Peer reviewing and manipulation under uncertainty

Alex Carver and Paolo Turrini

Abstract

We study the Borda peer reviewing system, where n participants submit one project proposal each and review m others, with the Borda count selecting k proposals be funded in the end. This method, used in a recent pilot project by the American National Science Foundation, has been shown to be manipulable, i.e., there are instances in which a reviewer is better off downgrading good projects in order to place themselves among the k winners. However, in peer reviewing settings, where reviewers are typically unaware of submissions other than the ones they are reviewing, this notion of manipulation is highly unrealistic. When imperfect information is part of the decision-making process, attempting manipulation involves a significant degree of risk, as it might end up downgrading the position of the manipulator in the final ranking. Estimating in which conditions manipulation is practically infeasible is therefore a major open issue left in the literature. Here we tackle this problem, establishing a number of results which combine the use of theoretical and computer-aided analysis. In particular we use brute-force computation to fully chart the effects of manipulation in committees up to six members and restricted cases of seven members, discovering novel configurations that resist strategic manipulation and identifying the parameters which make it a threat in theory but not in practice.

Department of Computing, Imperial College London Email: anc15@imperial.ac.uk

1 Introduction

Consider n project proposals simultaneously submitted for funding, each of which receives m reviews, and only k of which will be funded in the end.

The standard approach is to assign each proposal to m (external) reviewers and decide the relative ranking of the proposals based on those reviews. While this approach guarantees more independent evaluations - e.g., the reviewers are typically not proposers themselves - it increases the number of participants in the whole process and does not provide any incentive for reviewers to provide high quality reviews.

A different approach, made famous by an NSF peer reviewing pilot project [11, 10], is to have the PIs themselves rank their competitors, and aggregate the evaluations using the Borda count, i.e., each proposal gets x points each time is ranked in position $x + 1$ by some reviewer.

However, whilst the Borda count has the advantage of, e.g., forbidding the reviewers to assign the score of 0 to all other competitors, such system is still not *strategy-proof*; in other words, telling the truth is not always a good strategy, as shown by the following example.

Example 1. *Players A,B,C,D,E are evaluating one other. Each of them is assigned 3 different proposals. If everyone were sincere, they would rank their assigned proposal in alphabetical order. The reviewing process uses Borda count and only the top ranked gets funded. Table 1 shows an instance in which player B is better off misrepresenting their preferences and reversing the truthful ranking.*

As in the example, to establish that a peer reviewing system is not strategy-proof it is sufficient to find a review assignment and a player, such that, when everybody else is ranking their assigned proposals sincerely, that player has the incentive of not doing so. We note though that this *classical* strategy-proofness, when applied to peer reviewing setups, bears a very strong assumption: namely that the potential manipulator *knows* their competitors' ranking. But this can only be the case in

	A	B	C	D	E
A	-	3	-	2	1
B	3	-	2	-	1
C	3	-	-	2	1
D	3	2	1	-	-
E	-	3	2	1	-
	9	8	5	5	3

	A	B	C	D	E
A	-	3	-	2	1
B	1	-	2	-	3
C	3	-	-	2	1
D	3	2	1	-	-
E	-	3	2	1	-
	7	8	5	5	5

Table 1: On the left side, reviewers ranking truthfully. On the right side, player B reversing their ranking.

highly simplified settings, e.g., everyone is reviewing everyone else, or in rare cases in which a reviewer knows how proposals are assigned. This assumption is even more restrictive for cases in which only a few winners are allowed, with PIs having no incentive of sharing their pool.

In peer reviewing and similar settings (e.g., committee members' internal evaluations) we therefore need a notion of manipulation that takes uncertainty into account. Even more so with aggregation methods such as the Borda count, when the over-ranking of the worse proposals can come at the expense of the manipulator's own success.

Our contribution While acknowledging the fact that the Borda peer reviewing system lacks classical strategy-proofness, we study what happens when we incorporate uncertainty into the reviewers' decision-making. In particular, when the potential manipulators need to carefully assess the degree of risk involved in misrepresenting their preferences, as a function of the possible scenarios that could occur.

In this paper we assume the proposals to be distributed randomly, with a uniform probability distribution, and common knowledge thereof. To facilitate the analysis we also work with the assumptions that the number of participants, the number of reviews per proposal and the size of the funded set are common knowledge, together with the fact that reviewers are experts, i.e., everyone shares the same objective ranking of proposals.

We then lift the notion of strategy-proofness to incorporate imperfect information, and provide a classification of peer reviewing scenarios accordingly.

Using computer-aided analysis, we chart the effect of manipulation strategies in committees up to six members and constrained committees of seven members, showing that its expected utility depends on the number of funded proposals, the number of players involved and the number of reviews each one gets, and, for a number of configurations, can be unrewarding in expectation. As a side result, we obtain novel computer-generated proofs of classical strategy-proofness for a number of setups.

Related approaches The research on strategy-proof, or *impartial*, peer reviewing systems has gained increased attention in AI venues, since [9] and the observation that their framework is not strategy-proof [7]. Alternative methods have been proposed, from different angles, but all relying on classical strategy-proofness. Here we mention the ones that are most related to ours, other important references include: [4, 5, 6, 13, 12].

Credible Subset [7] In Credible Subset (CS), reviewers assign bounded scores to their reviews, and the potentially successful manipulators, i.e., the reviewers that could be within the k funded

ones, are also selected to be funded, with a given probability. While the system is strategy-proof, it has been shown to yield an empty set of funded proposals in a significant number of cases [2].

Dollar Raffle [2] The Dollar Raffle method (DR) consists of reviewers distributing a score from 0 to 1 to their reviews, rather than independently allocating them as in (CS). The system has been shown not to be strategy-proof and to perform significantly worse than other methods, e.g., CS [2].

Exact Dollar Partition [2] Exact Dollar Partition (EDP) features reviewers that get clustered at random and rank reviews coming from different clusters, with a Borda-like count. Scores determine how important a cluster is and the top proposals of each cluster are selected, depending on their clusters' importance. Dollar Partition has been demonstrated to be the best available in terms of optimality, and, although suboptimal, it preserves strategy-proofness [2].

As shown experimentally in [2], each of these methods underperforms when compared to truthful Borda, the outcome of Borda in which reviewers do not misreport their preferences.

This paper explores manipulation under uncertainty, going beyond classical strategy-proofness. Its take on manipulation relates to the *local dominance* framework in [8], which studies a general notion of uncertainty in voting games.

Paper structure In Section "The Borda peer reviewing system" we describe the formal background of our setup, establishing manipulability results for the basic case of everyone reviewing everyone else. We then move on to the case in which the review pool is too small, and incorporate the notion of uncertainty in strategy selection. We analyse the complexity of the problem and set up a computational framework in which to explore manipulation. We then fully chart the effects of manipulability in small committees. Finally, we discuss the future research directions.

2 The Borda Peer Reviewing System

Let $\mathcal{N} = \{1, 2, \ldots, n\}$ be a finite set of individuals, each submitting a project proposal and competing to have their own among the $k \leq n$ funded. Each proposal is randomly assigned to m individuals which, independently and concurrently, submit their evaluation. The evaluations received are then aggregated and the top k proposals are funded.

Let $1 \leq m \leq n-1$. A **(review) assignment** is an irreflexive m-regular directed graph over $\mathcal{N}$. Irreflexivity imposes no individual to review themselves, while m-regularity - meaning that vertices share the same indegree and outdegree of m - that all individuals review and are reviewed by an equal number of other individuals. We denote such assignment a_m and the set of all such assignments $\mathcal{A}_m$, where $a_m(c)$ is the m sized set of submissions assigned to c. We omit m when clear from the context. For $\mathbf{m} \subseteq \mathcal{N}$, with $|\mathbf{m}| = m$, we let $\Sigma_{\mathbf{m}}$ be the set of permutations over $\mathbf{m}$, i.e., bijections of the form $\sigma : \mathbf{m} \to \mathbf{m}$, and denote $\sigma(\mathbf{m})[k] \in \mathcal{N}$ the k-th element in the order induced by σ on $\mathbf{m}$. As by m-regularity the strategy space of the individuals only depends on the size of the reviewed pool, we dispense with mentioning the specific assignment explicitly whenever clear from the context, and in some cases, when no ambiguity can arise, also the size of the pool itself. So, when a_m is clear from the context, and thus $\mathbf{m}$ is understood to be the specific subset of $\mathcal{N}$ provided by a_m, we shall write σ (resp., Σ) instead of $\sigma(\mathbf{m})$ (resp., $\Sigma(\mathbf{m})$).

An **aggregator** for a is a function $F_a : \Sigma^{\mathcal{N}} \to \Sigma$ that associates to each profile of evaluations a final ranking of the individuals, and where the individuals permutations apply to the pool assigned to them by a.

In other words, an aggregator is a protocol to decide the final ranking out of individual evaluations. While the focus here is on the Borda count, the framework can easily be extended to different aggregation rules.

For $1 \leq k < n$, the **accepted set**, under profile $\rho = (\sigma_1, \sigma_2, \ldots, \sigma_n) \in \Sigma^{\mathcal{N}}$ and assignment a, is the set of the first k individuals in $F_a(\rho)$. We say that individual i **weakly prefers** ρ to ρ' (denoted $\rho \succeq_i \rho'$) whenever i is in the accepted set under ρ and that i **strictly prefers** it (denoted $\rho \succ_i \rho'$) when they weakly prefer it and they are not in the accepted set under ρ'.

Given profile ρ, the **Borda score** of player i is $\sum_{j \in N, \rho_j[k]=i} m - k$. In words, we add $m - k$ to i's final score, each time some individual has ranked i at the k-th place in their reviews pool. The **Borda ranking** following permutation σ (or σ-Borda ranking) is the order on $\mathcal{N}$ which follows the individuals' Borda scores, from the highest to the lowest, and where ties are broken following σ - i.e., the weak linear order where each indifference class is resolved with the ranking of σ. Unless otherwise specified we will use the identity permutation on $\mathcal{N}$ and simply omit σ. This, notice, implicitly assumes that, in case of ties, experts will be able to tell apart the better proposals.

Ground truth and truthful ranking We make the assumption that reviewers are both experts and committed, i.e., if they were to review all the proposals and truthfully report their evaluation they would come up with the same ranking $\mathcal{N}$. While aware that this is often a simplifying assumption, we do claim that it is perfectly reasonable in a number of cases. We refer to this ranking as the **ground truth**.

With sufficiently many proposals to be reviewed, reviewers will typically have smaller pool than $n - 1$ proposals. For an assignment a_m, we call a profile σ the **truthful ranking**, if it is the Borda ranking generated by permutations on each $\mathbf{m}$ that reflect the ground truth ordering. So if a reviewer were assigned proposals with ground truth ranking $\{7, 1, 15, 5\}$ then the unique permutation satisfying this property would be $(1, 5, 7, 15)$. Given a and m we denote σ_i^* as the permutation on the $\mathbf{m}$ reviews assigned to i by a which reflects $\mathcal{N}$, and call it the **truthful strategy**, while we call $\rho^* = (\sigma_1^*, \sigma_2^*, \ldots, \sigma_n^*,)$ the **truthful (strategy) profile**.

It is important to notice that the truthful ranking may change depending on how reviews are assigned. So it can very well happen that individual 3 only gets paired with individual 1, while individual 15 only with lower ranked ones. We observe that, with Borda, if $m = n - 1$ and $n > 2$ the **truthful ranking** coincides with the **ground truth**. While ideally we would like to fund the best possible proposals in the ground truth, sometimes we need to be happy with the ones given by the truthful ranking. Let $\rho_{-i} \in \Pi_{j \in \mathcal{N} \setminus \{i\}} \rho_j$. We call a strategy profile ρ a **Nash equilibrium** if for all i, and $\sigma \in \Sigma_i$ we have that $F_a(\rho) \succeq_i F_a(\sigma_i, \rho_{-i})$.

Observe the following property of the ideal but often infeasible case where $m = n - 1$, which also demonstrates the importance of how ties are broken.

Proposition 1. *Let F_a be the Borda aggregator for assignment a and let $m = n - 1$. Then, for all $k < n + 1$, ρ^* is a Nash equilibrium.*

Proof. By calculation, we have that in ρ^* the score of the r-th individual in $\mathcal{N}$ is $n^2 - nr - 2n + 2r$, where r is the rank of the individual in the ground truth ordering. To see this, notice that as each individual is gets $n - r - 1$ by the $n - r$ others who are lower in the ground truth ordering, and $n - r$ by the $r - 1$ ones who are higher. Since, once n is set, the only variable is r, the difference between the score of two neighbouring submissions in the ground truth ordering is constant with respect to n. By subtracting the formula with r replaced with $r - 1$ this difference amounts to $n - 2$ is all cases. Since a single manipulator cannot withhold more than $n - 2$ points from the final score of another paper, they can thus never strictly overtake. □

It is important to observe how the validity of Proposition 1 relies on the tie-breaking rule. While manipulators cannot improve upon their final ranking by misrepresenting their preferences it should be noted that they might indeed force ties. With all reviewers other than the manipulator acting 'truthfully' (not employing a manipulation strategy), these ties can only occur from the best and worst papers in the ground truth ordering (as only they can lose or gain respectively $n - 2$ in score). Thus only the second best and second worst papers can be put into a draw by their own strategy; the

second best by bottom ranking the top paper and the second worst by top ranking the worst paper. In light of this observation, we can establish a more general result.

Proposition 2. *Let σ be a permutation of $\mathcal{N}$, $m = n - 1$, and F_a be an aggregator returning the σ-Borda ranking. Then, for all $2 \leq k < n + 1$, ρ^* is a Nash equilibrium.*

Proof. Let c be a player ranked at most 3rd best in $\mathcal{N}$. With $m = n-1$ we have that $\{1, 2\} \subseteq a_m(c)$. As in σ^* the score of the r-th individual in $\mathcal{N}$ is $n^2 - nr - 2n + 2r$, and the only way c can improve their position given σ^*_c is to withhold $n - 2$ from 2. But since 1 is also a member of $a_m(c)$, 2 would be truthfully ranked second and so receive $n - 3$ points from c. Thus it is not possible for c to withhold at least $n - 2$ points from 2, nor any submission other than 1. As $k \geq 2$, this concludes the proof. □

From an inspection of the argument given in the proof one can in fact conclude that in some situations the second best player in $\mathcal{N}$ can, under some tie-breakings, overcome the best one. But as long as there are at least two funded proposals this is irrelevant, as truth-telling is still a sufficiently good strategy.

The results we have given establish how a situation in which everyone reviews everyone else is robust to manipulation. However, even in small committees, this might be an infeasible setup, due to time constraints and the length of the proposals themselves. The coming sections deal with the cases in which m is small enough, reintroducing the concrete threat of manipulation.

3 Manipulation Under Uncertainty

In situations in which $m << n$, individuals will typically lack access to the proposals' pool assigned to their fellow reviewers and will also be unable to have a precise estimate of their final position in the truthful ranking. So while we can come up with strategy profiles in which individuals can gain from manipulation (as in Table 1), such profiles assume an underlying knowledge of the review assignment.

With Borda this is a crucial point. Because of the way points are awarded, it is impossible to damage each reviewed submission, i.e., it is impossible to give 0 to everyone. Indeed, every reviewer must award the same total of points distributed across their reviewed submissions. Also, misrepresenting the score by, e.g., downgrading the submissions that are thought to be best, involves a degree of risk: giving a high score to submissions that are thought to be of low quality might actually make them jump to a favourable position in the final ranking, potentially at the expense of the manipulator themselves. Without being able to damage all submissions assigned, or even without the ability to give less reward than a truthful reviewer would, there is no clear or universally successful strategy with which to manipulate. Thus it is paramount for a manipulator to determine the expected success (or utility) of a manipulation strategy.

To do so, recall that the set A_m of review assignments of size m is the set of all m-regular irreflexive directed graphs over $\mathcal{N}$. When a reviewer i receives their poll $a_m(i)$, the set of the possible worlds is the set of all assignments in A_m compatible with $a_m(i)$. We denote this set $[a_m(i)]$. We assume, recall, that reviews are assigned randomly. This induces a natural notion of expected utility of a manipulation strategy with respect to the truthful profile: the number of times playing the strategy puts the manipulator from outside the top k into the top k, minus the number of times that it puts the manipulator out from inside the top k, divided by the number of possible assignments.

Formally, fixing m, n, k, we denote $[\sigma_i \uparrow]$ the set of assignments for which strategy σ_i is better than σ^*_i, i.e.,

$$\{x \in [a_m(i)] \text{ s.t. } F_x(\sigma_i, \rho^*_{-i}) \text{ accepts } i \text{ and } F_x(\rho^*) \text{ does not}\}$$

Likewise, we denote $[\sigma_i \downarrow]$ the set of assignments for which σ_i is worse than σ_i^*, i.e.,

$$\{x \in [a_m(i)] \text{ s.t. } F_x(\rho^*) \text{ accepts } i \text{ and } F_x(\sigma_i, \rho^*_{-i}) \text{ does not}\}$$

The expected utility is obtained combining the two.

$$\mathbb{E}(\sigma_i(a_m(i))) = \frac{|[\sigma_i \uparrow]| - |[\sigma_i \downarrow]|}{|[a_m(i)]|}$$

Again fixing m, n, k, we call a (peer reviewing) system:

very weakly safe if for all σ_i we have that $\text{avg}_{i,a_m(i)} \mathbb{E}(\sigma_i(a_m(i))) \leq 0$, i.e., the average expected utility is at most 0 for all strategies, across all positions, and all received assignments.

weakly safe if there is no i and σ_i for which $\text{avg}_{a_m(i)} \mathbb{E}(\sigma_i(a_m(i))) > 0$, i.e., no strategy yields a positive expected utility, at no position, across all received assignments.

strongly safe if there is no i, $a_m(i)$ and $x \in [a_m(i)]$ such that $F_x(\sigma_i, \rho^*_{-i}) \succ_i F_x(\rho^*)$, i.e., no strategy yields a positive expected utility, at no assignment.

unsafe otherwise.

While strongly safe is the counterpart of (classically) strategy-proof in our setting, both very weakly and weakly safety seem appropriate requirements against which to evaluate a peer reviewing system.

We can for instance rely on a very weakly safe system when individuals do not know their relative position in the ground truth ordering nor their fellow reviewers' assignments, and on a weakly safe when they do not know their fellow reviewers' assignments. On the contrary, manipulating in a strongly safe system can never be profitable, even knowing the assignment.

3.1 Computing the odds for manipulation

As stated, the concern with manipulable peer reviewing systems is that reviewers can potentially increase their chances of being admitted into the top k by misrepresenting their evaluation. However, if we think of manipulators as rational agents, they would manipulate only if they expect it to improve their chances of being admitted into the top k, and therefore would factor the probability, not the the possibility, of successfully manipulating.

While the notion of expected utility of a strategy defined previously seems the natural one when reviews are assigned randomly, its concrete estimation is problematic, for two reasons: first, the problem of estimating the size of a_m remains unsolved to this day, even for undirected graphs (compare [3], which gives asymptotic estimates); second, a worst-case complexity analysis shows that the number grows rapidly with m and n. To see this notice how for each individual there are at most $\binom{n-1}{m}$ choices. So the obvious upper bound is $\binom{n-1}{m}^n$. If $m << n$ and assignments are random, there is a small (exponentially small but nontrivial) chance that we end up being m-regular, and the above upper bound is a relatively good estimate. [1]

What we do in the next section is to go around the complexity estimates and force-compute the expected utility for all strategies, all potential manipulators and all assignments, for all n up to 6, and for each of these with all $2 \leq m \leq n$, as well as the cases of $m = 2$ and $m = 6$ for $n = 7$. This allows us to give a precise estimate on the effects of manipulability in such committees.

[1] Generally, $\binom{n}{m}$ approximates $\exp(nh(m/n) \pm E)$, where $h(p) = -p\log(p) - (1-p)\log(1-p)$ and the error term $E\ \mathcal{O}(\log n)/2$. $\binom{n}{k}$ is always at least $(n/m)^m$ and at most $(ne/m)^m$.

4 Computer-Aided Analysis

To determine the expected utility of manipulation strategies two main procedures were created, to map out the results of all possible strategies, for each player i with $3 \leq n \leq 6$, for each assignment m with $2 \leq m < n$ and each size $1 \leq k < n$ of the accepted set. In the case of $n = 7$ we have only limited the analysis to $m = 2$ and $m = 6$, quantifying over the rest.

The first procedure takes as input an array of all the PIs, and the number m of reviews each PI is assigned to and then finds all possible assignments via a recursive depth-first search method (being fairly standard it has not been detailed for space reasons). Once all submissions have been assigned m valid reviews, i.e., satisfying the required constraints, we generate a valid assignment a_m for all PIs.

The next step is to calculate the expected utility of each strategy. In the `TestStrat` strategy testing algorithm (see algorithm 1), we generate all possible strategies a potential manipulator could employ. We refer to a strategy as a list of natural numbers inclusively between 1 and m, with 1 representing the best submission in the reviewer's true opinion (i.e. the PI with the best ground truth ranking), 2 representing the second best and so on, and m the worst. This list is then ordered according to how the reviewer plans to rank them, with the first position in the list denoting the PI to be given the top rank to the last position for the lowest rank.

To compare the results of manipulation, we first need the results of all reviewers ranking truthfully (i.e., the truthful ranking, which, it is important to recall, may well differ from the ground truth ranking, and indeed can differ for each assignment). After each reviewer has ranked their reviewed submissions truthfully, the submissions are ranked in a descending order based on their objective score. Then they are stably sorted by the Borda score they have received from their review rankings. The result is that they are sorted primarily by their total Borda review reward, and then ties are broken using the identity permutation.

With the truthful ranking created, we then iterate through each reviewer. Each reviewer in turn iterates through every untruthful strategy, updating the scores of those they reviewed, and the ranking is recalculated based upon this manipulation in the same manner. The reviewer then records the change in rank they achieved by employing the untruthful strategy, and attempts the next untruthful strategy. Once all untruthful strategies have been tested, the reviewer is reset to rank truthfully and the next reviewer tests the untruthful strategies, until all reviewers have tested all strategies. This is done for all compatible assignments.

Since we record each PI's truthful ranking and the rank change for each strategy, we can determine if a strategy was either successful (the strategy achieved rank $\leq k$ and truthful rank $> k$), neutral (the achieved rank and truthful rank were both $\leq k$ or they were both $> k$) or a failure (the truthful rank $\leq k$ but the achieved rank $> k$) by comparing these recorded values to a given k. The expected utility is then calculated aggregating the figures obtained.

5 Results

We present, as representative cases, samples of the tables of the utilities for each strategy. Each table presents how changing the values for n, m and k alters the utilities of strategies and so the safety of the system. We then present overview tables, which summarise the results as a function of n, m and k.

Strategies are formatted as an ordered list, with the number corresponding to the submissions' actual rank among the m submissions. Note that in these tables, A refers to the paper with the first position in the ground truth ordering of all n papers, B the second position and so on.

The entries in the matrix, relative to each strategy, determine their expected gain with respect to the truthful strategy, i.e., the difference in expected utility between the the strategy and ranking truthfully aggregated across all individual assignments: a positive gain indicates the strategy will

Algorithm 1 Procedure `TestStrat` is given a set of PIs that all have valid assignments and tests each strategy for each PI. PIs is an array of PI that each have a preset $objectiveScore$ and other PI instances assigned for them to review. They store their $reviewScore$ and $truthfulRank$, and can review PIs assigned to them with a given strategy

procedure TESTSTRAT(PIs, m)
  $honStrat \leftarrow \mathbb{N}$ **from** 1 **to** m
  $untruthfulStrats \leftarrow$ PERMUTATIONS($honStrat$)
    ▷ Get array of strategies by generating all permutations of $honStrat$
  REMOVE($dishonStrats, honStrat$)
  for all PI **in** PIs **do**
    REVIEW($PI, honStrat$)
      ▷ Initially all PIs review and reward according to truthful strategy
  end for
  $truthfulRanking \leftarrow$ SORT(PI **in** PIs **by** $PI.objectiveScore$ **descending**)
  $truthfulRanking \leftarrow$ SORT(PI **in** $truthfulRanking$ **by** $PI.reviewScore$ **descending**)
    ▷ Sort is stable, so submissions sorted first by total review reward, then by their true score
  for all $PI, rank$ **in** ENUMERATE(PIs) **do**
    $PI.truthfulRank \leftarrow rank$
      ▷ Each PI records their truthful ranking
  end for
  for all PI **in** PIs **do**
    for all $strat$ **in** $dishonStrats$ **do**
      REVIEW($PI, strat$)
        ▷ Each PI tries reviewing according to each untruthful strategy
      $manipulatedRanking \leftarrow$ SORT(PI **in** PIs **by** $PI.objectiveScore$ **descending**)
      $manipulatedRanking \leftarrow$ SORT(s **in** $manipulatedRanking$ **by** $PI.reviewScore$ **descending**)
      $rank \leftarrow$ INDEX($PI, manipulatedRanking$)
        ▷ Record in database the rank achieved with each strategy
    end for
    REVIEW($PI, honStrat$)
      ▷ Reset PI to have reviewed and rewarded truthfully before testing next
  end for
end procedure

$n = 6, m = 4, k = 2$						
Strategies	**A**	**B**	**C**	**D**	**E**	**F**
(4; 3; 1; 2)	0	0	0.2151	0	0	0
(2; 4; 3; 1)	0	0	0.034	0	0	0
(3; 4; 1; 2)	0	0	0.2151	0	0	0
(3; 1; 4; 2)	0	0	0.2151	0	0	0
(4; 2; 3; 1)	0	0	0.034	0	0	0
(4; 1; 3; 2)	0	0	0.2151	0	0	0
(1; 4; 3; 2)	0	0	0.2151	0	0	0
(3; 2; 4; 1)	0	0	0.034	0	0	0
(3; 4; 2; 1)	0	0	0.0679	0	0	0
(4; 3; 2; 1)	0	0	0.0679	0	0	0
(1; 3; 4; 2)	0	0	0.2151	0	0	0
(1; 4; 2; 3)	0	0	0.034	0	0	0
(2; 3; 4; 1)	0	0	0.034	0	0	0
(4; 1; 2; 3)	0	0	0.034	0	0	0
(3; 1; 2; 4)	0	0	0.034	0	0	0
(1; 3; 2; 4)	0	0	0.034	0	0	0

Table 2: $n = 6, m = 4, k = 2$ is an unsafe setup. Indeed no untruthful strategy is strictly worse than the truthful one, while many others are strictly better. Notice how player C is the only one able to successfully manipulate.

put the reviewer in the top k more than being truthful, whilst a negative utility indicates it will result in the reviewer being in the top k less than if they were truthful. A utility of 0 indicates that the strategy results in the reviewer being in the top k the same number of times compared to the truthful strategy. Only the strategies with non-zero utility at some position have been included; i.e, strategies that do not appear in these tables had the exact same utility and no change from the truthful strategy for all positions.

In Table 2, which presents the utilities for $n = 6$, $m = 4$ and $k = 2$ we can an example of an unsafe system; not only do there exist untruthful strategies that expect at least as much utility as the truthful strategy, but in fact there are no untruthful strategies that are strictly worse, at any position. No matter what their position may be, a potential manipulator can choose a risk-free strategy such as (4; 3; 1; 2).

Table 3 shows that by increasing k to 3 there do now exist untruthful strategies that offer a worse expected utility than being truthful. However, this is still unsafe, as strategies such as (1; 2; 4; 3) exist which offer at least as much utility as the truthful strategy and are strictly better in some cases. Again, they are risk-free manipulation strategies.

But with k increased again to 4, Table 4 shows a weakly safe configuration, as no strategy offers a better utility than the truthful strategy for any position. In fact this case is also strongly safe: there is no possible way for a manipulator to enter the accepted set by employing an untruthful strategy, for each possible assignment. This is an instance of a computer-generated classically strategy-proof configuration which was previously unknown.

The case of $m = 2$ is extremely instructive. This is a case for which each reviewer is asked to 'nominate' one of their two proposals assigned. With $m = 2$ we can see an example of a system that is weakly safe (Table 5), as without the knowledge of how PIs have been assigned the only strategy available has a strictly worse expected utility to being truthful. Note however that unlike when $m = 4$, this case is not also strongly safe; being untruthful can potentially result in success, but without knowledge of assignment it will more likely hinder than help, so would not rationally

$n = 6, m = 4, k = 3$						
Strategies	**A**	**B**	**C**	**D**	**E**	**F**
(4; 3; 1; 2)	0	0	0	0.0302	0	0
(2; 4; 3; 1)	0	0	-0.0755	0	0	0
(3; 4; 1; 2)	0	0	-0.0302	0.0302	0	0
(1; 2; 4; 3)	0	0	0	0.0755	0	0
(3; 1; 4; 2)	0	0	-0.0302	0.0302	0	0
(4; 1; 3; 2)	0	0	0	0.0302	0	0
(1; 4; 3; 2)	0	0	0	0.0302	0	0
(3; 2; 4; 1)	0	0	-0.0302	0	0	0
(3; 4; 2; 1)	0	0	-0.0302	0	0	0
(1; 3; 4; 2)	0	0	0	0.0302	0	0
(2; 4; 1; 3)	0	0	-0.0755	0.0755	0	0
(1; 4; 2; 3)	0	0	0	0.0755	0	0
(2; 3; 4; 1)	0	0	-0.0755	0	0	0
(3; 2; 1; 4)	0	0	-0.0302	0	0	0
(2; 1; 3; 4)	0	0	-0.0755	0	0	0
(4; 1; 2; 3)	0	0	0	0.0755	0	0
(2; 3; 1; 4)	0	0	-0.0755	0	0	0
(3; 1; 2; 4)	0	0	-0.0302	0	0	0
(4; 2; 1; 3)	0	0	0	0.0755	0	0
(2; 1; 4; 3)	0	0	-0.0755	0.0755	0	0

Table 3: An instructive unsafe configuration. Although the average expected utility for employing an untruthful strategy is 0, there exist strategies that yield a higher expected utility than the truthful one.

$n = 6, m = 4, k = 4$						
Strategies	**A**	**B**	**C**	**D**	**E**	**F**
(4; 3; 1; 2)	0	0	0	-0.0679	0	0
(3; 4; 1; 2)	0	0	0	-0.2151	0	0
(3; 1; 4; 2)	0	0	0	-0.2151	0	0
(4; 2; 3; 1)	0	0	0	-0.034	0	0
(4; 1; 3; 2)	0	0	0	-0.034	0	0
(3; 2; 4; 1)	0	0	0	-0.2151	0	0
(3; 4; 2; 1)	0	0	0	-0.2151	0	0
(4; 3; 2; 1)	0	0	0	-0.0679	0	0
(1; 3; 4; 2)	0	0	0	-0.034	0	0
(2; 3; 4; 1)	0	0	0	-0.034	0	0
(3; 2; 1; 4)	0	0	0	-0.2151	0	0
(4; 1; 2; 3)	0	0	0	-0.034	0	0
(2; 3; 1; 4)	0	0	0	-0.034	0	0
(3; 1; 2; 4)	0	0	0	-0.2151	0	0
(4; 2; 1; 3)	0	0	0	-0.034	0	0
(1; 3; 2; 4)	0	0	0	-0.034	0	0

Table 4: A configuration in which no strategy is ever strictly better than the truthful strategy, making it at least weakly safe. This case also happens to be strongly safe.

$n = 6, m = 2, k = 4$						
Strategy	**A**	**B**	**C**	**D**	**E**	**F**
(2; 1)	0	-0.008	-0.048	-0.122	-0.067	0

Table 5: This case is weakly safe as, even with knowledge of position, the average expected utility for the untruthful strategy is 0.

$n = 7, m = 2, k = 4$							
Strategy	**A**	**B**	**C**	**D**	**E**	**F**	**G**
(2; 1)	0	0	-0.003	-0.036	0.031	-0.005	0

Table 6: This case is very weakly safe as, without knowledge of position, the average expected utility for the untruthful strategy is 0.

be taken.

With $n = 7$, $m = 2$ and $k = 4$, as shown in Table 6, we have an example of a setup that is very weakly safe; on average across all positions, the only untruthful strategy available offers a worse expected utility than the truthful strategy. It is however not weakly safe, in that should a manipulator determine their likely position in the ground truth ordering to be fifth (position E), the untruthful strategy would offer a positive expected utility relative to the honest.

As can be observed, in some of the above tables there is a clear symmetry in the expected utilities (compare utilities between Tables 2 and 4, or between the utilities for positions C and D in Table 3). Our conjecture is that main factor determining such symmetry is the fact players' strategy spaces are permutations (see e.g., [1]).

What is key in determining systems' safety is that PIs are unaware of the full assignment and ground truth ordering. Without knowledge of this, as long as a system is weakly safe, they cannot ensure if a given strategy will be beneficial compared to being truthful. Even more so if they could also be kept unaware of k, n or if n, m and k were large enough values, making the system more complex and utility harder to predict.

In Table 7 we can see how the different values for m and k affect how safe we can consider a system when $n = 6$. As previously stated, when $m = n - 1$, no matter the value of k it will always be strongly safe. For other values of m the safety rating depends on k, though note in some cases, such as for $m \geq 3$, strongly, weakly and very weakly safe conditions may exactly coincide.

We also include similar findings for $n = 7$ for the completed cases of $m = 2$ and $m = 6$ in Table 8.

Table 7: Safety ratings of k for different m where $n = 6$

Safety	$m = 2$	$m = 3$	$m = 4$	$m = 5$
Strongly Safe	$k \geq 5$	$k \geq 5$	$k \geq 4$	$k \geq 1$
Weakly Safe	$k \geq 4$	$k \geq 5$	$k \geq 4$	$k \geq 1$
Very Weakly Safe	$k \geq 3$	$k \geq 5$	$k \geq 4$	$k \geq 1$
Unsafe	$k \leq 2$	$k \leq 4$	$k \leq 3$	None

Table 8: Safety ratings of k where $n = 7$

Safety	$m = 2$	$m = 6$
Strongly Safe	$k \geq 6$	$k \geq 0$
Weakly Safe	$k \geq 5$	$k \geq 0$
Very Weakly Safe	$k \geq 4$	$k \geq 0$
Unsafe	$k \leq 3$	None

6 Conclusion

We have studied the effect of manipulation in the Borda peer reviewing settings, taking the uncertainty of the potential manipulators into account. We have used computer-aided analysis to show novel cases of strategy proof configurations, but also gone beyond the classical approach and shown that in a number of cases manipulation can be a threat in theory but not in practice.

As our results show, which strategy is effective for a potential manipulator depends hugely on the system's chosen configuration, including the number of participants, how many each review, how many will be accepted, the distribution of submissions' quality and the potential manipulator's position in this distribution.

A number of further research directions are worth considering. First, the programme is expected to resolve the $n = 7$ case fully. With these results, we can further fine-grain the systems' classification and increase our understanding of manipulation in small committees. Second, generalisations of our findings can be used to construct safe peer reviewing systems beyond small committees. This can be obtained by clustering larger populations into small subgroups, for which safety properties have been already discovered. Third, the findings can be an input for formal analysis, in particular the mathematical understanding of the symmetry in some tables and link it to the use permutations in the players' strategy space ([1]). Finally, the research can be directed towards the analysis of more complex cases, such as the presence of multiple manipulators, or of potential collusion, and of peer reviewing with partial knowledge, e.g., manipulation over specific assignments or the effect of probabilistic beliefs rather than uniform distributions.

References

[1] M. Amstrong. *Groups and Symmetry*. Springer, 1988.

[2] H. Aziz, O. Lev, N. Mattei, J. S. Rosenschein, and T. Walsh. Strategyproof peer selection: Mechanisms, analyses, and experiments. In Schuurmans and Wellman [12], pages 397–403.

[3] B. Ballobás. The asymptotic number of unlabelled regular graphs. *J. London Math. Soc.*, 26(2):201–206, 1982.

[4] J. Douceur. Paper rating vs. paper ranking. *Operating Systems Review*, 43:117121, 2009.

[5] G. A. Hazelrigg. Dear colleague letter: Information to principal investigators (pis) planning to sub- mit proposals to the sensors and sensing systems (sss) program october 1, 2013, deadline. Retrieved on July 01, 2016, 2013.

[6] R. Holzman and H. Moulin. Impartial nominations for a prize. *Astronomy and Geophysics*, 81(1):173196, 2013.

[7] D. Kurokawa, O. Lev, J. Morgenstern, and A. D. Procaccia. Impartial peer review. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, (IJCAI-2015)*, pages 582–588, 2015.

[8] R. Meir, O. Lev, and J. S. Rosenschein. A local-dominance theory of voting equilibria. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, EC '14, pages 313–330, New York, NY, USA, 2014. ACM.

[9] M. Merrifield and D. Saari. Telescope time without tears: a distributed approach to peer review. *Astronomy and Geophysics*, 50(4):26, 2009.

[10] J. Mervis. Want a grant? first review someone else's proposal. *Science*, 2017.

[11] A. Procaccia. Nsf (actually) reviewing by social choice. Retrieved on July 01, 2016, 2013.

[12] D. Schuurmans and M. P. Wellman, editors. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA*. AAAI Press, 2016.

[13] T. Walsh. The peerrank method for peer assessment. In T. Schaub, G. Friedrich, and B. O'Sullivan, editors, *ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014)*, volume 263 of *Frontiers in Artificial Intelligence and Applications*, pages 909–914. IOS Press, 2014.