Homophily and the Glass Ceiling Effect in Social Networks

Chen Avin [*†]
Ben Gurion University of the
Negev, Israel
avin@cse.bgu.ac.il

Barbara Keller [‡]
ETH Zurich, Switzerland
barbara.keller@tik.ee.ethz.ch

Zvi Lotker [*]
Ben Gurion University of the
Negev, Israel
zvilo@cse.bgu.ac.il

Claire Mathieu
CNRS, École Normale
Supérieure, France
clairemmathieu@gmail.com

David Peleg [*]
Weizmann Institute, Israel
david.peleg@weizmann.ac.il

Yvonne-Anne Pignolet
ABB Corporate, Switzerland
yvonne-anne.pignolet@ch.abb.com

ABSTRACT

The glass ceiling effect has been defined in a recent US Federal Commission report as "the unseen, yet unbreakable barrier that keeps minorities and women from rising to the upper rungs of the corporate ladder, regardless of their qualifications or achievements". It is well documented that many societies and organizations exhibit a glass ceiling. In this paper we formally define and study the glass ceiling effect in social networks and propose a natural mathematical model, called the *biased preferential attachment* model, that partially explains the causes of the glass ceiling effect. This model consists of a network composed of two types of vertices, representing two sub-populations, and accommodates three well known social phenomena: (i) the "rich get richer" mechanism, (ii) a minority-majority partition, and (iii) homophily. We prove that our model exhibits a strong moment glass ceiling effect and that all three conditions are necessary, i.e., removing any one of them will prevent the appearance of a glass ceiling effect. Additionally, we present empirical evidence taken from a mentor-student network of researchers (derived from the DBLP database) that exhibits both a glass ceiling effect and the above three phenomena.

Categories and Subject Descriptors

G.2.2 [**Mathematics of Computing**]: DISCRETE MATHEMATICS—*Graph Theory*; J.4 [**Computer Applications**]: SOCIAL AND BEHAVIORAL SCIENCES

Keywords

social networks; homophily; glass ceiling

---

1. INTRODUCTION

Attaining *equality of opportunity* is a fundamental value in democratic societies, therefore existing inequalities present us with a major concern. A particularly sore example is that many highly-qualified women and members of minority groups are unable to realize their full potential in society (and specifically in the workforce) due to a phenomenon commonly referred to as the *glass ceiling*, a powerful visual image for an invisible barrier blocking women and minorities from advancing past middle management levels [20]. This concern was raised in a recent US Federal commission report [18]:

> The "glass ceiling"... is the unseen, yet unbreakable barrier that keeps minorities and women from rising to the upper rungs of the corporate ladder, regardless of their qualifications or achievements.

The existence of the glass ceiling effect is well documented [8, 16, 30]. In academia, for example, gender disparities have been observed in the number of professors [34], earnings [13, 34, 40] funding [29] and patents [10]. A recent study [26] analyzed gender differences in research output, research impact and collaborations based on Thomson Reuters Web of Science databases. When prominent author positions were analyzed by sole authorship, first-authorship and last-authorship, it was discovered that papers with women in those leading roles were less frequently cited. The question we focus on in this article concerns the causes of this phenomenon. What are the invisible mechanisms that combine to create the glass ceiling effect, and in particular, what is the role of the social network in creating this effect? Many papers discuss possible causes of the glass ceiling effect and potential solutions to it, e.g., [9, 15, 24], but to the best of our knowledge, the present work is the first attempt to study it in the context of the social network structure and to propose a *mathematical model* capturing this phenomenon.

The paper's main contributions are the following. (1) We propose a model for bi-populated social networks extending the classical preferential attachment model [2], and augment it by including two additional basic phenomena, namely, a minority-majority partition, and homophily. (2) We propose a formal definition for the glass ceiling effect *in social networks*. (3) We rigorously analyze this extended model and establish its suitability as a possible mechanism for the emergence of a glass ceiling effect. We also show that omitting any one of the three ingredients of our model prevents the occurrence of a glass ceiling effect. (4) We present empirical

evidence for a network exhibiting preferential attachment, minority-majority partition, homophily, and a glass ceiling effect.

In order to talk about the glass ceiling effect we have to agree on a measure of success in a social network. Following the traditional approach that sees network edges as the "social capital" of the network, we define successful members of a social network to be high degree vertices, namely, vertices that maintain a large number of connections, corresponding to high influence. We base our model on a bi-populated network augmented by three well-accepted observations on human behavior, namely (i) the "rich get richer" mechanism, (ii) minority-majority partition (with a slower growth rate of the minority group in the network), and (iii) homophily (affinity towards those similar to oneself). The main result of the paper is that under these three simple and standard assumptions the glass ceiling effect naturally arises in social networks. Let us first briefly describe these three social phenomena.

The "rich get richer" mechanism. This mechanism describes and explains the process of wealth concentration. It follows the basic idea that newly created wealth is distributed among members of society in proportion to the amount they have already amassed. In our setting, where the degree of the vertex captures its level of social wealth, this mechanism predicts that people may try to connect more often to people who already have many connections, either in order to profit from their social wealth or because they are more visible in the network.

Minority-majority partition. Many social groups exhibit unequal proportions of men and women. Certain occupations, such as construction, law enforcement, politics and computer science, tend to have a higher proportion of men. For example, the ratio of women taking up studies in the computing discipline varies per year and region between 10% and 35% [3, 21, 38, 44]. Other professions, such as elementary school teaching, nursing, and office administration, are occupied by a higher proportion of women. In fact, it is difficult to find an occupation with a balanced ratio of genders (this also holds for many other social partitions, e.g., ones based on ethnicity or family background). This imbalance is the second phenomenon underlying our model.

Homophily. It is a well established social phenomenon that people tend to associate with others who are similar to themselves. Characteristics such as gender, ethnicity, age, class background and education influence the relationships among human beings [27] and similarities make communication and relationship formation easier.

In summary, our model is obtained by applying the classical preferential attachment model (see Barabasi and Albert [2]) to a bi-populated minority-majority network augmented with homophily. The resulting model is hereafter referred to as the *Biased Preferential Attachment* Model.

Roadmap. The rest of the paper is organized as follows. In the next section we review related work, then in Section 3 we introduce the model and the formal definitions of the involved properties: glass ceiling, power inequality and homophily tests. In Section 4 we state our two main theorems, and in Section 5 we provide empirical evidence for the existence of all our necessary ingredients and for the glass ceiling effect in a student-mentor network of researchers in computer science. We conclude with a discussion.

2. RELATED WORK

Homophily in social networks. Different characteristics, such as gender, ethnicities, age, class background and education, influence the relationships human beings form with each other [27]. McPherson et al. [32] survey a variety of properties and how they lead to particular patterns in bonding. Gender-based homophily can already be observed in play patterns among children at school [31, 41]. Eder and Hallinan [12] discovered that young girls are more likely to resolve intransitivity by deleting friendship choices, while young boys are more likely to add them. Overall, children are significantly more likely to resolve intransitivity by deleting a cross-sex friendship than by adding another cross-sex friendship [45]. These results show that gender influences the formation of cliques and larger evolving network structures. These trends, displaying homophily and gender differences in resolving problems in the structure of relationships, mean that boys and girls gravitate towards different social circles. As adults, homophilic behavior persists, and men still tend to have networks that are more homophilic than women do. This behavior is even more pronounced in areas where they form the majority and in relationships exchanging advice and based on respect, e.g., mentoring [5, 22, 23, 39]. A homophilic network evolution model was studied in [4]. In this model new nodes connect to other nodes in two phases. First, they choose their neighbors with a bias towards their own type (the model allows a positive as well as a negative bias). In a second phase they make an unbiased choice of neighbors from among the neighbors of their biased neighbors. The authors show that the second phase overcomes the bias in the first phase and if the second phase is unbiased, then the network ends up in an integrated state. They illustrate their model with data on citations in physics journals.

Gender disparity in science and technology. Gender disparities have been observed in the number of professors [13, 34], earnings [40], funding [29] and patenting [10]. A related aspect is the "productivity puzzle": men are more successful when it comes to number of publications and name position in the author list [46], for reasons yet unclear. Some conjectures raised involve (unknown) biased perceptions related to pregnancy/child care [6]. E.g., it was observed in [34] that science faculty members of both sexes exhibit unconscious biases against women. Gender differences in research output, research impact and collaborations was analyzed in a study based on Thomson Reuters Web of Science databases [26]. It was not only revealed that papers with women in prominent author positions (sole authorship, first-authorship and last-authorship) were cited less frequently but the authors also found that age plays an important role in collaborations, authorship position and citations. Thus many of the trends observed therein might be explained by the under-representation of women among the elders of science. In other words, fixing the "leaky pipeline" [43] is key for a more equal gender distribution in science.

Minority of women in Computer Science. In the computing discipline, the ratio of women taking up studies varies by year and region between 10% and 35% [3, 21, 38, 44] (except in Malaysia, where women form a narrow majority [35]). This under-representation has been investigated [19, 42, 47] and remedial strategies have been propoesd [17, 37]. There is a positive feedback loop [25]: the lack of women leads to a strong male stereotype which

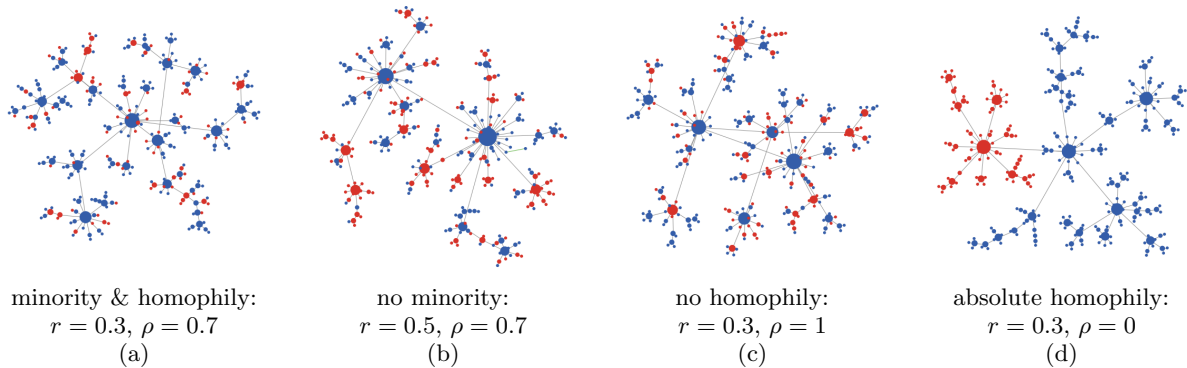| minority & homophily: | no minority: | no homophily: | absolute homophily: |
| $r = 0.3$, $\rho = 0.7$ | $r = 0.5$, $\rho = 0.7$ | $r = 0.3$, $\rho = 1$ | $r = 0.3$, $\rho = 0$ |
| (a) | (b) | (c) | (d) |

Figure 1: Examples of the Biased Preferential Attachment (BPA) model with various parameter settings. All examples depict a 300-vertex bi-populated network generated by our BPA model starting from a single edge connecting a blue and a red vertex (with vertex size proportional to its degree).

drives away even more women. It's been explained that the increase of the relative number of women in computer science was the best of the investigated remedial strategies, up to a "critical mass" of women. However, as pointed out by Etzkowitz [14], even achieving a representation of 15% women might not guarantee that the effects of a critical mass come into play.

3. MODEL AND DEFINITIONS

3.1 Biased preferential attachment model

Our first contribution is in proposing a simple bi-populated preferential attachment model. In a gist, our model is obtained by applying the classical preferential attachment model [2] to a bi-populated minority-majority network augmented with homophily. The resulting model is hereafter referred to as the *Biased Preferential Attachment* Model. Formally, for $r \leq 1/2$ and $0 \leq \rho \leq 1$ let $G(n, r, \rho)$ be a variant of the preferential attachment model in which the vertices are red or blue, n is the total number of nodes, and r represents the relative arrival rate of the red vertices (and hence the expected fraction of red vertices in the network converges to r as well, as the relative size of the initial population becomes smaller over time), and ρ represents the level to which homophily (incorporated by using rejection sampling) is expressed in the system: for $\rho = 1$ the system is uniform and exhibits no homophily, whereas for $\rho = 0$ the system is fully segregated, and all added edges connect vertex pairs of the same color.

Let us describe the model in more detail. Denote the social network at time t by $G_t = (V_t, E_t)$, where V_t and E_t, respectively, are the sets of vertices and edges in the network at time t, and let $\delta_t(v)$ denote the degree of vertex v at time t (we may omit the parameter t when it is clear from the context). The process starts with an arbitrary initial (connected) network G_0 in which each vertex has an arbitrary color, red or blue. (For simplicity we require that a minimal initial network consists of one blue and one red vertex connected by an edge, but this requirement can be removed if $\rho > 0$). This initial network evolves in time as follows. In every time step t a new vertex v enters the network. This vertex is red with probability r and blue with probability $1 - r$. On arrival, the vertex v chooses an existing vertex

$u \in V_t$ to attach to according to preferential attachment, i.e., with probability p proportional to u's degree at time t, i.e., $\mathbb{P}[u \text{ is chosen}] = \delta_t(u) / \sum_{w \in V_t} \delta_t(w)$. Next, if u's color is the same as v's color, then an edge is inserted between v and u; if the colors differ, then the edge is inserted with probability ρ, and with probability $1 - \rho$ the selection is *rejected*, and the process of choosing a neighbor for v is restarted. This process is repeated until some edge $\{v, u\}$ has been inserted. Thus in each time step, one new vertex and one new edge are added to the existing graph.

Figure 1 presents four examples of parameter settings for our model on a 300-vertex bi-populated social network. First, Figure 1(a) provides an example for the *minority & homophily* case with $r = 0.3$ and $\rho = 0.7$ so the red vertices are a strict minority in the network and there is some homophily in the edge selection. The next three sub-figures present special cases. Figure 1(b) illustrates the *no minority* case (equal-size populations, i.e., $r = 0.5$) with homophily ($\rho = 0.7$). Figure 1(c) considers the *no homophily* case ($\rho = 1$) with minority ($r = 0.3$). The last extreme case, shown in Figure 1(d), is *absolute homophily*, where $\rho = 0$, but the red vertices are still in the minority ($r = 0.3$). This case results in *fully segregated* societies, namely, societies where members connect *only* to members of their own color. In this extreme case, the society in effect splits into two separate networks, one for each of the two populations (except for the single edge connecting the initial red and blue vertices).

Consider as an example for our model the social network of mentor-student relationships in academia. With time, new PhD students arrive, but for some fields female students arrive at a lower rate than male students. Upon arrival, each student needs to select exactly one mentor, where the selection process is governed by the mechanisms of preferential attachment and homophily. Namely, initially the student selects the mentor according to the rules of preferential attachment and then homophily takes its role, rejecting the selection with some probability if their gender is different and forcing a re-selection. Over time, graduated students may become mentors and some mentors become more successful than others (in terms of the number of students they advise). A glass ceiling effect can be observed in this net-

work if, after a long enough time interval, the fraction of females among the most successful mentors tends to zero.

We would like to emphasize that the homophily effect that we look at is quite minor and "seemingly harmless", in two ways. First, it is "symmetric", i.e., it applies both to male students with respect to female mentors and to female students with respect to male mentors. Second, it does not adversely affect the student, in the sense that the student always gets admitted in our model. The only tiny (but ominous) sign for the potential dangers of this homophilic effect is that it does affect the professor: a male professor who rejects (or is rejected by) some fraction of the female candidates risks little, whereas a female professor who rejects (or is rejected by) some fraction of the male candidates will eventually have fewer students overall, since most of the applicants are male. In fact, as we show later on, this homophily-based consequence will only impact her if her future potential students use preferential attachment to select their mentors.

3.2 Power inequality and glass ceiling

Our second contribution is to propose formal definitions of the glass ceiling effect in social networks. Consider a *bi-populated* network $G(n)$ consisting of m edges and n nodes of two types, the groups R and B of red and blue nodes. We assume that the network size n tends to infinity with time. Let $n(\mathtt{R})$ and $n(\mathtt{B})$ denote the number of red and blue nodes, respectively, where $n(\mathtt{R}) + n(\mathtt{B}) = n$. The red nodes are assumed to be a minority in the social network, i.e., denoting the percentage of red nodes in the network by r, we assume $0 \le r < \frac{1}{2}$. Let $d(\mathtt{R})$ and $d(\mathtt{B})$ denote the sum of degrees of the red and blue nodes, respectively, where $d(\mathtt{R}) + d(\mathtt{B}) = 2m$. Let $\mathrm{top}_k(\mathtt{R})$ (respectively, $\mathrm{top}_k(\mathtt{B})$) denote the number of red (resp., blue) nodes that have degree *at least* k in G. When $G(n)$ is a random graph, we replace variables by their expectations in the definitions below, e.g., we use $\mathbb{E}[n(\mathtt{R})]$, $\mathbb{E}[d(\mathtt{R})]$, and $\mathbb{E}[\mathrm{top}_k(\mathtt{R})]$. Next we provide formal definitions for the social phenomena discussed in the introduction. *Power inequality* for the minority is defined in the following way.

DEFINITION 1 (POWER INEQUALITY). *A graph sequence $G(n)$ exhibits a* power inequality *effect for the red nodes if the average power of a red node is lower than that of a blue (or a random) node, i.e., there exists a constant $c < 1$ such that*

$$\lim_{n \to \infty} \frac{\frac{1}{n(\mathtt{R})} \sum_{v \in \mathtt{R}} \delta(v)}{\frac{1}{n(\mathtt{B})} \sum_{v \in \mathtt{B}} \delta(v)} = \frac{d(\mathtt{R})/n(\mathtt{R})}{d(\mathtt{B})/n(\mathtt{B})} \le c \ . \qquad (1)$$

The definition of the glass ceiling effect is more complex. We interpret the most powerful positions as those held by the highest degree nodes, and offer two alternative definitions. The first tries to capture the informal, "dictionary" definition, which describes a decreasing fraction of women among higher degree nodes, i.e., in the *tail* of the graph degree sequence. Formally:

DEFINITION 2 (TAIL GLASS CEILING). *A graph sequence $G(n)$ exhibits a* tail glass ceiling *effect for the red nodes if there exists an increasing function $k(n)$ (for short k) such that* $\lim_{n \to \infty} \mathrm{top}_k(\mathtt{B}) = \infty$ *and*

$$\lim_{n \to \infty} \frac{\mathrm{top}_k(\mathtt{R})}{\mathrm{top}_k(\mathtt{B})} = 0 \ .$$
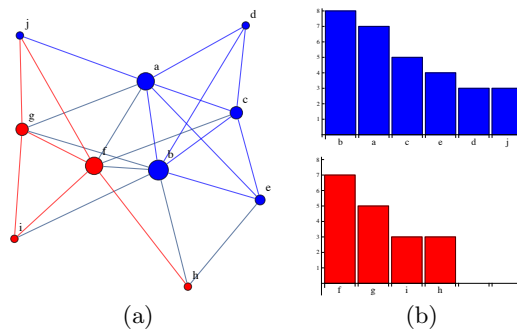

Figure 2: (a) An example bi-populated social network with blue and red populations of 6 and 4 vertices respectively. (b) The degree sequences of both populations (i.e., the sequence specifying for each vertex its degree in the network). Considering the *tail glass ceiling* definition, there are four blue vertices of degree greater or equal to 4, but only two such red vertices so $\mathrm{top}_4(\mathtt{R})/\mathrm{top}_4(\mathtt{B}) = 1/2$. For the *moment glass ceiling* definition, the second moment for the blue vertices is $\frac{1}{6}(8^2+7^2+5^2+4^2+3^2+3^2) = 28.6$, while for the red vertices it is $\frac{1}{4}(7^2 + 5^2 + 3^2 + 3^2) = 23$ and the ratio is $23/28.6$. To exhibit a glass ceiling, these ratios should converge to zero as the network size increases. Regarding homophily, in a random network with the same population, i.e., 60% blue vertices and 40% red vertices, one expects to find 36% blue-blue edges, 16% red-red edges and 48% mixed edge. If we take the degree sequences into account we would expect to see 46.8% mixed edge. In the above example network we observe only about 33% mixed edges, which indicates the effect of homophily.

The second definition considers a more traditional, distribution-oriented measure, the second moment of the two degree sequences. Formally:

DEFINITION 3 (MOMENT GLASS CEILING). *A graph sequence $G(n)$ exhibits a* moment glass ceiling *g for the red nodes where*

$$g = \lim_{n \to \infty} \frac{\frac{1}{n(\mathtt{R})} \sum_{v \in \mathtt{R}} \delta(v)^2}{\frac{1}{n(\mathtt{B})} \sum_{v \in \mathtt{B}} \delta(v)^2} \ .$$

When $g = 0$, we say that $G(n)$ has a *strong* glass ceiling effect. The intuition behind this definition is that a larger second moment (and assuming a similar average degree, i.e., no power inequality) will result in a larger variance and therefore a significantly larger number of high degree nodes. As we show in the full version of the paper, the above two definitions for the glass ceiling are independent, in the sense that neither of the effects implies the other.

Testing for *homophily* in a bi-populated network is based on checking whether the number of *mixed* (i.e., red-blue) edges is significantly lower than to be expected if neighbors were to be picked randomly and independently of their color. Formally:

DEFINITION 4 (HOMOPHILY TEST). *[11] A bi-populated social network exhibits homophily if the fraction of mixed edges is significantly less than $2r(1 - r)$.*

The above definition implicitly assumes that there is power equality between the colors and therefore is not always accurate. A more careful test should take the average degree of each gender into account.

DEFINITION 5 (NORMALIZED HOMOPHILY TEST). *A bi-populated social network exhibits homophily if the fraction of mixed edges is significantly less than* $2\frac{d(\text{R})}{2m}\left(1 - \frac{d(\text{R})}{2m}\right)$.

An illustration of these definitions can be found in Figure 2.

4. THEORETICAL RESULTS

4.1 Power inequality and glass ceiling

Our main theoretical result (Thm. 4.1) is that in the biased preferential attachment model, $G(n, r, \rho)$, the glass ceiling effect emerges naturally. Additionally, this process generates a *power inequality*, an independent property that is weaker than the glass ceiling effect. Power inequality describes the situation where the average degree of the minority is lower than that of the majority (although their members possess the same qualifications). Moreover, we also show (Thm. 4.2) that all three ingredients (unequal entry rate, homophily, preferential attachment) are necessary to generate what we call a *strong* glass ceiling effect, i.e., removing any one of them will prevent the appearance of a glass ceiling effect. One may suspect that the glass ceiling effect is in fact a byproduct of power inequality or unequal qualifications; we show in the full paper that this is not the case. Minorities can have a smaller average degree without suffering from a glass ceiling effect. We also note that our results are independent of the starting condition. Even if the network initially consisted entirely of vertices of one color, if a majority of the vertices being added are of the opposite color, then eventually the vertices that rise to the highest positions will be of the new color.

THEOREM 4.1. *Let* $0 < r < \frac{1}{2}$ *and* $0 < \rho < 1$. *For* $G(n, r, \rho)$ *produced by the Biased Preferential Attachment Model the following holds:*

1. *$G(n, r, \rho)$ exhibits power inequality, and*

2. *$G(n, r, \rho)$ exhibits both a tail and a strong glass ceiling effects.*

Moreover, all three ingredients are necessary to generate a *strong* glass ceiling effect.

THEOREM 4.2. *1. $G(n, r, \rho)$ will not exhibit a glass ceiling effect in the following cases:*

(a) *If the rate $r = \frac{1}{2}$ (no minority).*

(b) *If $\rho = 1$ (no homophily)*

(c) *If $\rho = 0$ (no heterophily).*

2. *$G(n, r, \rho)$ will not exhibit a strong glass ceiling effect if attachment is uniform rather than preferential, i.e., a new vertex at time t selects an existing vertex to attach to uniformly at random from all vertices present at time $t - 1$ (and for any value of r and ρ).*

Let us graphically illustrate the above results. Figure 3 presents the degree distributions of both the red and blue populations (as well as of the entire population) for four 1,000,000-vertex networks with parameters identical to the examples in Figure 1. The plots clearly show (and we prove this formally) that in all cases the degree distribution of both populations follows a power-law. (A subset W of vertices in a given network obeys a power-law degree distribution if the fraction $P(k)$ of vertices of degree k in W behaves for large values of k as $P(k) \sim k^{-\beta}$ for parameter β.) All figures present (in log-log scale) the cumulative degree distributions, so a power-law corresponds to a straight line (we present the samples together with the best-fit line). Theorem 4.1 corresponds to Figure 3(a) with the *minority & homophily* settings of $0 < r < \frac{1}{2}$ and $0 < \rho < 1$. In this case (and only in this case), the power-law exponents of the red and blue populations, $\beta(\text{R})$ and $\beta(\text{B})$ respectively, are *different*, where $\beta(\text{R}) > \beta(\text{B})$; we prove that this will eventually lead to both tail and strong glass ceiling effect for the red vertices. Theorem 4.2 corresponds to Figures 3(b) and 3(c). The figures show that in the case of *no minority* (i.e., $r = 0.5$) or *no homophily* (i.e., $\rho = 1$), both $\beta(\text{R})$ and $\beta(\text{B})$ are the same (in particular they are equal to 3 as in the classical Preferential Attachment model), and therefore there will be no glass ceiling effect. Figure 3(d) considers the last extreme case of *absolute homophily*. Perhaps surprisingly, in this case a glass ceiling effect also does not occur, as each sub-population forms an absolute majority in its own network (see again Figure 1(d)). The case of no preferential attachment (which does not lead to a glass ceiling) is more delicate and presented in the full version of the paper.

Proof Overview of Theorem 4.1. The basic idea behind the proof of Theorem 4.1 is to show that both populations in $G(n, r, \rho)$ have a power law degree distribution but with different exponents. Once this is established, it is simple to derive the glass ceiling effect for the population with a higher exponent in the degree distribution. To study the degree distribution of the red (and similarly the blue) population, we first define α_t to be the random variable that is equal to the ratio of the total degree of the red nodes (i.e., the sum of degrees of all red nodes) divided by the total degree (i.e., twice the number of edges). We show that the expected value of α_t converges to a fixed ratio independently of how the network started. The proof of this part is based on tools from dynamic systems. Basically, we show that there is only one fixed point for our system. However, determining the expectation of α_t is not sufficient for analyzing the degree distribution, and it is also necessary to bound the rate of convergence and the concentration of α_t around its expectation. We used Doob martingales for this part. Using the high concentration of the total degree, we were able to adapt standard techniques to prove the power law degree distribution. Next we give an overview of the proofs and the helping lemmas, but due to space limitations we defer the details to the full version of this paper.

4.2 Proof sketch of Theorem 4.1 Part 1

An urn process. The biased preferential attachment model $G(n, r, \rho)$ process can also be interpreted as a Polya's urn process, where each edge in the graph corresponds to two balls, one for each endpoint, and the balls are colored by the color of the corresponding vertices. When a new (red or blue) ball y arrives, we choose an existing ball c from the urn uniformly at random; if c is of the same color as y, then we add to the urn both y and another ball of the same

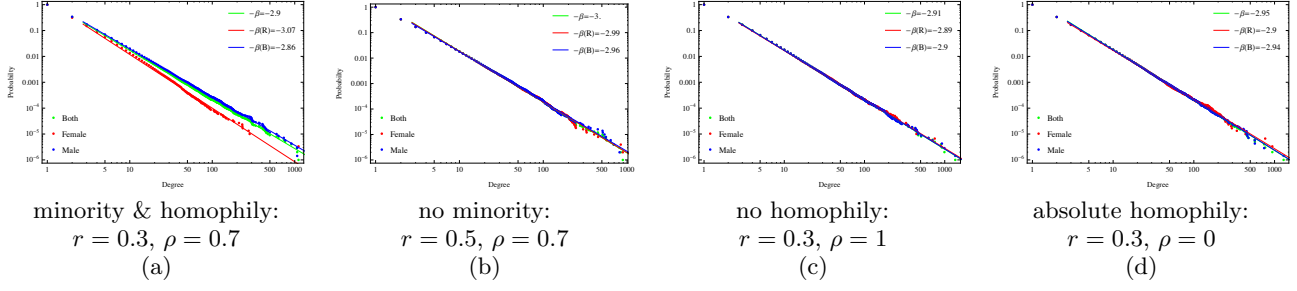| minority & homophily:	no minority:	no homophily:	absolute homophily:
$r = 0.3$, $\rho = 0.7$	$r = 0.5$, $\rho = 0.7$	$r = 0.3$, $\rho = 1$	$r = 0.3$, $\rho = 0$
(a)	(b)	(c)	(d)

Figure 3: Graphical illustrations of our formal claims concerning the glass ceiling effect in the Biased Preferential Attachment model. Each figure presents the degree distribution (on a log-log scale) of the red and blue populations from a 1,000,000-vertex network generated by the BPA model with the same parameters as the corresponding figure in Figure 1. In all cases both populations exhibit a power-law degree distribution but only in case (a) they with different exponents.

color as c; otherwise (i.e., if c is of a different color), with probability ρ we still add to the urn both y and another ball of the same color as c, and with probability $1 - \rho$ we reject the choice of c and repeat choosing an existing ball c' from the urn uniformly at random. To analyze power inequality, there is no need to keep track of the degrees of individual vertices; the sum of the degrees of all vertices of R is exactly the number of red balls in the urn.

Denote by $d_t(\texttt{R})$ (respectively, $d_t(\texttt{B})$) the number of red (resp., blue) balls present in the system at time $t \geq 0$. Altogether, the number of balls at time t is $d_t = d_t(\texttt{R}) + d_t(\texttt{B})$. Initially, the system contains d_o balls. Noting that exactly two balls join the system in each time step, we have $d_t = d_o + 2t$. Note that while $d_t(\texttt{R})$ and $d_t(\texttt{B})$ are random variables, d_t is not. Recall that balls represent degrees so $d_t = \sum_{w \in V_t} \delta_t(w)$.

Denote by α_t the random variable equal to $d_t(\texttt{R})/d_t$, the fraction of red balls in the system at time t.

Convergence of expectations. We first claim that the process of biased preferential attachment converges to a ratio of α red balls in the system. More formally, we claim that regardless of the starting condition, there exists a limit

$$\alpha = \lim_{t \to \infty} \mathbb{E}[\alpha_t] \ .$$

LEMMA 4.3. $\mathbb{E}[\alpha_{t+1}|\alpha_t] = \alpha_t + \dfrac{F(\alpha_t) - \alpha_t}{t + 1}, \quad where$

$$F(x) = \left(1 - (1-r)\frac{(1-x)}{1 - x(1 - \rho)} + r\frac{x}{1 - (1-x)(1-\rho)} \right)/2.$$

LEMMA 4.4. *1. F is monotonically increasing.*

2. *F has exactly one fixed point, denoted α^*, in $[0, 1]$.*

3. *The image of the unit interval by F is contained in the unit interval: $F([0,1]) = \left[\frac{r}{2}, \frac{1+r}{2}\right] \subset [0, 1]$*

4. *If $x < \alpha^*$ then $x < F(x) < \alpha^*$ and if $x > \alpha^*$ then $x > F(x) > \alpha^*$.*

5. *$\alpha^* < r$.*

Now assume $\alpha_t < \alpha^*$. By Lemma 4.4, $\alpha_t < F(\alpha_t) < \alpha^*$, so by Lemma 4.3 we obtain $\alpha_t < \mathbb{E}[\alpha_{t+1}|\alpha_t] < \alpha^*$.

Moreover we can bound the rate of convergence and show:

LEMMA 4.5. $|\alpha^* - \mathbb{E}[\alpha_t]| = O(1/\sqrt[3]{t})$.

THEOREM 4.6. *For any initial configuration, as t goes to infinity, the expected fraction of red balls in the urn, $\mathbb{E}(\alpha_t)$, converges to the unique α^* in $[0, 1]$ satisfying the equation*

$$2\alpha^* = 1 - (1-r)\frac{(1-\alpha^*)}{1 - \alpha^*(1-\rho)} + r\frac{\alpha^*}{1 - (1-\alpha^*)(1-\rho)}. \quad (2)$$

Hence the limit α^* is the solution of the cubic equation Eq. (3).

$$(4\rho - 2\rho^2 - 2)\alpha^3 + (2 + 3\rho^2 - 5\rho + 2r - 2r\rho)\alpha^2 \quad (3)$$
$$+ (2\rho - 2r + 2r\rho - \rho^2)\alpha - r\rho = 0$$

Note that this limit is independent of the initial values d_o and α_0 of the system.

We know that the expected degree of a random vertex is 2 and since the expected degree of a red vertex tends to $2\alpha^*/r$, which is strictly less than 2 (because of Lemma 4.4 Part 5), we can claim:

COROLLARY 4.7. *Let $0 < \rho < 1$, $0 < r < 1/2$. Then $G(n, r, \rho)$ has a power inequality effect.*

4.3 Proof sketch of Theorem 4.1 Part 2

Concentration. To prove the glass ceiling effect we first bound the degree distribution. To do this we need to bound the rate by which $d_t(\texttt{R})$ converge to $\alpha \cdot t$. Let $X_i \in \{0, 1, 2\}$ be the number of new red balls in the system at time i. Note that $d_t(\texttt{R}) = \sum_0^t X_i$. Let

$$\Psi_i = \mathbb{E}_{X_{i+1}, X_{i+2}, \dots, X_t}\Big[\sum_{j=0}^{t} X_j | X_1, X_2, \dots, X_i\Big].$$

Observe that $(\Psi_i)_i$ is a Doob Martingale [33], and note that $\Psi_0 = \mathbb{E}\Big[\sum_{i=0}^{t} X_i\Big] = \mathbb{E}\Big[d_t(\texttt{R})\Big]$.

THEOREM 4.8 (AZUMA'S INEQUALITY [1]). *Let Ψ_t be a martingale such that for all i, almost surely $|\Psi_i - \Psi_{i-1}| < c_i$. Then for all positive t and all positive reals x,*

$$\Pr(\Psi_t - \Psi_0 \geq x) \leq \exp\left(\frac{-x^2}{2\sum_i c_i^2}\right).$$

LEMMA 4.9. *Let $C_i = |\Psi_i - \Psi_{i-1}|$. Then $C_i = O(\sqrt{t/i})$.*

For simplicity of the description, let us assume hereafter that $d_o = 0$, hence $\alpha_t = d_t(\text{R})/(2t)$. By Theorem 4.8 and Lemma 4.9 we have

LEMMA 4.10. $\Pr\left[|d_t(\text{R}) - 2t\mathbb{E}(\alpha_t)| > O(2\sqrt{t}\log t)\right] \leq \dfrac{1}{t^4}$.

Combining Lemmas 4.5 and 4.10 yields:

COROLLARY 4.11.
$$\Pr\left[|\alpha^* - \alpha_t| > \max\left\{\frac{2\log t}{\sqrt{t}}, \frac{1}{\sqrt[3]{t}}\right\}\right] < \frac{1}{t^4}.$$

Degree distribution. We investigate the degree distribution of the red and blue vertices in a graph generated by the above described process, following the analysis outline of [7] for the basic preferential attachment model.

Let $m_{k,t}(\text{B})$ (resp., $m_{k,t}(\text{R})$) denote the number of blue (resp., red) vertices of degree k at time t. For $\mathbf{x} \in \{\text{R}, \text{B}\}$, define

$$M_k(\mathbf{x}) = \lim_{t\to\infty} \frac{\mathbb{E}(m_{k,t}(\mathbf{x}))}{t} \ . \tag{4}$$

THEOREM 4.12. *The expected degree distributions of the blue and red vertices follow a power law, namely, $M_k(\text{B}) \propto k^{-\beta(\text{B})}$ and $M_k(\text{R}) \propto k^{-\beta(\text{R})}$. If $0 < r < 1/2$ and $0 < \rho < 1$ then $\beta(\text{R}) > 3 > \beta(\text{B})$.*

Equipped with Theorem 4.12, Part 2 of Theorem 4.1 follows easily. Indeed, for the *tail glass ceiling* effect, let $k(n) = n^{\frac{1}{\beta(\text{R})}}$. Then

$$\mathbb{E}[\text{top}_k(\text{R})] = n(\text{R}) \sum_{k'\geq k} M_{k'}(\text{R}) \ ,$$

$$\mathbb{E}[\text{top}_k(\text{B})] = n(\text{B}) \sum_{k'\geq k} M_{k'}(\text{B}).$$

For $k' = n^{\frac{1}{\beta(\text{R})}}$ we have $nM_{k'}(\text{R}) = O(n \cdot n^{-\frac{\beta(\text{R})}{\beta(\text{R})}}) = O(1)$ while $nM_{k'}(\text{B}) = \Omega\left(n \cdot n^{-\frac{\beta(\text{B})}{\beta(\text{R})}}\right) = \Omega(n^{1-\frac{\beta(\text{B})}{\beta(\text{R})}}) = \Omega(n^\epsilon)$ for $\epsilon > 0$. The result then follows since $n(\text{R}) < n(\text{B})$ and $M_{k'}(\text{R}) < M_{k'}(\text{B})$ for $k' > k$.

For the *moment glass ceiling* effect we can show similarly:

$$\begin{aligned} g &= \lim_{n\to\infty} \frac{\sum k^2 M_k(\text{R})}{\sum k^2 M_k(\text{B})} = \lim_{n\to\infty} \frac{O(n^{3-\beta(\text{R})})}{\Omega(n^{3-\beta(\text{B})})} \\ &= \lim_{n\to\infty} O\left(\frac{1}{n^{\epsilon'}}\right) = 0 \end{aligned}$$

for some $\epsilon' > 0$.

The rest of this section sketches a proof of Theorem 4.12. Note that $m_{0,0}(\text{B}) = d_o(\text{B})$. We derive a recurrence for $\mathbb{E}(m_{k,t}(\text{B}))$. A blue vertex of degree k at time t could have arisen from three scenarios: (s1) at time $t-1$ it was already a blue vertex of degree k and no edge was added to it at time t. (s2) at time $t-1$ it was a blue vertex of degree $k-1$ and an edge was added to it at time t. (s3) in the special case where $k = 1$, at time $t-1$ it did not exist yet and it has arrived as a new blue vertex at time t. Thus letting $\mathcal{F}_t$ be the history of the process up to time t, for any $k > 1$, the expectation of $m_{k,t+1}(\text{B})$ conditioned on $\mathcal{F}_t$ satisfies

$$\mathbb{E}(m_{k,t+1}(\text{B})|\mathcal{F}_t) =$$
$$m_{k,t}(\text{B})\left(1 - \frac{rd_t(\text{B})\rho\frac{k}{d_t(\text{B})}}{d_t(\text{R})+d_t(\text{B})\rho} - \frac{(1-r)d_t(\text{B})\frac{k}{d_t(\text{B})}}{d_t(\text{R})\rho+d_t(\text{B})}\right)$$
$$+ m_{k-1,t}(\text{B})\left(\frac{rd_t(\text{B})\rho\frac{k-1}{d_t(\text{B})}}{d_t(\text{R})+d_t(\text{B})\rho} + \frac{(1-r)d_t(\text{B})\frac{k-1}{d_t(\text{B})}}{d_t(\text{R})\rho+d_t(\text{B})}\right).$$

For $k = 1$ we similarly have

$$\mathbb{E}(m_{1,t+1}(\text{B})|\mathcal{F}_t) =$$
$$m_{1,t}(\text{B})\left(1 - \frac{\rho r}{d_t(\text{R})+d_t(\text{B})\rho} - \frac{1-r}{d_t(\text{R})\rho+d_t(\text{B})}\right) + (1-r) \ .$$

Recalling again that $\alpha_t = d_t(\text{R})/(2t)$, the above can be rewritten as

$$\mathbb{E}(m_{k,t+1}(\text{B})|\mathcal{F}_t) =$$
$$m_{k,t}(\text{B})\left(1 - \frac{r\rho k}{2t(\alpha_t+(1-\alpha_t)\rho)} - \frac{(1-r)k}{2t(\alpha_t\rho+(1-\alpha_t))}\right)$$
$$+ m_{k-1,t}(\text{B})\left(\frac{r\rho(k-1)}{2t(\alpha_t+(1-\alpha_t)\rho)} + \frac{(1-r)(k-1)}{2t(\alpha_t\rho+(1-\alpha_t))}\right)$$

and for $k = 1$,

$$\mathbb{E}(m_{1,t+1}(\text{B})|\mathcal{F}_t) =$$
$$m_{1,t}(\text{B})\left(1 - \frac{\rho r}{2t(\alpha_t+(1-\alpha_t)\rho)} - \frac{1-r}{2t(\alpha_t\rho+(1-\alpha_t))}\right) + (1-r).$$

This can be expressed as

$$\mathbb{E}(m_{k,t+1}(\text{B})|\mathcal{F}_t) = m_{k,t}(\text{B})\left(1 - A_t\frac{k}{t}\right) \tag{5}$$
$$+ m_{k-1,t}(\text{B})A_t\frac{k-1}{t} \ ,$$

$$\mathbb{E}(m_{1,t+1}(\text{B})|\mathcal{F}_t) = m_{1,t}(\text{B})\left(1 - \frac{A_t}{t}\right) + (1-r), \tag{6}$$

using the notation

$$A_t = \frac{r\rho}{2\alpha_t + 2(1-\alpha_t)\rho} + \frac{(1-r)}{2\alpha_t\rho + 2(1-\alpha_t)} \ .$$

Note that A_t is a random variable so we next bound its divergence. Let

$$\begin{aligned} C_\text{B} &= \frac{r\rho}{2\alpha + 2(1-\alpha)\rho} + \frac{(1-r)}{2\alpha\rho + 2(1-\alpha)} \\ C_\text{R} &= \frac{(1-r)\rho}{2(\alpha\rho + 1 - \alpha)} + \frac{r}{2(\alpha + (1-\alpha)\rho))}. \end{aligned}$$

We have

LEMMA 4.13. $\Pr\left[|A_t - C_\text{B}| > \max\left\{\frac{2\log t}{\sqrt{t}}, \frac{1}{\sqrt[3]{t}}\right\}\right] < \frac{1}{t^4}$.

A similar claim can be made for C_R. This enables us to establish the following.

LEMMA 4.14.

- $M_1(\text{B})$ *exists and equals* $(1-r)/(1+C_\text{B})$,

- *For $k \geq 2$, $M_k(\text{B})$ exists and equals* $M_{k-1}(\text{B}) \cdot (k-1)C_\text{B}/(1+kC_\text{B})$,

- $M_1(\text{R})$ *exists and equals* $r/(1+C_\text{R})$, *and*

- *For $k \geq 2$, $M_k(\text{R})$ exists and equals* $M_{k-1}(\text{R}) \cdot (k-1)C_\text{R}/(1+kC_\text{R})$,

It is possible to show the following about C_B and C_R:

LEMMA 4.15.

- *If $0 < r < 1/2$ and $0 < \rho < 1$ then $C_\text{R} < \frac{1}{2} < C_\text{B}$*

- *If $r = 1/2$ then $C_\text{R} = C_\text{B} = 1/2$.*

- *If $\rho = 0$ or $\rho = 1$ then $C_\text{R} = C_\text{B} = 1/2$.*

To show that the degree distributions of both the red and the blue vertices follow power laws we recall that a power law distribution has the following property: $M_k \propto k^{-\beta}$ for large k, where β is independent of k. If $M_k \propto k^{-\beta}$, then

$$\frac{M_k}{M_{k-1}} = \frac{k^{-\beta}}{(k-1)^{-\beta}} = \left(1 - \frac{1}{k}\right)^{\beta} = 1 - \frac{\beta}{k} + O\left(\frac{1}{k^2}\right) .$$

Solving for the blue vertices, $M_k(\mathtt{B})$ and the blue exponent $\beta(\mathtt{B})$, and using Lemma 4.14, we get:

$$\frac{M_k(\mathtt{B})}{M_{k-1}(\mathtt{B})} = \frac{(k-1) \cdot C_{\mathtt{B}}}{1 + k \cdot C_{\mathtt{B}}} = 1 - \frac{C_{\mathtt{B}} + 1}{k \cdot C_{\mathtt{B}} + 1}$$
$$= 1 - \frac{1 + \frac{1}{C_{\mathtt{B}}}}{k} + O\left(\frac{1}{k^2}\right)$$

hence $\beta(\mathtt{B}) = 1 + 1/C_{\mathtt{B}}$. Similarly, for red vertices of degree k, $M_k(\mathtt{R})$ decays according to a power law with exponent $\beta(\mathtt{R}) = 1 + 1/C_{\mathtt{R}}$. Note that when $C_{\mathtt{R}} < \frac{1}{2} < C_{\mathtt{B}}$ we have $\beta(\mathtt{R}) > 3 > \beta(\mathtt{B})$ thus proving Theorem 4.12.

5. EMPIRICAL OBSERVATIONS

To provide empirical evidence illustrating the results of our analysis in real-life, we studied a *mentor-student network* of researchers in computer science, extracted from DBLP [28], a dataset recording most of the publications in computer science. A filtering process, described in detail in the full version of this paper, creates a list of edges connecting students to mentors. For each edge we determined the gender of the student and the mentor and the year in which the connection was established. The resulting network spans over 30 years and has 434232 authors and 389296 edges. As may be expected based on previously reported studies, our mentor-student network exhibits a minority-majority partition (namely, a low proportion of 21% females), homophily, power law distribution and a glass ceiling effect.

Figure 4(a) reveals that over time, the fraction of females in the network ($n(\mathtt{R})/n$, the shaded red area) has increased, but it is still below 21%. Also the average degree for females vertices is lower (1.48 vs 1.87). Figure 4(b) presents an indication for homophily in the mentoring selection process. This is done by the *homophily test* of [11], which compares the expected number of "mixed" (female-male) edges to the observed one (see also Section 3.2).

Figure 5 presents indications for the glass ceiling effect. Figure 5(a) shows that the fraction of females among the vertices of degree k or higher, namely, $\mathrm{top}_k(\mathtt{R})/\mathrm{top}_k(\mathtt{B})$, decreases continuously as k increases. The first major decrease occurs when moving from the group of "students" (i.e., degree 1 vertices) to the group of researchers of degree 2 or higher: the fraction of females drops from $\mathrm{top}_1(\mathtt{R}) \approx 21\%$ to $\mathrm{top}_2(\mathtt{R}) < 15\%$. It is important to note that the data indicates that even at the high end of the graph, a few female researchers with very high degrees are still present; however, our definitions for the glass ceiling ignore this extremal effect, which is caused by a few individuals, and concentrate on the averages over large samples. Indeed, when the sample size is large enough, the fraction of the female researchers decreases. Figure 5(b) shows a strong indication that the degree distribution of the vertices (females, males and combined) follows a power law. This in turn is associated with a preferential attachment mechanism that is known to result in a power law degree distribution. Note that the power-law exponent β for the graph of the female
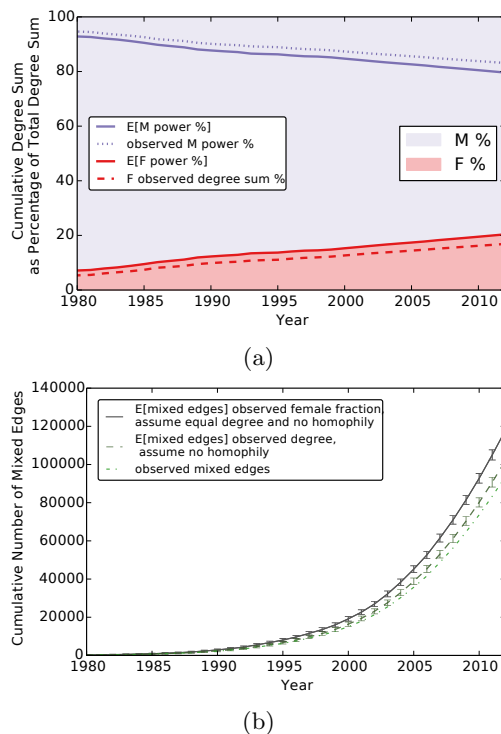

(a)

(b)

Figure 4: Female power and homophily in the computer science mentor graph. (a) The rate (i.e., percentage in the population) of females over time, compared with their normalized power, defined as $d(\mathtt{R})/(2m)$. Males have more power than expected by their rate, while females have less power than expected by their rate. (b) Evidence for homophily: a comparison of the observed number of "mixed" edges to the expected value assuming there is no homophily. We consider two cases: (i) the expected number of mixed edges ignoring the difference between the male and female average degree (expected: 127963.09 std: 293.08) and (ii) the expected number of mixed edges while considering the different degree sequences for males and females (expected: 110777.11 std: 281.52). In both cases the observed value (101607 edges) significantly deviates from the expectation (the error bars indicate the expected value $\pm$ 10 times the standard deviation) with extremely low p-values.

researchers is $\beta = 2.91$ (in the best fit), which is higher than the corresponding exponent in the graph for the male researchers, $\beta = 2.58$. Our analysis (presented in 4.2 and 4.3) establishes that if the degree distribution of both sub-populations follow a power law and the exponent for the minority sub-population is higher than that of the majority sub-population, then a strong moment glass ceiling effect will appear.

6. DISCUSSION

One obvious limitation of our model is that it is somewhat simplistic and captures only one possible mechanism for generating a glass ceiling effect. It ignores many impor-
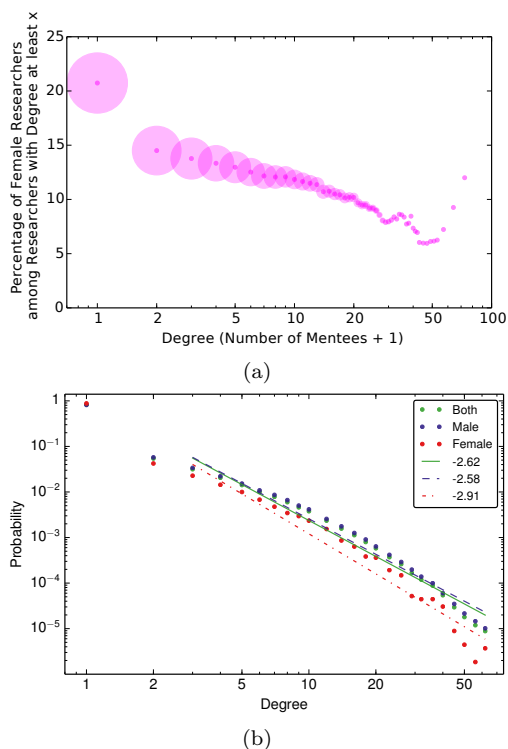
(a)

(b)

Figure 5: Glass ceiling effect in mentor graph: (a) percentage of females in the mentor population of degree at least k. Female start with 21% in the population and drop to below 15% when considering degree at least 2 (faculty members). It continues to decrease (ignoring small samples at the end, see text). Vertex size and darker color represent larger sample space. (b) The power-law-like degree distribution for both females and males. The exponent β for females is higher than for males, demonstrating the glass ceiling effect.

tant aspects of real life (such as sexual tension, fear, family responsibilities and jealousy, to name a few) and alternative (co-existing) mechanisms that contribute to the effect. For instance, our model cannot be used to explain the occurrence of a glass ceiling effect in contexts where pairwise individual interactions play a less dominant role than in academia. To account for the glass ceiling effect in such contexts as well as others, one may consider alternative explanations. In particular, a common possible explanation is the "leaky pipeline" phenomenon, namely, the phenomenon that women tend to quit or slow down their careers in order to invest more time in their families. This phenomenon can be modeled mathematically in several different ways. One such way is by introducing vertex departures in addition to vertex arrivals, with a bias in the form of increased departure rate of the minority group. But in fact, such a dynamic "leaky pipeline" model allows several reasonable sub-models that will not generate a glass ceiling effect, as well as some other sub-models that do. Moreover, the cause and effect relationships between glass ceiling and leaky pipeline are not necessarily one-directional; while the glass ceiling effect may indeed be the outcome of the "leaky pipeline" phenomenon

in certain settings, there are other settings where it may be its (partial) cause. An interesting direction for future work would be to describe a more complete model, most likely combining a number of different mechanisms contributing jointly to the glass ceiling effect. In any case, we find it remarkable that the simple mathematical mechanism presented here (based on homophily) is sufficient to explain (at least parts of) the glass ceiling effect, despite the fact that it does not utilize the "leaky pipeline".

Our Findings may suggest ways to deal with the glass ceiling phenomenon. By better understanding the roots of the glass ceiling effect, one can address each of the elements and attempt to mitigate them or deal with those elements that are easier to manage. Our research indicates that for certain mechanisms involved in the formation of a glass ceiling, removing one element may eliminate the glass ceiling effect. Hence, while it might be difficult to modify the human tendencies of homophily and preferential attachment one could attempt to balance the proportions of minorities within the population or impose a proportional representation of successful women at the top level. Both of these options may be classified as variants of affirmative action, but the latter, even if more common, seems to avoid the roots of the problem. In particular, a more equally represented society could be created by encouraging minorities to enter the system, as our findings indicate that increasing the ratio of minorities at the entry stage may mitigate the glass ceiling effect at least partially. This conclusion is in line with a common view [36, 43], which states that fixing the "leaky pipeline" is key for a more equal gender distribution in science. By determining and examining the causes of the glass ceiling effect, we can work on alleviating the glass ceiling effect, resulting in a richer and more diverse community.

Acknowledgements

7. REFERENCES

[1] N. Alon and J. H. Spencer. *The probabilistic method.* John Wiley & Sons, 2004.

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[3] S. J. Bock, L. J. Taylor, Z. E. Phillips, and W. Sun. Women and minorities in computer science majors: Results on barriers from interviews and a survey. *WOMEN*, 14(1):143–152, 2013.

[4] Y. Bramoulle, S. Currarini, M. O. Jackson, P. Pin, and B. W. Rogers. Homophily and long-run integration in social networks. *J. Economic Theory*, 147(5):1754–1786, 2012.

[5] D. J. Brass. Men's and women's networks: A study of interaction patterns and influence in an organization. *Academy of Management J.*, 28(2):327–343, 1985.

[6] S. J. Ceci and W. M. Williams. Understanding current causes of women's underrepresentation in science. *Proc. Nat. Acad. Sci.*, 108(8):3157–3162, 2011.

[7] F. R. K. Chung and L. Lu. *Complex graphs and networks*, volume 107 of *CBMS Regional Conf. Series in Mathematics*. AMS Bookstore, 2006.

[8] D. A. Cotter, J. M. Hermsen, S. Ovadia, and R. Vanneman. The glass ceiling effect. *Social forces*, 80(2):655–681, 2001.

[9] S. A. Davies-Netzley. Women above the glass ceiling perceptions on corporate mobility and strategies for success. *Gender & Society*, 12(3):339–355, 1998.

[10] W. W. Ding, F. Murray, and T. E. Stuart. Gender differences in patenting in the academic life sciences. *Science*, 313(5787):665–667, 2006.

[11] D. Easley and J. Kleinberg. Networks, crowds, and markets. *Cambridge Univ Press*, 6(1):6–1, 2010.

[12] D. Eder and M. T. Hallinan. Sex differences in children's friendships. *American Sociological Review*, pages 237–250, 1978.

[13] Education at a Glance. Organisation for economic co-operation and development. *OECD*, 2012.

[14] H. Etzkowitz, C. Kemelgor, M. Neuschatz, B. Uzzi, and J. Alonzo. The paradox of critical mass for women in science. *Science*, 266:51–54, 1994.

[15] A. Eyring and B. A. Stead. Shattering the glass ceiling: Some successful corporate practices. *J. Business Ethics*, 17(3):245–251, 1998.

[16] E. Falk and E. Grizard. The glass ceiling persists: The 3rd annual appc report on women leaders in communication companies. *The Annenberg Public Policy Center, University of Pennsylvania. Retrieved March*, 4:2005, 2003.

[17] W. Faulkner and M. Lie. Gender in the information society strategies of inclusion. *Gender, Technology and Development*, 11(2):157–177, 2007.

[18] Federal Glass Ceiling Commission. Solid investments: Making full use of the nation's human capital. US Government, Dept. of Labor. Washington DC, 1995.

[19] A. Fisher, J. Margolis, and F. Miller. Undergraduate women in computer science: Experience, motivation and culture. In *Proc 28th SIGCSE Technical Symp. on Computer Sci. Edu.*, pages 106–110. ACM, 1997.

[20] N. Frenkiel. The up and comers: Bryant takes aim at the settlers-in. *Adweek, March*, 1984.

[21] C. Hill, C. Corbett, and A. St Rose. *Why So Few? Women in Science, Technology, Engineering, and Mathematics.* ERIC, 2010.

[22] H. Ibarra. Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative science quarterly*, pages 422–447, 1992.

[23] H. Ibarra. Paving an alternative route: Gender differences in managerial networks. *Social Psychology Quarterly*, pages 91–102, 1997.

[24] A. P. Kottis. Women in management: The "glass ceiling" and how to break it. *Women in Management Review*, 8(4), 1993.

[25] V. A. Lagesen. The strength of numbers: Strategies to include women into computer science. *Social Studies of Science*, 37(1):67–92, 2007.

[26] V. Lariviere, C. Ni, Y. Gingras, B. Cronin, and C. R. Sugimoto. Bibliometrics: Global gender disparities in science. *Nature*, 504:211–213, 2013.

[27] P. F. Lazarsfeld, R. K. Merton, et al. Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society*, 18(1):18–66, 1954.

[28] M. Ley. Dblp: some lessons learned. *Proc. VLDB Endowment*, 2, 2009.

[29] T. J. Ley and B. H. Hamilton. The gender gap in NIH grant applications. *Science*, 322(5907):1472–1474, 2008.

[30] P. Longo and C. J. Straehley. Whack! i've hit the glass ceiling! women's efforts to gain status in surgery. *Gender medicine*, 5(1):88–100, 2008.

[31] E. E. Maccoby. *The two sexes: Growing up apart, coming together.* Harvard University Press, 1998.

[32] M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444, 2001.

[33] M. Mitzenmacher and E. Upfal. *Probability and computing: Randomized algorithms and probabilistic analysis.* Cambridge University Press, 2005.

[34] C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham, and J. Handelsman. Science faculty's subtle gender biases favor male students. *Proc. Nat. Acad. Sci.*, 109(41):16474–16479, 2012.

[35] M. Othman and R. Latih. Women in computer science: No shortage here! *Commun. ACM*, 49(3):111–114, Mar. 2006.

[36] A. N. Pell. Fixing the leaky pipeline: women scientists in academia. *J. Animal Sci.*, 74(11):2843–2848, 1996.

[37] E. S. Roberts, M. Kassianidou, and L. Irani. Encouraging women in computer science. *SIGCSE Bull.*, pages 84–88, 2002.

[38] S. Sassler, J. Glass, Y. Levitte, and K. Michelmore. The missing women in stem? accounting for gender differences in entrance into stem occupations. In *Annual meeting of the Population Association of America Presentation*, 2011.

[39] J. M. Sheltzer and J. C. Smith. Elite male faculty in the life sciences employ fewer women. *Proc. Nat. Acad. Sci.*, 2014.

[40] H. Shen. Mind the gender gap. *Nature*, 2013.

[41] L. Smith-Lovin and J. M. McPherson. You are who you know: A network approach to gender. *Theory on gender/feminism on theory*, pages 223–51, 1993.

[42] E. Spertus. Why are there so few female computer scientists? Technical report, Cambridge, MA, 1991.

[43] P. E. Stephan and S. G. Levin. Leaving careers in IT: gender differences in retention. *J. Technology Transfer*, 30(4):383–396, 2005.

[44] R. Stross. What has driven women out of computer science? *New York Times*, 15, 2008.

[45] N. B. Tuma and M. T. Hallinan. The effects of sex, race, and achievement on schoolchildren's friendships. *Social Forces*, 57(4):1265–1285, 1979.

[46] J. D. West, J. Jacquet, M. M. King, S. J. Correll, and C. T. Bergstrom. The role of gender in scholarly authorship. *PloS one*, 8(7):e66212, 2013.

[47] M. A. Whitecraft and W. M. Williams. *Why Aren't More Women in Computer Science? In Making Software: What Really Works, and Why We Believe.* 2011.