# A Two-Stage Mechanism for Ordinal Peer Assessment

Zhize Li, Le Zhang, Zhixuan Fang, and Jian Li(✉)

Institute for Interdisciplinary Information Sciences,
Tsinghua University, Beijing, China
{zz-li14,le-zhang12,fzx13}@mails.tsinghua.edu.cn,
lijian83@mail.tsinghua.edu.cn

**Abstract.** Peer assessment is a major method for evaluating the performance of employee, accessing the contributions of individuals within a group, making social decisions and many other scenarios. The idea is to ask the individuals of the same group to assess the performance of the others. Scores or rankings are then determined based on these evaluations. However, peer assessment can be biased and manipulated, especially when there is a conflict of interests. In this paper, we consider the problem of eliciting the underlying ordering (i.e. ground truth) of $n$ strategic agents with respect to their performances, e.g., quality of work, contributions, scores, etc. We first prove that there is no deterministic mechanism which obtains the underlying ordering in dominant-strategy implementation. Then, we propose a *Two-Stage Mechanism* in which truth-telling is the *unique* strict Nash equilibrium yielding the underlying ordering. Moreover, we prove that our two-stage mechanism is asymptotically optimal, since it only needs $n+1$ queries and we prove an $\Omega(n)$ lower bound on query complexity for any mechanism. Finally, we conduct experiments on several scenarios to demonstrate that the proposed two-stage mechanism is robust.

**Keywords:** Mechanism design · Peer assessment · Nash equilibrium

## 1 Introduction

Peer assessment is a commonly adopted solution for group evaluation without an independent arbiter, e.g., MOOC student assignments scoring [16], research proposal evaluation etc. Despite the pervasive success of peer assessment, there remain issues and controversies, especially on validity and reliability of peer review [21]. As the score of a participant is decided by the assessments given by others, one may manipulate the outcome by providing dishonest feedback. For example, students in a MOOC course usually conduct peer assessments by grading others' homeworks (e.g., percentile scores), and their scores are based on the average of all submitted assessments on their homeworks. In such cases, a student may be able to obtain a better ranking by giving worse evaluations of other students.


X. Deng (Ed.): SAGT 2018, LNCS 11059, pp. 176–188, 2018.
https://doi.org/10.1007/978-3-319-99660-8_16

Not surprisingly, professionals also suffer from unreliable, or "lottery-like" peer review results [30], e.g., irresponsible or derogatory comments appear in academic proposals, and even double-blind review cannot guarantee fairness [18]. In business or academic fields, it is almost inevitable that the reviews have an undisclosed conflict of interests. Under these circumstances, the fairness of peer assessments, even from experts, should be questioned [22,28].

The focus of this paper is to reveal the underlying ordering (ground truth) of the strategic agents. This work is mainly motivated by the applications in which the agents have a strong incentive to manipulate the system by not telling the truth. For example, a direct application is to rank the contributions in a relatively small team, where team members work collaboratively on a project and have a common opinion of the ranking of each member's contribution. Note that the bonus of each employee is indeed assessed by his/her team members in some companies. Thus the leader needs to know the ranking of all members. In this paper, we propose a two-stage mechanism to reveal the ground truth.

Previous work on the problem related to peer assessment, in particular to peer review, studied different setting in which their goal is to select a 'reasonable' aggregated ordering (or a subset) by partition-selection steps (see e.g., [2,10]). For example, although the mechanism given in [2,3] is strategyproof, it is not guaranteed to reveal the *true* underlying ordering even if all the agents share the same opinion of the ordering. Basically their algorithm divides the agents into disjoint clusters. Then the agents in one cluster give evaluations for the agents in other clusters. With these evaluations, the algorithm gets a kind of *value* for each cluster. Finally the top $k$ agents are drawn from the clusters with proportion to their values. Consider a simple example, where $n = 4, k = 1$ and all agents hold the same ordering $\langle 1, 2, 3, 4 \rangle$. Their algorithm divides the agents into two clusters (e.g., $C_1 = \{1, 2\}$ and $C_2 = \{3, 4\}$), and then both clusters get the same value 0.5 according to the Borda score adopted in their algorithm. Finally, their allocation algorithm selects an agent with the highest score in $C_1$ with probability 0.5 and selects an agent with the highest score in $C_2$ with probability 0.5, i.e., there is a 0.5 probability to return the wrong agent 3.

### 1.1 Our Contributions

In this paper, we make the following technical contributions:

1. We prove that there is no deterministic mechanism which obtains the underlying ordering in a dominant-strategy implementation.
2. Although mechanisms can be designed in which Nash equilibria exist, they do not guarantee to obtain the underlying ordering. Under a reasonable assumption, we show that there is a mechanism in which truth-telling is the *unique* strict Nash equilibrium and it leads to the underlying ordering, except that there is an arbitrarily small probability of disorder between the last two agents. Such a disorder is proved to be inevitable if a mechanism has a strict Nash equilibrium leading to the underlying ordering.
3. We prove a lower bound of query complexity for any mechanism, which indicates that our two-stage mechanism is asymptotically optimal.

4. The experimental results on several scenarios demonstrate that our two-stage mechanism is very robust.

### 1.2 Related Work

Extracting accurate grading results from non-strategic participants has been studied in previous work, where grading errors are treated as noise or systematic bias. Wilson [31] eliminated rater bias and error by regression. Ross et al. [27] calibrated rating bias by solving quadratic programming, while Piech et al. [24] used Gibbs sampling and expectation-maximization to infer parameters of assumed probabilistic grading models. In addition, ordinal methods have also been considered to obtain more robust ranking results instead of cardinal evaluations [19]. Raman and Joachims [25,26] used a maximum likelihood estimator based on the classic Mallows model [20] and Bradley-Terry model [6]. Mi and Yeung [23] used the probabilistic graphical models to boost the grading performance.

Although many mechanisms have been proposed to improve ranking accuracy in peer assessment, there still remains a critical challenge when the agents are strategic. Similar strategic cases have been studied in a variety of forms. Alon et al. [1], Kurokawa et al. [17] and Aziz et al. [2] considered it as a social choice or voting and designed strategyproof mechanisms. Jurca and Faltings [13,14] used monetary incentives to guarantee that truthful reporting is a Nash equilibrium. Gao et al. [11] also used rewards to incentivize truth-telling at equilibrium in peer-prediction mechanisms. Carbonara et al. [8] used a Stackelberg audit game [4,5], associating security games with punishment to incentivize honest reporting. Kahng et al. [15] designed an impartial rank aggregation rule which has a small relative error with some other (nonimpartial) rank aggregation rules. Note that this peer assessment problem is also related to crowdsourcing [7,9,32].

## 2 Model and Results

### 2.1 Model

The problem of ordinal peer assessment is formally defined as follows. Let $A = \{1, 2, \ldots, n\}$ denote the set of strategic agents. Let $r_i$ denote the ranking of agent $i$ with respect to its performance (e.g. contribution, score, etc.). W.l.o.g. assume that $r_i = i$ for all $i \in A$, i.e., the underlying ordering is $\langle 1, 2, \ldots, n\rangle$. The underlying ordering (ground truth) is a private common information shared among these $n$ agents. The problem is that a third party, who is not aware of the underlying ordering, wants to obtain it via a mechanism by adopting peer assessment. A mechanism will output an ordering (i.e., a permutation) of $A$ by asking some queries to the agents. We emphasize that our goal is to reveal the *exact* underlying ordering, not a 'reasonable' aggregated ordering (or a subset) as the previous work studied. Thus the strategic agents should share the same underlying ordering, but they may report untruthful answers. Now, we formally define the query operation and the mechanism as follows.

**Definition 1 (Query $q$).** *$q(i, A')$ : Ask agent $i$ to report the best agent in $A'$, where $A' \subseteq A$.*

Note that this defined query is sufficient to obtain the necessary information, e.g., the pairwise comparison query is such a query when $|A'| = 2$.

**Definition 2 (Mechanism $\mathcal{M}$).** *$\mathcal{M} : (A, \mathcal{Q}) \mapsto \mathcal{P}$, where $A$ denotes the set of all strategic agents, $\mathcal{Q}$ is the set of any sequence of queries asked by $\mathcal{M}$ and $\mathcal{P}$ denotes the outcome of $\mathcal{M}$ which is the set of all permutations of $A$. $\mathcal{M}$ outputs an ordering $O \in \mathcal{P}$ according to the sequence of queries $Q \in \mathcal{Q}$ and the corresponding reported answers to $Q$.*

**Definition 3 (Deterministic/Randomized Mechanism).** *$\mathcal{M}$ is a deterministic mechanism if it always outputs a deterministic ordering for a given $Q$ and its corresponding reported answers. $\mathcal{M}$ is a randomized mechanism if it may output a randomized ordering according to a distribution over $\mathcal{P}$.*

Naturally, mechanisms using fewer queries are more efficient. Now, we describe the actions of the agents. Every agent is self-interested, only caring about its own ranking in the output ordering of $\mathcal{M}$. We define the *payoff* of an agent to be its (expected) ranking in the outcome of $\mathcal{M}$ (if $\mathcal{M}$ is randomized). Their strategies are to report the answers for the queries asked by $\mathcal{M}$. Note that they may report untruthful answers.

**Definition 4 (Strict NE).** *A strategy profile is a strict Nash equilibrium if no agent can unilaterally switch to another strategy without reducing its payoff.*

### 2.2 Main Results

Our goal is to reveal the ground truth, i.e., obtain the underlying ordering. The first thing coming to mind is to design deterministic mechanisms in which all agents have dominant strategies. A *dominant strategy* means that it always achieves the best payoff no matter what the other agents do. Unfortunately, this is impossible, as shown in the following theorem.

**Theorem 1 (Impossibility Theorem).** *There is no deterministic mechanism which obtains the underlying ordering in a dominant-strategy implementation.*

Note that the Gibbard–Satterthwaite theorem [12,29] does not apply here since the condition is not satisfied, e.g., the preference relation induced by the utility function is *not* antisymmetric since two outcomes having ranked an agent at the same position have no difference for the agent. Also note that our theorem is different from [3] since they need randomization to guarantee the number of selected agents is *exact $k$*. The proof of our impossibility theorem can be found in the full version of this paper.

According to Theorem 1, the deterministic dominant-strategy implementation is too stringent to be achievable. Nonetheless, one can still obtain that the agents reporting the truth (truth-telling) is a Nash equilibrium, as in the following lemma.

**Lemma 1.** *There is a mechanism in which truth-telling is a Nash equilibrium.*

*Proof.* One such mechanism is the naive dictatorship, i.e., the mechanism returns a predefined ordering whatever the agents answer to the queries. Truth-telling obviously is a Nash equilibrium since no agent can affect its payoff by changing its strategy (answer). Note that this mechanism only outputs the true underlying ordering with a very small probability $1/n!$. □

According to the above lemma, although truth-telling and Nash equilibrium are easy to achieve, we still do not get the underlying order yet. Then, we provide a simple mechanism (randomly choose three agents from $A$, then let them report the best agent and take the majority answer (if it does not exist, we uniformly pick one from $A$) as the best agent $a^*$, finally let $a^*$ report the remaining ordering of $A - \{a^*\}$ to obtain a complete ordering (output)) and it is not hard to verify that truth-telling is a Nash equilibrium and this Nash equilibrium leads to the underlying ordering (the details can be found in our full version). However, there are many Nash equilibria in this simple mechanism, e.g., three random agents all report the second-best agent or other agents. Thus this mechanism will output wrong orderings with high probability since most Nash equilibria are bad as they lead to wrong orderings.

Hence, we want to consider the *strict* Nash equilibrium (more extremely, *unique*) that yields the underlying ordering since the strict Nash equilibrium (see Definition 4) implies stability. However, the following lemma indicates a negative result.

**Lemma 2.** *The underlying ordering cannot be obtained with probability 1 by any strict Nash equilibrium of any mechanism.*

*Proof.* Assume that there is a strict Nash equilibrium which leads to the underlying ordering with probability 1. Consider the last agent $n$, the expected ranking of agent $n$ is $n$ in this strict Nash equilibrium. However, if it changes its strategy, it does not get a strictly worse ranking than $n$ as the lowest ranking is $n$ anyway. This contradicts the definition of strict Nash equilibrium. □

Fortunately, if a tiny disorder between agents $n$ and $n-1$ is allowed, a strict Nash equilibrium is still possible. The point is that in the proof above, agent $n$ always ranks the last. As agent $n$ has no lower place to go down, the intuition is to let agent $n$ have a tiny probability to go upward. Besides, we need the following reasonable assumption to get a simple enough mechanism.

**Assumption 1.** *If the ranking of an agent is fixed in the outcome, then the agent will report truthfully thereafter.*

We propose a two-stage mechanism (Algorithm 1) in Sect. 3, satisfying the following main theorem. We analyze this mechanism and provide the proof sketch in Sect. 3. The detailed proof can be found in our full version.

**Theorem 2 (Main Theorem).** *Under Assumption 1, there is a mechanism in which truth-telling is the unique strict Nash equilibrium and it leads to the underlying ordering, except for an arbitrarily small probability of disorder between the last two agents. Further, the number of queries asked by the mechanism is* $n+1$.

Moreover, the proposed two-stage mechanism is asymptotically optimal as indicated by the following lower bound theorem (the proof can be found in our full version).

**Theorem 3 (Lower Bound).** *Any mechanism capable of retrieving the underlying ordering requires* $\Omega(n)$ *queries in the worst case.*

## 3 The Two-Stage Mechanism

Note that the problem is trivial if $n = 1$. When $n = 2$, it is impossible to distinguish the two agents. For $n > 3$, we propose the *Two-Stage Mechanism* which is described in Algorithm 1. We defer $n = 3$ to the end of this section. In Stage 1 of the two-stage mechanism, we use *self-loop* to denote the reported answer when an agent reports itself, and let *irrational answer* denote the answer when an agent $i$ reports another agent $j$ while agent $j$ does not report itself.

**Algorithm 1** Two-Stage Mechanism

**Input:** $A = \{1, 2, \ldots, n\}$, $n > 3$;
**Output:** An ordering (permutation) of $A$;

**Stage** 1 :
Randomly choose three agents from $A$, denoted as $i, j$ and $k$.
Ask $q_1(i, A)$, $q_2(j, A)$ and $q_3(k, A)$.
**if** At least two of them report themselves **then**
  Uniformly pick an agent from $A - \{i, j, k\}$ as $a^*$.
**else**
  Create a multi-set $C = \{i', j', k'\}$, where $i'$, $j'$ and $k'$ denote the answers reported by $i$, $j$ and $k$ respectively.
  Remove irrational answers from $C$ (e.g. $i' = i$, $j' = i$, and $k' = j$, $C$ changes from $\{i, i, j\}$ to $\{i, i\}$. $k' = j$ is an irrational answer since agent $j$ reporting $i$ does not report itself.).
  If there exists $o$ in $C$, where $o$ denotes an agent except $\{i, j, k\}$, then remove self-loop answers from $C$ (e.g. $i' = i$, $j' = o$, and $k' = j$, $C$ changes from $\{i, o, j\}$ to $\{i, o\}$ in the previous step, and again to $\{o\}$ in this step.).
  Randomly select $a^*$ from $C$ (candidate set) if $C \neq \emptyset$, otherwise uniformly pick from $A - \{i, j, k\}$.
**end if**
**Stage** 2 :
Let $a^*$ be ranked the first position and then let it report the remaining ordering of $A - \{a^*\}$, i.e., asking $q_4(a^*, A - \{a^*\})$, $q_5(a^*, A - \{a^*, \tilde{q}_4\})$ etc., where $\tilde{q}_i$ denotes the answer of $q_i$. Now we obtain a complete ordering $O$. Denote the ranking of agent $x$ as $\bar{r}_x$.
If the following answers were reported in Stage 1: $i' = i, j' = i, k' = j$: 

, then swap $k$ with $k+1$ if $k < n$ in $O$, i.e. punish $k$ to one position later.
Denote the last two agents in $O$ as $x$ and $y$. Slightly perturb $\bar{r}_y$ to $n-1$ (and hence $\bar{r}_x$ to $n$) with probability $\epsilon$ to obtain the final ordering $\widetilde{O}$ if $y$ but not $x$ is in $\{i, j, k\}$ and $i' = j' = k'$, where $\epsilon \in (0, 1)$ is arbitrarily small.
**return** $\widetilde{O}$.

**Intuition:** When an agent $i$ is asked to report the best agent ($a^*$) in Stage 1, the agent only has three possible answers described below, where the last two are wrong answers:

1. report $a^*$ (true answer);
2. report itself (self-loop answer);
3. or, report someone else (mostly irrational answers).

To enforce truth-telling strategy of all agents, we will ignore the wrong answers (Case 2 and 3) and punish the misreporting agents in our mechanism. Concretely, for the irrational answers, we ignore the answers in Line 8 of Algorithm 1 and punish the agents in Line 14. For the self-loops, we ignore the answers in Line 9 and indirectly punish the agents in Line 5. Note that there is an exception for agent $a^*$, i.e., $a^*$ reporting $a^*$ (itself) also belongs to Case 2. But this answer should only be considered as the Case 1 true answer.

**Analysis:** Now we analyze the two-stage mechanism in more details before moving to the proofs. The best agent $a^*$ is selected in Stage 1. This process can be distinguished as two modes:

1. Mode A: Uniformly pick
   (a) At least two self-loops (i.e. reporting themselves):
   Uniformly pick one from $A - \{i, j, k\}$. We denote this case as Case (1a), i.e., sub-item (a) in item 1. Similar denotations are used for other cases.
   (b) Candidate set $C = \emptyset$:
   Uniformly pick one from $A - \{i, j, k\}$. This case happens *exactly* when $i', j', k' \in \{i, j, k\}$, and no self-loop exists, i.e. all three answers are irrational.
2. Mode B: Pick candidates (Randomly select from $C$.)
   (a) $i', j', k' \in \{i, j, k\}$ and exactly one self-loop exists:
   W.l.o.g., let $i' = i$. $C$ must only contain $i$. Because agents $j$ and $k$ do not report themselves, those answers reporting $j$ or $k$ are irrational. Consequently an agent not reporting itself will not be contained in $C$.
   (b) There exists $o$ in $C$. Recall that $o$ denotes an agent except $\{i, j, k\}$:
   $C$ contains these $o$-type answers, and if it happens that a self-loop exists with another agent also reporting the "self-loop" agent, $C$ will contain the self-loop, e.g. $i' = o$, $j' = j$, $k' = j$, $C = \{o, j\}$.

The intuition is that Mode B usually allows us to select the correct agent $a^*$, while Mode A usually is a bad case but rarely happens. Besides, note that Line 14 and 15 of Algorithm 1 do not affect the ranking of $a^*$, i.e., $a^*$ (selected in Stage 1) always ranks the first in the outcome $\widetilde{O}$. This ensures that true answers to the queries of Stage 2 can be obtained according to Assumption 1.

Now, we move to the proof sketch part. The proof details can be found in our full version. To show truth-telling is the unique strict Nash equilibrium, it should at least be a strict Nash equilibrium as stated by the following lemma.

**Lemma 3.** *When $n > 3$, in Stage 1, the strategy profile consisting of the chosen agents reporting 1 (truth-telling) is a strict Nash equilibrium.*

Then, to show the uniqueness in the following theorem, we only need to show there are no other strict Nash equilibria. We prove this by contradiction, i.e., we show that any agent reporting 1 (true answer) is not strictly worse than others, and thus there is no other strict Nash equilibrium except truth-telling, otherwise the agent can change its answer to 1 without reducing its payoff which contradicts the definition of strict Nash equilibrium (Definition 4). Note that the number of queries is $(n+1)$ since Stage 1 uses 3 queries and Stage 2 using $n-2$ queries is sufficient since only one agent is remained.

**Theorem 4.** *When $n > 3$, the two-stage mechanism uses $n+1$ queries and yields that truth-telling is the unique strict Nash equilibrium and it leads to the underlying ordering, except that there is an arbitrarily small probability of disorder between the last two agents.*

The following corollary easily follows from the proof of Theorem 4 which is provided in our full version.

**Corollary 1.** *The only remaining strategy in Stage 1 for each agent is truth-telling using the iterated elimination of dominated strategies (IEDS) process from agent 1 to agent $n$.*

### 3.1 Two-Stage Mechanism for $n = 3$

For the $n = 3$ case, this situation follows the same paradigm as Algorithm 1. But note that formerly *Mode A uniformly pick* works perfectly when $n > 3$, but for $n = 3$, there is no agent for $A-\{i, j, k\}$. Thus, a slight difference is that we might need to identify the second-best agent, rather than the best agent, i.e. select the second-best agent, and fix its ranking as 2 in the outcome. This update obeys the "spirit" of Assumption 1.

We describe the selection process in Stage 1 by cases. Let $s$ denote the number of "self-loop" agents, i.e. the agents reporting themselves.

1. $s = 3$. Uniformly pick $a^*$ from $\{1, 2, 3\}$. We denote this case as Case ①.
2. $s = 2$. Select the only non-self-loop as the second-best agent.
3. $s = 1$. Select the only self-loop as $a^*$. Only two special sub-cases need *extra* processing.
   (a) The same as Line 14 of Algorithm 1, i.e. swap $k$ and $k+1$ if $k < 3$.
   (b) It is similar to Line 15 of Algorithm 1, while here $x \notin \{i, j, k\}$ is not required. The slight perturbation has be done if three agents report the same one.
4. $s = 0$. The same as Case ①.

The only slight modifications (differ from $n > 3$ case) are that the algorithm now uniformly picks from $\{i, j, k\}$ instead of from $A - \{i, j, k\}$, and the circumstance for *two self-loops* is tackled differently (i.e., identify the second-best agent now).

We have a similar result for this $n = 3$ case, as stated in the following lemma. The proof is not very hard and we provide it in our full version.

**Lemma 4.** *When $n = 3$, the two-stage mechanism with a slight difference uses $n + 1$ queries and yields that truth-telling is the unique strict Nash equilibrium and it leads to the underlying ordering, except that there is an arbitrarily small probability of disorder between the last two agents.*

Now, the main theorem, i.e. Theorem 2, easily follows from Theorem 4 and Lemma 4.

## 4 Experiments

In this section, we conduct experiments on several scenarios to show our two-stage mechanism is very robust. Recall truth-telling is the unique strict Nash equilibrium in our mechanism, and the *strict* Nash equilibrium implies stability. Intuitively, agents who adopt non-truth-telling strategy will eventually find it more beneficial to report the truth. The reason is that at least agent $a^*$ (the best agent) would like to adopt truth-telling. Consequently, other agents are more or less forced to be honest according to the mechanism. Hence the strategy of non-truth-telling naturally converges to the unique truth-telling equilibrium.

Concretely, we simulate the situation where there are many rounds for the agents to switch their strategies. Initially, a portion of agents is set to hold untruthful answers (randomly chosen). Let the $\mathsf{noise}$ factor denote the initial fraction of misreporting agents, hence equivalently, the initial *truthful ratio* is $1-\mathsf{noise}$. Experiments are conducted on the number of agents $n = 10, 40, 70, 100$ and $\mathsf{noise} = 0.1, 0.5, 1$. In each round, all agents sequentially switch their answers to their *best responses* with respect to the two-stage mechanism, under the condition of other agents keeping their answers unchanged. Thus the switching process in each round consists of $n$ iterations, i.e., in each iteration, exactly one agent switches to its best response. To compute the best response for an agent, as the two-stage mechanism is randomized, this is approximated by enumerating all answers of that agent and computing the average payoff for every enumerated answer (by running the algorithm 10000 times) and then chose the highest average payoff one.

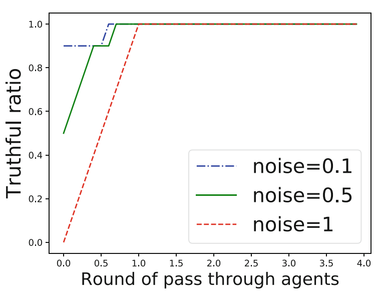


**Fig. 1.** $n = 10$

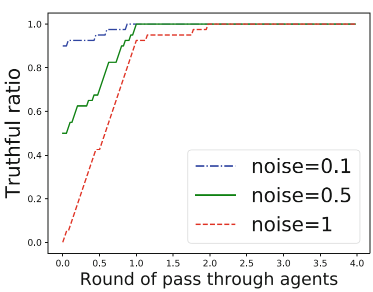


**Fig. 2.** $n = 40$

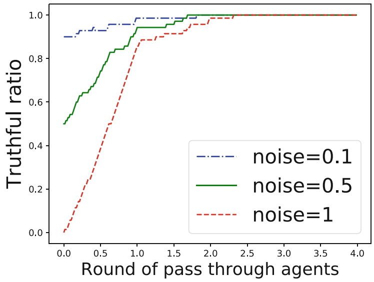


**Fig. 3.** $n = 70$

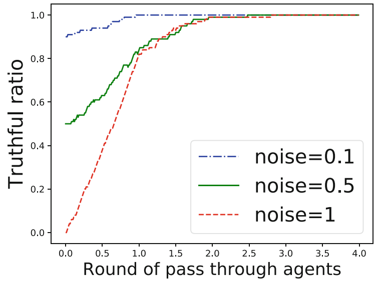


**Fig. 4.** $n = 100$

Figures 1, 2, 3 and 4 demonstrate the process of the strategies converging to the unique truth-telling equilibrium under different noise factors. Note that a truthful ratio equal to 1 means that all agents report true answers. The convergence processes in all figures are very similar, which means the number of agents does not affect the process much. Moreover, the total number of rounds in all figures are 4, which shows that this process converges very fast.

We use the following Figs. 5 and 6 to compare the convergence performance (speed) with respect to the number of agents. Particularly, if the initial fraction of misreporting agents (i.e., noise) is very large (see Fig. 6), the speed with which the strategies converge to the unique truth-telling equilibrium is almost independent of the number of agents (i.e., $n$). If the noise is not that large (see Fig. 5), the speed is inversely proportional to the number of agents. Nevertheless, they all converge quickly to the unique truth-telling equilibrium within four rounds.

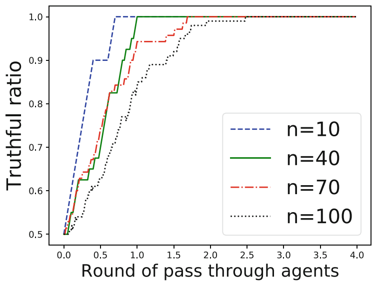


**Fig. 5.** noise $= 0.5$

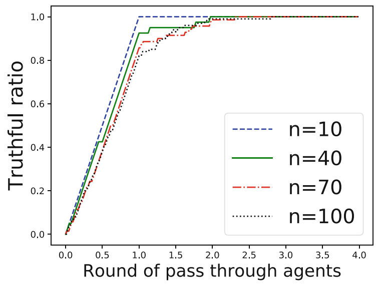


**Fig. 6.** noise $= 1$

In conclusion, the number of agents (i.e., $n$) and the initial fraction of misreporting agents (i.e., noise) both do not affect the performance much, and the strategies of all agents converge quickly to the unique truth-telling equilibrium. The experimental results validate that our two-stage mechanism is very robust.

## 5 Conclusions

In this paper, we consider the problem of obtaining the underlying ordering (ground truth) among $n$ strategic agents with respect to their performance by peer assessment. We first prove that there is no deterministic mechanism which obtains the underlying ordering in a dominant-strategy implementation. Then, we propose a two-stage mechanism in which truth-telling is the unique strict Nash equilibrium and it leads to the underlying ordering, except that there is an arbitrarily small probability of disorder between the last two agents. Note that such a disorder is proved to be inevitable if a mechanism has a strict Nash equilibrium leading to the underlying ordering. Moreover, our two-stage mechanism only needs $n + 1$ queries. We then prove an $\Omega(n)$ lower bound of query complexity for any mechanism, which indicates that our mechanism is asymptotically optimal. Finally, the experimental results demonstrate that the proposed query-optimal two-stage mechanism is also very robust.

**Discussions:** We discuss the applicability of our results in other query models. Theorem 1 (Impossibility Theorem) always holds for any query model since our proof is independent of the query model. For Lemma 2, it also always holds since the fact (the last-ranked agent can always cheat as we demonstrated in the proof) is irrelevant to the query model. As long as the number of possible answers to a single query is $O(n)$, Theorem 3 (Lower Bound) also holds. However, it might be violated if one allows a very powerful query model, e.g. requiring an agent to answer all the information it knows in just one query (the possible answers to a single query can be $n!$ in the worst case). For the extension, it is interesting to consider the situation of collusion.

**Acknowledgments.** This research is supported in part by the National Basic Research Program of China Grant 2015CB358700, the National Natural Science Foundation of China Grant 61772297, 61632016, 61761146003, and a grant from Microsoft Research Asia. The authors would like to thank Pingzhong Tang and the anonymous reviewers for their valuable comments.

## References

1. Alon, N., Fischer, F., Procaccia, A., Tennenholtz, M.: Sum of us: strategyproof selection from the selectors. In: Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge, pp. 101–110. ACM (2011)
2. Aziz, H., Lev, O., Mattei, N., Rosenschein, J.S., Walsh, T.: Strategyproof peer selection: mechanisms, analyses, and experiments. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI) (2016)
3. Aziz, H., Lev, O., Mattei, N., Rosenschein, J.S., Walsh, T.: Strategyproof peer selection using randomization, partitioning, and apportionment. arXiv preprint arXiv:1604.03632 (2016)
4. Blocki, J., Christin, N., Datta, A., Procaccia, A.D., Sinha, A.: Audit games. In: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI), pp. 41–47 (2013)

5. Blocki, J., Christin, N., Datta, A., Procaccia, A.D., Sinha, A.: Audit games with multiple defender resources. In: Proceedings of AAAI, pp. 791–797 (2015)
6. Bradley, R.A., Terry, M.E.: Rank analysis of incomplete block designs: I. The method of paired comparisons. Biometrika **39**(3/4), 324–345 (1952)
7. Cao, W., Li, J., Tao, Y., Li, Z.: On top-k selection in multi-armed bandits and hidden bipartite graphs. In: Advances in Neural Information Processing Systems (NIPS), pp. 1036–1044 (2015)
8. Carbonara, A.U., Datta, A., Sinha, A., Zick, Y.: Incentivizing peer grading in MOOCs: an audit game approach. In: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI) (2015)
9. Chai, C., Li, G., Li, J., Deng, D., Feng, J.: Cost-effective crowdsourced entity resolution: a partial-order approach. In: Proceedings of the 2016 International Conference on Management of Data, pp. 969–984. ACM (2016)
10. Fischer, F., Klimm, M.: Optimal impartial selection. In: Proceedings of the Fifteenth ACM Conference on Economics and Computation (EC), pp. 803–820. ACM (2014)
11. Gao, A., Wright, J.R., Leyton-Brown, K.: Incentivizing evaluation via limited access to ground truth: peer-prediction makes things worse. arXiv preprint arXiv:1606.07042 (2016)
12. Gibbard, A.: Manipulation of voting schemes: a general result. Econ. J. Econ. Soc. **41**, 587–601 (1973)
13. Jurca, R., Faltings, B.: Enforcing truthful strategies in incentive compatible reputation mechanisms. In: Deng, X., Ye, Y. (eds.) WINE 2005. LNCS, vol. 3828, pp. 268–277. Springer, Heidelberg (2005). https://doi.org/10.1007/11600930_26
14. Jurca, R., Faltings, B.: Mechanisms for making crowds truthful. J. Artif. Intell. Res. **34**, 209–253 (2009)
15. Kahng, A., Kotturi, Y., Kulkarni, C., Kurokawa, D., Procaccia, A.D.: Ranking wily people who rank each other. Technical report (2017)
16. Kulkarni, C., Wei, K.P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., Koller, D., Klemmer, S.R.: Peer and self assessment in massive online classes. In: Plattner, H., Meinel, C., Leifer, L. (eds.) Design Thinking Research, pp. 131–168. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-06823-7_9
17. Kurokawa, D., Lev, O., Morgenstern, J., Procaccia, A.D.: Impartial peer review. In: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI) (2015)
18. Lee, C.J., Sugimoto, C.R., Zhang, G., Cronin, B.: Bias in peer review. J. Am. Soc. Inf. Sci. Technol. **64**(1), 2–17 (2013)
19. Liu, T.Y.: Learning to rank for information retrieval. Found. Trends Inf. Retr. **3**(3), 225–331 (2009)
20. Mallows, C.L.: Non-null ranking models. I. Biometrika **44**(1/2), 114–130 (1957)
21. Marsh, H.W., Jayasinghe, U.W., Bond, N.W.: Improving the peer-review process for grant applications: reliability, validity, bias, and generalizability. Am. Psychol. **63**(3), 160 (2008)
22. Merrifield, M.R., Saari, D.G.: Telescope time without tears: a distributed approach to peer review. Astron. Geophys. **50**(4), 4–16 (2009)
23. Mi, F., Yeung, D.Y.: Probabilistic graphical models for boosting cardinal and ordinal peer grading in moocs. In: Proceedings of AAAI, pp. 454–460 (2015)
24. Piech, C., Huang, J., Chen, Z., Do, C., Ng, A., Koller, D.: Tuned models of peer assessment in MOOCs. In: Educational Data Mining 2013 (2013)

25. Raman, K., Joachims, T.: Methods for ordinal peer grading. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1037–1046. ACM (2014)
26. Raman, K., Joachims, T.: Bayesian ordinal peer grading. In: Proceedings of the Second (2015) ACM Conference on Learning @ Scale, pp. 149–156. ACM (2015)
27. Roos, M., Rothe, J., Scheuermann, B.: How to calibrate the scores of biased reviewers by quadratic programming. In: Proceedings of AAAI (2011)
28. Saidman, L.J.: Unresolved issues relating to peer review, industry support of research, and conflict of interest. Anesthesiology **80**(3), 491 (1994)
29. Satterthwaite, M.A.: Strategy-proofness and arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. J. Econ. Theory **10**(2), 187–217 (1975)
30. Smith, R.: Peer review: a flawed process at the heart of science and journals. J. R. Soc. Med. **99**(4), 178–182 (2006)
31. Wilson, H.G.: Parameter estimation for peer grading under incomplete design. Educ. Psychol. Meas. **48**(1), 69–81 (1988)
32. Zhou, Y., Chen, X., Li, J.: Optimal PAC multiple arm identification with applications to crowdsourcing. In: International Conference on Machine Learning (ICML), pp. 217–225 (2014)