Redner. Network growth by copying. *Physical Review E*, 71(3):036118, 2005.
- [16] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, D Sivakumar, Andrew Tomkins, and Eli Upfal. Stochastic models for the web graph. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, pages 57–65. IEEE, 2000.
- [17] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. American Mathematical Soc.
- [18] L. Lovász. Random walks on graphs: A survey. In *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, volume 2 of *Bolyai Soc. Math. Stud.*, pages 353–397. János Bolyai Math. Soc., Budapest, 1996.
- [19] Eli A Meirum, Shie Mannor, and Ariel Orda. Network formation games with heterogeneous players and the internet structure. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 735–752. ACM, 2014.
- [20] Mark Newman. *Networks: an introduction*. Oxford university press, 2010.
- [21] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
- [22] Derek de Solla Price. A general theory of bibliometric and other cumulative advantage processes. *Journal of the Association for Information Science and Technology*, 27(5):292–306, 1976.
- [23] Bruce Sagan. *The symmetric group: representations, combinatorial algorithms, and symmetric functions*, volume 203. Springer Science & Business Media, 2013.
- [24] Remco Van Der Hofstad. *Random graphs and complex networks*, 2016.
- [25] Alexei Vázquez. Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Physical Review E*, 67(5):056104, 2003.