Algorithmic Glass Ceiling in Social Networks

The effects of social recommendations on network diversity

Ana-Andreea Stoica
Columbia University
New York, USA
astoica@cs.columbia.edu

Christopher Riederer
Columbia University
New York, USA
mani@cs.columbia.edu

Augustin Chaintreau
Columbia University
New York, USA
augustin@cs.columbia.edu

ABSTRACT

As social recommendations such as friend suggestions and people to follow become increasingly popular and influential on the growth of social media, we find that prominent social recommendation algorithms can exacerbate the under-representation of certain demographic groups at the top of the social hierarchy. To study this imbalance in online equal opportunities, we leverage new Instagram data and offer for the first time an analysis that studies the effect of gender, homophily and growth dynamics under social recommendations. Our mathematical analysis demonstrates the existence of an algorithmic glass ceiling that exhibits all the properties of the metaphorical social barrier that hinders groups like women or people of color from attaining equal representation. What raises concern is that our proof shows that under fixed minority and homophily parameters the algorithmic effect is systematically larger than the glass ceiling generated by the spontaneous growth of social networks. We discuss ways to address this concern in future design.

KEYWORDS

social recommender, fairness, random walks, homophily

1 INTRODUCTION

Geographical distance and social prejudice have a long history of preventing equitable access to information and commercial transactions. One hope of an increasingly connected world is to at least take a step towards equal opportunities by removing some of those barriers. But, in the social networks now governing the knowledge, jobs and deals one can seek, what matters most today is one's position in a graph of advantageous connections. Those connections take time to set up and maintain. Recommender systems play a key role in assisting individuals in directing their social sharing

activity towards the most valuable sources and partners. But under which conditions can we ensure that those suggestions are not reproducing, or worse, *reinforcing* our historical biases, combined with a cloudy illusion of objectivity? One would *a priori* expect similarity metrics, usually the basis of recommender systems, to contribute to sustaining disparities among various groups. We show much more: using empirical evidence from newly collected data on Instagram and a rigorous analysis of mathematical models, we prove that prominent recommender algorithms reinforce the *rate at which disparity grows*.

While the issues introduced affect countless situations where suggestions are algorithmically produced, we focus for concreteness on an elementary case: the suggestion to build a new edge with another participant. Features like "People you may know" (Facebook), "Who to follow" (Twitter), "Suggested Accounts" (Instagram) are now pervasive across most popular social network services. Their impact on the evolution of connection graph has been demonstrated [16]. To quote the Chairs' opening statement from the FATREC workshop (2017), "If recommendation applications did not alter participants' behaviors, those applications would not exist". As a thought experiment, one may picture the *organic* growth of a connection graph (i.e., the addition of edges occurring in absence of any suggestions) as a reference point. An *algorithmic* growth, in contrast, results from suggestions. An algorithm to determine new edges to suggest is generally evaluated either via its *predictive power* (i.e., does it anticipate organic growth?), or its *hit rate* (i.e., are its suggestions well received?). We ask a different question: are there safe (respectively unsafe) conditions under which a well-know metric of disparate impact is no worse (respectively strictly worse) for algorithmic growth than for organic growth? Unfortunately, our study demonstrates that no safe conditions exist outside trivial cases for frequently used algorithms and three definitions of disparate impact. However, we quickly note that future suggestion algorithms could be designed to account for and correct that bias, and offer a short discussion of simple ways to address this in future design.

This paper presents the following contributions:

- To better investigate the role of gender in organic and artificial growth of social networks, we conduct a large crawl of the Instagram service and its active relationships (comments and likes). The graph has a moderate but distinct majority of female profiles. Nevertheless, with high statistical confidence, the organic growth we captured exhibits traces of a so-called *glass ceiling* preventing females from rising to the most commented and liked profiles. But that pales by comparison with the *algorithmic glass ceiling* which is observed when two well known algorithms are used: only in that case

do we observe a drastic gap arising in the degree distributions by gender. A sudden decrease of female representation at the top ensues (Section 3).

- We follow a mathematical analysis of the effect of random walks on dynamic graphs to investigate the conditions creating *algorithmic glass ceiling*. Previous works highlighted how combining an uneven population with homophily prepares the ground for glass ceiling in the typical organic growth of preferential attachment. Our method allows one to study the eventual perturbation effect of a recommendation algorithm as the result of a new polynomial fixed point equation. This is used to prove two major results: First, a simple recommendation algorithm leveraging random walks is not only subject to some bias, it also reinforces the rate at which a gender gap arises. Second, we prove numerically that this effect is always noticeable for observed parameters (Section 4).
- We provide two important theoretical follow ups. First, we address the apparent contradiction between empirical observations (showing glass ceiling against a female majority) and mathematical results (predicting glass ceiling at the expense of a minority). The two are reconciled by an extension of the above model we call *differentiated homophily* where the two groups exhibit varying propensity to favor their own peers. We prove that this can flip the glass ceiling effect to come at the expense of a majority. We document that, indeed, male accounts on Instagram exhibit much stronger homophily among each others when compared with female accounts. Finally, as a proof of concept we quickly describe a simple "rejection method" for an improved random walk recommendation algorithm. This approach makes stringent assumptions but offers the advantage of making the algorithmic growth more equitable than organic growth (Section 6).

As we detail immediately below, the sharp amplification of the glass ceiling effect by an algorithm is an entirely new result, which has no equivalent that we know of. It is a special case of a widely open problem: how to correct a seemingly neutral algorithm when the structure it exploits is not fair to begin with. Some might argue the algorithm is simply legitimate as it leverages sensible metrics and is not *solely* responsible for the unfairness found in the system. That remains however unsatisfactory for anyone wishing more equal opportunities for all. Those who abide by a strict fairness requirement (e.g., defined as statistical parity [18] or equality of opportunity), may suggest a correction step be applied. Unfortunately, without a deep understanding of the cause and reverberation of bias, any post-hoc correction can be harshly debated. Our paper offers an alternative way: identify some structural causes of the emerging unfairness, and require algorithms to be designed in a way that leverages structure while carefully avoiding those biases in the presence of the above conditions.

2 RELATED WORK

Our work addresses the fairness concerns recently raised for decisions informed by algorithms run on Big Data. Even with no explicit intention to discriminate and without access to sensitive attributes such as gender and race, such algorithms can potentially create undue biases [2]. Most related work on this issue centers on binary classification tasks such as loan approval, with a goal to maintain operational efficiency under various conditions of equal treatment or opportunity [4]. Similarly, fairness was studied for word embeddings [3] and nodes ranking [17]. But to our knowledge, an analysis of social recommendation and the effect of graph algorithms has yet to be conducted. The main difference is that one must analyze the effect of network structure.

We build on a growing body of evidence that online services (including Twitter[14], TaskRabbit [8], and Airbnb [7]) can reproduce well-known prejudices against historically under-represented groups. Issues raised include disparate treatment and evidence of a metaphorical *glass ceiling*. The latter denotes an invisible barrier preventing given demographics (most commonly females) to reach superior levels of recognition. This manifests itself when (1) chances of advancement to a higher level are uneven for members of that demographic group, (2) when it is not explained by task-relevant characteristics, and (3) when that inequality increases for higher and higher levels. We observe that these three conditions hold for the Instagram dataset, in particular for people who like and comment on others' posts. In addition, we collect for the first time the graph connecting individuals within our crawl. This leads us to observe whether algorithms (in this case, profile recommendation) would alleviate or reinforce this trend. Recent work suggested that the growth of Twitter was also impacted by recommendation [16]. Combining such previous results with ours raises concerns that the current engineered growth of social media may in the end systematically misrepresent entire groups of the population.

Homophily, a tendency of individuals to favor interactions with similar peers – widely established to hold for social networks and online communites [5, 12] – is at the core of our result. The result most relevant to ours is a recent proof that rich-get-richer dynamics combined with homophily naturally exacerbate the advantage of a majority group [1]. That model can explain why, as in other services, *organic* interaction on Instagram exhibits the aforementioned bias against some of its participants. We, on the other hand, introduce random walk dynamics, which closely capture the way recommender systems operate.

3 COMPARISON OF ORGANIC AND ALGORITHMIC GROWTH

3.1 Data Collection

Instagram activity graph. We performed a crawl on Instagram's public profiles starting from the founder of Instagram, Kevin Systrom. Through the Instagram API, we collected profile information and recursively followed the lists of followers. For each profile, we collected the name, username, unique id, and crawled the meta-data for each photo, which include: unique photo id, number of likes, number of comments, hashtags, location, timestamp, tagged users, and a random subset of 3-5 likes and comments with their authors. For each user, we aggregated this information in obtaining the total number of likes, comments, tags, first and last photo timestamp, and the set of used hashtags. In total, we collected $115,796,284$ photos from $245,038$ different users over multiple months of 2014 and 2015. While the data might not encompass the full diversity of Instagram users, it is sufficient to obtain statistically significant results most of the time.

Table 1: Male vs. Female Poisson Regression for degree, Instagram-activity, no control variables.

	Full	Bottom 25%	Bottom 25 − 50%	Top 50 − 25%	Top 25%	Top 10%	Top 1%	Top 0.1%
Male	0.011***	-3.6e-17***	4.4e-17***	-3.6e-17***	0.08***	-0.05***	-0.12***	0.13***
Male IRRs	1.11***	0.99***	1***	0.99***	1.09***	0.94***	0.88***	1.14***
Obs:	553,628	138,407	138,407	138,407	138,407	55,363	5,537	553

Table 2: Male vs. Female Poisson Regression for Adamic-Adar recommendation frequency, no control variables.

	Full	Bottom 25%	Bottom 25 − 50%	Top 50 − 25%	Top 25%	Top 10%	Top 1%	Top 0.1%
Male	0.08***	-2e-16***	-2e-16***	0***	0.11***	0.14***	0.14***	0.21***
Male IRRs	1.08***	0.99***	0.99***	1***	1.12***	1.15***	1.15***	1.23***
Obs:	295,919	73,980	73,980	73,980	73,979	29,592	2,959	295

Gender inference. Our gender inference method adapted from [13] leverages first name and social security data. Mislove et al. inferred the gender balance using a dataset of names from social security data ranging between 1890 and 2015. We only included name data between 1940 and 2007, as reported demographic statistics predict that less than 4% of Instagram users would be born before that [6].

We filtered reported names for those having at least 50 occurrences for either men or women, with at least 95% of a single gender. This created 32, 676 unique unambiguous first names, which we use to label 92, 935 Instagram users. That is approximately 38% of the population. Among those 49, 583, or 53%, are labeled as female, and 43, 352, or 47%, as male. We tested a small sample of classified names by hand, obtaining 97% accuracy. We call this graph the **Instagram-labeled-core**.

The graph of likes and comments. Users of **Instagram-labeled-core** collectively posted 44, 725, 840 photos, out of which we randomly sample 1 million photos. For each photo in this set, we looked at the 3-5 likes and comments received and extracted all authors, keeping those only when their gender could be labelled as above. We create an edge between the authors of the pictures and the authors of the likes or comments. We choose to focus on likes and comments as interactions, since they are a more involved way of connecting online than followers.

This created a graph of **553,628 nodes** and **652,931 edges**, where everyone has a labeled gender. Here, we obtained 54.43% users labeled as female, and 45.56% users labeled as male, showing that gender balance appears the same as above. This graph is called **Instagram activity** and is the main graph we will use in the rest of this article. Note that it is much larger, authors of comments may not have been in the original crawl and hence we do not have all meta-data about their activity and following. Building a graph implies that a user's degree count how many distinct (and labeled) users either received or gave a comment or a liked to one of her pictures.

DBLP graph. We validated our results against a selection of the DBLP co-authorship dataset [10], confirming the analysis from Avin et al. [1] that shows a gender discrepancy between male and female researchers in the numbers of mentees they have.

Table 3: Gender ratio for the initial graph

	Male	Female
Male	57%(1.2x)	43%(0.8x)
Female	40%(0.9x)	60%(1.1x)

We sample the first 200, 000 articles from the latest version of the DBLP dataset, from which we extract the authors and perform gender inference as above. We further create a graph of co-authorship between authors for whom our inference method finds a gender. The resulting graph, the *DBLP graph*, consists of 102, 263 nodes, out of which 80.7% are labeled as male and 19.3% as female, and 199, 679 edges. As an existing bias towards male researchers is proven to exist in this dataset [1], we analyze the effect of recommendation algorithm on such a gap.

3.2 Bias under organic growth

Instagram's "Suggested accounts" feature was introduced after our data crawl. The data we collected is hence a representative view of an *organic growth*, which is how the network evolves from interaction in the absence of algorithmic suggestions.

A rich body of literature focuses on the effect of gender in acquiring visibility and impact in the online world [14] [15] [1], where measures of wealth are redefined as volume of user-user interactions (such as the degree we defined above).

We conducted multiple analyses reproducing methodology from papers cited above, that we briefly summarize for space reasons. First, in the **Instagram-activity** graph, we use Poisson Regression with no control variables. This analysis reveals that in the last quartile containing most prominent accounts, male users have an incidence rate ratio (IRR) of 1.09*** (a 9% increase, where *** denotes statistical confidence $p < 0.001$) for comment/like degree. All other quartiles show no effect of gender. The effect is however relatively small, and most prominent at the top 0.1% which contains half a thousand nodes (see Table 1). In fact, among the top 5,000 or top 1%, women are statistically significant to have larger degrees up to 12%.

Figure 1 further illustrates this effect in both the Instagram activity graph and the DBLP graph. For the Instagram dataset (a), we

(a)

(b)
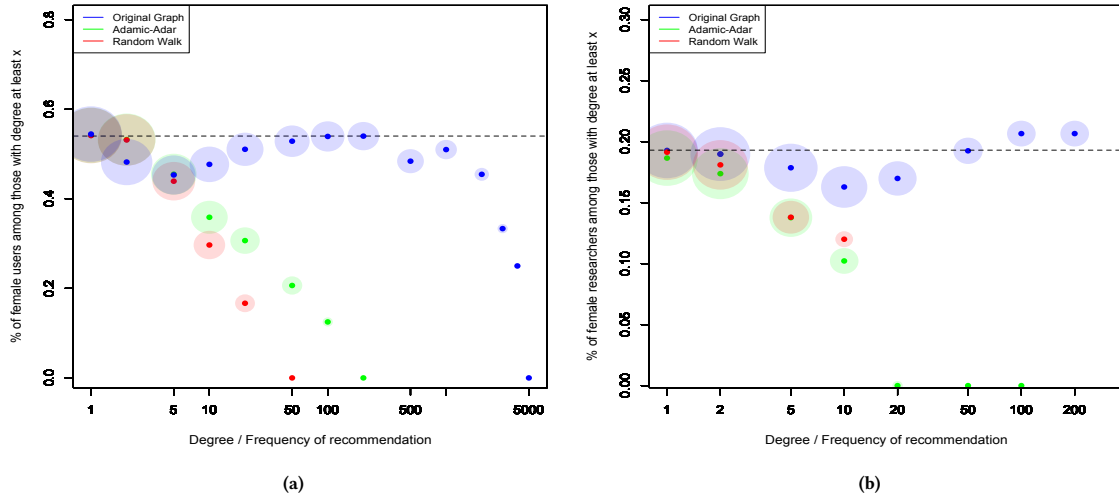
Figure 1: (a) Representation of nodes with degree or frequency of recommendation at least x in the original graph (blue) and the graph under recommendations (green and red), respectively, for the Instagram activity graph. The nodes' relative size is shown by vertex size, and where y-axis denotes the fraction of female among those. The dotted line reminds us the actual percentage of females in the dataset; (b) Idem for the DBLP graph.

observe that while in the initial graph the disparity between the two genders is not very large (blue dots and circles), it is still present. While females make up 54% of the total population, they represent only 48% of those with degree > 10 (which corresponds roughly to the top 1% most connected users), 54% of those with degree > 100 (a subset which exhibits no particular gender imbalance), and 45% of those with degree > 2000.

In summary, the organic growth we observe confirms a significant effect of gender. We obtain similar results (omitted here) when controlling for content productivity. Overall, women appear underrepresented at the top, although they form the largest group, yet not completely excluded from large levels of recognitions. This motivates us to study how those biases change when Instagram's organic growth is complemented by recommendations.

3.3 Recommendation algorithms

Recommendation algorithms typically leverage structure to find compatible matches among similar but not yet connected individuals. Many heuristics have been proposed and it is challenging to choose a representative subset of them. Two stand out from prior comprehensive comparisons:

The Adamic-Adar index. We first note that if algorithms were perfectly predictive and only suggested edges that nodes would have made in *organic growth*, we should not observe a significant effect. We therefore include social recommendation based on the Adamic-Adar index, as it is shown in [11] to be the most accurate in prediction and therefore a conservative choice from a perturbation standpoint. The Adamic-Adar index adapts a simple rule, "choose node with maximum-common-neighbors" to avoid always

recommending high degree nodes. Our own analysis of maximum-common-neighbors revealed that indeed that would be impractical and too simplistic to obtain sensible results, as suggestions are chosen in a restricted set. We define Adamic-Adar index between two nodes u and v as $(u, v) \rightarrow \sum_{w \in \Gamma(u) \cap \Gamma(v)} 1/\log|\Gamma(w)|$, where $|\Gamma(u)|$ denotes all neighbors of u [11]. Note that in this index common neighbors that are very popular would carry a smaller weight than common neighbors with a smaller degree. For each node, the social recommendations is computed by finding another node v with largest index that is not already a neighbor (breaking ties randomly).

Random Walk. Another interesting comparison point would be an algorithm that best reproduces suggestions that are *actually given* today. That is a difficult task since many of the algorithms currently deployed are complex and often kept secret to keep competitors and cheaters at bay. To best address this need, we choose an algorithm based on a random walk of length 2, as it was deemed similar to real recommendations in the only article we know that had access to this proprietary data [16]. This algorithm, in contrast with Adamic-Adar index, returns a random recommendation for each node u, by starting a random walk at a point $X_1 = u$ and returning v with probability

$$\mathbb{P}(X_3 = v | X_1 = u). \tag{1}$$

We note that the results we obtained seem to hold independently of the algorithms chosen. That could be because both algorithm, and indeed most recommendations, would more likely select a node with multiple common neighbors of small degrees, and differ only in how they combine those features.
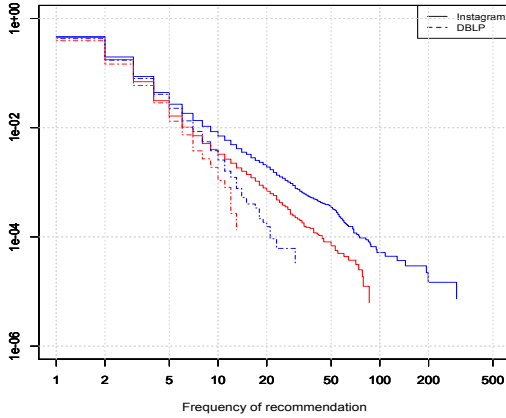
Figure 2: Distribution of recommendation frequency as CCDF under Adamic/Adar algorithm, separately drawn for female (red) and male (blue), and for the Instagram dataset (solid lines) and the DBLP dataset (dashed lines).

3.4 Bias under algorithmic growth

The natural equivalent of degree as a measure of recognition or wealth under recommendations is the number of users who get recommended to connect with you, i.e., *frequency of recommendation*. For every node in the Instagram-activity graph and the DBLP graph, we computed a single recommendation and collected the frequency of various nodes to appear as suggestions, which relates to degree growth under such dynamics.

We first present in Table 2 results allowing a direct comparison to organic growth. We use Poisson Regression with the dependent variable as the frequency of recommendation under Adamic-Adar for the Instagram activity graph (note that the total population in this case is $295, 919$, *i.e.* individuals who have been recommended at least once). Omitted results for the random walk algorithm are similar. Here, one sees a more noticeable and consistent glass ceiling effect. All the population in the last quartile sees male with larger frequencies, and this trend always grows with higher degrees.

While organic growth generates degree distributions for each gender that are hard to distinguish, it is immediately clear that the frequency of recommendation distribution for the Adamic-Adar algorithm exhibits a gap between males and females (Figure 2) for both the Instagram dataset and the selected DBLP dataset. This confirms and extends the results obtained by Avin et al [1], where a gender bias is observed in the DBLP network.

This gap, exhibited to grow in log-log scale, is a sign of different power coefficients governing the statistical chance to reach at least x recommendations, depending on gender. As one progressively selects to retain only the most successful individuals, the aforementioned effect translates into a sudden and accelerating drop of the observed fraction of females. That drop of female representation is illustrated on Figure 1 in red for Adamic-Adar recommendations and green for the Random Walk, where the size of the circle drawn

helps interpret the relative size of the selected group. The contrast with a similar statistic on organic growth shown in blue is remarkable.

As females are the majority even in the set of recommended users, the effect of gender disparity in recommendations amplifies, since simple homophily would suggest that females should be recommended more throughout. This suggests that degree alone is not sufficient in creating wealth, and that homophily also plays a role and can go as far as inverting inequality.

Table 3 shows this analysis for the initial graph, having as rows the whole set of nodes (split by Male and Female), and as columns the users to whom they are connected. For each row, we show the percentage of males and females, and in parenthesis the ratio to the actual percentage of males and females, respectively. We observe that users labeled as male are more homophilic than users labeled as females, showing an instance of differentiated homophily.

4 MODEL AND DEFINITIONS

We now introduce a model of algorithmic growth under a recommendation algorithm which reproduces the observed reinforcing bias. This serves two purposes: first, it validates that our results are not limited to activity and gender on Instagram, but are more broadly applicable. Second, it offers a starting point to investigate its causes, which may help to alleviate bias in future design. We build upon an extension of *preferential attachment* dynamics (or the linear copying model, equivalently) [1, 9].
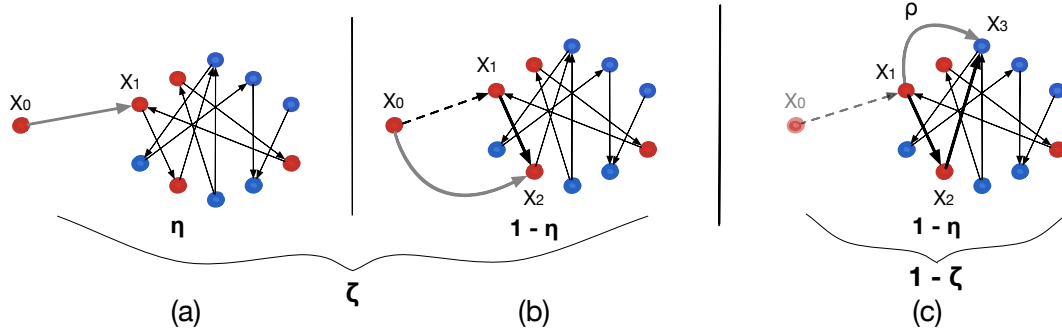
4.1 Organic growth

The first model we introduce, adapted from an existing model, reproduces a network's spontaneous addition of edges, where each node has a label, blue (B) or red (R), and at each point in time the network grows as:

- *Minority-majority partition*: a new node X_0 enters the network and receives label R with probability r and B with probability $1 - r$ (for $0 \leq r \leq 1/2$).
- *Randomness*: with probability η, the new node X_0 chooses an existing node X_1 at random (Figure 3 (a)).
- *Preferential attachment (rich get richer)*: with probability $1 - \eta$, X_0 chooses a node uniformly at random and *copies* one of its edges. This is equivalent to the new node connecting proportional to the ending node's degree, $\mathbb{P}(v \text{ is chosen}) = \delta_t(v)/\sum_{u \in V_t} \delta_t(u)$, where $\delta_t(x)$ denotes the degree of node x at time t and V_t is the set of nodes in the graph at time t (Figure 3 (b)).
- *Homophily:* if the new node has a different label than the node it chooses to connect to, the connection is accepted with probability ρ and the process is repeated until an edge is formed. The homophily parameter $0 \leq \rho \leq 1$ captures that a person less similar is less likely to be eventually chosen that one of the same kind.

Thus, as the network grows according to this model, exactly one node and one directed edge are added at each timestep. We call this process the "organic growth" of the network, as individuals enter the network and form new connections without the influence of any external forces.

Figure 3: Growth of organic and recommendation model. In each of the three cases, a new node X_0 appears (red in this example). In the organic growth model, with probability η, X_0 connects to a random item X_1 (a), and with probability $1 - \eta$ it chooses X_1 according to preferential attachment (b). In the recommendation growth model, with probability ζ it follows the organic growth model, and with probability $1 - \zeta$, a phantom node X_0 is created, chooses a node X_1 at random and follows a random walk of length 2 as $X_1 - X_2 - X_3$ from that item (c) through the bolded arrows. In the end, X_1 connects to X_3. In each case, if the connecting nodes have the same color, they connect; if not, they connect with probability ρ.

4.2 Recommendation algorithm

We extend the organic growth model by adding sparse recommendations between existing nodes. Most recommendation algorithms [11] rely on connecting nodes with one of distance-2 neighbors or friends-of-friends (*e.g.,* common neighbors, Adamic-Adar, Jaccard, random walks). For the purpose of this theoretical analysis, we focus on the random walk of length 2 (see Eq.1 above), as it encompasses the general properties of such algorithms and makes theoretical analysis tractable. We assume that the network grows under the recommendation algorithm in the following way:

- *Minority-majority partition*: a new node X_0 enters the network and receives label R with probability r and B with probability $1 - r$ (for $0 \leq r \leq 1/2$).
- *Organic growth:* with probability ζ, X_0 connects through organic growth.
- *Recommendation:* with probability $1 - \zeta$, X_1 is chosen uniformly at random, and follows a random walk of length 2: $X_1 \rightarrow X_2 \rightarrow X_3$. At the end, X_1 connects to X_3 (if same color, connect; if different color, connect with probability ρ). X_0 can be thought as a virtual node in this case, that disappears as it does not receive any connection (Figure 3 (c)).

We analyze the effects of recommendation asymptotically, assuming that the network size tends to infinity.

4.3 Definitions

Definition 4.1 (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.

$$\exists c < 1 \text{ such that} \quad \lim_{n \to \infty} \frac{\frac{1}{n(R)} \sum_{v \in R} \delta(v)}{\frac{1}{n(B)} \sum_{v \in B} \delta(v)} \leq c. \quad (2)$$

where $n(R)$ is the number of red nodes, and $\delta(u)$ the u's degree.

Definition 4.2 (Tail glass ceiling effect). 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

$$\lim_{n \to \infty} top_k(B) = \infty \text{ and } \lim_{n \to \infty} \frac{top_k(R)}{top_k(B)} = 0, \quad (3)$$

where $top_k(R)$, $top_k(B)$ denotes the number of red and blue nodes with degree at least k, respectively.

Avin et al. [1] proves that organic growth leads to the two subpopulations having their degree distributions following two power laws, with different coefficients, which in turn proves the existence of a glass ceiling effect as per the definitions above.

5 PROOF OF ALGORITHMIC GLASS CEILING

Our theoretical contribution is a novel analysis of how recommendation algorithms can produce an even higher bias than organic growth. In our analysis, we note that the algorithms we analyzed do not create bias *ex nihilo*, but simply amplify bias in networks where it exists. For that reason, we focus on the *Biased Preferential Attachment Model* with the added effect of randomness and recommendation through random walks. We first provide a novel analysis of the way a random walk of length 2 affects the evolution of the in-degree distribution of nodes in the network, and show that the in-degree distribution follows a power law with different coefficients for the red and blue nodes, respectively. We analyze the in-degree since the graph is directed and the out-degree is constant by construction.

We use the formal definitions of the glass ceiling effect from Avin et al. [1], which rely on defining individual wealth as the degree of a node, and collective wealth as the sum of individual wealth values. Thus, we denote by R the set of red nodes and by B the set of blue nodes, $\beta_{i,t}$ the number of red nodes, $u_{i,t}(R)$ as the sum of degrees of all R nodes after t recommendations following random walk (and similarly for blue nodes) for the organic model ($i = 1$) and the recommendation model ($i = 2$).

Since an edge is added at each step, it makes sense to introduce $u_t(R) = t\alpha_t$ and $u_t(B) = t(1 - \alpha_t)$ for the organic model, where α_t is equivalently defined as the sum of degree of the red nodes

divided by the total sum of degree, or simply the fraction of edges created towards R nodes under the recommendation dynamics. We define similarly $\alpha_{2,t}$ for the recommendation model.

Finally, we introduce some useful notation: $\mathcal{P}_i(A|B)$ for $i = 1, 2, 3$ as the conditional probability that a node of color B connects to a node of color A through *randomness*, *preferential attachment*, and *recommendation*, respectively. We observe that

$$
\begin{aligned}
\mathcal{P}_1(R|R) &= \frac{\beta_{2,t}}{\beta_{2,t} + \rho(1 - \beta_{2,t})}, \mathcal{P}_1(R|B) = \frac{\rho\beta_{2,t}}{\rho\beta_{2,t} + 1 - \beta_{2,t}}, \\
\mathcal{P}_2(R|R) &= \frac{\alpha_{2,t}}{\alpha_{2,t} + \rho(1 - \alpha_{2,t})}, \mathcal{P}_1(R|B) = \frac{\rho\alpha_{2,t}}{\rho\alpha_{2,t} + 1 - \alpha_{2,t}}, \\
\mathcal{P}_3(R|R) &= \frac{\mathbb{P}(X_3 \in R | X_1 \in R)}{\mathbb{P}(X_3 \in R | X_1 \in R) + \rho \cdot \mathbb{P}(X_3 \in B | X_1 \in R)}, \\
\mathcal{P}_3(R|B) &= \frac{\rho \cdot \mathbb{P}(X_3 \in R | X_1 \in B)}{\mathbb{P}(X_3 \in B | X_1 \in B) + \rho \cdot \mathbb{P}(X_3 \in R | X_1 \in B)}.
\end{aligned}
\tag{4}
$$

Note that $\mathcal{P}_i(B|R) = 1 - \mathcal{P}_i(R|R)$ and $\mathcal{P}_i(B|B) = 1 - \mathcal{P}_i(R|B)$ for all $i \in \{1, 2, 3\}$. The denominator in each case represents the competition that a node faces when receiving a connection through each dynamic.

The important step in computing each of these probabilities is to analyze the probability for a red (respectively, blue) node to receive a connection, which depends on the dynamics of connection. Figure 4 shows an instance of this analysis, where the random walk $X_1 - X_2 - X_3$ is decomposed into edges $X_1 - X_2$ and $X_2 - X_3$ To compute the probability for $X_1 - X_2$ to have formed, we see that it could have been created through random connection or through preferential attachment, enclosed in circles in Figure 4.

As the random walk can be decomposed in two independent edges, $\mathbb{P}(X_3 \in R | X_1 \in R)$ may be written

$$
\begin{aligned}
&\mathbb{P}(X_3 \in R | X_2 \in R, X_1 \in R) \cdot \mathbb{P}(X_2 \in R | X_1 \in R) \\
&+ \mathbb{P}(X_3 \in R | X_2 \in B, X_1 \in R) \cdot \mathbb{P}(X_2 \in B | X_1 \in R).
\end{aligned}
\tag{5}
$$

Then, denote by $\tilde{\mathcal{P}}(A|B) = \eta \cdot \mathcal{P}_1(A|B) + (1 - \eta) \cdot \mathcal{P}_2(A|B)$, for any two colors A and B. Using a similar reasoning we obtain

$$
\begin{aligned}
\mathbb{P}(X_3 \in R | X_1 \in R) &= \tilde{\mathcal{P}}(R|R)^2 + \tilde{\mathcal{P}}(R|B) \cdot \tilde{\mathcal{P}}(B|R), \\
\mathbb{P}(X_3 \in B | X_1 \in R) &= \tilde{\mathcal{P}}(R|R) \cdot \tilde{\mathcal{P}}(B|R) + \tilde{\mathcal{P}}(B|B) \cdot \tilde{\mathcal{P}}(B|B), \\
\mathbb{P}(X_3 \in R | X_1 \in B) &= \tilde{\mathcal{P}}(R|B) \cdot \tilde{\mathcal{P}}(R|R) + \tilde{\mathcal{P}}(B|B) \cdot \tilde{\mathcal{P}}(R|B), \\
\mathbb{P}(X_3 \in B | X_1 \in B) &= \tilde{\mathcal{P}}(B|B)^2 + \tilde{\mathcal{P}}(R|B) \cdot \tilde{\mathcal{P}}(B|R).
\end{aligned}
\tag{6}
$$

5.1 Glass ceiling in the organic growth model

To show that the organic growth model exhibits a glass ceiling effect, we employ similar techniques to [1], with the additional effect of random connections apart from preferential attachment. Randomness, however, does not affect the degree distribution coefficients (except for nodes of in-degree 0), but only the rate of growth of edges towards the red population.

We firstly note that the rate of growth of edges towards the red population converges towards a constant $\alpha < r$. To see this, we make use of the following lemma, where the function F models the growth of edges towards R:
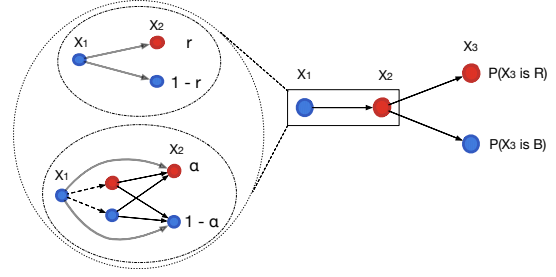
Figure 4: Evolution of the recommendation model. Each edge of the random walk of length 2 (right) is assumed to have been created through organic growth: either through a random connection (top left) or preferential attachment (bottom left). The end node can be red or blue with different probabilities, $(r, 1 - r)$ in random connections, $(\alpha, 1 - \alpha)$ in preferential attachment, and $(\mathbb{P}(X_3 \in R), 1 - \mathbb{P}(X_3 \in R))$ in the random walk.

LEMMA 5.1. $\mathbb{E}[\alpha_{t+1}|\alpha_t] = \alpha_t + \frac{F(\alpha_t) - \alpha_t}{t + 1}$, where

$$
\begin{aligned}
F(\alpha_t) = {}&\eta r \cdot \mathcal{P}_1(R|R) + \zeta\eta(1 - r) \cdot \mathcal{P}_1(R|B) + \\
&(1 - \eta)r \cdot \mathcal{P}_2(R|R) + \zeta(1 - \eta)(1 - r) \cdot \mathcal{P}_2(R|B).
\end{aligned}
$$

We omit the proofs here due to space constraint and similarity in methods, reserving space for the analysis of the recommendation model. We note that α is the fixed point of the function F.

Degree distribution. We note that the in-degree distribution of the two populations follows a power law with different coefficients, similar as in [1], with a slight addition of the randomness factor.

Denote by $m_{k,t}(B)$ and $m_{k,t}(R)$ the number of blue and red nodes of degree k at time t, and define

$$
M_k(x) = \lim_{t \to \infty} \frac{\mathbb{E}(m_{k,t}(x))}{t}.
\tag{7}
$$

THEOREM 5.2. *For the organic growth model, the in-degree distribution of the red and blue populations also follows a power law, i.e., $M_k(R) \sim k^{-\beta(R)}$ and $M_k(B) \sim k^{-\beta(B)}$ for coefficients $\beta(R) = 1 + \frac{1}{C^R}$ and $\beta(B) = 1 + \frac{1}{C^B}$, where*

$$
\begin{aligned}
C^R &= (1 - \eta) \cdot \left(\frac{r}{\alpha} \cdot \mathcal{P}_2(R|R) + \frac{1 - r}{\alpha} \cdot \mathcal{P}_2(R|B) \right) \text{ for } k > 0, \\
C^R &= \eta \cdot \left(\frac{r}{\beta_t} \cdot \mathcal{P}_1(R|R) + \frac{1 - r}{\beta_t} \cdot \mathcal{P}_1(R|B) \right) \text{ for } k = 0.
\end{aligned}
\tag{8}
$$

A similar analysis follows for the blue population.

5.2 Glass ceiling in the recommendation model

Convergence of wealth of red nodes. We begin by showing that the sum of degrees of the red nodes converges to a constant that is smaller than r, and even smaller than in the organic case:

$$
\lim_{t \to \infty} \alpha_{2,t} = \alpha_2 < \alpha < r.
\tag{9}
$$

This first shows that power inequality is even stronger under recommendation, since the average degree of a red node is α/r in the organic model and α_2/r in the recommendation model.

To show this, we define a new function F_2 that represents the rate of growth of the number of edges towards R. We make a few important assumptions:

- edges considered in the random walk are assumed to have formed through organic growth[1].
- edges considered in the random walk are assumed to exist independently of each other[2].
- edges that already exist are naturally formed at different steps in time; however, when computing an evolution equation, we set their time of formation to the current time, t.

LEMMA 5.3. $\mathbb{E}[\alpha_{2,t+1}|\alpha_{2,t}] = \alpha_{2,t} + \frac{F_2(\alpha_{2,t})-\alpha_{2,t}}{t+1}$, where

$$F_2(\alpha_{2,t}) = \zeta \cdot \eta \cdot r \cdot \mathcal{P}_1(R|R) + \zeta \cdot \eta \cdot (1-r) \cdot \mathcal{P}_1(R|B) +$$
$$\zeta \cdot (1-\eta) \cdot r \cdot \mathcal{P}_2(R|R) + \zeta \cdot (1-\eta) \cdot (1-r) \cdot \mathcal{P}_2(R|B) +$$
$$(1-\zeta)\left(r \cdot \mathcal{P}_3(R|R) + (1-r) \cdot \mathcal{P}_3(R|B)\right).$$

PROOF. We analyze the number of edges towards R that appear at each step. In doing so, we note that

$$\mathbb{E}(u_{2,t+1}(R)|\alpha_{2,t}) = 0 \cdot \mathbb{P}(u_{2,t+1}(R) = u_{2,t}(R)) +$$
$$1 \cdot \mathbb{P}(u_{2,t+1}(R) = u_{2,t}(R) + 1). \quad (10)$$

The formula for F_2 encapsulates $\mathbb{P}(u_{2,t+1}(R) = u_{2,t}(R) + 1)$ as it is defined as the rate of growth of edges towards R. □

The following lemma summarizes properties of F_2 (proof omitted here due to space constraints), which exploits the polynomial property of the function. The final property, the hardest to prove, exploits a proof by contradiction as $\alpha^* < \alpha_2^*$ (Figure 5 (a)) implies that from a given fraction of edges towards R, the fraction increases under the random walk dynamics while it decreases under organic growth, which can be shown to be impossible. Moreover, α^* and α_2^* cannot be equal unless $\rho = 1$ or $r = 1/2$, given the polynomial function they satisfy. Considering F_2 as a function of x, the following holds:

LEMMA 5.4. For $r < 1/2$ and $\rho \in [0,1]$, we have

(1) F_2 is monotonically increasing.
(2) F_2 has exactly one fixed point in $[0,1]$, denoted by α_2^*, which is stable.
(3) $F_2[0,1] \subset [0,1]$.
(4) $\alpha_2^* < \alpha^*$ where α^* denotes the fraction of edges going towards red nodes in the organic growth.

Degree distribution. We show that the degree distribution of the two populations also follows a power law, and that its coefficients for the blue and red populations are even further apart than the ones stemming from organic growth (Theorem 5.6). Following similar notation as for organic growth, denote by $m_{2,k,t}(B)$ and $m_{2,k,t}(R)$ the number of blue and red nodes of degree k at time t in the

---

[1]Note that this may be a feature of the recommendation system, as it is desirable to avoid its own previous suggestions and feed solely on the organic growth behavior.
[2]Edges lead to older nodes, the choice of the first edge may differ owing to (among all others) how the second edge was chosen. That effect is very small and we neglect it.
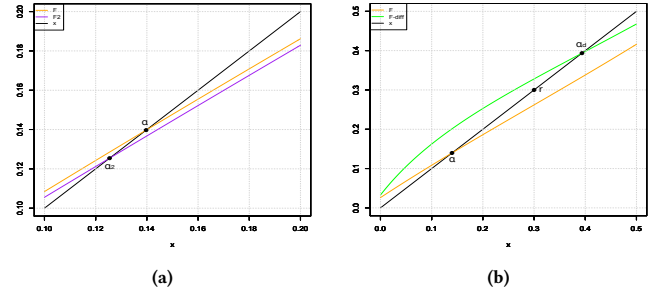

(a)   (b)

Figure 5: (a) Functions F (orange) and F_2 (purple) for $(\zeta, \eta, r, \rho) = (0.9, 0.1, 0.3, 0.5)$; (b)Functions F (orange) and F_d (green) for $(\eta, r, \rho_R, \rho_B) = (0.1, 0.3, 0.2, 0.5)$.

recommendation model and $M2, k(x)$ as the limiting expectation, similarly.

THEOREM 5.5. *In the case of the random walk of length 2, the degree distribution of the red and blue populations also follows a power law, i.e. $M_{2,k}(R) \sim k^{-\beta_2(R)}$ and $M_{2,k}(B) \sim k^{-\beta_2(B)}$ for coefficients $\beta_2(R)$ and $\beta_2(B)$.*

PROOF. We first derive a recurrence relation for $m_{2,k,t}(R)$ and $m_{2,k,t}(B)$, respectively. Conditioning on the history of the growth until time t, $\mathcal{F}_{2,t}$, we find that the evolution can be written as

$$\mathbb{E}[m_{2,k,t+1}(R)|\mathcal{F}_{2,t}] = \frac{m_{2,k,t}(R)}{t}\left(1 - k \cdot C_{2,t}^R - T_{2,t}^R\right)$$
$$+ \frac{m_{2,k-1,t}(R)}{t} \cdot \left((k-1) \cdot C_{2,t}^R + T_{2,t}^R\right), \quad (11)$$

$T_{2,t}^R$ is the rate for red nodes to receive edge through *randomness*:

$$T_{2,t}^R := \zeta \cdot \eta \cdot \left(\frac{r}{\beta_t} \cdot \mathcal{P}_1(R|R) + \frac{1-r}{\beta_t} \cdot \mathcal{P}_1(R|B)\right), \quad (12)$$

$k \cdot C_{2,t}^R$ is the rate at which nodes connect to red nodes of degree k through *preferential attachment* or *recommendations*:

$$C_{2,t}^R := \zeta(1-\eta)\left(\frac{r}{\alpha_{2,t}} \cdot \mathcal{P}_2(R|R) + \frac{1-r}{\alpha_{2,t}} \cdot \mathcal{P}_2(R|B)\right)$$
$$+ (1-\zeta)\left(\frac{r}{\alpha_{2,t}} \cdot \mathcal{P}_3(R|R) + \frac{1-r}{\alpha_{2,t}} \cdot \mathcal{P}_3(R|B)\right). \quad (13)$$

Note that the additional factors of $\frac{1}{\beta_{2,t}}$ and $\frac{1}{\alpha_{2,t}}$ come from the fact that the node receiving the connection must be of degree k (a similar analysis for degree $k-1$), and so the probability to choose it between all other red nodes is $\frac{m_{2,k,t}(R)\cdot k}{t\cdot\beta_{2,t}}$ when the choice is random and $\frac{m_{2,k,t}(R)\cdot k}{t\cdot\alpha_{2,t}}$ when the choice is through preferential attachment or recommendation.

We note that a similar concentration argument applies as in [1], showing that $\beta_{2,t} \to r$ and $\alpha_{2,t} \to \alpha_2$. Thus, we may define $C_2^R := \lim_{t\to\infty} C_{2,t}^R$ and $T_2^R := \lim_{t\to\infty} T_{2,t}^R$. From here, one can reuse equations 7

and 11 and Lemma 4.14 from Avin et al. [1] to write, for large k,

$$\frac{M_k(R)}{M_{k-1}(R)} = \frac{(k-1)C_2^R + T_2^R}{1 + kC_2^R + T_2^R} = 1 - \frac{1 + 1/C_2^R}{k} + O\left(\frac{1}{k^2}\right), \quad (14)$$

thus proving that the degree distribution follows a power law (the factor pertaining to random choices vanishes). A similar analysis follows for the blue population, and we can deduce that degrees quickly converge to a power law as t gets large, where coefficient $\beta_2(R) = 1 + \frac{1}{C_2^R}$ and $\beta_2(B) = 1 + \frac{1}{C_2^B}$. □

THEOREM 5.6. *For $r \in [0, 1/2]$ and $\rho \in [0, 1]$, we have*

$$\beta_2(R) > \beta(R) > 3 > \beta(B) > \beta_2(B). \quad (15)$$

where $\beta(R)$ and $\beta(B)$ are the equivalent power law coefficients under organic growth. This implies tail and strong glass ceiling effects that are even more pronounced under algorithmic recommendations.

PROOF. Due to space constraints we present a sketch of the proof. It leverages the previous comparison $\alpha_2^* < \alpha^*$ together with an *invariant* equation that is, to the best of our knowledge, novel.

LEMMA 5.7. *Given α and α_2 the limits of the fraction of edges towards R as defined above, the following holds:*

$$\alpha \cdot C^R + r \cdot T^R = \alpha, \text{ and } \alpha_2 \cdot C_2^R + r \cdot T_2^R = \alpha_2. \quad (16)$$

This lemma follows from the observation that at equilibrium, when t is large, the rate at which red edges appear must equal the current fraction of red edges, as it does not evolve anymore. Revisiting Eq. (11-13), we find that the rate at which edges towards red nodes of degree k are created is $\frac{m_{k,t}(R)}{t} \cdot (k \cdot C_t^R + T_t^R)$. Summing up after all nodes, the rate at which edges towards R are created is:

$$\sum_{k \geq 0} \frac{m_{k,t}(R)}{t} \cdot (k \cdot C_t^R + T_t^R) = C_t^R \cdot \underbrace{\sum_{k \geq 0} \frac{k \cdot m_{k,t}(R)}{t}}_{\alpha_t} + T_t^R \cdot \underbrace{\sum_{k \geq 0} \frac{m_{k,t}(R)}{t}}_{\beta_t}$$

When t is large, this must be equal to the rate at which edges towards R are created, which is by definition α. An identical analysis follows for the recommendation model, leading to Theorem 5.6. Using $\alpha_2 < \alpha$ and the invariant equation above, $C_2^R < C_R$ follows as ζ is close to 1 and T_R and T_R^2 differ only by a factor of ζ.

A similar proof indicates that $(1-\alpha) \cdot C_B + (1-r)T_B = 1-\alpha$, shows that the coefficient for the blue population degree distribution after the random walk is larger than the one after organic growth ($C_{2,B} > C_B \Leftrightarrow \beta_2(B) > \beta(B)$). □

6 EXTENSIONS

As we have seen, homophily plays a central role in the creation of a glass ceiling effect, yet characterizing its contribution precisely becomes challenging in complex models. For populations exhibiting equal rates of homophily, our results show that the majority have an advantageous position. But what happens in the case of *differential homophily*, where some populations may accept to connect with more diverse individuals than others? A natural question is: can the homophily parameter be used advantageously by the minority population in reversing the glass ceiling effect? We explore this question below, showing that in cases where the minority has a

higher homophily parameter, the hierarchy between the majority and minority populations can be reversed.

For the organic growth model, define ρ_R and ρ_B as the accept probability for incoming red and blue nodes, respectively. A new growth function F (denoted F_d) then occurs as $\mathcal{P}_1$ and $\mathcal{P}_2$ change:

$$F_d(x) = \eta \cdot \left(r \cdot \frac{r}{r + \rho_R(1-r)} + (1-r) \cdot \frac{\rho_B r}{\rho_B r + 1 - r} \right)$$
$$+ (1-\eta) \cdot \left(r \cdot \frac{x}{x + \rho_R(1-r)} + (1-r) \cdot \frac{\rho_B x}{\rho_B x + 1 - x} \right). \quad (17)$$

Setting $\rho_R < \rho_B$, we effectively allow the red population to fewer more incoming edges from B than the blue nodes accept from R, in a sense "protecting" themselves. We show an instance of this in Figure 5 (b), where we set $\rho_R = 0.2$ and $\rho_B = 0.5$ for F_d (green) and $\rho_R = \rho_B = \rho = 0.5$ for the original F (orange). Note that the fixed point of F_d, α_d is now larger than r, which shows that the accumulated wealth of the red nodes has surpassed the expectation (their fraction, r), effectively reversing the hierarchy between the red and the blue nodes.

Thus, in alleviating the algorithmic glass ceiling effect, we propose that the homophily parameter ρ be accounted for in designing structure-based recommendation algorithms. In our case, given a network with labels, a majority-minority partition, and a recommendation model based on random walks, one can infer and use the value of ρ in counteracting the bias, i.e., conduct a random walk which "accepts" with probability ρ a transition that leads to same label connection. This would be able to cancel the effect of homophily on the random walk and ought to equalize influence and power law coefficients among labels. It is not evident how to infer ρ from the dataset, as it is not simply equal to the observed fraction of mixed-label edges, since rejected edges are lost. However, as one can approximate the fraction of minority population r and the fraction of their total power, α, we can replace them as parameters in the function F from Lemma 4.3 from [1], knowing that F has α as its fixed point. Thus, considering $F(\alpha) - \alpha$ as a function of ρ, we know that it must be equal to 0 from the fixed-point equation, so we can solve for ρ and obtain an approximate value.

7 CONCLUSIONS

Our theoretical and empirical analysis unravel subtle dynamics of network growth, showing the way homophily and preferential attachment tie into the creation of a glass ceiling effect in a network that undergoes recommendations. We have thus formalized the algorithmic glass ceiling effect, and showed that while algorithms do not create disparity out of nowhere, they can worsen pre-existing inequality.

Our empirical results reverse the hierarchy projected by the theoretical analysis in an unequal network in a surprising but not contradictory manner. Indeed, we show how a subtle mechanism of network growth, namely, differential homophily, can contribute in reversing the glass ceiling effect against the majority population. We hope these results will help future design of social recommendations that optimize relevance while preventing glass ceiling.

REFERENCES

[1] C Avin, B Keller, Z Lotker, C Mathieu, and David Peleg. 2015. Homophily and the glass ceiling effect in social networks. *ITCS* (2015).

[2] S Barocas and A D Selbst. 2016. Big data's disparate impact. *California Law Review* (2016).

[3] Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. *Advances in Neural Information Processing Systems (NIPS)* (2016), 4349–4357.

[4] Alexandra Chouldechova. 2017. Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *Big data* 5, 2 (2017), 153–163.

[5] Y Dong, R A Johnson, J Xu, and N V Chawla. 2017. Structural Diversity and Homophily: A Study Across More Than One Hundred Big Networks. In *KDD '17*. ACM Press, New York, New York, USA, 807–816.

[6] Maeve Duggan. 2015. Mobile Messaging and Social Media 2015. *Pew Research Center* (2015).

[7] Benjamin G Edelman, Michael Luca, and Dan Svirsky. 2017. Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American Economic Journal: Applied Economics* 9, 2 (2017), 1–22.

[8] A Hannak, C Wagner, D Garcia, Alan Mislove, Markus Strohmaier, and Christo Wilson. 2017. Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr. *Proceedings of CSCW* (2017).

[9] Ravi Kumar, Prabhakar Raghavan, S Rajagopalan, D Sivakumar, Andrew Tomkins, and E Upfal. 2000. Stochastic models for the Web graph. *IEEE FOCS* (2000), 57–65.

[10] Michael Ley. 2009. DBLP: some lessons learned. *Proceedings of the VLDB Endowment* 2, 2 (2009), 1493–1500.

[11] David Liben-Nowell and Jon Kleinberg. 2007. The link-prediction problem for social networks. *journal of the Association for Information Science and Technology* 58, 7 (2007), 1019–1031.

[12] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a Feather: Homophily in Social Networks. *Annual review of sociology* 27 (2001), 415–444.

[13] Alan Mislove, S Lehmann, Y Y Ahn, and J-P Onnela. 2011. Understanding the Demographics of Twitter Users. In *ICWSM*.

[14] Shirin Nilizadeh, Anne Groggel, Peter Lista, Srijita Das, Yong-Yeol Ahn, Apu Kapadia, and Fabio Rojas. 2016. Twitter's Glass Ceiling: The Effect of Perceived Gender on Online Visibility.. In *ICWSM*. 289–298.

[15] Flávio Souza, Diego de Las Casas, Vinícius Flores, SunBum Youn, Meeyoung Cha, Daniele Quercia, and Virgílio Almeida. 2015. Dawn of the selfie era: The whos, wheres, and hows of selfies on Instagram. In *Proceedings of the 2015 ACM on conference on online social networks*. ACM, 221–231.

[16] J Su, A Sharma, and S Goel. 2016. The Effect of Recommendations on Network Structure. *WWW '18: Proceeding of the 25th International Conference on World Wide Web* (2016).

[17] Ke Yang and Julia Stoyanovich. 2016. Measuring Fairness in Ranked Outputs. *Proceedings of Workshop FATML* (2016).

[18] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rodriguez, and Krishna Gummadi. 2017. Fairness Beyond Disparate Treatment & Disparate Impact: Learning Classification without Disparate Mistreatment. In *WWW '17 Proceedings of the 26th International Conference on World Wide Web*.