Mechanism Design for Paper Review

Yichong Xu
Machine Learning Department
Carnegie Mellon University
yichongx@cs.cmu.edu

Han Zhao
Machine Learning Department
Carnegie Mellon University
han.zhao@cs.cmu.edu

Xiaofei Shi
Department of Mathematical Sciences
Carnegie Mellon University
xiaofeis@andrew.cmu.edu

1 Introduction

The reviewing process of scientific conferences decides whether submissions to the conference will be accepted or not. For important conferences, the reviewing process could potentially make an influential impact on the scientific career of practitioners in those fields. Hence the problem of ensuring a fair and impartial reviewing process cannot be over emphasized [Shah et al., 2017]. However, often the case, the reviewing processes in real-world scenarios are complicated, and reviewers tend to have strategic behaviours during the process. This is especially the case when the reviewers themselves are also the authors, where the reviewers have an incentive to lower the evaluation scores for other papers in order to maximize the chance that their own papers get accepted.

One of the relevant works to paper reviewing is peer reviewing [Kahng et al., 2017], where peer students grade homework submissions of each other so that the final grade of one student is determined by her grades given by the peers. Such scenario is ubiquitous, especially due to the popularization of massive open online courses (MOOC), e.g., coursera [Severance, 2012], Edx, etc. Specifically, the authors [Kahng et al., 2017] propose two different algorithms for peer reviewing, including the k-Partite algorithm, which is a direct extension of the Naive-Bipartite scheme from 2 groups into k, and the COMMITTEE scheme, where a group of seeding reviewers are first selected to make the final ranking together. They show that under certain assumptions these two algorithms are *impartial*, meaning that peers could not influence their own position in the aggregated ranking by strategic behaviors. They also conduct experimental simulations to demonstrate the efficacy and practicability of their algorithms.

Formally, using graph theoretic terminology, we can formalize the authorship relationships as a bipartite graph $G_A = (R, P; E)$ where $R = \{r_1, \dots, r_n\}$ is the set of n reviewers, $P = \{p_1, \dots, p_n\}$ is the set of n papers, $E = \{(r_i, p_j)\}$ is the set of edges and $(r_i, p_j) \in E$ iff reviewer r_i authors paper p_j. During the paper reviewing process, each reviewer r_i has a ranking over the papers, denoted as $\pi_i \in \Pi_n$ where Π_n is the set of all possible permutations over $[n] = \{1, 2, \dots, n\}$. We use $\pi(p), \forall p \in [n]$ to denote the pth element of π, i.e., the rank of the pth paper in π. A ranking profile $\boldsymbol{\pi} = \{\pi_1, \dots, \pi_n\}$ is a collection of all reviewers' rankings, and the goal is to define an aggregation rule $f : \Pi_n^n \to \Pi_n$ that maps from ranking profile to a single ranking, i.e., the output ranking.

One of the key assumptions in [Kahng et al., 2017] is that in the authorship graph G_A, each reviewer authors exactly one paper and each paper is authored by exactly one reviewer. Using the notation we just introduced, this means that the edge set in the authorship graph G_A has the following form: $E = \{(r_i, p_i)\}_{i=1}^n$. This structural assumption hardly holds in practice. In fact, it usually happens that one reviewer may author multiple papers and most papers have at least two authors. In this project we attempt to generalize the above scenario in the following two ways: in the first part we consider

Project report for course CMU 10709, Fall 2017. The formatting style file is borrowed from the NIPS conference.

a randomized setting where the authorship graph is created at random, with the only constraint on the maximum number of papers one reviewer can author. Under this assumption we show a positive result that over the randomness of the graph generation process, with high probability the generated graph has a unique largest component with size $\Theta(n)$, where the Naive-Bipartite scheme can be deployed to prevent strategic reviews. In the second part, we prove an impossibility theorem which states that there is no review aggregation rule f that simultaneously satisfies both the Pareto efficiency (or the unanimity) property and the impartial property (both defined formally in Sec. 2). This negative result underlies the essential hardness in designing perfect reviewing rules. As a future direction, we plan to relax the requirement of reviewing rule by not asking it to output a rank, but only a set of $k > 1$ papers that are selected as accepted papers. We are interested to see whether we can design a reviewing rule that satisfy certain desirable properties, or is there similar intrinsic hardness (by means of impossibility theorem) in the relaxed setting as well.

2 Main Results

We provide detailed statements and their proofs in this section. In the first part, we show that if the authorship graph G_A is generated randomly with bounded expected degree, then with high probability we expect to see an unconnected graph where the largest component has size $\Theta(n)$. This result allows us to adapt the Naive-Bipartite algorithm to our paper reviewing setting, and provides a basic algorithm that satisfies the impartial property. In the second part, we show that no algorithm can simultaneously satisfy both the unanimity and the impartial properties. This impossibility theorem highlights the essential hardness in designing a fair reviewing system.

Before giving our algorithms, we first formally define two properties that we hope any fair reviewing system could achieve:

Definition 2.1 (Unanimity)**.** A review aggregation rule f is called *unanimous*, or *Pareto efficient*, if $\forall j_1, j_2 \in [n]$, $j_1 \neq j_2$, $\forall \pi_i(j_1) > \pi_i(j_2)$, $\forall i \in [n]$, then $f(\boldsymbol{\pi})(j_1) > f(\boldsymbol{\pi})(j_2)$, where $\boldsymbol{\pi} = \{\pi_1, \ldots, \pi_n\}$ is the ranking profile collected from all the n reviewers.

In words, the unanimity property requires a reviewing rule to be consistent with all the reviewers if they reach a consensus: for any pair of papers j_1 and j_2, $j_1 \neq j_2$, if all the reviewers agree with ranking j_1 lower than j_2, then in the output ranking given by the aggregation rule, j_1 should also be ranked lower than j_2.

Definition 2.2 (Impartiality)**.** A review aggregation rule f is called *impartial*, if $\forall i \in [n]$, $\exists j \in [n]$, such that $\forall \boldsymbol{\pi} = \{\pi_1, \ldots, \pi_i, \ldots, \pi_n\}$ and $\boldsymbol{\pi}' = \{\pi_1, \ldots, \pi_i', \ldots, \pi_n\}$, $f(\boldsymbol{\pi})(j) = f(\boldsymbol{\pi}')(j)$.

The impartiality property guarantees that under the review aggregation rule f, for any reviewer, there is at least one paper whose final ranking does not depend on the ranking from that reviewer, given that all the other reviewers keep their preferences fixed. In other words, one cannot hope to change the ranking of at least one paper if all the other rankings are kept the same. We argue that the impartiality is a desired property to have, otherwise there will be a reviewer who can change the ranking of any paper in the system by strategic behavior, including her own paper!

2.1 Randomized Setting

In this section we consider a randomized setting where we assume the authorship graph G_A is generated at random. We then provide a simple algorithm that is impartial.

Consider the following scenario where reviewers write papers in a random way. Since every reviewer, as a author, has only finite time and energy, we require that the expected number of papers each reviewer can author is Δ:

1. For each review i, uniformly at random draw an integer from $t \sim \text{Bin}(n, p)$ with $p = \Delta/n$.
2. Draw t papers uniformly at random without replacement from $[n]$, denoted as $\{j_{i_1}, \ldots, j_{i_t}\}$.
3. Create links between i and $\{j_{i_1}, \ldots, j_{i_t}\}$ and add them to G_A.

Equivalently, the above generation process of the authorship graph could be understood in an alternative way: for each possible edge $(i, j), \forall i, j \in [n]$, edge (i, j) is drawn to present with probability $p = \Delta/n$, i.e., each edge is drawn with probability Δ/n. Clearly, this leads to a generalization of the authorship graph in [Kahng et al., 2017] where each reviewer authors exactly one paper and each paper is owned by exactly one reviewer.

Algorithm 1 Naive-Bipartite

Input: Voting rule f and profile $\boldsymbol{\pi}$

Output: Ranking π

1: Randomly split the n papers into two sets X and Y, where $|X| = \lceil n/2 \rceil$ and $|Y| = \lfloor n/2 \rfloor$

2: $\pi_X \leftarrow f(\boldsymbol{\pi}_Y)$, $\pi_Y \leftarrow f(\boldsymbol{\pi}_X)$

3: π interlaces between π_X and π_Y:

$$\pi^{-1}(j) = \begin{cases} \pi_X^{-1}((j+1)/2), & j = 1 \ (\mathrm{mod} 2) \\ \pi_Y^{-1}(j/2), & j = 0 \ (\mathrm{mod} 2) \end{cases}$$

4: **return** π

In order to guarantee an algorithm is impartial, the simplest possible requirement is to ask that the reviewer could not change the ranking of her own papers. This is easily achievable using the Naive-Bipartite algorithm when there is a one-to-one correspondence between reviewers and papers. To make our discussion more complete and self-contained, we list the pseudo-code of the Naive-Bipartite algorithm in Alg. 1.

Note that in Alg. 1 we use $\boldsymbol{\pi}_X(\boldsymbol{\pi}_Y)$ to mean the subprofile of $\boldsymbol{\pi}$ restricted on the subset $X(Y)$, and $\pi^{-1}(j)$ represents the index of the paper that is ranked in the jth place of π. Intuitively, the Naive-Bipartite algorithm simply partitions all the papers into two disjoint sets with roughly even sizes, and then each subset of reviewers is asked to review and rank papers in the other subset. At the end, the output ranking is provided by merging the two rankings given by two subsets of reviewers, one from each subset at a time. If each reviewer authors exactly one paper, then Alg. 1 clearly satisfies the impartiality property, as no matter how one reviewer changes her own ranking, she cannot influence the rankings of all the papers in the other subset.

On the other hand, even in this simpliest form, Naive-Bipartite does not satisfy the unanimous property, as illustrated in the following example: suppose there are four reviewers, namely r_1, r_2, r_3 and r_4, where r_1, r_2 are assigned into X and r_3, r_4 are assigned into Y. Suppose all the reviewers share the ranking $p_1 \succ p_2 \succ p_3 \succ p_4$, but due to the partitioning in the algorithm, the final aggregated ranking is $p_3 \succ p_1 \succ p_4 \succ p_2$, clearly violating the unanimous property. Essentially, the problem happens because papers are not evenly partitioned based on their intrinsic qualities: good papers p_1, p_2 are assigned into the same subset and bad papers p_3, p_4 are also assigned into the same subset. Because of the last interlacing step in Alg. 1, in the worst case roughly half of the bad papers will be ranked higher than half of the good papers.

Note that the success of the Naive-Bipartite algorithm on impartiality crucially depends on the fact that there is a one-to-one correspondence between reviewers and papers. One may be interested in a more general and realistic case whether it still holds in our randomized setting. In what follows we will provide an affirmative answer to this question, using tools from random graph theory [Frieze and Karoński, 2015], provided that Δ is not too large:

Theorem 2.1. Let $G_{n,n,p}$ denote the random bipartite graph derived from the complete bipartite graph $K_{n,n}$ where each edge is included independently with probability at most p. Then if $p = c/n$ where $c > 1$ is a constant then w.h.p. $G_{n,n,p}$ has a unique largest component of size $2G(c)n$ where $G(c) = 1 - x/c + o(1)$, and x is the nontrivial solution of the following equation: $x\exp(-x) = c\exp(-c)$. The remaining components are of order at most $O(\mathrm{poly}\log n)$.

Thm. 2.1 is stated and left as an exercise in the book by Frieze and Karoński [2015]. We plot the function $y = x\exp(-x)$ in Fig. 1. Hence provided that the expected degree of G_A, i.e., the expected number of papers each reviewer can author, is a small bounded constant, e.g., 2 or 3, with high probability the authorship graph G_A will not be connected, and the largest connected component will have size $c'n$ where $c' \in (0.5, 1)$.

Proof Sketch of Thm. 2.1. Given the random bipartite graph $G_{n,n,p}$, we can split it into two independent copies of $G_{n,p}$ by separating the "upward" and "downward" edges and smash the vertices together. In each copy, by Thm. 2.14 in Frieze and Karoński [2015], there exists a giant component of size $G(c)n$, and every other small components will be of order at most $O(\log n)$. Therefore, the upper bound of the giant component will be $2G(c)n$. To show the lower bound, an similar argument

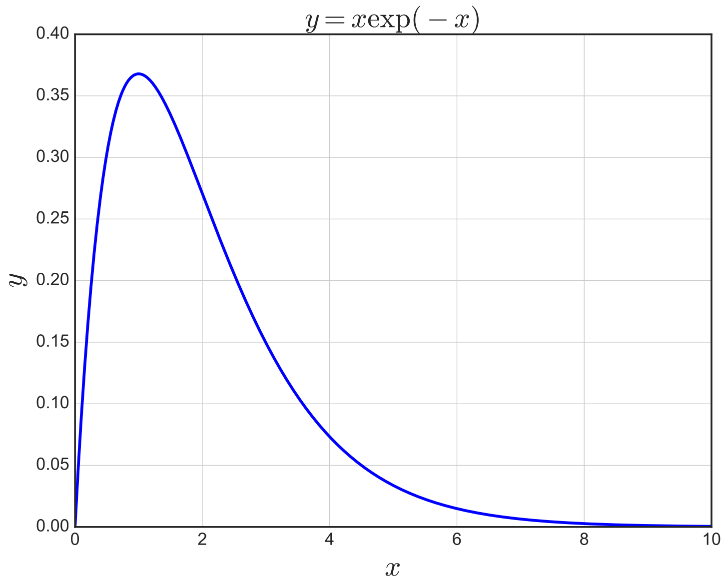

Figure 1: $y = x\exp(-x)$. The maximum value is achieved at $x = 1$ with $y = \exp(-1) \approx 0.368$.

as in the proof Thm. 2.14 in Frieze and Karoński [2015] will show that with high probability, the giant components in the two $G_{n,p}$ copies can only come from $O(1)$ numbers of connected components with size of $O(n)$. Therefore, with high probability, the size of the giant component in $G_{n,n,p}$ will still of size $2G(c)n$. The uniqueness of the giant components follows directly from the uniqueness of the two independent copies in the two independent copies. Similar arguments show that with high probability the sub-linear components from the two independent copies could form components of size at most $O(\text{poly}\log n)$. ■

Once we know that w.h.p. G_A is not connected, we immediately have the following corollary hold:

Corollary 2.1. Let G_A be the authorship graph where $\Delta = 3$, then w.h.p. the largest component of G_A contains at most $0.64n$ reviewers.

Proof. Simply plug in $c = 3$ into Thm. 2.1. By Thm. 2.1, the size (number of vertices) in the largest component will $\approx 2 \cdot 0.95n$. Let d be the number of reviewers in this largest component. Because $\Delta = 3$, then in expectation there will be $d(1 + \Delta)/2 = 2d$ papers covered by this set of d reviewers. Hence we have $d + 2d \leq 2 \cdot 0.95n$, which implies $d \leq 0.64n$. ■

Algorithm 2 Graph-Bipartite

Input: Authorship graph G_A, voting rule f and profile $\boldsymbol{\pi}$
Output: Ranking π
1: Partition the reviewers of G_A into R_c and $R_{\bar{c}}$.
2: $\pi_c \leftarrow f(\boldsymbol{\pi}_{\bar{c}})$, $\pi_{\bar{c}} \leftarrow f(\boldsymbol{\pi}_c)$
3: π interlaces between π_c and $\pi_{\bar{c}}$:

$$\pi^{-1}(j) = \begin{cases} \pi_c^{-1}((j+1)/2), & j = 1 \ (\text{mod} 2) \\ \pi_{\bar{c}}^{-1}(j/2), & j = 0 \ (\text{mod} 2) \end{cases}$$

4: **return** π

Let R_c and P_c be the subset of reviewers and the subset of papers in the largest component. Then by definition of connected component there will not have any edges between $R \backslash R_c$ and P_c, nor R_c and $P \backslash P_c$. In this case we can adapt the Naive-Bipartite algorithm such that the partitioning of reviewers/papers is based on the structure of the authorship graph, and as a result the following algorithm is still impartial:

Alg. 2 always works whenever the reviewers in G_A can be partitioned into two disjoint components. As we just discussed above, over the randomness of graph generation, this holds with high probability.

2.2 An Impossibility Theorem

Both Alg. 1 and Alg. 2 are impartial, but not unanimous. A natural question to ask is then: does there exist any review aggregation rule f that satisfy both properties simultaneously? Unfortunately the answer is no, as we will show now.

Definition 2.3 (Influence graph)**.** For any review aggregation rule f, we first define the following influence graph G_f induced by f as follows:

1. $G_f = (R, P; E)$ is a bipartite graph, where R is the set of reviewers and P is the set of papers. $|R| = |P| = n$.
2. For each pair of r_i, p_j, $(r_i, p_j) \in E$ iff $\exists \boldsymbol{\pi} = \{\pi_1, \ldots, \pi_i, \ldots, \pi_n\}$ and $\boldsymbol{\pi}' = \{\pi_1, \ldots, \pi_i', \ldots, \pi_n\}$ such that $f(\boldsymbol{\pi})(j) \neq f(\boldsymbol{\pi}')(j)$.

Intuitively, under the aggregation rule f, r_i is connected to p_j iff there exists a certain profile $\boldsymbol{\pi}$ such that r_i is able to change the output ranking of p_j by changing her own preference. By definition, it is clear to see that the following lemma hold:

Lemma 2.1. Let f be a review aggregation rule and $G_f = (R, P; E)$ be the influence graph induced by f. f is impartial iff $\forall r \in R$, $\deg(r) < n$.

Proof. The proof follows directly from the definition of the influence graph. For each reviewer r, if $\deg(r) < n$, then $\exists p \in P$ such that $(r, p) \notin E$, i.e., $\forall \boldsymbol{\pi} = \{\pi_1, \ldots, \pi_i, \ldots, \pi_n\}$ and $\boldsymbol{\pi}' = \{\pi_1, \ldots, \pi_i', \ldots, \pi_n\}$, we have $f(\boldsymbol{\pi})(j) = f(\boldsymbol{\pi}')(j)$. Since this holds for every reviewer $r \in R$, f is unanimous. On the other hand, if f is unanimous, then by definition it is easy to see that $\deg(r) \leq n - 1, \forall r \in R$. ■

For the ease of discussion, we call an influence graph G_f is *non-full* if $\deg(r) < n, \forall r \in R$. We proceed to show that no f can be unanimous and impartial simultaneously by showing that every non-full G_f cannot be unanimous:

Theorem 2.2. Let f be any review aggregation rule, then f cannot be both unanimous and impartial.[1]

Proof. For the sake of contradiction assume that f is a reviewing rule that is both unanimous and impartial. Let G_f be the corresponding influence graph defined as Def. 2.3. Without loss of generality, we first assume that $\deg(r) > 0, \forall r \in R$ and $\deg(p) > 0, \forall p \in P$. If $\exists r \in R, \deg(r) = 0$, then there is no paper that will be influenced by the preference of review r, i.e., f is blind to reviewer r. In this case we can safely remove r from the graph and no change will happen in the output ranking given by f. On the other hand, if $\exists p \in P, \deg(p) = 0$, then the output ranking of paper p will not depend on the preference from any reviewer, i.e., f is dictatorial on p. In this case we can safely remove p from the graph as well because now we only need to deal with the output ranking of the other $n - 1$ papers.

Now for any reviewer $r_i \in R$, let $e(r_i) \in P$ be the paper with the least alphabetical order in P such that $(r_i, e(r_i)) \notin E$. Define the set of such papers as $P' := \{e_1, \ldots, e_m\} = \{e(r_i) : r_i \in R\}$ and $m = |P'|$. Note that we must have $m \leq n$ and in fact m can be small than n becauseof the possible overlap between $e(r_i), \forall i \in [n]$. Since we also have $\deg(p) > 0, \forall p \in P$, it follows that $m > 1$, otherwise $\deg(e_1) = 0$.

We partition all the reviewers R into m groups $\{R_1', \ldots, R_m'\}$ based on P', such that all the reviewers in R_j' are not connected to e_j. With a slight abuse of notation, we use $(R_j', e_j) \notin E, \forall j \in [m]$ to actually mean that $(u, e_j) \notin E, \forall u \in R_j', \forall j \in [m]$. Restrict the output ranking to the subset of

[1] Note that Thm. 2.2 also holds when the number of reviewers does not equal to the number of papers.

papers in P', and consider the following two preferences over P'[2]:

$$\pi = e_1 \succ e_2 \succ \cdots \succ e_m, \quad \pi' = e_2 \succ e_3 \succ \cdots \succ e_m \succ e_1$$

Using π and π', define the following $m+1$ different profiles:

$$\boldsymbol{\pi}^{(0)} = (\pi, \ldots, \pi), \quad \boldsymbol{\pi}^{(k)} = (\pi', \ldots, \pi', \pi, \ldots, \pi), \forall k \in [m]$$

where in $\boldsymbol{\pi}^{(k)}$ the first k preferences are π' and the last $m-k$ preferences are π. In what follows, we will use the diagonalization argument [Cantor, 1891] to generate a contradiction using the condition that f is impartial. We first present the following lemma, which we shall prove later:

Lemma 2.2. Lef f be both unanimous and impartial, then $f(\boldsymbol{\pi}^{(k)})(k) = e_k$, i.e., under $\boldsymbol{\pi}^{(k)}$, the k-th position in the output ranking is e_k.

To better understand the main idea, let us consider the $m \times m$ table shown in Table 1, where rows index profiles and columns index the ranking of papers in the output ranking given by f.

From Lemma 2.2, since $f(\boldsymbol{\pi}^{(k)})(k) = e_k$, we know that diagonal of Table 1 is $(e_1, e_2, \ldots, e_m)$. Furthermore, noticing that since both π and π' agree with the ranking $e_2 \succ e_3 \succ \cdots \succ e_m$, by the unanimous property of f, once the diagonal element of the kth row is fixed, all the elements after it at the kth row can also be determined. As shown in Table 1, applying this argument to the whole table can help us to fully determine the upper triangular part of Table 1.

Table 1: Ranking table of the constructed m profiles: $\boldsymbol{\pi}^{(1)}, \ldots, \boldsymbol{\pi}^{(m)}$. Rows correspond to the output ranking of each profile, and columns give paper of a specific position in the output ranking. Assuming f is unanimous and impartial, we can fully determine the values of the upper triangular table, which are highlighted in blue.

f	1	2	3	$\cdots$	m
$\boldsymbol{\pi}^{(1)}$	e_1	e_2	$\cdots$	$\cdots$	e_m
0$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{\pi}^{(m)}$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	e_m

Now check the last row of Table 1, which shows that e_m should be positioned as the last paper in the output ranking of $\boldsymbol{\pi}^{(m)}$. However, on the other side, since $\boldsymbol{\pi}^{(m)} = (\pi', \ldots, \pi')$ and $\pi' = e_2 \succ e_3 \succ \cdots \succ e_m \succ e_1$, then again, by the unanimous property, e_1 should be positioned as the last paper in the output ranking instead. This leads to a contradiction, hence f cannot be both impartial and unanimous. ∎

We are only left to show that Lemma 2.2 is correct.

Lemma 2.2. Lef f be both unanimous and impartial, then $f(\boldsymbol{\pi}^{(k)})(k) = e_k$, i.e., under $\boldsymbol{\pi}^{(k)}$, the k-th position in the output ranking is e_k.

Proof. We prove by induction on k.

Base case. Since f is unanimous, and $\boldsymbol{\pi}^{(0)} = e_1 \succ e_2 \succ \cdots \succ e_m$, the output ranking $f(\boldsymbol{\pi}^{(0)})$ must be:

$$f(\boldsymbol{\pi}^{(0)}) = e_1 \succ e_2 \succ \cdots \succ e_m$$

Consider $k = 1$. Note that π and π' differs only at the position of e_1, and in $\boldsymbol{\pi}^{(1)}$, only R'_1 changes their preference and all the other preferences are kept fixed. Then by the impartiality of f, the output ranking of e_1 will not be changed due to R'_1, so we must have:

$$f(\boldsymbol{\pi}^{(1)}) = e_1 \succ e_2 \succ \cdots \succ e_m$$

i.e., $f(\boldsymbol{\pi}^{(1)})(1) = e_1$.

[2]It is justifiable as equivalently we can understand it as positioning all the other $n - m$ papers at the end of each preference.

Induction step. Suppose Lemma 2.2 hold for any integer $\leq k$. We now consider the case of $k+1$. Notice that $\boldsymbol{\pi}^{(k)}$ and $\boldsymbol{\pi}^{(k+1)}$ only differs at the $k+1$th preference given by R'_{k+1}. Since $f(\boldsymbol{\pi}^{(k)})(k) = e_k$, and the following preference hold in both π and π':

$$e_k \succ e_{k+1} \succ \cdots \succ e_m$$

Then by unanimity of f, we claim that the last $m-k+1$ positions in the output ranking of $f(\boldsymbol{\pi}^{(k)})$ must be given by $e_k \succ e_{k+1} \succ \cdots \succ e_m$, i.e., $f(\boldsymbol{\pi}^{(k)})(k+1) = e_{k+1}$. But now since $(R'_{k+1}, e_{k+1}) \notin E$, by the impartiality of f, changing the preference given by R'_{k+1} will not influence the output ranking of e_{k+1}, hence we conclude that $f(\boldsymbol{\pi}^{(k+1)})(k+1) = e_{k+1}$, which finishes our proof. ■

3 Discussion & Future directions

Somewhat surprisingly, there is no reviewing aggregation rule f that can be both impartial and unanimous, while arguably both properties are desirable for any fair reviewing aggregation rule to have. Our result highlights the intrinsic hardness in designing a perfect reviewing system, and can be naturally extended to the case when the number of reviewers and papers are different. Theorem 2.2 can be seen as a "paper review" version of Arrow's impossibility theorem (ref) for voting and Gibb-S (ref) theorem for social choice. The result also applies to other areas where a group of people want to rank a set of items.

We give some possible directions for future work in the following subsections.

3.1 Limited reviews

One important extension is the limited reviews setting: in practice, usually reviewers only give comments to a small subset of all the papers, instead of ranking all the papers. Formally, suppose we are given a set of papers $S_i \subseteq P$ for each reviewer r_i, that represents the set of papers that r_i is going to review. Each π_i is a permutation over S_i instead of P. To resemble the actual review process, we assume that $|S_i| = c$ for some small constant c.

3.1.1 Impartiality and Unanimity under Limited Reviews

The property of impartiality (Definition 2.2) naturally extend to the limited reviews setting. Note that a reviewer r_i might be able to influence the ranking of some paper p_j even if $p_j \notin S_i$; for example, suppose $n=3$, and $S_1 = \{p_2, p_3\}, S_2 = \{p_1, p_3\}, S_3 = \{p_1, p_2\}$, and f is Borda count. If $\pi_2 = p_1 \succ p_3, \pi_3 = p_2 \succ p_1$, then p_1 has a $1/3$ chance of being the top if $\pi_1 = p_3 \succ p_2$, while always being the 2nd if $\pi_1 = p_2 \succ p_3$. So r_1 can influence the position of p_1 even $p_1 \notin S_1$.

On the other hand, the property of unanimity does not extend quite well to the limited review setting. A most natural extension of Definition 2.1 under limited reviews setting is to restrict the set of reviewers for each pair (j_1, j_2) to be those reviewers that have reviewed both of them: I.e., we define an aggregation rule to be unanimous if for any pairs $j_1, j_2 \in [n], j_1 \neq j_2$, for every r_i such that $j_1, j_2 \in S_i$ we have $\pi_i(j_1) > \pi_i(j_2)$, then we have $f(\boldsymbol{\pi})(j_1) > f(\boldsymbol{\pi})(j_2)$. Generally this property cannot be satisfied for any aggregation rule. For example, in the same case above where $n=3$, $S_1 = \{p_2, p_3\}, S_2 = \{p_1, p_3\}, S_3 = \{p_1, p_2\}$, if $\pi_1 = p_3 \succ p_2, \pi_2 = p_1 \succ p_3, \pi_3 = p_2 \succ p_1$, then $f(\boldsymbol{\pi})$ have to satisfy $p_3 \succ p_2 \succ p_1 \succ p_3$, which is impossible. This extension is also not so justifiable either, since if only r_1 has given some opinion on the best between p_2, p_3, we might not want to follow this signal since we may get different messages from other reviewers.

There are other possible modification of unanimity under limited reviewers setting. One is to think of the set of *compatible* total rankings ; if there exists some total ranking (i.e., ranking on all papers) that is compatible with everyone's (partial) ranking, we should use this specific ranking. We formally state this as below:

Condition 3.1. If there exists some total ranking π such that π is compatible with π_i on S_i for every reviewer r_i, then $f(\boldsymbol{\pi})$ should also be compatible with every π_i on S_i (there might be several possible such rankings).

Notice Condition 3.1 is much weaker than the previous definition of unanimity: In the case where $S_i = P$ (everyone ranks all the papers), this corresponds to the requirement that when everyone gives a same ranking, then the aggregation ranking is this specific same ranking. Although Condition 3.1 is weaker than unanimity, we still suspect giving an aggregation rule satisfying Condition 3.1 and impartiality is impossible.

Another possible modification is to think of the ranking π_i as a partial observation of r_i's total ranking over P. Specifically, suppose each reviewer r_i has a total ranking over P, which is consistent with r_i's reviews π_i. Now consider all such possible total rankings for every reviewer (we call this a total profile), if in every possible total profile we have $j_1 \succ j_2$, then we should have $f(\boldsymbol{\pi})(p_{j_1}) > f(\boldsymbol{\pi})(p_{j_2})$. Formally:

Condition 3.2. For each reviewer r_i, consider the set of total rankings π'_i that π' is compatible with π_i. If for every possible total profile $\boldsymbol{\pi}' = (\pi'_1, \pi'_2, ..., \pi'_n)$ we have $\pi'_i(j_1) > \pi'_i(j_2)$, then we have $f(\boldsymbol{\pi})(p_{j_1}) > f(\boldsymbol{\pi})(p_{j_2})$.

Condition 3.2 is as strong as unanimity when $S_i = P$. However, it is also quite weak, since that it actually only gives a constraint on the ordering of (j_1, j_2) when everyone will review both paper j_1 and j_2. Such conditions is not quite possible if we restrict $|S_i| = c$ for some small constant c. However, this will be different when we use other forms of limited reviews, which we will illustrate below.

3.1.2 Alternative Definition of Limited Reviews

The definition of limited reviews might be too weak and not resembling the actual review process well. Actually, if we only have partial ordering on (say) 5 papers from each reviewer, then each paper is only pairwise compared with around 20 other papers. It will be hard to derive a total ranking that compare this paper to the other $n - 20$ papers. In the actual review process, reviewers instead give scores to fully illustrate their opinions. So instead we propose the limited reviews with scores setting: Each reviewer r_i gives still gives comments on papers in S_i, but now with a slight abuse of notations, let $\pi_i(p_j) \in [n]$ be the score of reviewer i on paper j. We assume that r_i gives an estimate on p_j's position in the total ranking, so the score range is $[n]$. We require that $\pi_i(p_j) \neq \pi_i(p'_j)$ for $p_j, p'_j \in S_i$, i.e., reviewers do not give the same score to two papers.

Under the new setting with scores, Condition 3.2 becomes more interesting. Suppose some paper p_j is given a score of 1, then in the set of possible rankings we will always have $p_j \succ p'_j$ for every other paper p'_j. This is different than the limited reviews without scores case, where we cannot judge the position of p_j and p'_j if p'_j is not reviewed.

3.2 Other directions

We also plan to investigate whether we can have a positive result by relaxing the output of f to be a subset of papers that are going to be accepted, rather than asking f to output a total ordering of all the papers. Another possible direction is to consider sub-fields that are present in scientific papers.

References

G. Cantor. Uber eine elementare frage der mannigfaltigkeitslehre. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 1(1):75–78, 1891.

A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, 2015.

A. B. Kahng, Y. Kotturi, C. Kulkarni, D. Kurokawa, and A. D. Procaccia. Ranking wily people who rank each other. *Technical Report*, 2017.

C. Severance. Teaching the world: Daphne koller and coursera. *Computer*, 45(8):8–9, 2012.

N. B. Shah, B. Tabibian, K. Muandet, I. Guyon, and U. von Luxburg. Design and analysis of the nips 2016 review process. *arXiv preprint arXiv:1708.09794*, 2017.