

Quantifying Gendered Citation Imbalance in Computer Science Conferences

Kazuki Nakajima¹, Yuya Sasaki², Sohei Tokuno³, George Fletcher⁴

¹Tokyo Metropolitan University

²Osaka University

³Nara Institute of Science and Technology

⁴Eindhoven University of Technology

nakajima@tmu.ac.jp, sasaki@ist.osaka-u.ac.jp, tokuno.sohei@cdg.naist.jp, g.h.l.fletcher@tue.nl

Abstract

The number of citations received by papers often exhibits imbalances in terms of author attributes such as country of affiliation and gender. While recent studies have quantified citation imbalance in terms of the authors' gender in journal papers, the computer science discipline, where researchers frequently present their work at conferences, may exhibit unique patterns in gendered citation imbalance. Additionally, understanding how network properties in citations influence citation imbalances remains challenging due to a lack of suitable reference models. In this paper, we develop a family of reference models for citation networks and investigate gender imbalance in citations between papers published in computer science conferences. By deploying these reference models, we found that homophily in citations is strongly associated with gendered citation imbalance in computer science, whereas heterogeneity in the number of citations received per paper has a relatively minor association with it. Furthermore, we found that the gendered citation imbalance is most pronounced in papers published in the highest-ranked conferences, is present across different subfields, and extends to citation-based rankings of papers. Our study provides a framework for investigating associations between network properties and citation imbalances, aiming to enhance our understanding of the structure and dynamics of citations between research publications.

Introduction

Citation analysis serves as a standard bibliometric tool for assessing the impact of different papers, researchers, institutions, and countries (Waltman 2016; Zeng et al. 2017; Fortunato et al. 2018). The outcomes of such analyses are pivotal, as they influence various facets of research evaluation, including the hiring and promotion of researchers (Moed 2006; Moher et al. 2018). However, achieving unbiased citation analysis remains a challenge due to the diverse citation practices adopted by individual researchers and across research disciplines (Radicchi, Fortunato, and Castellano 2008). Consequently, citation imbalances persist, such as the over-citation of certain journal groups (Kojaku, Livan, and Masuda 2021) and country groups (Gomez, Herman, and Parigi 2022).

Motivation. Gender imbalance persists in computer science across various aspects, including educational attainment (OECD 2017), faculty representation (Laberge et al. 2022), and career progression (Lietz et al. 2024). Addressing this imbalance is crucial for promoting gender equality in education, academia, industry, and society (Jaccheri, Pereira, and Fast 2020). While recent studies have highlighted citation imbalances related to authors' gender in neuroscience and physics (Dworkin et al. 2020; Teich et al. 2022), the extent of gendered citation imbalance in computer science remains largely unknown. Previous findings from other disciplines may not be directly applied to the computer science discipline due to unique publication practices. Unlike many disciplines where journals are the primary publication venue, computer scientists often choose conferences to present their work (Freyne et al. 2010; Vrettas and Sanderson 2015). In addition, conference ranks, rather than journal impact factors, may influence citation practices of computer scientists (Freyne et al. 2010; Vrettas and Sanderson 2015). These distinct research practices, combined with the technical challenges of inferring authors' gender, make it difficult to quantify gendered citation imbalances in computer science using existing data sources alone.

We represent citations between papers as a directed network to quantify gendered citation imbalance. Citation networks are valuable for understanding the scientific knowledge space, as they exhibit various structural and dynamical properties (Zeng et al. 2017). Notably, citation networks often exhibit *homophily in citations*, where papers tend to cite others with similar characteristics (Ciotti et al. 2016). Moreover, there exists *heterogeneity in the number of citations received per paper*, with most papers garnering few citations while a small fraction receives a substantial number (Eom and Fortunato 2011). Given that these two properties often drive other structural and dynamical properties in empirical citation networks (Zeng et al. 2017), we hypothesize that they influence gender imbalance in citations received by papers. Investigating this hypothesis necessitates reference models that randomize citations while preserving these properties in the given citation network. However, existing reference models are not intended to investigate associations between citation imbalances and these properties.

Contribution. The present study makes three significant contributions. First, we construct a new citation network

based on the OpenAlex (Priem, Piwowar, and Orr 2022) and DBLP (The DBLP team 2024b) databases. Our dataset encompasses rich metadata of papers, such as authors' gender and conference ranks. Second, we develop a family of reference models for quantifying citation imbalances. While the existing reference model preserves the number of citations made by each paper (Dworkin et al. 2020), our models additionally preserve up to homophily in citations and heterogeneity in the number of citations received per paper. Third, we quantify gender imbalance in citations between papers published in computer science conferences and examine how network properties influence this gendered citation imbalance. Our key findings are as follows.

- Conference papers written by female authors as the first and/or last authors are less likely to receive citations than expected by reference models.
- Homophily in citations exhibits a strong association with gendered citation imbalance, whereas heterogeneity in the number of citations received per paper shows a minor association.
- Gendered citation imbalance is most pronounced among conference papers published in the highest-ranked conferences and persists across different subfields.
- Conference papers written by female authors as the first and/or last authors are less likely to appear at the top of citation-based rankings of papers than expected by reference models.

Related Work

Quantifying Gender Imbalance in Academia Gender imbalance in academia has been quantitatively investigated using bibliometric data (Sugimoto and Larivière 2023). Previous studies have focused on gender imbalances within individual disciplines because the extent of gender imbalance and research practices, including authorship, publication venue, and citation, may depend on the discipline. Dworkin et al. investigated gender imbalance in citations between journal papers within the neuroscience discipline (Dworkin et al. 2020). Teich et al. conducted similar analyses within the physics discipline (Teich et al. 2022). Jadidi et al. examined gender imbalance in research careers, productivity, and collaboration within the computer science discipline (Jadidi et al. 2018). Avin et al. proposed a mathematical model that partially explains the causes of the so-called glass ceiling effect and empirically observed gender imbalance in collaborations between mentors and students in the computer science discipline (Avin et al. 2015). In our study, we investigate gendered citation practices within the computer science discipline, aiming to expand upon previous findings regarding gender imbalance in academia.

Various aspects of gender imbalance in academia have been investigated. Huang et al. found that gender imbalance in the productivity of individual researchers is sufficiently explained by that in the length of their publishing careers (Huang et al. 2020). Li et al. found that gender imbalance in the research performance of researchers with publishing careers of at least 15 years is largely explained by that in the number of collaborators (Li et al. 2022). Gender imbalance

in citations is partially explained by gendered patterns in collaborators of authors and the proximity of research subfields between papers (Dworkin et al. 2020; Teich et al. 2022). In contrast to these studies, we investigate the extent to which homophily in citations and heterogeneity in the number of citations received per paper explain gendered citation imbalance in computer science.

Applications of Reference Models for Citation Networks

In general, reference models generate instances that preserve certain properties of the original network and randomize other properties (Orsini et al. 2015; Nakajima, Shudo, and Masuda 2022). Standard practice in network analysis involves comparing the structure and dynamics of the original network with those of randomized instances generated by reference models (Orsini et al. 2015; Nakajima, Shudo, and Masuda 2022). Previous studies have deployed reference models to characterize the structure of citation networks. For example, Uzzi et al. investigated atypical combinations of citations associated with the impact of a paper by randomizing citations via a Monte Carlo algorithm (Uzzi et al. 2013). Kojaku et al. examined anomalous citation patterns among journals using a reference model that accounts for scientific communities and journal size (Kojaku, Livan, and Masuda 2021). These reference models preserve the number of citations received by each paper and, thus, are not intended to quantify imbalances in citations received by papers. Dworkin et al. introduced a reference model for quantifying gender imbalance in citations received by papers. This reference model, however, does not preserve properties other than the number of citations made by each paper. In light of these limitations, we develop a family of reference models that preserve up to the number of citations made by each paper, homophily in citations, and heterogeneity in the number of citations received per paper.

New Dataset

We construct a citation network composed of conference papers within the computer science discipline. To this end, we use two open databases of publication records: OpenAlex and DBLP. OpenAlex provides extensive coverage of publications and citations between them, while DBLP offers precise conference metadata for computer science publications. By integrating data from both sources, we construct a publication dataset with enriched metadata.

Existing Data Sources

OpenAlex OpenAlex provides hundreds of millions of publication records with rich metadata across various disciplines (Priem, Piwowar, and Orr 2022). We used publication data from a snapshot of OpenAlex released on February 27th, 2024, extracting 103,335,085 research articles published between 2000 and 2020. For each paper v , the title, publication date, primary research topic, authors' information (i.e., OpenAlex ID, name, and affiliations), papers cited by v , and author position (i.e., first, middle, or last) are available. Each paper is assigned one of the 4,516 research topics as the primary research topic based on its metadata (e.g., title, abstract, and citations). Each research topic belongs to

one of the 252 research subfields (OpenAlex 2024). OpenAlex uses a proprietary algorithm to assign a unique ID to each author (Priem, Piwowar, and Orr 2022), and the algorithm identified 45,191,646 authors. While OpenAlex covers a large extent of publication metadata (Delgado-Quirós and Ortega 2024), we observed that conference names in OpenAlex data are often missing or not disambiguated for conference papers in computer science.

DBLP DBLP provides millions of publication records in computer science (The DBLP team 2024b). We used publication data from a snapshot of DBLP released on April 1st, 2024, and extracted 2,583,968 conference papers published between 2000 and 2020. Each paper has the title, publication year, authors’ names, and conference names. While DBLP covers a large extent of conference names and curates them (The DBLP team 2024a), we observed that it offers less metadata for publications compared to OpenAlex. Specifically, DBLP does not provide research subfields and topics for publications.

Construction of a New Dataset

We integrate publication data from OpenAlex and DBLP to complement their respective limitations. We constructed a citation network composed of 99,329 papers (i.e., nodes) and 152,598 citations (i.e., directed edges) between them. Table 1 shows basic statistics of the dataset. See Appendix A in Ref. (Nakajima et al. 2024) for additional statistics. Below, we describe the process of constructing our dataset.

Country of Affiliation and Gender of Authors We assign the country of affiliation and gender to authors in the OpenAlex data using the methods outlined in (Nakajima et al. 2023). First, we assigned a country of affiliation to each author based on the most frequently appearing country in the affiliations of their publication records. Next, we assigned gender to each author based on their country of affiliation, first publication year, and first name. We used the Gender API¹ to infer the gender of authors. We set the same parameters as (Nakajima et al. 2023) for assigning gender based on the outputs of the API. We then categorized each paper into one of four gender categories (i.e., ‘MM’, ‘MW’, ‘WM’, and ‘WW’) based on the gender of the first and last authors, where the first letter, M or W, indicates that the first author is a man or woman, respectively, and the second letter indicates the gender of the last author. Note that we categorized sole-author papers by men and women as MM and WW papers, respectively.

Conference Rank We used the CORE ranking data in 2021² to identify the rank of conferences. This conference ranking is based on citation statistics of papers and authors presented at the conference and the acceptance rate for the conference. The conferences are classified into four ranks: ‘A*’, ‘A’, ‘B’, and ‘C’, with A* being the highest and C the lowest. The ranking includes 768 conference names in our dataset. We manually assigned one of these four ranks to each of the conference names appearing in the DBLP data.

Meaning	Count
Papers	99,329
Citations	152,598
Gender-assigned authors	83,231
Conferences	637
Countries of affiliation	102
Research topics	1,351
Research subfields	187
A*-ranked conferences	59
A-ranked conferences	124
B-ranked conferences	257
C-ranked conferences	197
Female authors	13,097
Male authors	70,134
MM papers	76,562
MW papers	8,281
WM papers	10,479
WW papers	4,007

Table 1: Summary of our dataset.

Matching Publications Between OpenAlex and DBLP

We complete the conference metadata for publications in the OpenAlex data by matching publications between OpenAlex and DBLP based on the following criteria: (i) both papers are published in the same year, (ii) both papers have identical sets of author last names³, and (iii) the titles of both papers differ by no more than 25% in terms of the Levenshtein distance between the two titles divided by the length of the longer title (Huang et al. 2020). If the metadata between OpenAlex and DBLP is inconsistent, we use OpenAlex metadata except for the conference name.

Citations Given an observed citation from paper u to paper v , we say that u made the citation to v , and v received the citation from u . We removed citations made by u to v if either of the following holds true: (i) the publication date of v is ten years older than that of u , (ii) both first and last authors of v are either the first or last author of u . The first criterion addresses the different dynamics of citations received by different papers (Sinatra et al. 2016); the second criterion avoids self-citations (Nakajima et al. 2023). We also removed papers that neither make nor receive any citations.

Reference Models for Citation Networks

We aim to characterize gender imbalance in citations received by conference papers in computer science. To this end, we compare gendered citation patterns in the original network with those in reference models. Our reference models randomize citations made by each paper while preserving certain structural properties of the original network.

Below, we first describe an existing reference model that preserves the number of citations made by each paper (Dworkin et al. 2020). Then, we extend this model to reference models that preserve up to the number of citations

¹<https://gender-api.com/en/> (Accessed April 2024)

²<https://portal.core.edu.au/conf-ranks/> (Accessed April 2024)

³We regarded the last space-separated word of a given author’s name as their last name (Nakajima et al. 2023).

made by each paper, homophily in citations, and heterogeneity in the number of citations received per paper. We list the properties that each reference model preserves in Table 2. See Appendix B in Ref. (Nakajima et al. 2024) for the pseudocodes of the three reference models.

Notation

We denote by $V = \{v_1, \dots, v_N\}$ the set of all the papers (i.e., nodes), E the set of citations (i.e., directed edges) in the network, and $M = |E|$ the number of citations, where N is the number of papers. We denote by $(A_{ij})_{1 \leq i \leq N, 1 \leq j \leq N}$ an adjacency matrix of the network, where $A_{ij} = 1$ if (v_i, v_j) is in E and $A_{ij} = 0$ otherwise. We denote by $k_i = \sum_{j=1}^N A_{ij}$ the number of citations made by $v_i \in V$, and we denote by $c_j = \sum_{i=1}^N A_{ij}$ the number of citations received by $v_j \in V$.

Random-Draws Model

We first describe the random-draws model (Dworkin et al. 2020). Consider an observed citation made by paper $v_i \in V$ to paper $v_{i'} \in V$. This model assumes that $v_{i'}$ is drawn uniformly at random from a set of papers that v_i could potentially cite, which we denote by $\bar{V}_{\text{RD}}(i)$. The subscript ‘RD’ indicates the ‘random-draws’ model. According to our definition of citations between papers, we define $\bar{V}_{\text{RD}}(i)$ as the set of papers that meet the following criteria: (i) the publication date of each paper is at most ten years older than that of v_i , and (ii) both the first and last authors of each paper are neither the first nor last authors of v_i . Then, any paper in $\bar{V}_{\text{RD}}(i)$ receives a citation from v_i with the probability $1/|\bar{V}_{\text{RD}}(i)|$ under the random-draws model. Therefore, the probability of a citation made by v_i to $v_j \in V$ under the model is given by

$$\bar{w}_{ij,\text{RD}} = \begin{cases} k_i/|\bar{V}_{\text{RD}}(i)| & \text{if } v_j \in \bar{V}_{\text{RD}}(i), \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Then, aggregating the probability of a citation received by v_j from other papers yields the expected number of citations received by v_j under the model:

$$\bar{c}_{j,\text{RD}} = \sum_{i=1}^N \bar{w}_{ij,\text{RD}}. \quad (2)$$

The random-draws model preserves the expected number of citations made by each paper in the original network. Indeed, it holds true that $\sum_{j=1}^N \bar{w}_{ij,\text{RD}} = k_i$ for any $v_i \in V$.

Homophilic-Draws Model

Empirical citation networks often exhibit homophily, where papers tend to cite those with similar properties (Ciotti et al. 2016). Homophily in citations may be associated with gender imbalance in citations received by papers (Nakajima et al. 2023). To investigate this possibility, we need to compare gendered citation patterns between a reference model that destroys homophilic citation patterns in the original network and one that preserves them. The random-draws model destroys homophilic citation patterns. Therefore, we extend the random-draws model to a reference model that preserves

homophilic citation patterns in the original network, which we refer to as the homophilic-draws model.

We denote by S a set of properties of the paper, other than the gender category, that can be relevant to homophilic citations. In this study, we define the set of the following three properties as S : (i) conference rank, (ii) country of affiliation, and (iii) research topic. Note that S is not limited to these properties. For example, one may focus on journal names for citation networks in other disciplines (Kojaku, Livan, and Masuda 2021).

We describe the algorithmic procedure. Consider an observed citation made by $v_i \in V$ to $v_{i'} \in V$. We define $\bar{V}_{\text{HD}}(i, i', S)$ as the set of papers in $\bar{V}_{\text{RD}}(i)$ that belong to the same category⁴ as $v_{i'}$ for any property in S . Note that $\bar{V}_{\text{HD}}(i, i', S)$ includes $v_{i'}$. The subscript ‘HD’ indicates the ‘homophilic-draws’ model. We assume that any paper in $\bar{V}_{\text{HD}}(i, i', S)$ receives a citation from v_i with the probability $1/|\bar{V}_{\text{HD}}(i, i', S)|$ under the homophilic-draws model. Then, the probability of a citation made by v_i to $v_j \in V$ under the model is given by

$$\bar{w}_{ij,\text{HD}} = \sum_{\substack{i'=1 \\ (v_i, v_{i'}) \in E \\ v_j \in \bar{V}_{\text{HD}}(i, i', S)}}^N 1/|\bar{V}_{\text{HD}}(i, i', S)|. \quad (3)$$

The expected number of citations received by $v_j \in V$ under the model is given by

$$\bar{c}_{j,\text{HD}} = \sum_{i=1}^N \bar{w}_{ij,\text{HD}}. \quad (4)$$

The homophilic-draws model preserves the number of citations made by each paper. In addition, it preserves the number of citations between papers in terms of any pair of conference ranks, any pair of countries of affiliation, and any pair of research topics (see Figs. 1(b)–1(d) for numerical evidence).

Preferential-Draws Model

Empirical citation networks may exhibit a greater extent of heterogeneity in the number of citations received per paper than expected under the homophilic-draws model. One possible explanation is that citation dynamics largely adhere to the preferential attachment mechanism (i.e., papers are more likely to cite papers that receive more citations in a research topic) (Eom and Fortunato 2011). We investigate the extent to which heterogeneity in the number of citations received per paper is associated with gender imbalance in citations received by papers. To this end, we extend the homophilic-draws model to a reference model that approximately preserves heterogeneity in the number of citations received per paper, which we refer to as the preferential-draws model.

We examine citations made by older papers in sequential order. We sort the papers by publication date in ascending

⁴For example, in our analyses, $\bar{V}_{\text{HD}}(i, i', S)$ is defined as the set of papers in $\bar{V}_{\text{RD}}(i)$ that belong to the same conference rank, the same country of affiliation, and the same research topic as $v_{i'}$.

Model	Properties to be preserved
Random-draws model	<ul style="list-style-type: none"> • Number of citations made by each paper
Homophilic-draws model	<ul style="list-style-type: none"> • Number of citations made by each paper • Homophilic citation patterns in terms of given properties of the paper
Preferential-draws model	<ul style="list-style-type: none"> • Number of citations made by each paper • Homophilic citation patterns in terms of given properties of the paper • Heterogeneity in the number of citations received per paper

Table 2: Properties to be preserved in each reference model.

order: v_{x_1}, \dots, v_{x_N} , where x_l is the index of the l -th paper in the sorted result; v_{x_1} and v_{x_N} are the oldest and newest papers, respectively. We randomize citations made by paper v_{x_l} by accounting for the following properties for the papers $v_{x_1}, \dots, v_{x_{l-1}}$: (i) any property in S and (ii) the expected number of citations received before v_{x_l} is published under the model. Below, we denote by $\bar{c}_{i,l,\text{PD}}$ the expected number of citations received by $v_i \in V$ from the papers v_{x_1}, \dots, v_{x_l} under the model, where the subscript ‘PD’ indicates the ‘preferential-draws’ model.

We describe the algorithmic procedure. First, we initialize $\bar{c}_{i,l,\text{PD}}$ to zero for any $i = 1, \dots, N$ and any $l = 1, \dots, N$. Second, we perform the following procedure from $l = 2$ to N in sequential order. Consider an observed citation made by paper v_{x_l} to $v_{i'} \in V$. We define $\bar{V}_{\text{PD}}(l, i', S)$ as the set of papers in $\bar{V}_{\text{HD}}(x_l, i', S)$ that receives the same expected number of citations from the papers $v_{x_1}, \dots, v_{x_{l-1}}$ as $v_{i'}$, i.e., $\bar{V}_{\text{PD}}(l, i', S) = \{v_j \mid v_j \in \bar{V}_{\text{HD}}(x_l, i', S) \wedge \bar{c}_{j,l-1,\text{PD}} = \bar{c}_{i',l-1,\text{PD}}\}$. Note that $\bar{V}_{\text{PD}}(l, i', S)$ includes $v_{i'}$. We assume that any paper in $\bar{V}_{\text{PD}}(l, i', S)$ receives a citation from v_{x_l} with the probability $1/|\bar{V}_{\text{PD}}(l, i', S)|$. Therefore, for any paper $v_j \in \bar{V}_{\text{PD}}(l, i', S)$ and any $z = l, \dots, N$, we increase $\bar{c}_{j,z,\text{PD}}$ by $1/|\bar{V}_{\text{PD}}(l, i', S)|$. Finally, the probability of a citation made by v_i to $v_j \in V$ under the model is given by

$$\bar{w}_{ij,\text{PD}} = \sum_{l=2}^N \sum_{\substack{i'=1 \\ (v_i, v_{i'}) \in E \\ v_j \in \bar{V}_{\text{PD}}(l, i', S)}}^N 1/|\bar{V}_{\text{PD}}(l, i', S)|. \quad (5)$$

The expected number of citations received by $v_j \in V$ under the model is given by

$$\bar{c}_{j,\text{PD}} = \sum_{i=1}^N \bar{w}_{ij,\text{PD}}. \quad (6)$$

As with the homophilic-draws model, the preferential-draws model preserves the number of citations made by each paper and homophilic citation patterns in terms of the conference rank, country of affiliation, and research topic. In addition, it approximately preserves the heterogeneity in the number of citations received per paper in the original network (see Fig. 1(e) for numerical evidence).

Characterizing Gender Imbalance in Citations

We quantify gender imbalance in citations received by papers using the family of reference models. Consider two

subsets of papers, denoted by $V_{\text{from}} \subseteq V$ and $V_{\text{to}} \subseteq V$. We measure the extent to which the papers in V_{from} over- or under-cite the papers in V_{to} and in a given gender category $g \in \{\text{MM}, \text{MW}, \text{WM}, \text{WW}\}$ (Dworkin et al. 2020; Teich et al. 2022). We first count the citations made by the papers in V_{from} to those in V_{to} and in gender category g , which we denote by $n_{g,\text{obs}}$. Then, we compare $n_{g,\text{obs}}$ with the expectation obtained by a reference model. The expectation under the random-draws model is given by

$$\bar{n}_{g,\text{RD}} = \sum_{v_i \in V_{\text{from}}} k'_i p_{i,g,\text{RD}}, \quad (7)$$

where k'_i is the number of citations made by v_i to the papers in V_{to} , and $p_{i,g,\text{RD}}$ is the fraction of the papers that are in both $\bar{V}_{\text{RD}}(i)$ and g . The expectation under the homophilic-draws model is given by

$$\bar{n}_{g,\text{HD}} = \sum_{v_i \in V_{\text{from}}} \sum_{\substack{v_j \in V_{\text{to}} \\ (v_i, v_j) \in E}} p_{i,j,S,g,\text{HD}}, \quad (8)$$

where $p_{i,j,S,g,\text{HD}}$ is the fraction of the papers that are in both $\bar{V}_{\text{HD}}(i, j, S)$ and g . The expectation under the preferential-draws model is given by

$$\bar{n}_{g,\text{PD}} = \sum_{l=2}^N \sum_{\substack{v_j \in V_{\text{to}} \\ (v_{x_l}, v_j) \in E}} p_{l,j,S,g,\text{PD}}, \quad (9)$$

where $p_{l,j,S,g,\text{PD}}$ is the fraction of the papers that are in both $\bar{V}_{\text{PD}}(l, j, S)$ and g . For given reference model $r \in \{\text{RD}, \text{HD}, \text{PD}\}$, we calculate the over/under-citation made by the papers in V_{from} to those in V_{to} as $(n_{g,\text{obs}} - \bar{n}_{g,r})/\bar{n}_{g,r}$.

We also perform bootstrap resampling to estimate confidence intervals for the over/under-citations made by papers in our dataset (see Appendix C in Ref. (Nakajima et al. 2024) for details).

Characterizing Gender Imbalance in Citation-Based Rankings

We quantify gender imbalance in citation-based rankings of papers. To this end, we compute the fraction of W|W papers (i.e., MW, WM, and WW papers) among the top $d\%$ of papers with the largest values of a given impact measure (Karimi et al. 2018). We compare this fraction in the original network with that in the reference models.

We first use the number of citations received by the paper, c_j , as the impact measure (Zeng et al. 2017). We employ

Eqs. (2), (4), and (6) to compute the ranking of the papers based on the expectations for the reference models.

The quantity c_j does not account for which papers cited v_j and how v_j influenced the impact of those citing papers. Alternative metrics exist for measuring the impact of a paper (Zeng et al. 2017). One such metric is PageRank, originally introduced for ranking web pages (Brin and Page 1998), and it has been deployed in various applications, including the ranking of papers in several disciplines (Ma, Guan, and Zhao 2008; Masuda, Porter, and Lambiotte 2017). We also use PageRank as the impact measure for the papers in our study.

We compute the PageRank of the papers in the original network as follows. Consider a random walk with teleportations on the network: an individual reading paper $v_i \in V$ follows one of the citations made by v_i uniformly at random with the probability α and teleports to one of the N research papers with probability $1 - \alpha$. If the individual reads a paper that makes no citations, they also teleport to one of the N research papers. We assume that the paper to which the individual teleports is selected from the probability distribution proportional to the number of citations received by the paper (Lambiotte and Rosvall 2012). The PageRank of v_i , denoted by $p_{i,\alpha}^*$, is the probability that the individual reads paper v_i after a sufficient number of time steps. We define $p_{i,\alpha}^*$ by the following equation (Lambiotte and Rosvall 2012)

$$p_{i,\alpha}^* = (1 - \alpha) \frac{c_i}{M} + \alpha \sum_{j=1}^N T_{ji} p_{j,\alpha}^*, \quad (10)$$

where we define

$$T_{ij} = \begin{cases} A_{ij}/k_i & \text{if } k_i > 0, \\ c_j/M & \text{otherwise.} \end{cases} \quad (11)$$

We refer to $(T_{ij})_{1 \leq i \leq N, 1 \leq j \leq N}$ as a transition-probability matrix for the original network. We set $\alpha = 0.85$ (Brin and Page 1998). We compute the PageRank of the papers iteratively as follows. We assume that the initial probability distribution is $p_i(0) = c_i/M$ for any $i = 1, \dots, N$. The computation at each step $t > 0$ yields

$$p_i(t+1) = (1 - \alpha) \frac{c_i}{M} + \alpha \sum_{j=1}^N T_{ji} p_j(t). \quad (12)$$

The computation ends when t reaches t_{\max} or it holds that

$$\frac{1}{N} \sum_{i=1}^N |p_i(t+1) - p_i(t)| < \epsilon, \quad (13)$$

where we set $t_{\max} = 100$ and $\epsilon = 10^{-6}$ (NetworkX 2024).

To compute the PageRank of the papers for a given reference model, we assume the following: (i) an individual reading paper $v_i \in V$ switches to paper $v_j \in V$ with the probability of a citation made by v_i to v_j under the reference model; (ii) the paper to be teleported is selected from the probability distribution proportional to the expected number of citations received by the paper under the reference model. Based on these assumptions, we define the PageRank of v_i

for the reference model, denoted by $\bar{p}_{i,\alpha,r}^*$, as

$$\bar{p}_{i,\alpha,r}^* = (1 - \alpha) \frac{\bar{c}_{i,r}}{M} + \alpha \sum_{j=1}^N \bar{T}_{ji,r} \bar{p}_{j,\alpha,r}^*, \quad (14)$$

where $r \in \{\text{RD}, \text{HD}, \text{PD}\}$ and we define

$$\bar{T}_{ij,r} = \begin{cases} \bar{w}_{ij,r}/k_i & \text{if } k_i > 0, \\ \bar{c}_{j,r}/M & \text{otherwise.} \end{cases} \quad (15)$$

For any $r \in \{\text{RD}, \text{HD}, \text{PD}\}$, it holds true that $\sum_{j=1}^N \bar{w}_{ij,r} = k_i$ for any $v_i \in V$ and $\sum_{j=1}^N \bar{c}_{j,r} = M$. We refer to $(\bar{T}_{ij,r})_{1 \leq i \leq N, 1 \leq j \leq N}$ as a transition-probability matrix for the reference model. Note that $(\bar{T}_{ij,r})_{1 \leq i \leq N, 1 \leq j \leq N}$ for any $r \in \{\text{RD}, \text{HD}, \text{PD}\}$ is typically much denser than $(T_{ij})_{1 \leq i \leq N, 1 \leq j \leq N}$. We assume that the initial probability distribution is $\bar{p}_{i,r}(0) = \bar{c}_{i,r}/M$ for any $v_i \in V$ and any $r \in \{\text{RD}, \text{HD}, \text{PD}\}$. We set $\alpha = 0.85$. The iterative computation procedure is the same as that for the original network.

For both impact measures, we use the score (i.e., the number of citations received by the paper or PageRank) of each paper v_i divided by the average score of the papers that were published in the same year and belong to the same subfield as v_i . This normalization addresses the time-and-subfield dependence of the average number of citations received per paper (Radicchi, Fortunato, and Castellano 2008).

Results

Comparison of Reference Models

We first examine which structural properties in the original citation network are preserved by the three reference models. First, all models exactly preserve the distribution of the number of citations made by the paper in the original network (see Fig. 1(a)). Second, while the random-draws model destroys homophilic citation patterns in the original network, both the homophilic-draws and preferential-draws models exactly preserve (see Figs. 1(b)–(d)). Third, while the random-draws and homophilic-draws models hardly preserve the heterogeneity in the number of citations received per paper in the original network, the preferential-draws model approximately preserves (see Fig. 1(e)). These results align with our expectations.

Gendered Citation Imbalance in Conference Papers

We now quantify gender imbalance in citations between papers published in computer science conferences. Each conference paper is categorized into one of four gender categories (i.e., MM, MW, WM, or WW) based on the gender of the first and last authors. We first count the number of citations made by a subset of papers to those in each gender category. Then, we compare the obtained citation counts with the expectations derived from the reference models.

Figure 2(a) shows the over/under-citation made by any papers to the papers in each gender category, computed using the random-draws model. We found that any papers tend to over-cite MM papers and under-cite MW, WM, and WW

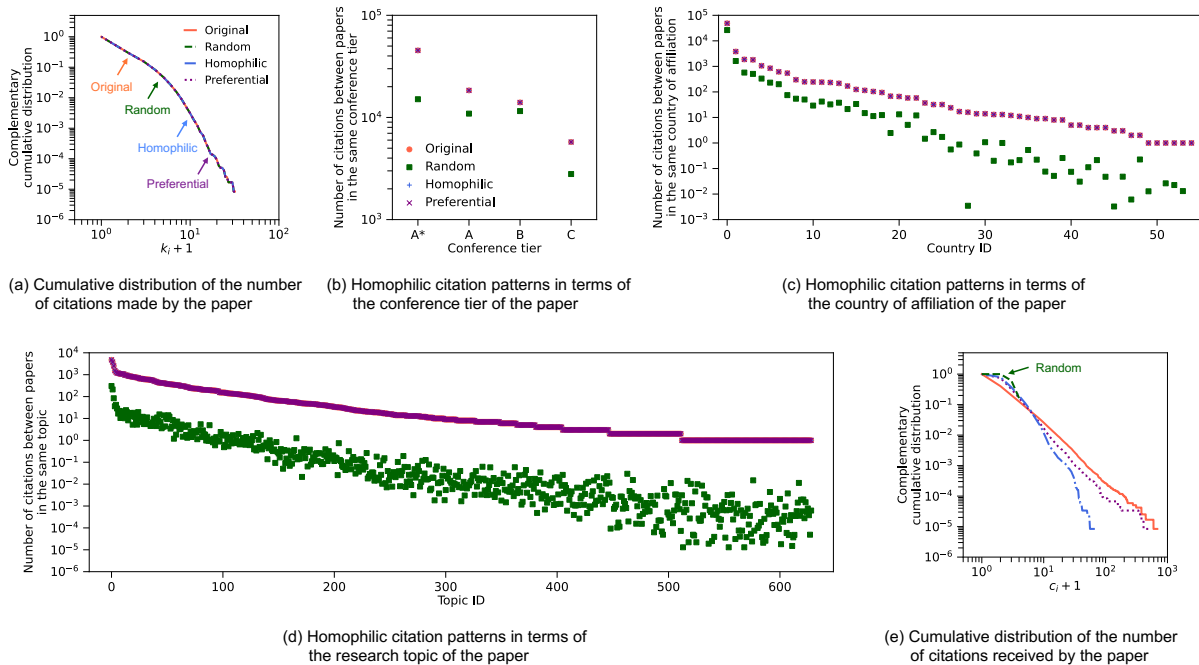


Figure 1: Comparison of structural properties between the original citation network and reference models. The legend ‘Original’ indicates the result for the original network, ‘Random’ indicates that for the random-draws model, ‘Homophilic’ indicates that for the homophilic-draws model, and ‘Preferential’ indicates that for the preferential-draws model. In Figs. 1(c) and 1(d), we focused on the 55 countries of affiliation and the 625 research topics, each of which has one or more homophilic citations in the original network. We sorted the IDs in descending order by the number of homophilic citations in each case. Curves that completely or heavily overlap are indicated by arrows and labels.

papers. Using the random-draws model, we also computed the over/under-citation made by MM and WW papers. The over/under-citation patterns made by MM papers are qualitatively similar to those made by any papers (see Fig. 2(b)). This result is expected because the fraction of citations made by MM papers is dominant (i.e., 77.7%). In contrast, WW papers exhibit the opposite over/under-citation patterns: they tend to under-cite MM papers and over-cite MW, WM, and WW papers (see Fig. 2(c)). These results are largely consistent with previous findings on gendered citation imbalances in journal papers in neuroscience and physics (Dworkin et al. 2020; Teich et al. 2022).

Comparisons of the over/under-citation between the random-draws and homophilic-draws models will reveal the influence of homophilic citation patterns. This is because the random-draws model destroys homophilic citation patterns in the original network, whereas the homophilic-draws model preserves them. Similarly, comparisons of the over/under-citation between the homophilic-draws and the preferential-draws models will reveal the influence of the heterogeneity in the number of citations received per paper.

Figures 2(d)–2(f) show the results for the homophilic-draws model. We found that homophilic citation patterns are associated with the under-citation received by MW, WM, and WW papers. Specifically, the over/under-citation received by MW papers is reduced from -4.8% for the random-draws model (see Fig. 2(a)) to -3.3% for the

homophilic-draws model (see Fig. 2(d)). Similarly, the over/under-citations received by WM and WW papers are reduced from -12.9% and -31.4% for the random-draws model (see Fig. 2(a)) to -8.2% and -21.2% for the homophilic-draws model (see Fig. 2(d)), respectively. We make qualitatively similar observations for the over/under-citation made by MM papers (see Figs. 2(b) and 2(e)). Furthermore, we found that gendered citation patterns made by WW papers are sufficiently explained by homophilic citation patterns. Indeed, the over/under-citation made by WW papers is almost eliminated in the homophilic-draws model (see Figs. 2(c) and 2(f)).

Figures 2(g)–2(i) show the results for the preferential-draws model. We found that the heterogeneity in the number of citations received per paper is slightly associated with the under-citation received by MW, WM, and WW papers. We also make qualitatively similar observations for MM papers (see Fig. 2(h)). In contrast, the heterogeneity in the number of citations received per paper contributes little to describing gendered citation patterns made by WW papers (see Figs. 2(f) and 2(i)).

In summary, gender imbalance in citations received by conference papers in computer science exists. Homophily in citations is strongly associated with the gender imbalance in citations, whereas heterogeneity in the number of citations received per paper shows a minor association with it. Gendered citation patterns made by MM papers persist

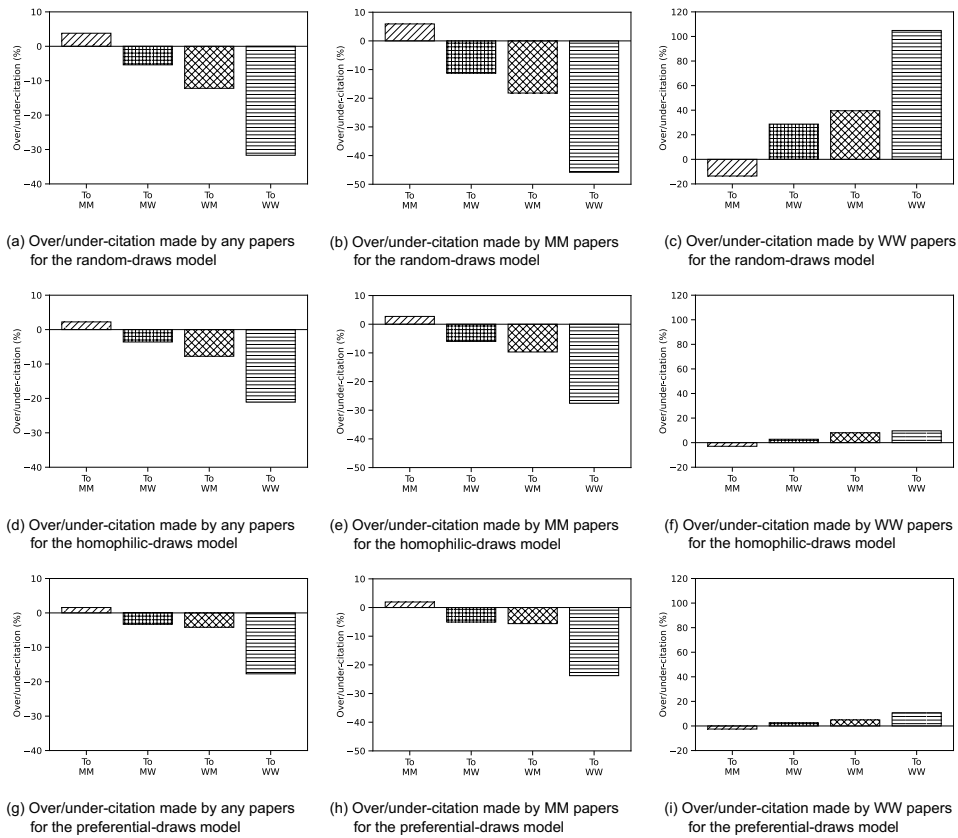


Figure 2: Gender imbalance in citations received by conference papers in computer science.

even after controlling for both of these properties, whereas those made by WW papers are sufficiently explained by homophilic citation patterns.

Gendered Citation Imbalance Across Conference Ranks and Subfields

Gender imbalance in academia has been investigated in terms of prestige, including institutional prestige (Huang et al. 2020). Journals and conferences are often ranked based on citation statistics of papers published there (Garfield 1972; Freyne et al. 2010). Indeed, several ranking systems exist for conferences in computer science, with the highest-ranked conferences often regarded as prestigious (Li et al. 2018). We hypothesize that the gendered citation imbalance in computer science is associated with the prestige of the conference. To investigate this possibility, we computed the over/under-citation received by MM and WW papers within each conference rank using the preferential-draws model. It quantifies the over/under-citation that cannot be immediately explained by the number of citations made by each paper, homophily in citations, and the heterogeneity in the number of citations received per paper.

Figure 3 compares the over/under-citation received by MM and WW papers within each conference rank. The gender imbalance persists across conference ranks. Notably, the degree of under-citation received by WW papers pub-

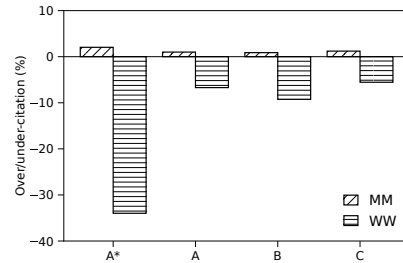


Figure 3: Gendered citation imbalance across conference ranks.

lished in A*-ranked conferences is at least four times larger than in A or lower-ranked conferences. In contrast, the degree of over-citation received by MM papers does not vary greatly across different conference ranks. The difference in the over/under-citation received by WW papers by conference rank is not immediately explained by the fraction of WW papers within each rank: 2.9% for A* rank, 4.5% for A rank, 3.7% for B rank, and 5.8% for C rank.

The degree of gender imbalance often varies by subfield within a single discipline. Indeed, gendered citation imbalance exists across different subfields in physics (Teich et al. 2022). We compare the degree of gender imbalance in ci-

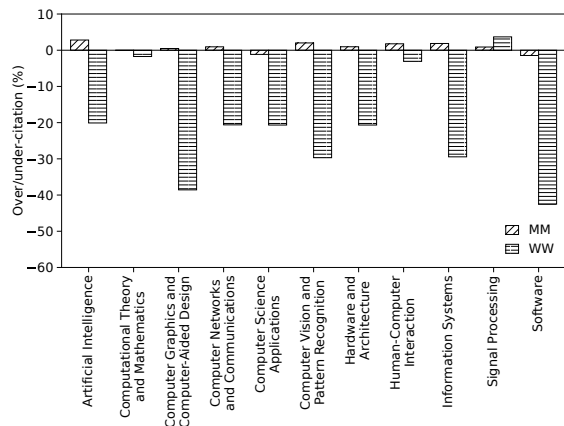


Figure 4: Gendered citation imbalance across subfields.

tations across different subfields in computer science. To this end, we focus on the 11 subfields categorized under the ‘Computer Science’ discipline in the OpenAlex data. We then compute the degree of over/under-citation received by MM and WW papers within each subfield using the preferential-draws model.

We found that the under-citation received by WW papers exists in all subfields except for the ‘signal processing’ subfield (see Fig. 4). The variation in the degree of the under-citation received by WW papers across subfields cannot be directly explained by that in the fraction of WW papers in each subfield: the Spearman’s rank correlation coefficient is 0.14 (P -value is 0.69).

To sum up, gender imbalance in citations received by conference papers persists across conference ranks and subfields in computer science. This imbalance is most pronounced for papers published in the highest-ranked conferences.

Gender Imbalance in Citation-Based Rankings

We have observed gender imbalance in citations received by conference papers in computer science. We further examine how this gendered citation imbalance influences gender imbalance in citation-based rankings of papers. To this end, we first rank the conference papers based on the number of citations received. We compute the number of citations received by each paper in the original network and compare it with the expectations in a given reference model (see Methods Section for details). For a given d value, we then compute the fraction of W|W papers (i.e., MW, WM, and WW papers) among the top $d\%$ of papers with the highest number of citations received, comparing the original network with the reference model.

Figure 5(a) compares the fraction of W|W papers in the top $d\%$ papers between the original network and the reference models. We found that the fraction of W|W papers in the top $d\%$ for the original network is lower than that for any reference model across most d values in the range $[0.1, 20]$. Additionally, the fraction of W|W papers for the homophilic-draws model is smaller than that for the random-

draws model across most d values in the range $[0.1, 100]$. This comparison reveals that homophily in citations is associated with gender imbalance in the ranking of papers, consistent with previous findings in the physics discipline (Karimi et al. 2018). In contrast, the fraction of W|W papers for the preferential-draws model is comparable to that for the homophilic-draws model, indicating that heterogeneity in the number of citations received per paper has a minor influence on gender imbalance in the rankings of papers.

The number of citations received by a given paper v does not account for which papers cited v or how v influenced the impact of the citing papers. To address this, we use PageRank, which measures the impact of a paper based on the citation network (Zeng et al. 2017), to rank the conference papers. We compute the PageRank of the papers in both the original network and a given reference model (see Methods Section for details). Then, we compare the fraction of W|W papers in the top $d\%$ of papers with the largest PageRank between the original network and the reference model.

Figure 5(b) compares the fraction of W|W papers in the top $d\%$ papers between the original network and the reference models. The fraction for the original network is lower than that for both reference models across most d values in the range $[0.1, 2]$. Comparisons among the reference models indicate that homophily in citations moderately influences gender imbalance in the ranking of papers based on PageRank, whereas heterogeneity in the number of citations received per paper does little.

To sum up, gender imbalance in citation-based rankings of conference papers in computer science exists. Homophily in citations is associated with this gender imbalance, whereas heterogeneity in the number of citations received per paper has little association.

Discussion

Gender Imbalance in Citations Received by Papers

We found that conference papers written by female authors as the first and/or last authors are less likely to receive citations across different subfields in computer science. This result is largely consistent with previous findings on gendered citation imbalance in other disciplines (Dworkin et al. 2020; Fulvio, Akinnola, and Postle 2021; Wang et al. 2021b; Teich et al. 2022). Additionally, we make significant insights into these previous findings. First, we found that homophily in citations is strongly associated with gender imbalance in citations, whereas the heterogeneity in the number of citations received per paper shows a minor association. Second, we found that the prestige of the highest-ranked conferences is strongly associated with the gender imbalance in citations. Our results indicate the presence of gendered citation patterns in computer science that cannot be explained by these structural properties alone.

Gender homophily has been observed in citations, and it is often associated with gender imbalance in citations (Tekles, Auspurg, and Bornmann 2022; Zhou, Chai, and Freeman 2024). Tekles et al. found that gendered citation patterns in the biological and medical disciplines are sufficiently explained by homophilic citations in terms of the research

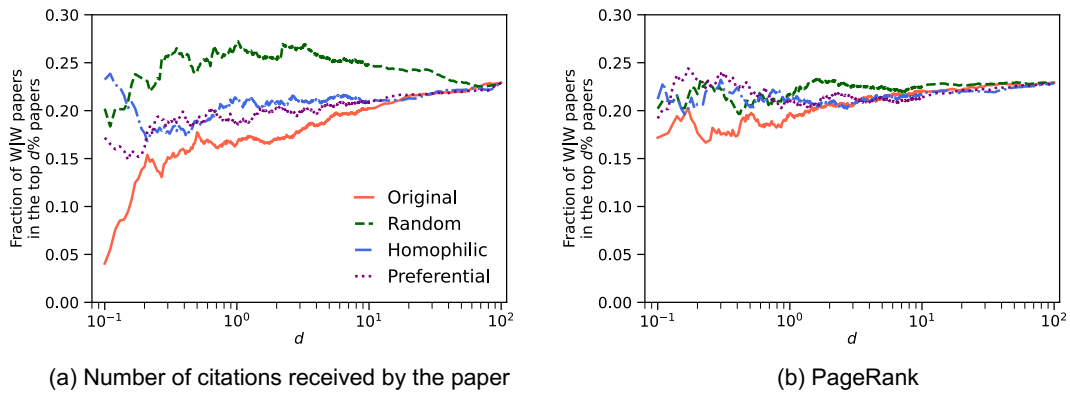


Figure 5: Gender imbalance in citation-based rankings of conference papers.

topic (Tekles, Auspurg, and Bornmann 2022). Consistent with this finding, we found that gendered citation patterns for papers written by female authors as the first and last authors in computer science are also largely explained by homophilic citations. Therefore, our results support the previous finding that extensively controlling for properties of the paper is important for assessing the degree of gender imbalance in citations (Tekles, Auspurg, and Bornmann 2022).

Gender Imbalance in Citation-Based Rankings

Karimi et al. found that papers categorized within the “minority” group of the research topic are less likely to appear at the top of citation-based rankings than expected by random chance (Karimi et al. 2018). In contrast, we proposed the use of reference models to assess imbalances in citation-based rankings. Our findings indicate that papers written by female authors as the first and/or last authors are less likely to appear at the top of citation-based rankings, even after accounting for homophily in citations and heterogeneity in citation counts per paper.

We expect that modifying network structures using reference models could potentially ameliorate gender imbalances in citation-based rankings. For example, one could utilize the preferential-draws model to calculate the PageRank of each paper in the given citation network. Indeed, our results suggest that adjusting a transition-probability matrix and a probability distribution of teleportation in the PageRank using the preferential-draws model improves gender imbalance in the ranking of papers (see Fig. 5(b)). A similar approach has been employed to enhance the fairness of PageRank among nodes in a network (Tsioutsoulklis et al. 2021). In contrast to this previous study (Tsioutsoulklis et al. 2021), our reference models do not require any algorithmic parameters to enhance the fairness in node rankings.

The development of fairness-aware methods for network analysis has emerged as a recent and challenging topic (Saxena, Fletcher, and Pechenizkiy 2024). The prospect of ameliorating imbalances in node rankings by modifying network structures using reference models remains an open area for exploration and discussion. This approach may be extended to encompass other centrality measures and network types

and is expected to contribute to the establishment of fairness-aware methods in node rankings.

Conclusion

In this study, we quantified gendered citation imbalance in conference papers in computer science. Previous studies have identified gender imbalance in the number of researchers (Way, Larremore, and Clauset 2016; Jadidi et al. 2018; Huang et al. 2020; Laberge et al. 2022), research career (Jadidi et al. 2018; Huang et al. 2020; Morgan et al. 2021; Lietz et al. 2024), and authorship (Holman, Stuart-Fox, and Hauser 2018; Wang et al. 2021a) within the computer science discipline. Additionally, studies have quantified gendered citation imbalance in journal papers across other disciplines (Dworkin et al. 2020; Fulvio, Akinnola, and Postle 2021; Wang et al. 2021b; Teich et al. 2022). By developing a family of reference models for citation networks, we extended previous findings on gender imbalance in computer science in terms of citation practices.

Our study presents several future challenges. Our reference models do not consider the attractiveness or fitness of the paper (Eom and Fortunato 2011; Wang, Song, and Barabási 2013) and the authors’ reputation (Petersen et al. 2014), both of which significantly impact citation dynamics. Extending our reference models to include these factors could help us understand how these properties are associated with gendered citation imbalance. Furthermore, citation imbalances may also be present concerning other author attributes, such as nationality (Nakajima et al. 2023) and race (Kozłowski et al. 2022; Liu, Rahwan, and AlShebli 2023), in addition to gender, within the computer science discipline. Investigating citation imbalances in relation to these other relevant author attributes warrants further work.

Acknowledgments

This work was supported by Japan Science and Technology Agency (JST) as part of Adopting Sustainable Partnerships for Innovative Research Ecosystem (ASPIRE), Grant Number JPMJAP2328. KN thanks the financial support by JSPS KAKENHI Grant Number 24K21056. YS thanks the support by JST Presto Grant Number JPMJPR21C5.

References

- Avin, C.; Keller, B.; Lotker, Z.; Mathieu, C.; Peleg, D.; and Pignolet, Y.-A. 2015. Homophily and the Glass Ceiling Effect in Social Networks. In *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, 41–50.
- Brin, S.; and Page, L. 1998. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30: 107–117.
- Ciotti, V.; Bonaventura, M.; Nicosia, V.; Panzarasa, P.; and Latora, V. 2016. Homophily and missing links in citation networks. *EPJ Data Science*, 5: 7.
- Delgado-Quirós, L.; and Ortega, J. L. 2024. Completeness degree of publication metadata in eight free-access scholarly databases. *Quantitative Science Studies*, 1–19.
- Dworkin, J. D.; Linn, K. A.; Teich, E. G.; Zurn, P.; Shinohara, R. T.; and Bassett, D. S. 2020. The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience*, 23: 918–926.
- Eom, Y.-H.; and Fortunato, S. 2011. Characterizing and Modeling Citation Dynamics. *PLOS ONE*, 6: 1–7.
- Fortunato, S.; Bergstrom, C. T.; Börner, K.; Evans, J. A.; Helbing, D.; Milojević, S.; Petersen, A. M.; Radicchi, F.; Sinatra, R.; Uzzi, B.; Vespignani, A.; Waltman, L.; Wang, D.; and Barabási, A.-L. 2018. Science of science. *Science*, 359: eaao0185.
- Freyne, J.; Coyle, L.; Smyth, B.; and Cunningham, P. 2010. Relative status of journal and conference publications in computer science. *Communications of the ACM*, 53: 124–132.
- Fulvio, J. M.; Akinola, I.; and Postle, B. R. 2021. Gender (im)balance in citation practices in cognitive neuroscience. *Journal of Cognitive Neuroscience*, 33: 3–7.
- Garfield, E. 1972. Citation analysis as a tool in journal evaluation. *Science*, 178: 471–479.
- Gomez, C. J.; Herman, A. C.; and Parigi, P. 2022. Leading countries in global science increasingly receive more citations than other countries doing similar research. *Nature Human Behaviour*, 6: 919–929.
- Holman, L.; Stuart-Fox, D.; and Hauser, C. E. 2018. The gender gap in science: How long until women are equally represented? *PLOS Biology*, 16: 1–20.
- Huang, J.; Gates, A. J.; Sinatra, R.; and Barabási, A.-L. 2020. Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences of the United States of America*, 117: 4609–4616.
- Jaccheri, L.; Pereira, C.; and Fast, S. 2020. Gender issues in computer science: lessons learnt and reflections for the future. In *2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, 9–16.
- Jadidi, M.; Karimi, F.; Lietz, H.; and Wagner, C. 2018. Gender disparities in science? Dropout, productivity, collaborations and success of male and female computer scientists. *Advances in Complex Systems*, 21: 1750011.
- Karimi, F.; Génois, M.; Wagner, C.; Singer, P.; and Strohmaier, M. 2018. Homophily influences ranking of minorities in social networks. *Scientific Reports*, 8: 11077.
- Kojaku, S.; Livan, G.; and Masuda, N. 2021. Detecting anomalous citation groups in journal networks. *Scientific Reports*, 11: 14524.
- Kozłowski, D.; Larivière, V.; Sugimoto, C. R.; and Monroe-White, T. 2022. Intersectional inequalities in science. *Proceedings of the National Academy of Sciences of the United States of America*, 119: e2113067119.
- Laberge, N.; Wapman, K. H.; Morgan, A. C.; Zhang, S.; Larremore, D. B.; and Clauset, A. 2022. Subfield prestige and gender inequality among U.S. computing faculty. *Communications of the ACM*, 65: 46–55.
- Lambiotte, R.; and Rosvall, M. 2012. Ranking and clustering of nodes in networks with smart teleportation. *Physical Review E*, 85: 056107.
- Li, W.; Zhang, S.; Zheng, Z.; Cranmer, S. J.; and Clauset, A. 2022. Untangling the network effects of productivity and prominence among scientists. *Nature Communications*, 13: 4907.
- Li, X.; Rong, W.; Shi, H.; Tang, J.; and Xiong, Z. 2018. The impact of conference ranking systems in computer science: a comparative regression analysis. *Scientometrics*, 116: 879–907.
- Lietz, H.; Jadidi, M.; Kostic, D.; Tsvetkova, M.; and Wagner, C. 2024. Individual and gender inequality in computer science: A career study of cohorts from 1970 to 2000. *Quantitative Science Studies*, 1–25.
- Liu, F.; Rahwan, T.; and AlShebli, B. 2023. Non-White scientists appear on fewer editorial boards, spend more time under review, and receive fewer citations. *Proceedings of the National Academy of Sciences of the United States of America*, 120: e2215324120.
- Ma, N.; Guan, J.; and Zhao, Y. 2008. Bringing PageRank to the citation analysis. *Information Processing & Management*, 44: 800–810.
- Masuda, N.; Porter, M. A.; and Lambiotte, R. 2017. Random walks and diffusion on networks. *Physics Reports*, 716–717: 1–58.
- Moed, H. F. 2006. *Citation analysis in research evaluation*, volume 9. Springer Science & Business Media.
- Moher, D.; Naudet, F.; Cristea, I. A.; Miedema, F.; Ioannidis, J. P. A.; and Goodman, S. N. 2018. Assessing scientists for hiring, promotion, and tenure. *PLOS Biology*, 16: 1–20.
- Morgan, A. C.; Way, S. F.; Hofer, M. J. D.; Larremore, D. B.; Galesic, M.; and Clauset, A. 2021. The unequal impact of parenthood in academia. *Science Advances*, 7: eabd1996.
- Nakajima, K.; Liu, R.; Shudo, K.; and Masuda, N. 2023. Quantifying gender imbalance in East Asian academia: Research career and citation practice. *Journal of Informetrics*, 17: 101460.
- Nakajima, K.; Sasaki, Y.; Tokuno, S.; and Fletcher, G. 2024. Quantifying gendered citation imbalance in computer science conferences. *arXiv preprint arXiv:2408.01900*.

- Nakajima, K.; Shudo, K.; and Masuda, N. 2022. Randomizing Hypergraphs Preserving Degree Correlation and Local Clustering. *IEEE Transactions on Network Science and Engineering*, 9: 1139–1153.
- NetworkX. 2024. pagerank. https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.link_analysis.pagerank_alg.pagerank.html (Accessed April 2024).
- OECD. 2017. The under-representation of women in STEM fields. <https://dblp.org/faq/How+accurate+is+the+data+in+dblp.html> (Accessed April 2024).
- OpenAlex. 2024. OpenAlex: End-to-End Process for Topic Classification. <https://docs.google.com/document/d/1bDopkhuGieQ4F8gGNj7sEc8WSE8mvLZS/edit#heading=h.5w2tb5fcg77r> (Accessed April 2024).
- Orsini, C.; Dankulov, M. M.; Colomer-de Simón, P.; Jamakovic, A.; Mahadevan, P.; Vahdat, A.; Bassler, K. E.; Toroczka, Z.; Boguñá, M.; Caldarelli, G.; Fortunato, S.; and Krioukov, D. 2015. Quantifying randomness in real networks. *Nature Communications*, 6: 8627.
- Petersen, A. M.; Fortunato, S.; Pan, R. K.; Kaski, K.; Penner, O.; Rungi, A.; Riccaboni, M.; Stanley, H. E.; and Pammolli, F. 2014. Reputation and impact in academic careers. *Proceedings of the National Academy of Sciences of the United States of America*, 111: 15316–15321.
- Priem, J.; Piwowar, H.; and Orr, R. 2022. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833*.
- Radicchi, F.; Fortunato, S.; and Castellano, C. 2008. Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America*, 105: 17268–17272.
- Saxena, A.; Fletcher, G.; and Pechenizkiy, M. 2024. FairSNA: Algorithmic fairness in social network analysis. *ACM Computing Surveys*, 56: 213.
- Sinatra, R.; Wang, D.; Deville, P.; Song, C.; and Barabási, A.-L. 2016. Quantifying the evolution of individual scientific impact. *Science*, 354: aaf5239.
- Sugimoto, C. R.; and Larivière, V. 2023. *Equity for Women in Science*. Harvard University Press.
- Teich, E. G.; Kim, J. Z.; Lynn, C. W.; Simon, S. C.; Klishin, A. A.; Szymula, K. P.; Srivastava, P.; Bassett, L. C.; Zurn, P.; Dworkin, J. D.; and Bassett, D. S. 2022. Citation inequity and gendered citation practices in contemporary physics. *Nature Physics*, 18: 1161–1170.
- Tekles, A.; Auspurg, K.; and Bornmann, L. 2022. Same-gender citations do not indicate a substantial gender homophily bias. *PLOS ONE*, 17: e0274810.
- The DBLP team. 2024a. DBLP computer science bibliography. How accurate is the data in dblp? <https://dblp.org/faq/How+accurate+is+the+data+in+dblp.html> (Accessed April 2024).
- The DBLP team. 2024b. DBLP computer science bibliography. Monthly snapshot release of April 2024. <https://dblp.org/xml/release/dblp-2024-04-01.xml.gz> (Accessed April 2024).
- Tsioutsoulouklis, S.; Pitoura, E.; Tsaparas, P.; Kleftakis, I.; and Mamoulis, N. 2021. Fairness-aware PageRank. In *Proceedings of the Web Conference 2021*, 3815–3826.
- Uzzi, B.; Mukherjee, S.; Stringer, M.; and Jones, B. 2013. Atypical Combinations and Scientific Impact. *Science*, 342: 468–472.
- Vrettas, G.; and Sanderson, M. 2015. Conferences versus journals in computer science. *Journal of the Association for Information Science and Technology*, 66: 2674–2684.
- Waltman, L. 2016. A review of the literature on citation impact indicators. *Journal of Informetrics*, 10: 365–391.
- Wang, D.; Song, C.; and Barabási, A.-L. 2013. Quantifying long-term scientific impact. *Science*, 342: 127–132.
- Wang, L. L.; Stanovsky, G.; Weihs, L.; and Etzioni, O. 2021a. Gender trends in computer science authorship. *Communications of the ACM*, 64: 78–84.
- Wang, X.; Dworkin, J. D.; Zhou, D.; Stiso, J.; Falk, E. B.; Bassett, D. S.; Zurn, P.; and Lydon-Staley, D. M. 2021b. Gendered citation practices in the field of communication. *Annals of the International Communication Association*, 45: 134–153.
- Way, S. F.; Larremore, D. B.; and Clauset, A. 2016. Gender, productivity, and prestige in computer science faculty hiring networks. In *Proceedings of the 25th International Conference on World Wide Web*, 1169–1179.
- Zeng, A.; Shen, Z.; Zhou, J.; Wu, J.; Fan, Y.; Wang, Y.; and Stanley, H. E. 2017. The science of science: From the perspective of complex systems. *Physics Reports*, 714-715: 1–73.
- Zhou, S.; Chai, S.; and Freeman, R. B. 2024. Gender homophily: In-group citation preferences and the gender disadvantage. *Research Policy*, 53: 104895.