Poisoning Attack against Estimating from Pairwise Comparisons

Ke Ma, Member, IEEE, Qianqian Xu, Senior Member, IEEE, Jinshan Zeng, Xiaochun Cao, Senior Member, IEEE, and Qingming Huang, Fellow, IEEE

Abstract—As pairwise ranking becomes broadly employed for elections, sports competitions, recommendation, information retrieval and so on, attackers have strong motivation and incentives to manipulate or disrupt the ranking list. They could inject malicious comparisons into the training data to fool the target ranking algorithm. Such a technique is called “poisoning attack” in regression and classification tasks. In this paper, to the best of our knowledge, we initiate the first systematic investigation of data poisoning attack on the pairwise ranking algorithms, which can be generally formalized as the dynamic and static games between the ranker and the attacker, and can be modeled as certain kinds of integer programming problems mathematically. To break the computational hurdle of the underlying integer programming problems, we reformulate them into the distributionally robust optimization (DRO) problems, which are computational tractable. Based on such DRO formulations, we propose two efficient poisoning attack algorithms and establish the associated theoretical guarantees including the existence of Nash equilibrium and the generalization ability bounds. The effectiveness of the suggested poisoning attack strategies is demonstrated by a series of toy simulations and several real data experiments. These experimental results show that the proposed methods can significantly reduce the performance of the ranker in the sense that the correlation between the true ranking list and the aggregated results with toxic data can be decreased dramatically.

Index Terms—Adversarial Learning, Poisoning Attack, Pairwise Comparison, Rank Aggregation, Robust Game, Distributionally Robust Optimization.

1 INTRODUCTION

Rank aggregation, in particular estimating a ranking based on comparisons between pairs of objects, arises in a variety of disciplines, including the social choice theory [3], psychology [16], statistics [34], machine learning [39], bioinformatics [57] and others. The convenience of these rank aggregation methods relies on their utilization of the ordinal data. Without features, the comparisons only contain

- K. Ma is with the School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China, and with the Artificial Intelligence Research Center, Peng Cheng Laboratory, Shenzhen 518055, China. E-mail: name@ucas.ac.cn
- Q. Xu is with the Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China. E-mail: qianqian.xu@ict.ac.cn, xuxianjian@ict.ac.cn.
- J. Zeng is with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang, Jiangxi, 330022, China. E-mail: jishanzeng@jxnu.edu.cn, jsh.zeng@gmail.com
- X. Cao is with the State Key Laboratory of Information Security (SKLOIS), Institute of Information Engineering, Chinese Academy of Sciences, Beijing, 100083, China, and with the School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China. E-mail: caoxiaochun@ie.ac.cn
- Q. Huang is with the Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100049, China, and with the School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China, and with the Key Laboratory of Big Data Mining and Knowledge Management, the School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100049, China, and with the Artificial Intelligence Research Center, Peng Cheng Laboratory, Shenzhen 518055, China. E-mail: qnhuang@ucas.ac.cn.

However, the existing work ignores the security issue. Beyond statistical property and computational complexity, situations become complicated when the pairwise ranking algorithms are utilized in high-stakes applications, e.g. elections, sports competitions, and recommendation. In pursuit of huge economic benefits, the potential attackers have strong motivations and incentives to manipulate or disrupt the aggregated results. When the victims are ranking algorithms, a profit-oriented adversary could try his/her best to manipulate or disrupt the ranking list which will favor his/her demands-say, the attacker could place the special object at the top of the recommendation list, help the particular candidate to win an election or just defeat the candidate who should have won the election. If the attackers compromise the integrity of ranking results, the fairness and rationality will be lost in these high-stakes applications. Un-
Unfortunately, to the best of our knowledge, adversarial poisoning attack on pairwise comparison has not been comprehensively examined yet. Can rank aggregation algorithms with pairwise comparisons be easily manipulated or disrupted? How reliable are their results in the high-stakes applications?

To the best of our knowledge, the adversarial arsenal for pairwise ranking methods has never been seriously studied. On one hand, the pairwise comparisons are the most simple data in the literature as just binary variables can represent them. Due to the absence of features, modifying these binary data is an easy job. On the other hand, any single comparison does not dictate the aggregated result. Even manipulating a small quantity of binary data could not affect the final global ranking. Such a contradiction inspires us to initiate an adversarial investigation of pairwise ranking problem.

To execute the attack strategy in the scenario, the adversary must analyze the characteristics of pairwise ranking problems. Unlike the supervised learning tasks (e.g., regression, classification, multi-arm bandit and reinforcement learning), the rank aggregation does not need the test protocol. This means that the evasion attacks (a.k.a. adversarial examples) are not realistic. Evasion attack causes the fixed model to misbehave by well-crafted test data. But there is no test phrase to implement such a kind of attack. To archive his/her goal, the adversary needs to inject the manipulated data into the training data. Thus, rank aggregation in an adversarial setting is inherently related to the challenging poisoning attacks. Next, the adversary should consider the discrete property of the pairwise comparisons. Unlike the data consisting of features in continuous space, the input of pairwise ranking only consists of binary data. The adversary could only add, delete or flip pairwise comparisons to execute the poisoning attacks. Such limitations make the substantial attack operations on pairwise ranking even harder. How to design efficient algorithms that are able to inject toxic data in a discrete domain? It is the distinguishable characteristic of our work which is different with the existing poisoning attack approaches.

Given these challenges, we propose a principle framework for adversarial perturbations of pairwise comparisons that aims to break the integrity of rank aggregation result. In particular, we focus on the parametric model solved by maximum likelihood estimation. We make the following contributions:

- Different statistical attacks corresponding to the dynamic and static threat models are formulated into the bi-level optimization problem and distributionally robust optimization problem. In the bi-level optimization problem, we adopt $\chi^2$ divergence to describe the uncertainty set around the original data. The optimal attack strategy can be obtained by the projection onto a simplex. In the distributionally robust optimization problem, the uncertainty set is a Wasserstein ball. Based on the strong duality, the optimal attack behavior is obtained by a least square problem with a special regularization.
- We prove the existence of robust optimization equilibrium and establish a minimax framework for pairwise ranking under adversarial setting.

To the best of our knowledge, this is the first systematic study of attacking rank aggregation under different adversarial models. The extensively evaluations are conducted on several datasets from different high-stake domains, including election, crowdsourcing, and recommendation. Our experiments demonstrate that the proposed poisoning attack could significantly decrease the correlation between the true ranking list and the aggregated result.

Notations

Let $V$ be a finite set. We will adopt the following notation from combinatorics:

$$\binom{V}{k} := \text{set of all } k \text{ element subset of } V.$$

In particular $\binom{V}{2}$ would be the set of all unordered pairs of elements of $V$. The sets of ordered pair will be denoted $V \times V$. Ordered and unordered pairs will be delimited by parentheses $(i,j)$ and braces $\{i,j\}$ respectively. We will use positive integers to indicate alternatives and voters. Henceforth, $V$ will always be the set $\{n\} = \{1, \ldots, n\}$ and will denote a set of alternatives to be ranked. $\mathcal{U} = \{1, \ldots, m\}$ will denote a set of voters. For $i, j \in V$, we write $i \succ j$ to mean that alternative $i$ is preferred over alternative $j$. If we wish to emphasize the preference judgment of a particular voter $u \in \mathcal{U}$, we will write $i \succ_u j$. Suppose that $\Omega \subset \mathbb{R}^n$ is the data space, we denote $\langle \cdot, \cdot \rangle$ as a metric space equipped with some metric $d : \Omega \times \Omega \to \mathbb{R}$.

2 Ranking with Pairwise Comparisons

Given a collection $V$ of $n$ alternatives, we suppose that each $i \in V$ has a certain numeric quality score $\theta^*_i$. We represent the quality scores of $V$ as a vector $\theta^* \in \mathbb{R}^n$. Suppose that a comparison of any pair $[i, j] \in \binom{V}{2}$ is generated via the comparison of the corresponding scores $\theta^*_i, \theta^*_j$ in the presence of noise. Let $y^*_{ij}$ be the true direction of a pair $[i, j]$ as

$$y^*_{ij} = \begin{cases} 1, & \theta^*_i > \theta^*_j, \\ -1, & \theta^*_i < \theta^*_j. \end{cases}$$

Let $C$ be a collection of $N$ pairwise comparisons

$$C = \{c = [i, j] \mid y_{ij} = 1, \ i, j \in V, \ i \neq j\},$$

and $y_{ij}$ is the label of pair $[i, j]$ which could not be consist with $y^*_{ij}$. It is worth noting that $C$ is always a multi-set. For any pair $[i, j]$, it could be labeled by multiple users. Given a
set of voter $\mathcal{U} = \{u_1, \ldots, u_m\}$, let $y_{ij}^{u}$ be the judgment of pair $[i, j]$ given by voter $u \in \mathcal{U}$. We can aggregate $y_{ij}^{u_1}, \ldots, y_{ij}^{u_m}$ into a weight $w_{ij}^0$. Define $w(i, j, u)$ as the indicator of $y_{ij}^u$:

$$w(i, j, u) = \begin{cases} 1, & \text{if } y_{ij}^u = 1, \ u \in \mathcal{U} \\ 0, & \text{otherwise} \end{cases}$$

and the weight $w_{ij}^0$ of $y_{ij}$ is

$$w_{ij}^0 = \sum_{u \in \mathcal{U}} w(i, j, u).$$

Moreover, we introduce the comparison matrix $A$. If there exists a comparison $c \in C$, it can be described by its label $y_{ij}$ and a row of $A$ as $a^c = \{a^c_1, \ldots, a^c_{|C|}\}$:

$$a^c_k = \begin{cases} 1, & k = i \\ -1, & k = j \\ 0, & \text{otherwise} \end{cases}$$

Then the data of pairwise ranking problem can be represented by $C_U = \{A, y, w_0\}$ where $w_0 = \{w_{ij}^0\}, y = \{y_{ij}\}$ is a $(n(n-1))/2$-d binary vector.

In statistical ranking or estimation from pairwise comparison, our goal is to obtain a score vector $\theta$ to minimize a loss of a global ranking on the given data $C_U$.

$$\hat{\theta} \in \arg \min_{\theta} \ell(\theta; C_U).$$

In particular, let the estimation of $y_{ij}$ be

$$\hat{y}_{ij} = \text{sgn}\left( (a^c, \theta) + \varepsilon_c \right), \ \forall c \in C,$n

where $\text{sgn}(\cdot)$ is the sign function, $\varepsilon_c$ is the independent and identically distributed (i.i.d) noise variable and has a cumulative distribution function (c.d.f) $F$. Actually, (6) minimizes the derivation between the observed label $y$ and its estimation $\hat{y} = \{\hat{y}_{ij}\}$ based on the observing data $C_U$. In addition, the random variable $\varepsilon_c$ plays the role of a noise parameter, with a higher magnitude of $\varepsilon_c$ leading to more uncertainty in the comparisons and the higher probability of sign inconsistency occurred between $y_{ij}$ and $\theta_i - \theta_j$. The event that object $i$ dominating object $j$ ($y_{ij} = 1$) is generally independent of the order of the two items being compared, thus, the following holds:

$$\Pr(y_{ij} = 1) = 1 - \Pr(y_{ij} = -1)$$

and $F$ is a symmetric c.d.f whose continuous inverse is well-defined. Some typical examples of (7) are the uniform model [63], the Bradley-Terry- Luce (BTL) model [14], [46], and the Thurstone model with Gaussian noise (Case V) [67], which have been extensively studied in literature (e.g., [17], [28]). In this paper, we focus on the Uniform Model: one can adopt the symmetric c.d.f $F(t) = \frac{t+1}{2}$, and the general set-up (7) turns to be a linear model. Furthermore, the loss function in (6) can be specialized as the weighted sum-of-squares function:

$$\ell(\theta; C_U) = \frac{1}{2|C_U|} \| y - A\theta \|_2^2 w_0 = \frac{1}{2|C_U|} \sum_{(i, j)} w_{ij}^0 (y_{ij} - \theta_i + \theta_j)^2.$$ 

### 3.1 Poisoning Attack on Pairwise Ranking

We provide here a detailed adversarial framework for poisoning attacks against pairwise ranking algorithms. The framework consists of defining the adversary’s goal, knowledge of the attacked method, and capability of manipulating the pairwise data, to eventually define the optimal poisoning attack strategies.

**The Goal of Adversary.** If an adversary executes the poisoning attack, he/she will provide the ranker with the toxic data. This action will mislead its opponent into picking parameters to generate a different ranking result from $\theta$ obtained by the original data $C_U$ in (6). Let $\theta$ be the solution of (6) with the toxic data, it satisfies

$$d(\pi_{\theta^*}, \pi_\theta) \leq d(\pi_{\theta^*}, \pi_\theta),$$

where $\theta^*$ is the true quality scores of the objects, $\pi_\theta$ is the ranking list decided by $\theta$ and $d(\pi_1, \pi_2)$ measures the similarity of two ordered lists $\pi_1$ and $\pi_2$.

**The Knowledge of Adversary.** We assume two distinct attack scenarios which are distinguished by the knowledge of adversary, referred to as dynamic and static attacks in the following. The adversaries in the two scenarios have different knowledge of the victims.

- In **dynamic** attacks, the attacker is assumed to know the observed data $C_U$, the ranking algorithm, and even the ranking parameters $\theta$ obtained by the original data $C_U$ in (6). If a dictator wants to sabotage the election which will subvert his/her predominant, he/she would not need to manipulate the results of the election. Making the most competitive opponent lose the advantage in the key districts will achieve the purpose. The dictator could execute the dynamic strategies as the aggregation process is a “white-box” to him/her. This adversarial mechanism can be implemented by establishing the hierarchical relationship between the ranker and the attacker. The attacker is assumed to anticipate the reactions of the ranker; this allows him/her to choose the best—or optimal—strategy accordingly. Such a hierarchical interaction results in the fact that the mathematical program related to the ranking process is part of the adversary’s constraints. It is also known as the dynamic or Stackelberg (leader-follower) game [7] in the literature: the two agents take their actions in a sequential (or repeated) manner. Moreover, the hierarchical relationship is the major feature of bilevel optimization. The bi-level program includes two mathematical programs within a single instance, one of these problems being part of the constraints of the other one.
• In static attacks, the attacker could not grasp \( \hat{\theta} \) but is still aware of the observed data \( C_U \) and the ranking algorithm. This scenario comes from the fact that the ranking aggregation problem does not need the test protocol. Once the adversary provides the modified data, the victim would generate the ranking list immediately. There is no chance to monitor the ranker’s behavior. In most cases, the adversary can not obtain \( \theta \). There is no feedback for the adversary to update his/her strategies. A competitor of the e-commerce platform, who wants to disrupt the recommendation results and destroy the user experience, would execute the static strategies. Promoting the rank of specific goods is challenging. Disrupting the normal ranking result is sufficient to archive his/her purpose. The competitor could only execute the static strategies as the aggregation process is a “gray-box”. The leading e-commerce platform is the only one who could access the ranking parameters. The objective function and the pairwise comparisons for recommendation can be perceivable to the adversary. This adversarial mechanism should be modeled as a static game. A static game is one in which a single decision is made by each player, and each player has no knowledge of the decision made by the other players before making their own decision. In other words, decisions or actions are made simultaneously (or the order is irrelevant).

The Capability of Adversary. To modify the original data \( C_U \) in poisoning attacks, the adversary will inject an arbitrary pair \( [i, j] \in (V)^2 \) with any directions into \( C_U \), delete the existing comparison \( c = (i, j) \in C_U \) or just flip the label of \( c \). The three kinds of operations require some new representations of the observed set. We augment the observed data \( C_U \) with the comparisons which are not labeled by users in \( U \). Let \( D = V \times V \) be the set of all ordered pairs, and \( |D| = N = n(n - 1) \). The weights of all possible comparisons are \( w' \) and there exist \( 0 \) entries in \( w' \).

As \( D \) is the complete comparison set, the comparison matrix \( B \) will be fixed and we can adopt a \( n(n-1) \times d \) single-value vector to represent the labels, saying that \( y' \) is a vector with all entries are \( 1 \). Now all attack operations (adding, deleting and flipping) can be executed by increasing or decreasing the corresponding weight \( w' \).

\[
B = \begin{bmatrix} b_{1,2} \\ b_{1,3} \\ \vdots \\ b_{n,n-2} \\ b_{n,n-1} \end{bmatrix} \subset \{-1, 0, 1\}^{N \times n}, \tag{11}
\]

\[
y' = \begin{bmatrix} y_{1,2}, y_{2,1}, \ldots, y_{n-1,n}, y_{n,n-1} \end{bmatrix}^\top = \begin{bmatrix} 1, 1, \ldots, 1, 1 \end{bmatrix}^\top.
\]

Besides injecting the toxic data, the attacker also needs to disguise himself/herself. It means that the adversary needs to coordinate a poisoned \( w = \{w_{ij}\} \) associated with \( w' \). Intuitively, the adversary could not obtain \( w \) through the drastic changes, neither on each \( w_{ij} \) nor \( \sum_{i,j} w_{ij} \). Such limitations lead to the following constraints for the adversary’s action. First, the total difference between \( w_0 \) and \( w \) would be smaller than \( b \), namely,

\[
\| w - w_0 \|_1 \leq b, \ b \in \mathbb{N}_+.
\]  

(12)

Here the positive integer \( b \) bounds the total number of malicious samples thereby limiting the capabilities of the attacker. Furthermore, the adversary could not alter the number of votes on each pairwise comparison \( c \in D \) obviously. This constraint on the adversary leads to the following condition:

\[
\| w - w_0' \|_\infty \leq l, \ l \in \mathbb{N}_+, \ l \leq \min(\{\max(w_0'), b\}).
\]

(13)

The positive integer \( l \) leads the conservative perturbations on the observed samples. To summarize, the adversary’s action set \( \Omega_1 \) is

\[
\Omega_1 = \left\{ w \left| \begin{array}{l} w \in \mathbb{N}^N, \ l, b \in \mathbb{N}, \\
\| w - w_0' \|_1 \leq b, \\
\| w - w_0' \|_\infty \leq l,
\end{array} \right. \right\}, \tag{14}
\]

Furthermore, the attacker must pay for his/her malicious behaviors. Let \( s : \mathbb{N}^N \times \mathbb{N}^N \to \mathbb{R} \) is a “cost” function measured the overhead of the perturbation as changing \( w_0 \) into \( w \). The attacker hopes that the toxic weight \( w \) will represent the lowest cost option. Let \( \Omega_2 \) be the bound set of the adversary

\[
\Omega_2 = \left\{ w \left| w \in \operatorname{arg} \min_{w, w_0'} s(w, w_0') \right. \right\}.
\]

(15)

Finally, the action set is \( \Omega_0 = \Omega_1 \cap \Omega_2 \), which figures out the capability of the adversary.

Poisoning Attack Strategies. Here we specify the different poisoning strategies for the two attack scenarios.

• Dynamic attack strategy. Consider the goal and knowledge of attacker, we formulate the interaction between ranker and the adversary with full knowledge as a dynamic game. In this game, information is assumed to be complete (i.e., the players’ payoffs functions, as well as the constraint set \( \Omega_{l0} \) and the flexible set of ranking parameter \( \Theta \) are common knowledge) and perfect (i.e., the attacker knows the ranker’s decision). Having received the ranker’s decision \( \hat{\theta} \), the attacker chooses a feasible decision \( w \in \Omega_0 \) that maximizes the ranker’s loss function to increase the risk of the ranker’s estimation based on \( \{w, B, y\} \). Such a dynamic game can be formulated into the following bi-level optimization problem:

\[
\max_{w \in \Omega_0} \ell(w; \hat{\theta}, B, y'), \tag{16a}
\]

subject to \( \hat{\theta} \in \operatorname{arg} \min_{\theta \in \Theta} \ell(\theta; w_0', B, y'). \tag{16b}
\]

The upper level optimization (16a) amounts to selecting the toxic data \( w \) to maximize the loss function of the ranker, while the lower level optimization (16b) corresponds to calculate the ranking parameter \( \hat{\theta} \) with original data \( \{w_0, B, y\} \). Once the adversary generates \( w \), he/she will deliver the toxic data to the ranker. Then the poisoned parameter \( \hat{\theta} \) will be obtained by

\[
\hat{\theta} = \operatorname{arg} \min_{\theta \in \Theta} \ell(\theta; w, B, y').
\]

(17)
• **Static attack strategy.** This strategy is represented such a type of adversary whose ability is to inflict the highest possible risk of the ranker when no information about his/her interests is available. It means that the two players make decisions simultaneously, and the attacker does not knows the ranker’s decision. Such a static game can be formulated into the following min-max optimization problem:

\[
\min_{\theta} \max_{w \in \Omega_0} \ell(\theta, w; B, y').
\]

The poisoned parameter \(\theta\) will be solved by (17).

However, solving the dynamic and static attack strategies from (16) and (18) are challenging. On one hand, the bi-level optimization (16) and the min-max problem (18) are both mixed-integer programming problem as the variable \(w\) is restricted to be positive integers. On the other hand, the feasible set \(\Omega_0\) corresponds to a non-linear constraint as it requires to find the perturbation in the neighborhood of \(w_0\) with the lowest cost. It is well-known that linear integer programings are NP-complete problems [36]. Such a non-linear constraint makes these problems even more complex. Obviously, adopting the heuristic methods to solve the optimal attack strategies (16) and (18) is sub-optimal. In this part, we will develop the other model based on idea from distributionally robust optimization [12], [22], [54] that provides the tractable convex formulations for solving the optimal strategies in the dynamic and static scenarios.

### 3.2 Distributional Perspective and Robust Game

In the above formulations (16) and (18), the attacker modifies the number of votes on each pairwise comparison with constraints \(\Omega_0\). This formulation leads to the mixed-integer programming problem. Here we introduce a distributional perspective to establish the tractable optimization problem. Generally speaking, the attacker and the ranker both access the original data \(D\) to play the dynamic or static game. The non-toxic pairwise comparisons \(D = \{w_0, B, y'\}\) are actually drawn from an empirical distribution \(P_N\)

\[
P_N = \frac{1}{N} \sum_{e \in E} \delta(w_{ij}^0, b_{ij}, y_{ij}), \quad c = (i, j),
\]

where \(\delta(w_{ij}^0, b_{ij}, y_{ij})\) is the Dirac probability measure on \(\{w_{ij}^0, b_{ij}, y_{ij}\}\). With \(B\) and \(y'\) as (11), the marginal distribution of \(w_0\) plays a vital role in the sequel. With some abuse of symbol, we treat the marginal distribution of \(w_0\) as the distribution of the original data and

\[
P_N = \frac{1}{N} \sum_{e \in E} \delta(w_{ij}^0), \quad c = (i, j).
\]

The attacker chooses a perturbation function \(\psi: \mathbb{N}^N \rightarrow \mathbb{N}^N\) that changes the weight \(w_0^i\) to \(w \in \Omega_0\). Such a perturbation \(\psi\) induces a transition from the empirical distribution \(P_N\) to a poisoned distribution \(Q\). If the attacker selects \(Q\) in a sufficiently small neighborhood of \(P_N\), namely, the “distance” between the poisoned distribution \(Q\) and the empirical distribution \(P\) would be sufficiently small, the attacker could obtain a “local” solution and \(Q\) is a “good” approximation of \(P_N\) in the sense of such a “distance”. Therefore, the poisoned sample \(w\) would satisfy the constraints (12) and (13). Here we directly work with the empirical distribution \(P_N\) (or other nominal distribution) and consider \(Q\) is close to the nominal distribution in terms of a certain statistical distance.

There exists some popular choices of the statistical distance, such as \(\phi\)-divergences [8], [9], [18], [32], [53], [54], [75], Prokhorov metric [21], Wasserstein distances [11], [23], [40], [51], [77] and maximum mean discrepancy [66].

For dynamic attack strategy (15), we adopt the \(\phi\)-divergence [41] as the discrepancy measure between the empirical distribution \(P_n\) and the toxic distribution \(Q\).

**Definition 1** (\(\phi\)-divergence and \(\chi^2\)-divergence). Let \(\phi: \mathbb{R}_+ \rightarrow \mathbb{R}\) be a convex function with \(\phi(1) = 0\). Then the \(\phi\)-divergence between distributions \(Q\) and \(P\) defined on a measurable space \(X\) is

\[
d\phi(Q \parallel P) = \int \phi \left( \frac{dQ}{dP} \right) dP = \int_X \phi \left( \frac{q(x)}{p(x)} \right) p(x) dq(x),
\]

where \(\mu\) is a \(\sigma\)-finite measure on \(X\) satisfying \(Q\) and \(P\) are absolutely continuous with respect to \(\mu\), and \(q = \frac{dQ}{d\mu}\), \(p = \frac{dP}{d\mu}\) are the Radon–Nikodym derivative with respect to \(\mu\). If \(\phi\) is adopted as \(\phi(t) = \frac{1}{2} (t - 1)^2\), it is known as the \(\chi^2\)-divergence.

Suppose that \(X(\mathbb{P}_N)\) is a set of probability distributions from the empirical distribution with \(\chi^2\)-divergence. This \(\chi^2\) ball with radius \(\alpha\) is given by

\[
X^\alpha(\mathbb{P}_N) = \left\{ Q \in \mathbb{P}(\Omega_1) \left| d_{\chi^2}(Q \parallel \mathbb{P}_N) \leq \alpha \right. \right\},
\]

where \(\mathbb{P}(\Omega_1)\) denotes the set of all Borel probability measures on \(\Omega_1\). With carefully chosen \(\alpha\), the adversary chooses \(w\) from the toxic distribution \(Q \in X^\alpha(\mathbb{P}_N)\). \(w\) could satisfy the neighborhood constraints as (12) and (13). Replacing the minimal ‘cost’ constraint (5) by the neighborhood constraint defined with the \(\chi^2\) ball, we formulate the following bi-level optimization to obtain the dynamic attack strategy

\[
\max_{Q \in X^\alpha(\mathbb{P}_N)} \mathbb{E}_{w \sim Q} [\ell(w; \hat{\theta})],
\]

subject to \(\hat{\theta} = \arg \min_{\theta} \ell(\theta; w').\)

The \(\chi^2\)-divergence and the “local” neighborhood constraint \(Q \in X^\alpha(\mathbb{P}_N)\) will help us to develop a tractable algorithm for the dynamic attack strategy.

Different with the dynamic attack strategy, the ranking parameter \(\theta\) would be unknown for the adversary in the static attack strategy. The \(\chi^2\) divergence will not help to simplify the min-max problem (18). To sum up, we adopt the \(p\)-Wasserstein distance [22] as the discrepancy measure between the empirical distribution \(P_n\) and the toxic distribution \(Q\) for the static attack strategy. The \(p\)-Wasserstein distance will help us to reformulate the min-max problem (18) into a single regularized problem.

**Definition 2** (\(p\)-Wasserstein distance). Let \(p \in [1, \infty]\). The \(p\)-Wasserstein distance between distributions \(P, Q \in \mathbb{P}(\Omega)\) is defined as

\[
W_p(P, Q) = \left( \min_{\gamma \in \Pi(P, Q)} \left\{ \int_{\Omega \times \Omega} d(w, w')^p \gamma(dw, dw') \right\} \right)^{\frac{1}{p}}
\]
\[ p = \infty \]
\[ \mathcal{W}_p(\mathbb{P}, \mathbb{Q}) = \inf_{\gamma \in \Gamma(\mathbb{P}, \mathbb{Q})} \sup_{\Omega \times \Omega} \gamma\text{-ess sup } d(w, w') \] (22)

where \( \Gamma(\mathbb{P}, \mathbb{Q}) \) denotes the set of all Borel probability distributions on \( \Omega \times \Omega \) with marginal distributions \( \mathbb{P} \) and \( \mathbb{Q} \), \( d : \Omega \times \Omega \to \mathbb{R}_+ \) is a nonnegative function, and \( \gamma\text{-ess sup} \) expresses the supremum of \( d(\cdot, \cdot) \) with respect to the measure \( \gamma \).

The Wasserstein distance (21) and (22) arise in the problem of optimal transport (22), (23): for any coupling \( \gamma \in \Gamma(\mathbb{P}, \mathbb{Q}) \), the conditional distribution \( \gamma_{w|w'} \) can be viewed as a randomized overhead for ‘transporting’ a unit quantity of some material from a random location \( w \sim \mathbb{P} \) to another location \( w' \sim \mathbb{Q} \). If the cost of transportation from \( w \in \Omega \) to \( w' \in \Omega \) is given by \( d(w, w') \), \( \mathcal{W}_p(\mathbb{P}, \mathbb{Q}) \) will be the minimum expected transport cost (60).

Suppose that \( \mathcal{W}_p(\mathbb{P}_N) \) is a set of probability distributions constructed from the empirical distribution \( \mathbb{P}_N \) with \( p \)-Wasserstein distance. This Wasserstein ball of radius \( \alpha \) is given by

\[ \mathcal{W}_p^\alpha(\mathbb{P}_N) = \left\{ Q \in \mathcal{P}(\Omega) \mid \mathcal{W}_p(\mathbb{P}_N, Q) \leq \alpha \right\}. \] (23)

With local uncertainty set \( \mathcal{W}_p^\alpha(\mathbb{P}_N) \), the min-max optimization (15) could be expressed as the following distributionally robust optimization (DRO) problem:

\[ \min_{\varrho} \sup_{Q \in \mathcal{W}_p^\alpha(\mathbb{P}_N)} \mathbb{E}_{\varrho \sim Q} \left[ \ell(\varrho, Q) \cdot \mathbb{W}_p(\mathbb{P}_N, Q) \right], \] (24)

where the supremum operation w.r.t. \( Q \) means that all players’ optimal decision is based on the worst expected value of \( \ell \) from the set of distributions \( \mathcal{W}_p^\alpha(\mathbb{P}_N) \). Here we replace the minimal ‘cost’ constraint in (15) by the neighborhood constraint on the worst-case expectation. With the local constraint \( Q \in \mathcal{W}_p^\alpha(\mathbb{P}_N) \), the Wasserstein distance between the empirical distribution \( \mathbb{P}_N \) and the perturbed distribution \( Q \) must be smaller than a given budget \( \alpha \) as \( \mathcal{W}_p(\mathbb{P}_N, Q) \leq \alpha \). It means that the attacker has a budget \( \alpha \) to implement his/her perturbation on the original data for ranking aggregation. The robust game formulation (24) would relax the coarse-grid constraint as (14), and the analysis in the sequel reveals the central role played by this relaxation.

Actually, the bi-level problem (20) and the DRO problem (24) relate to a general robust game (1), (44), (45) between the attacker and the ranker as

\[ \min_{x_r \in X_r} \sup_{q \in \mathcal{U}} \mathbb{E}_{\xi \sim q} \left[ f_r(x_r, x_{-r}, q) \right], \quad r = 1, 2 \] (25)

where \( r \) indicates the role of the agent in the robust game, \( x_r \) is the decision variable of the special player \( r \), and \( x_{-r} \) denotes the decision variables of its rivals, and \( X_r \) is the action set of player \( r \). The random variable \( \xi \) illustrates the uncertainty or inaccuracy of distributional information to the players, and \( \mathcal{U} \) is the uncertainty set of distribution of random variable \( \xi \) for all players (i.e., \( X(\mathbb{P}_N) \) and \( \mathcal{W}_p^\alpha(\mathbb{P}_N) \)). The pay-off function \( f_r \) could be different for each player and the corresponding game is a non-zero sum game. Comparing the general case (25) with (20) and (24), all players in (24) focus on the same pay-off function as \( f_1 = f_2 = \ell \). Moreover, the decision variable of the ranker \( \theta \) equals to \( x_1 \). The random variable \( \xi \) represents the distribution of pairwise comparison as \( w \). So the decision variable of the attacker \( x_2 \) will be the constant (its role has been replaced by \( \xi \)). The robust game problem is first proposed by Bertsimas and Aghassi in (1). It expands the boundaries of research of the classical Nash game (55), (56), (73) and the Bayesian game (26), (27), (28). Different form the Nash and the Bayesian game (1), the only common knowledge of all participants in robust game is that all players being aware about an uncertainty set like \( X^\alpha(\mathbb{P}_N) \) and \( \mathcal{W}_p^\alpha(\mathbb{P}_N) \). All possible parameters of payoff function are related to this set. Here we investigate the existence of the equilibrium for distributionally robust Nash equilibrium of the proposed model (25). First, we give the definition of the distributionally robust Nash equilibrium.

**Definition 3.** A pair of different players’ action \( \{ x_1^*, x_2^* \} \) is called a distributionally robust Nash equilibrium (DRNE) of (25) if they satisfy the following

\[ x_r^* \in \arg \min_{x_r \in X_r} \sup_{q \in \mathcal{U}} \mathbb{E}_{\xi \sim q} \left[ f_r(x_r, x_{-r}, q) \right], \quad r = 1, 2. \] (26)

Next, we can prove the existence of DRNE for the general robust game (25).

**Theorem 1.** Let the pay-off function \( f_r, r = 1, 2 \) be the weighted sum-of-squared loss \( \ell(q) \) in (25). If the uncertainty set is \( X^\alpha(\mathbb{P}_N) \) or \( \mathcal{W}_p^\alpha(\mathbb{P}_N) \), the general robust game (25) has a DRNE.

To prove this existence result, we reformulate the problem (25) into a single optimization problem and show that the single problem has an optimal solution. The detailed proof is provided in the Appendix A.

### 3.3 Optimization

In this part we show our algorithms for computing the adversarial strategies. Suppose the total number of pairwise comparison without perturbation is \( M^0 \), and the frequencies of each type of the observed comparisons are

\[ p = \frac{1}{M^0} \cdot w_0^l, \quad M^0 = \sum_{(i,j)} w_0^{ij} \cdot. \] (27)

Let the maximum toxic dosage be \( \kappa \). It suggests that the number of toxic pairwise comparisons \( M \) satisfies

\[ M = \sum_{(i,j)} w_{ij} \leq (1 + \kappa) \cdot M^0. \] (28)

We replace the toxic weight \( w \) with its frequency \( q = \{ q_{ij} \} \in \mathbb{R}_+^n \) when analyzing the equilibrium, studying the statistical nature of the worst-case estimator and solving the corresponding optimization problem. We relax the integer programming problem into a general optimization by such a variable substitution. Thus, the pay-off function (9) turns to be

\[ \ell(\varrho, q) = \frac{1}{2N} \sum_{(i,j)} q_{ij} (y_{ij} - \theta_i + \theta_j)^2, \] (29)

and we still adopt \( \mathbb{P}_N \) and \( Q \) as the distribution of the empirical data and the toxic data. Furthermore, we can implement the integer attack with the optimal \( q \) and \( M \). Now we come to solve the bi-level optimization (20) and
the distributionally robust optimization problem \([24]\) with the variable substitution:

\[
\max_{\hat{Q} \in \mathcal{X}^n(P_N)} E_{\tilde{q} \sim \hat{Q}} \left[ \ell(q; \hat{\theta}) \right],
\]

\[
\text{subject to } \hat{\theta} = \arg \min_{\theta \in \Theta} \left( \ell(\theta; w'_0) \right),
\]

and

\[
\min_{\theta \in \Theta} \sup_{\hat{Q} \in \mathcal{X}^n(P_N)} E_{\tilde{q} \sim \hat{Q}} \left[ \ell(\theta; q) \right]
\]

\((31)\)

For the dynamic attack strategy \([30]\), a similar formulation has been studied for archiving a better variance-bias trade-off in maximum likelihood estimation \([34]\). Based on the \(\chi^2\)-divergence, the bi-level problem \((30)\) turns to be a convex problem. We provide a detailed process of solving \((30)\) in the supplementary material.

The distributionally robust optimization formulation \((31)\) involves optimizing over the uncertainty set \(\mathcal{W}_2^2(P_N)\), which contains joint probability measures. However, recent strong duality results of distributionally robust optimization involving Wasserstein uncertainty set \([23]\) Theorem 1 and \([12, \text{Theorem 1}]\) ensure that the inner supremum in \((31)\) admits an equivalent reformulation which would be a tractable, univariate optimization problem. In the adversarial scenario of pairwise ranking, we have the following result. The DRO problem \((31)\) could be reformulated as a regularized regression problem.

**Theorem 2.** Let \(Z = \{p, B, y', y\}'\) be the observed data set, where \(B\) and \(y'\) are defined as \([11]\). \(p\) is the frequency of each type of pairwise comparison as \([27]\). Consider the loss function of \(z\) and the distance function between \(z_c, z'_c\) are based on the \(\ell_2\)-norm. In other words, we take \(\ell(\theta, z)\) as \([29]\)

\[
d(z_c, z'_c) = \| (p_{ij}, b_{ij}, y'_ij) - (q_{ij}, b_{ij}, y_{ij}) \|_2
\]

\((32)\)

Then, the DRO problem \((31)\) has an equivalent form:

\[
\min_{\theta \in \Theta} \sup_{\hat{Q} \in \mathcal{X}^n(P_N)} E_{\tilde{q} \sim \hat{Q}} \left[ \ell(\theta; q) \right]
\]

\((33)\)

where

\[
\mathcal{L}(\theta) = \frac{1}{2N} \sum_{(i,j)} p_{ij} (y'_{ij} - \theta^T b_{ij})^2,
\]

\((34)\)

and

\[
\mathcal{R}(\theta) = \sqrt{\frac{\alpha}{4N} \sum_{(i,j)} (y'_{ij} - \theta^T b_{ij})^2}.
\]

\((35)\)

We provide a detailed proof in the Appendix \([8]\). The example of linear regression with Wasserstein distance based uncertainty sets has been considered in \([11]\). The representation for regularized linear regression in Theorem 2 can be seen as an extension of \([11]\). We adopt the weighted sum-of-squared loss and the "regularization" \((35)\) here is not the \(\ell_2\)-norm of \(\theta\). \([35]\) can be treated as a "regularization" which is the square root of the residual between \(y_{ij}\) and its estimation. It represents a 'worst' case in pairwise ranking; all possible comparisons appear and they have the same number of votes. In this case, the pairwise ranking algorithm could not generate a reasonable ranking result. The uncertainty budget \(\alpha\) play the role as the regularization parameter. As \(\alpha\) increase, the ranking scores \(\theta\) obtained by \([33]\) would come closer to the solution of \([35]\). The validity of the analysis above will be illustrated in the empirical studies.

With Theorem 2, we will have the following corollary which gives a tractable method to obtain the worst-case distribution \(q^*_w\). If we have the worst-case solution, we can solve the corresponding dual variable from the optimal value of the original DRO problem.

**Corollary 1.** For \(\lambda \geq 0\) and the weighted least square loss \([29]\), we define \(\psi : \mathbb{R}^N \rightarrow \mathbb{R}\)

\[
\psi_{\lambda}(\ell(p; \theta)) = \sup_{z' \in \mathbb{R}^{n+2}} \frac{1}{N} \sum_{(i,j)} \left\{ \ell(\theta; q_{ij}) - \lambda (p_{ij} - q_{ij})^2 \right\}
\]

\((36)\)

where

\[
\ell(\theta; q_{ij}) = \frac{q_{ij}}{2} \left( y'_{ij} - \theta_i + \theta_j \right)^2.
\]

\((37)\)

Let

\[
I_{\text{primal}} = \sup_{q \in \mathcal{W}_2^2(P_N)} E_{\tilde{z}' \sim Q} \left[ \ell(\theta; z') \right]
\]

\((38)\)

we have

\[
I_{\text{primal}} = \inf_{\lambda \geq 0} \left\{ \lambda \alpha + E_{\tilde{z}' \sim Q} \left[ \psi_{\lambda}(\ell(p; \theta)) \right] \right\}.
\]

\((39)\)

Moreover, let \(\theta^{*}_{\alpha}\) be the optimal solution of the right hand side of \((33)\) and the dual variable of \(\theta^{*}_{\alpha}\) is \(\lambda^{*}_{\alpha}\) will be a solution of \((39)\).

\[
\lambda^{*}_{\alpha} = \left\lbrack \frac{1}{16N\alpha} \cdot \sum_{(i,j)} (y'_{ij} - (\theta^{*}_{\alpha})^T b_{ij})^2 \right\rbrack.
\]

\((40)\)

The optimal static attack strategy \(q^{*}_{\alpha}\) is a solution of \((36)\) corresponding to \(\theta^{*}_{\alpha}\) and \(\lambda^{*}_{\alpha}\).

Finally, we describe the whole optimization of the static poisoning attack on pairwise ranking with Algorithm 1. First, the adversary changes the original weight \(w'_0\) into the frequency \(p\) as the initialization (line 1). By Theorem 2, the attacker could obtain the worst-case estimation \(\theta^{*}_{\alpha}\) through \((33)\) (line 2). But the attacker cannot adopt \(\theta^{*}_{\alpha}\) as the attack operation. Here we solve the dual variable \(\lambda^{*}_{\alpha}\) (line 4) to find the toxic distribution. Then the toxic distribution \(q^{*}_{\alpha}\) with uncertainty budget \(\alpha\) is obtained by Corollary 1 (line 4). With some rounding operation (line 5 & 6), the adversary prepares the poisoned data \(\{w^{*}_{\alpha}, B, y'\}\). Then the poisoned data is provided to the ranker who solved the ranking parameter by \((17)\). Then the whole poisoning process will be completed.

### 3.4 Theoretical Analysis

In this section, we come back to \((31)\) and give a couple of inequalities relating the local worst-case (or local minimax) risks and the usual statistical risks of the pairwise ranking under adversarial conditions. In the traditional paradigm
Algorithm 1: Static Poisoning Attack on Pairwise Ranking.

**Input**: the original data \( \{ w'_\alpha, B, y' \} \), the maximum toxic dosage \( \kappa \), the uncertainty budget \( \alpha \).

1. Initialize the frequency of weights \( p \) by \[^{\text{[27]}}\].
2. Obtain the worst-case ranking scoring \( \theta^*_\alpha \) under the uncertainty budget \( \alpha \):

\[
\theta^*_\alpha \in \arg \min_{\theta \in \Theta} \left( \frac{1}{N} \sum_{(i,j)} \left( y'_{ij} - \theta b_{ij} \right)^2 + \frac{1}{N} \sum_{(i,j)} p(y'_i - \theta^*_\alpha b_{ij})^2 \right).
\]

3. Calculate the optimal dual variable through \[^{\text{[40]}}\]:

\[
\lambda^*_\alpha = \sqrt{\frac{1}{N} \sum_{(i,j)} \left( y'_{ij} - \left(\theta^*_\alpha\right)^T b_{ij} \right)^2}.
\]

4. Obtain the toxic distribution \( q^*_\alpha \) corresponding to \( \theta^*_\alpha \) and \( \lambda^*_\alpha \):

\[
q^*_\alpha \in \arg \max_{q \in \mathbb{Q}} \sum_{(i,j)} \left( q_{ij} y'_{ij} - \left(\theta^*_\alpha\right)^T b_{ij} \right)^2 - \lambda^*_\alpha \left| q_{ij} - q_{ij}' \right|^2.
\]

5. Assign the toxic weights with \( q^*_\alpha \):

\[
w_\alpha = M_0 (1 + \kappa) q^*_\alpha.
\]

6. Round \( w_\alpha \) to obtain the \( w^*_\alpha \) as integer vector:

\[
w^*_\alpha = \text{rounding}(w_\alpha).
\]

**Output**: the poisoned data \( \{ w^*_\alpha, B, y' \} \).

We assume that the instance space \( Z \) is a Polish space (i.e., a complete separable metric space) with metric \( d_Z \). We denote by \( \mathcal{P}(Z) \) the space of all Borel probability measures on \( Z \), and by \( \mathcal{P}_m(Z) \) with \( m \geq 1 \) the space of all \( \mathcal{P} \in \mathcal{P}(Z) \) with finite \( m \)-th moments. The metric structure of \( Z \) can be used to define a family of metrics on the spaces \( \mathcal{P}_m(Z) \). We then define the local worst-case risk of \( \ell \) at \( \mathcal{P} \),

\[
R_{\mathcal{P},\alpha,p}(\ell) := \sup_{Q \in \mathcal{P}_p} R_Q(\ell)
\]

and the local minimax risk of \( \mathcal{P} \),

\[
R^*_{\mathcal{P},\alpha,p}(F) := \inf_{\ell \in F} R_{\mathcal{P},\alpha,p}(\ell).
\]

Next, we analyze the performance of the local minimax ERM procedure of the pairwise ranking, namely,

\[
\ell \in \arg \min_{\ell \in F} R_{\mathcal{P},\alpha,2}(\ell).
\]

**Theorem 3.** Consider the setting of pairwise ranking problem with the sum-of-squared loss, for any \( t > 0 \), it holds

\[
\Pr \left( \exists \ell \in F : R_{\mathcal{P},\alpha,2}(\ell) > \varsigma_1 \right) \leq e^{-2t^2} \quad \text{and} \quad \Pr \left( \exists \ell \in F : R_{\mathcal{P},\alpha,2}(\ell) > \varsigma_2 \right) \leq 2e^{-2t^2}
\]

where

\[
\varsigma_1 = \min_{\lambda \geq 0} \left\{ \lambda \alpha^2 + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda,\ell}(z) \right] \right\} + \frac{24 \mathcal{J}(F) + t}{\sqrt{N}},
\]

and

\[
\varsigma_2 = \min_{\lambda \geq 0} \left\{ (\lambda + 1) \alpha^2 + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda,\ell}(z) \right] + \frac{\sqrt{\log(\lambda + 1)}}{\sqrt{N}} \right\} + \frac{24 \mathcal{J}(F) + t}{\sqrt{N}}.
\]

where \( \mathcal{J}(F) \) is the Dudley’s entropy integral \[^{\text{[20]}}\] which is served as the complexity measure of the hypothesis class \( F \).

Theorem 3 is a type of data-dependent generalization bounds which is proposed for margin cost function class \[^{\text{[38], [40]}}\]. By the strong duality results, we can establish this result from the dual representation of the Wasserstein DRO problem. The detailed proof is provided in the Appendix \[^{\text{[D]}}\]. Here we note that the hypothesis selected by the minimax ERM procedure \[^{\text{[45]}}\] are uniform smoothness with respect to the underlying metric \( d_Z(\cdot, \cdot) \). Further, we have the following result. Proofs are relegated to the Appendices \[^{\text{[E]}}\].

**Theorem 4.** Consider the setting of pairwise ranking problem with the sum-of-squared loss, the following holds with probability at least \( 1 - \eta \)

\[
R_{\mathcal{P},\alpha,2}(\ell) - R^*_{\mathcal{P},\alpha,2}(F) \leq \frac{48 \mathcal{J}(F)}{\sqrt{N}} + \frac{48L \left[ \text{diam}(Z) \right]^2}{\alpha \sqrt{N}} + 3 \sqrt{\frac{\log(\frac{2}{\eta})}{2N}},
\]

where \( \text{diam}(Z) \) is the diameter of \( Z \)

\[
\text{diam}(Z) = \sup_{z, z' \in Z} d_Z(z, z').
\]
4 Experiments

In this section, four examples are exhibited with both simulated and real-world data to illustrate the validity of the proposed poisoning attack on pairwise ranking. The first example is with simulated data while the latter three exploit real-world datasets involved crowdsourcing, election and recommendation.

4.1 Simulated Study

Settings. We first validate our poisoning attack framework on simulated data. We create a random total ordering on set $V$ with $n$ candidates as the ground-truth ranking and generate the comparison matrix $B$ and the labels $y'$ as $[11]$. Next, we generate the ground-truth weight of each comparison $w_0$. Notice that the original data $\{w_0, B, y'\}$ consists of some noisy comparisons. In the simulation study, we can specify the percentage of noisy comparisons, denoted as $\rho$. We validate the proposed attack framework when $n$, $w_0$ and $\rho$ vary. Moreover, the maximum toxic dosage $\kappa$ and the uncertainty budget $\alpha$ are the hyper-parameters of the Algorithm $[1]$. Since the annotations of pairwise data are usually collected via crowdsourcing platforms where the attacker could produce hundreds of pairwise data are usually collected via crowdsourcing, we also vary $\kappa$ and $\alpha$ in our experiments. At last, there exists a rounding operator in the Algorithm $[1]$ and we explore the results of different rounding functions, e.g. ceiling, floor, and the nearest integer of each element in $w_0$.

Competitors. To the best of our knowledge, the proposed method is the first poisoning attack on pairwise ranking. To see whether our proposed method could provide efficient perturbation data for misleading the pairwise ranking algorithm, we implement the random perturbation attack (referred to as ‘Random’) and the Stackelberg or dynamic game attack (referred to as ‘Dynamic’) as the competitors.

- The random perturbation attack modifies $w_0$ as $w_{\text{random}}$ to manipulate the ranking result. The random perturbation attack generates $w_{\text{random}}$ and obeys the constraints $[12]$ and $[13]$ to hide his/her behaviors. We vary both $b$ and $l$ to explore the ability of random attack. The random perturbation data is noted as $Z_{\text{random}} = \{A, y, w_{\text{random}}\}$. We assume this attacker is also lack of prior knowledge on the true ranking. So the random perturbation attack also adopts the fixed label set $y$.

- The Stackelberg (dynamic) game attack comes from $[16]$. To execute this type of poisoning attack, the adversary would have the full knowledge of original training data $w_0$ and the corresponding relative ranking score $\hat{\theta}_{\text{original}}$. With these advantages, the adversary can adjust his/her strategies to provide the optimal malicious action with the bi-level optimization like $[16]$. Without a doubt, the adversary endues with the privilege by such a hierarchical relation. For the fair competition, we only perform one round of the leader-follower game as the other competitors. Notice that this kind of attack is also proposed by this paper. Due to the length limitation, we provide the details of this attack in the supplementary materials.

It is worth noting that the poisoning attack with dynamic game is not a practical attack method. $[16]$ is a bi-level optimization and the maximization process needs the solution of the minimization problem. In other words, the attacker must obtain the relative ranking score $\hat{\theta}$ estimated from the original training data without perturbation. This operation is much harder than injecting some modified training samples into the victim’s training set. Only the so-called “white-box” setting would satisfy its necessary requirements. As the ‘Dynamic’ method needs more exorbitant conditions, the ‘Dynamic’ method only reflects the vulnerability of ranking aggregation algorithms but can not show the superiority of the ‘Static’ method.

Evaluation Metrics. We adopt the following measures for evaluating the ranking results aggregated by the different sets of pairwise comparisons.

- Kendall $\tau$ Distance (Kendall-$\tau$). The Kendall rank correlation coefficient evaluates the degree of similarity between two sets of ranks given the same objects. This coefficient depends upon the number of inversions of pairs of objects which would be needed to transform one rank order into the other. Let $V = [n]$ be a set of $n$ candidates and $\pi_1$, $\pi_2$ are two total orders or permutations on $V$, the Kendall $\tau$ distance is defined to be

$$d_K(\pi_1, \pi_2) = \frac{2}{n(n-1)} \cdot \vartheta,$$

where

$$\vartheta = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \vartheta(\pi_1(i), \pi_1(j), \pi_2(i), \pi_2(j))$$

is the number of different pairs between these two ordered sets $\pi_1, \pi_2$ as

$$\vartheta(\pi_1(i), \pi_1(j), \pi_2(i), \pi_2(j)) = \begin{cases} 1, & \text{if } (\pi_1(i) - \pi_1(j))(\pi_2(i) - \pi_2(j)) > 0, \\ -1, & \text{if } (\pi_1(i) - \pi_1(j))(\pi_2(i) - \pi_2(j)) < 0, \\ 0, & \text{otherwise,} \end{cases}$$

and $\pi_1(i)$ represents the ranking score of the $i^{th}$ object in ranking list $\pi_1$. Kendall $\tau$ distance counts the number of pairwise mismatches between two rank orders. Then this metric considers all candidates of $V$. However, Kendall-$\tau$ ignores the importance of the top objects in a ranking list.

- Reciprocal Rank (R-Rank). The reciprocal rank is a statistic measure for evaluating any process that produces an ordered list of possible responses to a sample of queries, ordered by the probability of correctness or the ranking scores. The reciprocal rank of a rank order is the multiplicative inverse of the rank of the first correct object:

$$RR = \frac{1}{\text{rank}_i},$$

where $\text{rank}_i$ refers to the rank position of the first candidates of the ground-truth ranking in the other list.

- Precision at $K$ ($P@K$). Precision at $K$ is the proportion of the top-$K$ objects in the other rank order that are consistent with the true ranking. In this case, the precision and recall will be the same. So we do not report

 "$\ldots\$"
Comparative Results. We display the comparative results of different attack methods in Table 1. There the number of candidates ranges from 10 to 100 (n = 10, 20, 50, 100). The percentage of noisy comparisons is θ = 0 in the four cases. We let the maximum toxic dosage to be 0 as κ = 0 to verify the effectiveness of the worst-case distribution in the Wasserstein ball with uncertainty budget α. We show the attack effect of ‘Static’ and ‘Dynamic’ methods with different budgets. The performance of ‘Random’ are affected by two parameters: the percentage of the new comparisons injected into the original training set, and the percentage of the existed comparisons deleted from the original training set. Here we set these two parameters be s1 = s2 = 0.05. We obtain the following observations from Table 1.

The ‘Static’ method can decrease the Kendall-τ when the uncertainty budget α increases. Looking back on the Algorithm 1, the uncertainty budget α is the weight of the second term in (35) and the two parts of (35) have the same monotonic respect to θ. With the increasing of α, the impact of the second term (35) to the solution (33) becomes gradually. The solution of (35) means that the algorithm will adopt all possible pairwise comparisons with same number of voting to aggregate the final ordered list. There is no doubt that this case would be far away from the ground-truth ranking. If α approaches ∞, we would obtain this confusing solution. This explains the behaviors of the ‘Static’ methods when the Kendall-τ is larger than 0. In Figure 1, we see that the ‘Static’ method does two things to perturb the training set: adding pairwise comparisons which conflict with the ground-truth ranking and removing the pairwise comparisons which is consistent with the ground-truth ranking. The total amount of change enlarge when the uncertainty budget α increase. If the Kendall-τ is smaller than 0, it means that the poisoned training dataset would support an opposite ranking list. In Figure 2, each group corresponds to a poisoned data set by ‘Static’ method with a certain uncertainty budget. When the Kendall-τ is smaller than 0 (α ≥ 10−5), we observe that the number of comparisons which conflict with the ground-truth ranking is larger that the number of comparisons which is consistent with the ground-truth ranking. Such training data could generate an arbitrarily ordered list. If it happens, the Kendall-τ could not monotonically decrease when we increase the uncertainty budget continuously. Moreover, the uncertainty budget α plays a totally different role in the ‘Dynamic’ method. The existing work [18], [54] reveal that such kind of min-max problem is a new type of regularization. This regularization also carries out the ‘bias-variance’ trade-off like the classical approaches like Tikhonov regularization. In this case, the uncertainty budget α can be explained as a regularization coefficient. The Kendall-τ of ‘Dynamic’ method presents a ‘U’-type curve in our experiments.

Visualization. We visualize the ranking list in Figure 3. The visualization shows the same phenomenons as the numeric results in Table 1. As the target ranking aggregation algorithm does not emphasize the top-K results and the adversary has no prior knowledge of the ranking results, the untrustworthy results of ‘Static’ method only depend on the original data and the uncertainty budget. So the proposed method is the non-target attack for pairwise ranking.

- **Average Precision at K (AP@K).** Average precision at K is a weighted average of the precision. If the top objects in the new ranking list are consistent with the true ranking, they will contribute more than the tail objects in this metric.

- **Normalized Discounted Cumulative Gain at K (NDCG@K).** Using a graded relevance scale of objects in ranking result, discounted cumulative gain (DCG) measures the usefulness, or gain, of the objects based on its position in the order list when recovering to the true ranking. The gain is accumulated from the top to the bottom, with the gain of each result discounted at lower ranks. Compared to DCG, NDCG will be normalized by the ideal DCG.

**Fig. 1:** The amount of changed pairwise comparisons by the poisoning attack with static game. The x-axis is the index of pairwise comparisons and the y-axis is the amount of change. Note that the ranges of y-axis in each sub-figure are different.

**Fig. 2:** The number of correct pairwise comparisons and comparisons which conflict with the ground-truth ranking in the poisoned training set by ‘Static’ method.

the recall and F score for our poisoning attack method.
algorithm. Manipulating the ranking list with specific goals, a.k.a. the target attack, is the future work.

4.2 Human Age

Description. 30 images from human age dataset FGNET are annotated by a group of volunteer users on ChinaCrowds platform. The ground-truth age ranking is known to us. The annotator is presented with two images and given a binary choice of which one is older. Totally, we obtain 8,017 pairwise comparisons from 94 annotators.

Comparative Results. Notice that the real-world data has a high percentage of outliers (about 20% comparisons conflict with the correct age ranking). We observe similar phenomena as the simulation experiments. When the uncertainty budget increase, the ‘Static’ method would inject more comparisons which conflict with the true age ranking and delete the original comparisons which indicate the true ordered list. Once the ‘wrong’ samples overwhelm the ‘correct’ samples, the ranking aggregation algorithm would like to generate a reversed list. As there are only the ‘wrong’ samples in the toxic training set by ‘Static’ method, the final result could be arbitrary.

Fig. 3: The ranking generated from the original data (Original), random attack data (Random), static poisoning attack data (Static) and dynamic poisoning attack data (Dynamic).
Fig. 4: The ranking generated from the original data (Original), random perturbation data (Random), poisoned data (Static and Dynamic) on Human Age dataset. When the Kendall-τ is smaller than 0 (α ≥ 10^-4), we observe that the aggregated results would put the younger people at the top of the lists. Moreover, the same phenomenon in the simulation are still observed. The training data with more than 50% outliers could generate an arbitrarily ordered list. If it happens, the Kendall-τ could not monotonically decrease when we increase the uncertainty budget continuously for the static attack strategies.

### 4.3 Dublin Election

**Description.** The Dublin election data set contains a complete record of votes for elections held in county Meath, Dublin, Ireland on 2002. This set contains 64,081 votes over 14 candidates. These votes could be a complete or partial list over the candidate set. The ground-truth ranking of 14 candidates are based on their obtained first preference votes. The five candidates who receive the most first preference votes will be the winner of the election. We are interested in the top-5 performance of the pairwise rank aggregation method. Then these votes are converted into the pairwise comparisons. The total number of the comparisons is 652,817.

**Comparative Results.** In this experiment, we evaluate the ability of poisoning attack in election. The election result is not obtained by pairwise ranking aggregation. However, the ordered list aggregated from induced comparisons still shows positive correlation with the actual election result. Different from the manipulation or strategic voting setting in election, the adversary could control the whole votes but with some constraints. As a consequence, the poisoning attack could break the barrier of computational complexity [69, 74]. The proposed method focuses on the ‘non-target’ attack on pairwise ranking aggregation. The ‘Static’ method could perturb the ranking list generated by the original algorithm with a sufficient uncertainty budget. But the adversary is not able to manipulate the order with her/his preference as she/he can not decide the winner of election. We call the problem as the ‘target’ attack, where the adversary manipulates the order with her/his preference. Our future work will study the ‘target’ poisoning attack on pairwise ranking. Moreover, the ‘Dynamic’ method does not completely destroy the election result. It indicates that the inaccurate supervision would mislead the adversary and the corresponding Nash equilibrium could show partiality for the ranking aggregation algorithm.

### 4.4 Sushi Preference

**Description.** This dataset contains the results of a series surveys which involves 5000 individuals for their preferences about various kinds of sushi. The original survey provides 10 complete strict rank orders of 10 different kinds of sushi as 1) ebi (shrimp), 2) anago (sea eel), 3) maguro (tuna), 4) ika (squid), 5) uni (sea urchin), 6) sake (salmon roe), 7) tamago (egg), 8) toro (fatty tuna), 9) tekka-maki (tuna roll), and 10) kappa-maki (cucumber roll). The complete

---

1. http://www.preflib.org/data/election/irish/
2. https://electionsireland.org/result.cfm?election=2002&cons=178&sort=first

---

### Table 3: Comparative results of different attack methods on Dublin election data.

| Method     | Budget | Kendall-τ | Tendency | R-Rank | P@3 | AP@3 | NDCG@3 |
|------------|--------|-----------|----------|--------|-----|------|--------|
| Original   | -      | -         | -        | 1.000  |
| Random     | 0.05/0.05 | 0.4736    | -        | 1.000  |
| Static 1e-6| 10^-6  | 0.4723    | -        | 0.0769 |
| Static 1e-5| 10^-5  | 0.4723    | -        | 0.0769 |
| Static 1e-4| 10^-4  | 0.9821    | -        | 0.0769 |
| Static 1e-3| 10^-3  | -0.3864   | -        | 0.0769 |
| Static 1e-2| 10^-2  | -0.4725   | -        | 0.0769 |
| Static 1e-1| 10^-1  | -0.0330   | -        | 0.0769 |
| Dynamic 1e-6| 10^-6  | 0.4826    | -        | 0.0769 |
| Dynamic 1e-5| 10^-5  | 0.5085    | -        | 0.0769 |
| Dynamic 1e-4| 10^-4  | 0.5085    | -        | 0.0769 |
| Dynamic 1e-3| 10^-3  | 0.4723    | -        | 0.0769 |
| Dynamic 1e-2| 10^-2  | 0.5085    | -        | 0.0769 |
| Dynamic 1e-1| 10^-1  | 0.1648    | -        | 0.0769 |

### Table 4: Comparative results of different attack methods on Sushi election data.

| Method     | Budget | Kendall-τ | Tendency | R-Rank | P@5 | AP@5 | NDCG@5 |
|------------|--------|-----------|----------|--------|-----|------|--------|
| Original   | -      | -         | -        | 1.000  |
| Random     | 0.05/0.05 | 0.4736    | -        | 1.000  |
| Static 1e-6| 10^-6  | 0.4267    | -        | 0.0769 |
| Static 1e-5| 10^-5  | 0.4267    | -        | 0.0769 |
| Static 1e-4| 10^-4  | 0.4267    | -        | 0.0769 |
| Static 1e-3| 10^-3  | -0.0330   | -        | 0.0769 |
| Static 1e-2| 10^-2  | -0.0330   | -        | 0.0769 |
| Static 1e-1| 10^-1  | -0.0330   | -        | 0.0769 |
| Dynamic 1e-6| 10^-6  | 0.4222    | -        | 0.0769 |
| Dynamic 1e-5| 10^-5  | 0.4222    | -        | 0.0769 |
| Dynamic 1e-4| 10^-4  | 0.4222    | -        | 0.0769 |
| Dynamic 1e-3| 10^-3  | 0.4222    | -        | 0.0769 |
| Dynamic 1e-2| 10^-2  | 0.4222    | -        | 0.0769 |
| Dynamic 1e-1| 10^-1  | 0.4222    | -        | 0.0769 |
strict rank orders are converted into the pairwise graph by \cite{50}. We adopt the whole 221,670 comparisons and the Hodgerank \cite{34} method to aggregate a ranking list as the ground-truth. Then 20 percent of pairwise comparisons are chosen to consist of the observation set. The different attack approaches can manipulate the subset of data and induce the pairwise ranking algorithm to generate a different order list.

**Comparative Results.** This experiment is a classic setting in recommendation and computational advertisement. With the selected subset, the ranking aggregation method can produce a same ranking list as adopting with the whole preference data. The random attack would not change this list in this experiment. In addition, the ‘Dynamic’ method is trapped with the inaccurate supervision and only shows a moderate destructive effect. The ‘Static’ method could generate a promise perturbation to mislead the ranking aggregation method as the Kendall-$\tau$ would be $-1$.

### 4.5 Computational Complexity Analysis

The computational complexity of the dynamic strategy depends on the number of turns of \cite{20}. Given $n$ candidates, the complexity of the ranker is $O(n^6)$ for solving a least square problem and the complexity of the adversary is $O(n^2 \log(n^2) + \log \frac{1}{\epsilon} \cdot \log(n^2))$ where $\epsilon$ is the solution accuracy, $n^2 \log(n^2)$ is for sorting and the last part corresponds to the projection onto the $l_2$ ball. The computational complexity of the static strategy depends on the subroutines of Line 2 and Line 4 in Algorithm 1. We solve the sub-routine of Line 2 by gradient descent and evaluating the gradient needs $O(n^4)$ each time. The complexity of Line 4 is $O(n^3 + n^2 \log(n^2) + n^2)$ where $n^3$ is for the closed form, $n^2 \log(n^2)$ is for the sorting and $n^2$ for the projection onto the simplex. We also display the computational complexity comparisons on the synthetic and the real-world datasets in Table 5 and 6. The results are mean of 100 trials with different initialization.

### Table 5: Computational complexity (ms) comparisons on the synthetic dataset. The results are the mean of 100 trials with different pairwise comparisons.

| Method | Budget | No. of Candidates |
|--------|--------|-------------------|
|        | 1e-$6$ | 1e-$5$ | 1e-$4$ | 1e-$3$ | 1e-$2$ | 1e-$1$ | 1 |
| Static | 0.0736 | 0.0727 | 0.0692 | 0.0715 | 0.0712 | 0.0662 | 0.0680 |
| Dynamic| 0.0779 | 0.0738 | 0.0751 | 0.0743 | 0.0743 | 0.0717 | 0.0701 |

### Table 6: Computational complexity (ms) comparisons on the real-world datasets. The results are the mean of 100 trials with different initialization.

| Method | Budget | Age | Dublin | Sushi |
|--------|--------|-----|--------|-------|
| Static | 1.2012 | 0.1018 | 0.0548 |
| Dynamic| 1.1174 | 0.0798 | 0.0645 |

5 Conclusion

We initiate the first study of data poisoning attacks in the context of pairwise ranking. We formulate the attack problem as a robust game between two players, the ranker and the adversary. The attacker’s strategies are modeled as the distributionally robust optimization problems and some theoretical results are established, including the existence of distributionally robust Nash equilibrium and the generalization bounds. Our empirical studies show that our attack strategies significantly break the performance of pairwise ranking in the sense that the correlation between the true ranking list and the aggregated result with toxic data can be decreased dramatically.

There are many avenues for further investigation – such as, providing the finite-sample and asymptotic results characterizing the theoretical performance of the estimator with adversarial learning, extending our attacks to more pairwise ranking algorithms such as spectral ranking, and trying to attack the ranking algorithms with defense paradigm. We believe that a very interesting open question is to expand our understanding to better understand the role and capabilities of adversaries in pairwise ranking.
REFERENCES

[1] Michele Aghassi and Dimitris Bertsimas. Robust game theory. Mathematical Programming, 107(1):231–273, 2006.
[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Gradient flows: in metric spaces and in the space of probability measures. Springer, 2008.
[3] K.J. Arrow and E.S. Maskin. Social Choice and Individual Values: Third Edition. Yale University Press, 2012.
[4] Bernd Bank, Jürgen Guddat, Diethard Klatte, Bernd Kummer, and Klaus Tammer. Non-linear Parametric Optimization. Springer, 1982.
[5] Jonathan Bard. Some properties of the bilevel programming problem. Journal of Optimization Theory and Applications, 68(2):371–379, 1991.
[6] Jonathan F. Bard. Practical Bilevel Optimization: Algorithms and Applications, volume 30. Springer, 2013.
[7] Tamer Basar and Geert J. Olsder. Dynamic Non-Cooperative Game Theory. SIAM, 1999.
[8] Güzin Bayraksan and David K. Love. Data-Driven Stochastic Programming Using ø-Divergences, chapter 1, pages 1–19.
[9] Aharon Ben-Tal, Dick den Hertog, Anja De Waegenaere, Bertrand Melenberg, and Gijs Rennen. Robust solutions of optimization problems affected by uncertain probabilities. Management Science, 59(2):341–357, 2013.
[10] Battista Biggio, Blaine Nelson, and Pavel Laskov. Poisoning attacks against support vector machines. In International Conference on Machine Learning, pages 1467–1474, 2012.
[11] Jose Blanchet, Yang Kang, and Karthyek Murthy. Robust waterstein profile inference and applications to machine learning. Journal of Applied Probability, 56(03):830–857, 2019.
[12] Jose Blanchet and Karthyek Murthy. Quantifying distributional model risk via optimal transport. Mathematics of Operations Research, 44(2):565–600, 2019.
[13] Stephen Boyd and Leven Vandenberghe. Convex Optimization. Cambridge University Press, 2004.
[14] Ralph Allan Bradley and Milton E. Terry. The Method of Paired Comparisons. Cambridge University Press, 1952.
[15] Yiding Chen and Xiaojin Zhu. Optimal attack against autoregressive models by manipulating the environment. In AAAI Conference on Artificial Intelligence, pages 3543–3552, 2020.
[16] Douglas E Critchlow, Michael A Flinger, and Joseph S Verducci. Probability models on rankings. Journal of Mathematical Psychology, 35(3):294 – 318, 1991.
[17] Herbert Aron David. Social Choice and Individual Values. Yale University Press, 2012.
[18] Tamer Basar and Geert J. Olsder. Dynamic Non-Cooperative Game Theory. SIAM, 1999.
[19] Jianfeng Wang. Robust solutions of optimization problems affected by uncertain probabilities. Management Science, 59(2):341–357, 2013.
[20] Battista Biggio, Blaine Nelson, and Pavel Laskov. Poisoning attacks against support vector machines. In International Conference on Machine Learning, pages 1467–1474, 2012.
[21] Jose Blanchet, Yang Kang, and Karthyek Murthy. Robust waterstein profile inference and applications to machine learning. Journal of Applied Probability, 56(03):830–857, 2019.
[22] Jose Blanchet and Karthyek Murthy. Quantifying distributional model risk via optimal transport. Mathematics of Operations Research, 44(2):565–600, 2019.
[23] Stephen Boyd and Leven Vandenberghe. Convex Optimization. Cambridge University Press, 2004.
[24] Ralph Allan Bradley and Milton E. Terry. The Method of Paired Comparisons. Cambridge University Press, 1952.
[25] Yiding Chen and Xiaojin Zhu. Optimal attack against autoregressive models by manipulating the environment. In AAAI Conference on Artificial Intelligence, pages 3543–3552, 2020.
[26] Douglas E Critchlow, Michael A Flinger, and Joseph S Verducci. Probability models on rankings. Journal of Mathematical Psychology, 35(3):294 – 318, 1991.
Hongseok Namkoong and John C. Duchi. Stochastic gradient methods for distributionally robust optimization with f-divergences. In Annual Conference on Neural Information Processing Systems, pages 2208–2216, 2016.

Hongseok Namkoong and John C. Duchi. Variance-based regularization with convex objectives. In Annual Conference on Neural Information Processing Systems, pages 2975–2984, 2017.

John Nash. Non-cooperative games. Annals of Mathematics, 54(2):286–295, 1951.

John F. Nash. Equilibrium points in n-person games. Proceedings of the National Academy of Sciences, 36(1):48–49, 1950.

Sahand Negahban, Sewoong Oh, and Devavrat Shah. Rank centrality: Ranking from pairwise comparisons. Operation Research, 65(1):266–287, 2017.

Sahand Negahban, Sewoong Oh, Kiran Koshy Thekumparampil, and Jiaming Xu. Learning from comparisons and choices. Journal of Machine Learning Research, 19(40):1–95, 2018.

Ashwin Pananjady, Cheng Mao, Vidya Muthukumar, Martin J. Wainwright, and Thomas A. Courtade. Worst-case vs average-case design for estimation from fixed pairwise comparisons. Annals of Statistics, 48(2):1072–1097, 2020.

Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport. Foundations and Trends® in Machine Learning, 11(5-6):355–607, 2019.

Arun Rajkumar, Suprovat Ghoshal, Lek-Heng Lim, and Shivani Agarwal. Ranking from stochastic pairwise preferences: Recovering condorcet winners and tournament solution sets at the top. In International Conference on Machine Learning, pages 665–673, 2015.

J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. Econometrica, 33(3):520–534, 1965.

Nihar B. Shah, Sivaraman Balakrishnan, Joseph K. Bradley, Abhay Parekh, Kannan Ramchandran, and Martin J. Wainwright. Estimation from pairwise comparisons: Sharp minimax bounds with topology dependence. Journal of Machine Learning Research, 17(58):1–47, 2016.

Nihar B. Shah, Sivaraman Balakrishnan, Aditya Guntuboyina, and Martin J. Wainwright. Stochastically transitive models for pairwise comparisons: Statistical and computational issues. In International Conference on Machine Learning, pages 11–20, 2016.

Nihar B. Shah and Martin J. Wainwright. Simple, robust and optimal ranking from pairwise comparisons. Journal of Machine Learning Research, 18(199):1–38, 2017.

Matthew Staib and Stefanie Jegelka. Distributionally robust optimization and generalization in kernel methods. In Annual Conference on Neural Information Processing Systems, pages 2438–2446, 2019.

Louis L. Thurstone. A law of comparative judgment. Psychological Review, 34(4):273–286, 1927.

Alexandre B. Tsybakov. Introduction to Nonparametric Estimation. Springer, 2008.

Rohit Vaish, Neeldhara Misra, Shivani Agarwal, and Avrim Blum. On the computational hardness of manipulating pairwise voting rules. In International Conference on Autonomous Agents & Multiagent Systems, pages 358–367, 2016.

Aad W. van der Vaart and Jon A. Wellner. Weak Convergence and Empirical Processes: With Applications to Statistics. Springer New York, 1996.

Vladimir N. Vapnik. The Nature of Statistical Learning Theory. Springer, 1995.

Cédric Villani. Optimal Transport: Old and New. Springer, 2008.

John von Neumann and Oskar Morgenstern. Theory of Games and Economic Behavior. Princeton University Press, 1947.

Toby Walsh. Is computational complexity a barrier to manipulation? Annals of Mathematics and Artificial Intelligence, 62(1-2):7–26, 2011.

Zizhuo Wang, Peter W Glynn, and Yinyu Ye. Likelihood robust optimization for data-driven problems. Computational Management Science, 13(2):241–261, 2016.

Fabian L. Wauthier, Michael I. Jordan, and Nebojsa Jojic. Efficient ranking from pairwise comparisons. In International Conference on Machine Learning, pages 109–117, 2013.

David Wozabal. Robustifying convex risk measures for linear portfolios: A nonparametric approach. Operations Research, 62(6):1302–1315, 2014.

Qianqian Xu, Jiechao Xiong, Xiaochun Cao, Qingming Huang, and Yuan Yao. From social to individuals: A parsimonious path of multi-level models for crowdsourced preference aggregation.
APPENDIX A

PROOF OF THEOREM 1

Property 1. Let the pay-off function \( f_r, r = 1, 2 \) be the weighted sum-of-squared loss \( \ell \) in (25). If the uncertainty set is \( \mathcal{X}^\alpha(\mathbb{P}_N) \) or \( 2\mathbb{B}_p(\mathbb{P}_N) \), we have

1) \( f_r \) is a continuous function, and for any fixed \( \{ x_r, \xi \} \), \( f_r(x_r, x_r, \xi) \) is convex over \( \mathcal{X}_r \).
2) \( \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \) is a compact set.
3) \( \mathbb{E}_{\xi \sim Q}[f_r(x_r, x_r, \xi)] \) is finite-valued, \( \forall \ x \in \mathcal{X}, \ Q \in \mathfrak{U} \).
4) \( \mathfrak{U} \) is a weakly compact set.

Proposition 1. Let \( x = \{ x_1, x_2 \}, v = \{ v_1, v_2 \} \), we define \( \phi : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+ \) as

\[
\phi(v, x) = \sup_{Q \in \mathfrak{U}} \mathbb{E}_{\xi \sim Q} \left[ f_1(v_1, x_2, \xi) \right] + \sup_{Q \in \mathfrak{U}} \mathbb{E}_{\xi \sim Q} \left[ f_2(x_1, v_2, \xi) \right]
\]

(59)

With Property 1, \( x^* = \{ x^*_1, x^*_2 \} \) is a distributional robust Nash equilibrium of (26) if and only if

\[
\{ x^*_1, x^*_2 \} \in \arg \min_{v \in \mathcal{X}} \phi(v, x^*)
\]

(60)

Proof. The reformulation \( \phi \) is well known for deterministic Nash equilibrium, see for example (62). The “if” part follows from the fact that if \( \{ x^*_1, x^*_2 \} \) is not an equilibrium of (25), there exists some \( x_r, r = 1, 2 \), such that

\[
\sup_{Q \in \mathfrak{U}} \mathbb{E}_{\xi \sim Q} \left[ f_r(x_r, x^*_r, \xi) \right] < \sup_{Q \in \mathfrak{U}} \mathbb{E}_{\xi \sim Q} \left[ f_r(x^*_r, x^*_r, \xi) \right]
\]

Let \( \bar{x} = \{ x^*_1, x^*_2 \} \), we have \( \phi(\bar{x}, x^*) < \phi(x^*, x^*) \). This is a contradiction.

The “only if” part is obvious as

\[
\sup_{Q \in \mathfrak{U}} \mathbb{E}_{\xi \sim Q} \left[ f_r(x_r, x^*_r, \xi) \right] > \sup_{Q \in \mathfrak{U}} \mathbb{E}_{\xi \sim Q} \left[ f_r(x^*_r, x^*_r, \xi) \right]
\]

Summing up each \( r \) on both sides, the inequality shows that \( \{ x^*_1, x^*_2 \} \) is a global minimizer.

Based on the Proposition 1, we have the following existence result for distributional robust Nash equilibrium of (25). Theorem 1. Let the pay-off function \( f_r, r = 1, 2 \) be the weighted sum-of-squared loss \( \ell \) in (25). If the uncertainty set is \( \mathcal{X}^\alpha(\mathbb{P}_N) \) or \( 2\mathbb{B}_p(\mathbb{P}_N) \), the general robust game (25) has a DRNE.

Proof. Based on the Proposition 1, each \( \mathbb{E}_{\xi \sim Q}[f_r(x_r, x_r, \xi)] \) is continuous and convex for any \( Q \in \mathfrak{U} \). The supremum preserves the convexity of \( f_r \) and, under weakly compactness of \( \mathfrak{U} \), the continuity of \( f_r \) will hold. Therefore \( \phi(v, x) \) is continuous and convex w.r.t. \( v \) on \( \mathcal{X} \) for any fixed \( x \in \mathcal{X} \).

The existence of an optimal solution to

\[
\min_{v \in \mathcal{X}} \phi(v, x)
\]

(61)

follows from compactness of \( \mathcal{X} \) under the third condition in Assumption 1. To complete the proof, we are left to show the existence of \( x^* \in \mathcal{X} \) such that

\[
x^* \in \arg \min_{v \in \mathcal{X}} \phi(v, x^*)
\]

(62)

Let \( \Phi(x) \) be the set of optimal solutions to \( \min \phi(v, x) \) for each fixed \( x \in \mathcal{X} \). Then \( \Phi(x) \subset \mathcal{X} \) holds. By the convexity of \( \phi \), \( \Phi(x) \) is a convex set. Obviously, \( \Phi(x) \) is closed, namely, there exists a sequence \( \{ x_k \} \) with \( \lim_{k \to \infty} x_k = \bar{x} \) and \( v_k \in \Phi(x_k) \), if \( \lim_{k \to \infty} v_k = \bar{v} \), we have \( \bar{v} \in \Phi(\bar{x}) \). Further, following by Theorem 4.2.1 in (4), \( \Phi \) is upper semi-continuous on \( \mathcal{X} \). By Kakutani’s fixed point theorem (35), there exists \( x^* \in \mathcal{X} \) such that \( x^* \in \Phi(x^*) \).

APPENDIX B

PROOF OF THEOREM 2

The following proposition shows the strong duality result for Wasserstein DRO (12), which ensures that the inner supremum in (31) admits a reformulation which is a simple, univariate optimization problem. Note that there exists the other strong duality result of Wasserstein DRO (23).

Proposition 2. Let \( d : \mathbb{R}^{n+2} \times \mathbb{R}^{n+2} \to [0, \infty] \) be a lower semi-continuous cost function satisfying \( d(z, z') = 0 \) whenever \( z = z', z = (p, b, y), z' = (p', b', y') \). For \( \lambda \geq 0 \) and loss function \( \ell \) in (29) that is upper semi-continuous in \( (p, b, y) \) for each \( \theta \), define

\[
\psi_{\lambda, \ell}(z; \theta) := \sup_{z' \in \mathbb{R}^{n+2}} \sum_{(i,j)} \left\{ \ell(\theta; z') - \lambda d(z, z') \right\}.
\]

(63)
Then
\[
\sup_{Q \in \mathbb{P}_2^d} \mathbb{E}_{z' \sim Q} \left[ \ell(\theta, z') \right] = \min_{\lambda \geq 0} \left\{ \lambda \alpha + \frac{1}{N} \sum_z \psi_{\lambda, \ell}(z; \theta) \right\}
\]  
(64)

**Theorem 2.** Let \( Z = \{ p, B, y' \} \) be the observed data set, where \( B \) and \( y' \) are defined as (11), \( p \) is the frequency of each type of pairwise comparison as (27). Consider the loss function of \( z \), and the distance function between \( z_c, z'_c \) are based on the \( \ell_2 \)-norm. In other words, we take \( \ell(\theta, z) \) as (29) and
\[
d(z_c, z'_c) = \|[p_{ij}, b_{i,j}, y'_{ij}] - (q_{ij}, b_{i,j}, y'_{ij})\|_2
= |p_{ij} - q_{ij}|.
\]  
(32)

Then, the DRO problem (31) has an equivalent form:
\[
\min_{\theta \in \Theta} \sup_{Q \in \mathbb{P}_2^d} \mathbb{E}_{q \sim Q} \left[ \ell(\theta; q) \right] = \min_{\theta \in \Theta} \mathcal{L}(\theta) + \mathcal{R}(\theta),
\]  
(33)

where
\[
\mathcal{L}(\theta) = \frac{1}{2N} \sum_{(i,j)} \sum_{c} p_{ij} (y'_{ij} - \theta^\top b_{i,j})^2,
\]  
(34)

and
\[
\mathcal{R}(\theta) = \sqrt{\frac{\alpha}{4N} \sum_{(i,j)} (y'_{ij} - \theta^\top b_{i,j})^2}.
\]  
(35)

**Proof.** Let \( \Delta_{ij} = q_{ij} - p_{ij} \). The \( \psi_{\lambda, \ell} \) function (36) has a new formulation as
\[
\psi_{\lambda, \ell}(\theta, p)
= \sup_{q \in \mathbb{R}_c^d} \frac{1}{N} \sum_{c \in C} \left\{ \ell(\theta; q_{ij}) - \lambda [d(p_{ij}, q_{ij})]^2 \right\}
= \sup_{q \in \mathbb{R}_c^d} \frac{1}{N} \sum_{c \in C} \left\{ \frac{q_{ij}}{2} \left( \frac{1}{\theta^\top b_{ij}} \right)^2 \right\} - \lambda |p_{ij} - q_{ij}|^2
= \frac{1}{N} \sum_{c \in C} \sup_{\Delta_{ij} \in \mathbb{R}} \left( \Delta_{ij}^2 b_{ij} - \lambda \Delta_{ij}^2 + p_{ij} b_{ij} \right),
\]  
(65)

where \( b_{ij} = (y_{ij} - \theta^\top a_c)^2 / 2 \), and the third equality holds due to \( \psi_{\lambda, \ell} \) is a decomposable function. Expanding (65), we can simplify \( \psi_{\lambda, \ell} \) as below:
\[
\psi_{\lambda, \ell}(\theta, p)
= \frac{1}{N} \langle p, b \rangle + \frac{1}{N} \sum_{c \in C} \sup_{\Delta_{ij} \in \mathbb{R}} (\Delta_{ij}^2 b_{ij} - \lambda \Delta_{ij}^2)
= \left\{ \begin{array}{ll}
\langle p, b \rangle / N + \|b\|_2^2 / (4\lambda N), & \text{if } \lambda > 0,
\infty, & \text{if } \lambda = 0.
\end{array} \right.
\]  
(66)

Next, we investigate the duality of (31) with Proposition 2. As \( \psi_{\lambda, \ell}(\theta, z) = \infty \) when \( \lambda = 0 \), the dual formulation of the supremum in (38) would be
\[
\sup_{Q \in \mathbb{P}_2^d} \mathbb{E}_{z' \sim Q} \left[ \ell(\theta, z') \right] = \min_{\lambda \geq 0} \left\{ \lambda \alpha + \psi_{\lambda, \ell}(\theta, p) \right\}
= \min_{\lambda > 0} \left\{ \lambda \alpha + \frac{1}{N} \langle p, b \rangle + \frac{1}{4\lambda N} \|b\|_2^2 \right\}. \tag{67}
\]  

By the definition of \( b \), we know that
\[
\ell(\theta, p) = \frac{1}{N} \langle p, b \rangle \tag{68}
\]
Moreover, notice that the right hand side of (67) is a convex function which approaches infinity when \( \lambda \to \infty \), the global optimal of it can be obtained uniquely via the first order optimality condition as
\[
\frac{\partial}{\partial \lambda} \left\{ \lambda \alpha + \frac{1}{N} \langle p, b \rangle + \frac{1}{4\lambda N} \|b\|_2^2 \right\} = 0, \tag{69}
\]
and the optimal dual variable is
\[ \lambda^* = \frac{\|b\|_2}{2\sqrt{\alpha N}}. \] (70)
Substituting \( \lambda^* \) and \( b \) into (67), we have
\[
\sup_{Q \in \mathcal{M}_p^N(\mathcal{P}_N)} E_{z \sim Q} \left[ \ell(\theta, z') \right] \leq \sqrt{\frac{\alpha}{N}} \cdot \|b\|_2 + \frac{1}{N} \cdot (p, b)
\] (71)

**APPENDIX C
SOME PROPOSITIONS FOR GENERALIZATION ANALYSIS.**

**Proposition 3.** Suppose that \( \ell \) is L-Lipschitz function, i.e., \( |\ell(z) - \ell(z')| \leq L \cdot d_Z(z, z') \) for all \( z, z' \in Z \). Then, for any \( Q \in \mathcal{M}_p^N(\mathcal{P}_N) \),
\[ R_Q(\ell) \leq R_{\mathcal{F}, \alpha, p}(\ell) \leq R_Q(\ell) + 2L\alpha. \] (72)

**Proof.** For \( p = 1 \), the result follows immediately from the Kantorovich dual representation of \( \mathcal{W}_1(\cdot, \cdot) \) [72]:
\[ \mathcal{W}_1(\mathcal{P}, Q) = \sup \left\{ \left| E_{z \sim \mathcal{F}}[h(z)] - E_{z \sim Q}[h(z)] \right| \left| \sup_{z, z' \in Z, z \neq z'} \frac{h(z) - h(z')}{d_Z(z, z')} \right| \leq 1 \right\} \] (73)
with the triangle inequality:
\[ \mathcal{W}_1(\mathcal{P}, Q) \leq 2\alpha, \quad \forall \ P, Q \in \mathcal{M}_p^N(\mathcal{P}_N). \] (74)
For \( p > 1 \), the result follows from the fact that
\[ \mathcal{W}_1(\mathcal{P}, Q) \leq \mathcal{W}_p(\mathcal{P}, Q), \quad \forall \ P, Q \in \mathcal{P}_p(Z). \] (75)

Next we consider the case when the function \( \ell \) is smooth but not Lipschitz-continuous. Since we are working with general metric spaces that may lack an obvious differentiable structure, we need to first introduce some concepts from metric geometry [2].

**Definition 4 (Geodesic Space).** A metric space \( (Z, d_Z) \) is a geodesic space if for every pair of points \( z, z' \in Z \) there exists a constant-speed geodesic path \( \varrho : [0, 1] \to Z \), such that \( \varrho(0) = z, \varrho(1) = z' \), and for all \( 0 \leq s \leq t \leq 1 \)
\[ d_Z\left[ \varrho(s), \varrho(t) \right] = (t - s) \cdot d_Z\left[ \varrho(0), \varrho(1) \right]. \] (76)

**Definition 5 (Geodesic convexity).** A functional \( \ell : Z \to \mathbb{R} \) is geodesically convex if for any pair of points \( z, z' \in Z \) there is a constant-speed geodesic \( \varrho \), so that
\[ \ell(\varrho(t)) \leq (1 - t) \cdot \ell(\varrho(0)) + t \cdot \ell(\varrho(1)) = (1 - t) \cdot \ell(z) + t \cdot \ell(z'). \] (77)

**Definition 6 (Upper Gradient).** Suppose that \( \ell : Z \to \mathbb{R} \) is a Borel function. The upper gradient of \( \ell \) is a functional \( G_\ell : Z \to \mathbb{R}_+ \) satisfies that: for any pair of points \( z, z' \in Z \), there exist a constant-speed geodesic path \( \varrho \):
\[ |\ell(z') - \ell(z)| \leq \int_0^1 G_\ell(\varrho(t))dt \cdot d_Z(z, z'). \] (78)

**Proposition 4.** Suppose that \( \ell \) has a geodesically convex upper gradient \( G_\ell \), we have
\[ R_Q(\ell) \leq R_{\mathcal{F}, \alpha, p}(\ell) \leq R_Q(\ell) + 2\alpha\mu, \] (79)
where
\[ \mu = \sup_{Q \in \mathcal{M}_p^N(\mathcal{P})} \left( E_{z \sim Q}\left[ \left| G_\ell(z) \right|^q \right] \right)^\frac{1}{q} \] (80)
and \( 1/p + 1/q = 1 \).
Proposition 5. Consider the setting of pairwise ranking problem with the sum-of-squared loss: let \( \ell(z', z) \leq \int_0^1 G_\ell(q(t))dt \cdot d_Z(z, z') \)
\[
\leq \frac{1}{2} \left( G_\ell(z) + G_\ell(z') \right) \cdot d_Z(z, z'),
\]
where the first inequality is from the definition of the upper gradient \((78)\) and the second one is by the assumed geodesic convexity of \(G_\ell\). Taking expectations of both sides with respect to \( \gamma \) and using Hölder inequality, we obtain
\[
R_Q(\ell) - R_{Q'}(\ell) \leq \frac{1}{2} \left( \mathbb{E}_{(z, z') \sim \gamma} \left[ |G_\ell(z) + G_\ell(z')|^q \right] \right)^{\frac{1}{q}} \left( \mathbb{E}_{(z, z') \sim \gamma} \left[ |d_Z(z, z')|^p \right] \right)^{\frac{1}{p}}
\]
\[
= \frac{1}{2} \left( \mathbb{E}_{(z, z') \sim \gamma} \left[ |G_\ell(z) + G_\ell(z')|^q \right] \right)^{\frac{1}{q}} \cdot W_p(Q, Q'),
\]
where we adopt the \(p\)-Wasserstein optimality of \( \gamma \) for \( Q \) and \( Q' \). By the triangle inequality, and since \( z \sim Q \) and \( z' \sim Q' \),
\[
\left( \mathbb{E}_{(z, z') \sim \gamma} \left[ |G_\ell(z) + G_\ell(z')|^q \right] \right)^{\frac{1}{q}} \leq \left( \mathbb{E}_{z \sim Q} \left[ |G_\ell(z)|^q \right] \right)^{\frac{1}{q}} + \left( \mathbb{E}_{z' \sim Q'} \left[ |G_\ell(z')|^q \right] \right)^{\frac{1}{q}}
\]
\[
\leq 2 \sup_{\mathbb{W}_p^q(P)} \left( \mathbb{E}_{z \sim Q} \left[ |G_\ell(z)|^q \right] \right)^{\frac{1}{q}}.
\]
Interchanging the roles of \( Q \) and \( Q' \) and proceeding with the same argument, we obtain the following estimation
\[
\sup_{Q, Q' \in \mathbb{W}_p^q(P)} \left| R_Q(\ell) - R_{Q'}(\ell) \right| \leq 2\alpha \sup_{Q \in \mathbb{W}_p^q(P)} \left( \mathbb{E}_{z \sim Q} \left[ |G_\ell(z)|^q \right] \right)^{\frac{1}{q}}.
\]
Then
\[
R_Q(\ell) \leq R_{\bar{\tau}, o, p}(\ell)
\]
\[
= 2\alpha \sup_{Q' \in \mathbb{W}_p^q(P)} \left[ R_{Q', o, p}(\ell) - R_Q(\ell) + R_Q(\ell) \right]
\]
\[
\leq R_Q(\ell) + 2\alpha \sup_{Q \in \mathbb{W}_p^q(P)} \left( \mathbb{E}_{z \sim Q} \left[ |G_\ell(z)|^q \right] \right)^{\frac{1}{q}}
\]
\[
\tag{84}
\]
\[
\square
\]

Proposition 5. Consider the setting of pairwise ranking problem with the sum-of-squared loss: let \( \mathcal{A} \) be a convex subset of \( \mathbb{R}^n \), \( \mathcal{Y} = [-1, 1] \), and equip \( \mathcal{Z} = \mathcal{A} \times \mathcal{Y} \) with the Euclidean metric
\[
d_Z(z, z') = \sqrt{||a - a'||^2 + |y - y'|^2}, \quad z = (a, y).
\]
It means that we do not aggregate the pairwise comparisons into the same type and the weight. Then, it holds that
\[
R_Q(\ell) \leq R_{\bar{\tau}, o, z}(\ell) \leq R_Q(\ell) + 4\alpha(1 + C)\tau,
\]
where
\[
\tau = \left( 1 + L \sup_{Q \in \mathbb{W}_p^2(P_N)} \mathbb{E}_Q \| A \|_2 \right), \quad z = (a, y) \sim Q,
\]
and \( \mathcal{A} = [a_1^T, \ldots, a_N^T] \).

Proof. As \( \mathcal{Z} \subseteq \mathbb{R}^{n+1} \), \( \mathcal{Z} \) is a geodesic space as
\[
\gamma(t) = (1 - t) \cdot z + t \cdot z', \quad \forall \ z, z' \in \mathcal{Z}
\]
is the unique constant-speed geodesic path.
Moreover, the geodesically convex upper gradient of \( \ell \) is
\[
G_\ell(z) = G_\ell(a, y) = 2(B + C)(1 + L\| \nabla h(a) \|_2), \quad \forall \ z \in \mathcal{Z}.
\]
where $\ell(z) = \ell(a, y) = (y - h(a))^2$. In such a flat Euclidean setting, geodesic convexity coincides with the usual definition of convexity, and the map $z \to G_\ell(z)$ is convex evidently: for all pair $z, z' \in Z$

$$G_\ell((1 - t) \cdot z + t \cdot z') \leq (1 - t) \cdot G_\ell(z) + t \cdot G_\ell(z').$$

(90)

With the mean-value theorem

$$\ell(z) - \ell(z')$$

$$\leq \int_0^1 \langle z - z', \nabla \ell((1 - t) \cdot z + t \cdot z')\rangle dt$$

$$\leq \int_0^1 \|\nabla \ell((1 - t) \cdot z + t \cdot z')\|_2 dt \cdot \|z - z'\|_2$$

$$= \int_0^1 \|\nabla \ell((1 - t) \cdot z + t \cdot z')\|_2 dt \cdot d_G(z, z')$$

(91)

and a simple calculation

$$\|\nabla \ell(z)\|_2^2 = 4\ell(z)(1 + \|\nabla h(a)\|_2^2) \leq 4(B + C)^2(1 + L^2\|a\|_2^2),$$

(92)

we have $\|\nabla \ell(z)\|_2 \leq G_\ell(z)$ for any $z \in Z$. Thus, by Proposition 3 we have

$$R_\ell(\ell) \leq R_{P, \alpha, 2}(\ell)$$

$$\leq R_\ell(\ell) + 2\alpha \sup_{Q \in \mathcal{W}_1(P_N)} \left( \mathbb{E}_{z \sim Q}\left[ |G_\ell(z)|^2 \right] \right)^{\frac{1}{2}}$$

$$= R_\ell(\ell) + 4\alpha(B + C) \left( 1 + L \sup_{Q \in \mathcal{W}_1(P_N)} \mathbb{E}_{z \sim Q} \|A\|_2 \right).$$

(93)

**APPENDIX D**

**PROOF OF THEOREM 3.**

**Assumption 1.** $d : Z \times Z \to \mathbb{R}_+$ in (21) and (22) is a nonnegative lower semi-continuous function satisfying $d(w, w') = 0$ if and only if $w = w'$.

**Assumption 2.** The loss function $\ell \in \mathcal{F} \subseteq L^1(dQ)$ are upper semi-continuous, where $L^1(dQ)$ denote the collection of Borel measurable functions $\ell : \mathcal{Z} \to \mathbb{R}$ such that

$$\int |\ell| dQ < \infty, \ \forall \ Q \in \mathcal{P}(Z).$$

**Assumption 3.** The instance space $Z$ is bounded, namely,

$$\text{diam}(Z) = \sup_{z, z' \in Z} d_G(z, z') < \infty.$$

(94)

**Assumption 4.** $\ell \in \mathcal{F}$ is uniformly bounded as

$$0 \leq \ell(z) \leq B < \infty, \ \forall \ \ell \in \mathcal{F}, \ \text{and} \ z \in Z.$$  

(95)

**Definition 7.** Let $(Z, d_Z)$ be a metric space. For a function $\ell : Z \to \mathbb{R}$ and a point $s \in \mathbb{R}$, the upper contour set defined by $s$ is

$$\ell^{-1}([s, \infty)) = \{z \in Z : \ell(z) \geq s\},$$

(96)

and the corresponding lower contour set is

$$\ell^{-1}((\infty, s]) = \{z \in Z : \ell(z) \leq s\}.$$

(97)

We call a function $\ell : Z \to \mathbb{R}$ is upper semi-continuous if and only if for any $s \in \mathbb{R}$, $\ell^{-1}((\infty, s])$ is an open set.

We adopt the Dudley’s entropy integral [20] as the complexity measure of the hypothesis class $\mathcal{F}$,

$$\mathcal{F}(\mathcal{F}) = \int_0^\infty \sqrt{\log \mathfrak{N}(\mathcal{F}, \|\cdot\|_\infty, \nu)} d\nu,$$

(98)

where $\mathfrak{N}(\mathcal{F}, \|\cdot\|_\infty, \nu)$ is $\nu$-covering number of $\mathcal{F}$ with respect to the uniform metric $\|\cdot\|_\infty$, defined as the size of the smallest $\nu$-cover of $\mathcal{F}$

$$\mathfrak{N}(\mathcal{F}, \|\cdot\|_\infty, \nu)$$

$$= \min_{m \in \mathbb{N}} \left\{ \exists \{\ell_1, \ldots, \ell_m\} \subseteq \mathcal{F} \subseteq \bigcup_{k=1}^m B_{\nu \|\cdot\|_\infty}(\ell_k) \right\}$$

(99)
and $\bigcup_{\ell=1}^m B_\nu^\|\cdot\|_\infty(\cdot)$ is a $\nu$-cover of $F$ with respect to $\|\cdot\|_\infty$

$$\|\ell - l'\|_\infty = \sup_{z \in Z} |\ell(z) - l'(z)|.$$

(100)

Theorem 3. Consider the setting of pairwise ranking problem with the sum-of-squared loss, for any $t > 0$, it holds

$$\Pr \left( \exists \ell \in F : R_{F,\alpha,p}(\ell) > \varsigma_1 \right) \leq e^{-2t^2}$$

(49)

and

$$\Pr \left( \exists \ell \in F : R_{F_N,\alpha,p}(\ell) > \varsigma_2 \right) \leq 2e^{-2t^2}$$

(50)

where

$$\varsigma_1 = \min_{\lambda \geq 0} \left\{ \lambda \alpha^2 + \mathbb{E}_{z \sim \mathcal{Z}} \left[ \psi_{\lambda,\ell}(z) \right] \right\} + \frac{24\mathcal{J}(F) + t}{\sqrt{N}}.$$  

(51)

and

$$\varsigma_2 = \min_{\lambda \geq 0} \left\{ (\lambda + 1)\alpha^2 + \mathbb{E}_{z \sim \mathcal{Z}} \left[ \psi_{\lambda,\ell}(z) \right] \right\} + \frac{\sqrt{\log(\lambda + 1)} + 24\mathcal{J}(F) + t}{\sqrt{N}},$$

(52)

where $\mathcal{J}(F)$ is the Dudley’s entropy integral [20], which is served as the complexity measure of the hypothesis class $F$.

Proof. This proof is a specialization of data-dependent generalization bounds for margin cost function class [38]. From the definition of the local minimax risk [47] and its duality form,

$$R_{F,\alpha,p}(\ell) = \min_{\lambda \geq 0} \left\{ \lambda \alpha^2 + \mathbb{E}_{z \sim \mathcal{Z}} \left[ \psi_{\lambda,\ell}(z) \right] \right\} \leq \min_{\lambda \geq 0} \left\{ \lambda \alpha^2 + \mathbb{E}_{z \sim \mathcal{Z}} \left[ \psi_{\lambda,\ell}(z) \right] + V_{\lambda} \right\}$$

(101)

where

$$V_{\lambda} = \sup_{\ell \in F} \left\{ \mathbb{E}_{z \sim \mathcal{Z}} \left[ \psi_{\lambda,\ell}(z) \right] - \mathbb{E}_{\mathcal{Z} \sim \mathcal{F}} \left[ \psi_{\lambda,\ell}(z) \right] \right\}$$

(102)

is a data-dependent random variable for any $\lambda \geq 0$. As $F$ and $P$ satisfy the Assumption [3] and [4], we have

$$0 \leq \psi_{\lambda,\ell}(z) \leq B, \forall z \in \mathcal{Z}.$$  

(103)

Furthermore, known from McDiarmid’s inequality that, for any fixed $\lambda \geq 0$

$$\Pr \left( V_{\lambda} \geq \mathbb{E} V_{\lambda} + \frac{Bt}{\sqrt{N}} \right) \leq 2e^{-2t^2}.$$  

(104)

Using a standard symmetrization argument, we have

$$\mathbb{E} V_{\lambda} \leq 2 \cdot \mathbb{E} \left[ \sup_{\ell \in F} \frac{1}{N} \sum_{i=1}^{N} \epsilon_i \psi_{\lambda,\ell}(z_i) \right]$$

(105)

where $\epsilon_1, \ldots, \epsilon_N$ are i.i.d. Rademacher random variables independent of $z_1, \ldots, z_N$.

To bound (105), we define the $F$-indexed process $\beta_F = \{\beta_{\ell}\}_{\ell \in \mathcal{F}}$ as

$$\beta_{\ell} = \frac{1}{N} \sum_{i=1}^{N} \epsilon_i \psi_{\lambda,\ell}(z_i).$$

(106)

This is a zero-mean and sub-Gaussian process [70] with respect to the metric $\|\cdot\|_\infty$ as

$$\mathbb{E} \left[ \exp(t(\beta_{\ell} - \beta_{l'})) \right] = \mathbb{E} \left[ \exp \left( \frac{t}{\sqrt{N}} \sum_{i=1}^{N} \epsilon_i \left( \psi_{\lambda,\ell}(z_i) - \psi_{\lambda,l'}(z_i) \right) \right) \right]$$

$$= \left\{ \mathbb{E} \left[ \exp \left( \frac{t}{\sqrt{N}} \epsilon_i \sup_{z' \sim \mathcal{Z}} \left\{ \ell(z') - \lambda d(z_1, z') p - \ell'(z'') + \lambda d(z_1, z'') p' \right\} \right) \right] \right\}^N$$

$$\leq \left\{ \mathbb{E} \left[ \exp \left( \frac{t}{\sqrt{N}} \epsilon_i \sup_{z} \left\{ \ell(z') - \ell'(z') \right\} \right) \right] \right\}^N$$

$$\leq \exp \left( \frac{t^2 \|\cdot - l'\|_\infty^2}{2} \right),$$

(107)
where the second equation comes from the independence of \( \{z_i\}_{i \in [N]} \) and the definition of \( \psi_{\lambda,\ell}(\cdot) \). The last inequality follows the Hoeffding's lemma \[30\].

With the \( \mathcal{F} \)-indexed process \( \beta_{\mathcal{F}} \) and invoking Dudley's entropy integral \[98\] for the right-hand side of (105), we obtain
\[
\mathbb{E}V_{\lambda} \leq 2 \cdot \mathbb{E} \left[ \sup_{\ell \in \mathcal{F}} \beta_{\ell} \right] \leq \frac{24}{\sqrt{N}} \mathcal{J}(\mathcal{F}), \quad \forall \lambda \geq 0
\]
and
\[
\Pr \left( V_{\lambda} \geq \frac{24\mathcal{J}(\mathcal{F}) + Bt}{\sqrt{N}} \right) \leq 2e^{-2t^2}.
\]
In addition, the first part of the claims holds with ant fixed \( \lambda \geq 0 \):
\[
\Pr \left( \exists \ell \in \mathcal{F} : R_{\mathcal{F},\alpha,p}(\ell) > \varsigma_1 \right) \leq e^{-2t^2}, \quad \forall t > 0,
\]
where
\[
\varsigma_1 = \min_{\lambda \geq 0} \left\{ \lambda \alpha^p + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda,\ell}(z) \right] \right\} + \frac{24\mathcal{J}(\mathcal{F}) + Bt}{\sqrt{N}}.
\]
For the second part, we start with two sequences: \( \{\lambda_k\} \) and \( \{t_k\} \)
\[
\lambda_k = k, \quad t_k = t + \sqrt{\log(k)}, \quad k = 1, 2, \ldots
\]
and \[49\] also holds as
\[
\Pr \left( \exists \ell \in \mathcal{F} : R_{\mathcal{F},\alpha,p}(\ell) > \min_k \left\{ \lambda_k \alpha^p + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda_k,\ell}(z) \right] \right\} + \frac{24\mathcal{J}(\mathcal{F}) + Bt_k}{\sqrt{N}} \right) 
\leq \sum_k e^{-2t_k^2} 
\leq \sum_k e^{-2\log(k)} \cdot e^{-2t^2} 
\leq 2e^{-2t^2}.
\]
Moreover,
\[
\min_k \left\{ \lambda_k \alpha^p + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda_k,\ell}(z) \right] \right\} + \frac{24\mathcal{J}(\mathcal{F}) + Bt_k}{\sqrt{N}} 
= \min_k \left\{ k \alpha^p + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda_k,\ell}(z) \right] \right\} + \frac{24\mathcal{J}(\mathcal{F}) + Bt}{\sqrt{N}} + \frac{B\sqrt{\log(k)}}{\sqrt{N}} 
\leq \min_{\lambda \geq 0} \left\{ (\lambda + 1) \alpha^p + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda,\ell}(z) \right] \right\} + \frac{24\mathcal{J}(\mathcal{F}) + Bt}{\sqrt{N}} + \frac{B\sqrt{\log(\lambda + 1)}}{\sqrt{N}}
\]
where the last inequity holds since, for any \( \lambda \geq 0 \), there exists \( k \in \mathbb{N}_+ \) such that \( \lambda \leq k \leq \lambda + 1 \), and \( \psi_{\lambda_1,\ell} \leq \psi_{\lambda_2,\ell} \) holds whenever \( \lambda_1 \geq \lambda_2 \) as \[63\].

Notice that
\[
R_{\mathcal{F},\alpha,p}(\ell) \leq \min_{\lambda > 0} \left\{ \lambda \alpha^p + \mathbb{E}_{z \sim \mathcal{F}N} \left[ \psi_{\lambda,\ell}(z) \right] + W_{\lambda} \right\},
\]
where
\[
W_{\lambda} = \sup_{\ell \in \mathcal{F}} \left\{ \mathbb{E}_{z \sim \mathcal{F}N} \left[ \psi_{\lambda,\ell}(z) \right] - \mathbb{E}_{z \sim \mathcal{F}} \left[ \psi_{\lambda,\ell}(z) \right] \right\}.
\]
Following the similar analysis of \( R_{\mathcal{F},\alpha,p}(\ell) \), the second part of the claims holds
\[
\Pr \left( \exists \ell \in \mathcal{F} : R_{\mathcal{F},\alpha,p}(\ell) > \varsigma_2 \right) \leq 2e^{-2t^2}, \quad \forall t > 0
\]
where
\[
\varsigma_2 = \min_{\lambda \geq 0} \left\{ (\lambda + 1) \alpha^p + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda,\ell}(z) \right] + \frac{B\sqrt{\log(\lambda + 1)}}{\sqrt{N}} \right\} + \frac{24\mathcal{J}(\mathcal{F}) + Bt}{\sqrt{N}}.
\]
APPENDIX E
PROOF OF THEOREM 4

The common choice of the smoothness assumption is Lipschitz smoothness. Next, we explore the behavior of the dual variable $\lambda$ in (64) when the (64) archives the minimal. The following lemma enables the control of its upper bound.

Assumption 5. The functions in $F$ are $L$-Lipschitz, if they satisfy

$$\sup_{z, z' \in Z, z \neq z'} \frac{f(z') - f(z)}{d_Z(z', z)} \leq L, \forall \ell \in F.$$  \hspace{1cm} (115)

Lemma 1. Suppose that $Q \in \mathcal{M}_p^\alpha(\mathbb{P}_n) \subset \mathcal{P}_m(Z)$ and $\hat{\ell}$ is the optimal solution of local worst-case risk with distribution $Q$

$$\hat{\ell} \in \arg \min_{\ell \in F} R_{Q, \alpha, p}(\ell),$$

$\bar{\lambda}$ is the infimum-archiving dual variable corresponding to $\hat{\ell}$

$$\bar{\lambda} \in \min_{\lambda \geq 0} \left\{ \lambda \alpha + \mathbb{E}_{z \sim Q} \left[ \psi_{\lambda, \hat{\ell}}(z) \right] \right\}.$$ \hspace{1cm} (117)

Then under Assumption 3-5, $\bar{\lambda}$ satisfies

$$\bar{\lambda} \leq L\alpha^{-(p-1)}.$$ \hspace{1cm} (118)

Proof. With the fixed $Q$ and the estimator $\hat{f}$, we have

$$\bar{\lambda} \alpha^p \leq \bar{\lambda} \alpha^p + \mathbb{E}_{z \sim Q} \left[ \sup_{z' \in Z} \left( \hat{f}(z') - \hat{f}(z) - \bar{\lambda} \left[ d_Z(z, z') \right]^p \right) \right]$$ \hspace{1cm} (119)

and the equality holds with $z' = z$. Due to the optimality of $\bar{\lambda}$ and the dual formulation of local worst-case risk (64), (119) can be further bounded as

$$\bar{\lambda} \alpha^p \leq \lambda \alpha^p + \mathbb{E}_{z \sim Q} \left[ \sup_{z' \in Z} \left( \hat{f}(z') - \hat{f}(z) - \lambda \left[ d_Z(z, z') \right]^p \right) \right],$$

$$\leq \lambda \alpha^p + \mathbb{E}_{z \sim Q} \left[ \sup_{z' \in Z} \left( L \cdot d_Z(z, z') - \lambda \left[ d_Z(z, z') \right]^p \right) \right],$$ \hspace{1cm} (120)

and the second line comes from the Lipschitz smoothness of $\ell \in F$ and the third line holds by substituting $v = d_Z(z, z')$.

When $p = 1$, the result can be obtained by taking $\lambda = L$

$$\bar{\lambda} \alpha \leq L \alpha + \sup_{z' \in Z} \left\{ L v - L v \right\} = L \alpha.$$ \hspace{1cm} (121)

If $p > 1$, we can take $v^* = \left( \frac{L}{L} \right)^{\frac{1}{p-1}}$ which satisfies first-order optimal condition for

$$\sup_{v \geq 0} \left\{ L v - \lambda v^p \right\}$$ \hspace{1cm} (122)

and

$$\bar{\lambda} \alpha^p \leq \lambda \alpha^p + (p - 1) L \alpha^{\frac{p-1}{p}} \lambda^{\frac{1}{p}}.$$ \hspace{1cm} (123)

Treating $\lambda$ as a variable and minimizing the right-hand side of (123) by choosing $\lambda = L \alpha^{\frac{1}{p-1}}$, the claim holds. \hfill $\square$

Theorem 4. Consider the setting of pairwise ranking problem with the sum-of-squared loss, the following holds with probability as least $1 - \eta$

$$R_{p, \alpha, 2}(\hat{\ell}) - R_{p, \alpha, 2}^*(F) \leq \frac{48 f(F)}{\sqrt{N}} + \frac{48 L \left[ \text{diam}(Z) \right]^2}{\alpha \sqrt{N}} + 3 \sqrt{\frac{\log \left( \frac{2}{\eta} \right)}{2N}},$$ \hspace{1cm} (53)

where $\text{diam}(Z)$ is the diameter of $Z$

$$\text{diam}(Z) = \sup_{z, z' \in Z} d_Z(z, z').$$ \hspace{1cm} (54)

Proof. Suppose that $\ell^* \in F$ can archive the local minimax risk $R_{p, \alpha, p}(F)$, we decompose the excess risk

$$R_{p, \alpha, p}(\hat{\ell}) - R_{p, \alpha, p}^*(F) = R_{p, \alpha, p}(\hat{\ell}) - R_{p, \alpha, p}^*(F) \leq R_{p, \alpha, p}(\hat{\ell}) - R_{p, \alpha, p}(\hat{\ell}) + R_{p, \alpha, p}(\ell^*) - R_{p, \alpha, p}(\ell^*),$$ \hspace{1cm} (124)
where the last equality stands by the optimality of \( \hat{\ell} \).

Next, we introduce \( \hat{\lambda} \) and \( \lambda^* \) as the corresponding dual variables of \( \ell \) and \( \ell^* \) as

\[
\hat{\ell} \in \min_{\lambda \geq 0} \left\{ \lambda \alpha^p + \mathbb{E}_{z \sim \mathcal{P}_N} \left[ \psi_{\lambda, \ell}(z) \right] \right\},
\]

(125)

and

\[
\ell^* \in \min_{\lambda \geq 0} \left\{ \lambda \alpha^p + \mathbb{E}_{z \sim \mathcal{P}} \left[ \psi_{\lambda, \ell^*}(z) \right] \right\}.
\]

(126)

By the first part of Theorem 2, the right-hand side of (124) can be further bounded by

\[
R_{\theta, \alpha, p}(\hat{\ell}) - R_{\theta, N, \alpha, p}(\hat{\ell}) = \min_{\lambda \geq 0} \left\{ \lambda \alpha^p + \int \psi_{\lambda, \ell}(z) \mathbb{P}(dz) \right\} - \left( \hat{\lambda} \alpha^p + \int \psi_{\lambda, \ell}(z) \mathbb{P}_N(dz) \right)
\]

\[
\leq \int \psi_{\lambda, \ell}(z)(\mathbb{P} - \mathbb{P}_N)(dz),
\]

(127)

and

\[
R_{\theta, N, \alpha, p}(\ell^*) - R_{\theta, N, \alpha, p}^*(\ell^*) \leq \int \psi_{\lambda^*, \ell^*}(z)(\mathbb{P}_N - \mathbb{P})(dz).
\]

(128)

By Lemma 1 we know

\[
\hat{\lambda} \in \Lambda := [0, \lambda \alpha^{-1}]
\]

(129)

and define the function class

\[
\Psi = \left\{ \psi_{\lambda, \ell} \mid \lambda \in \Lambda, \ell \in \mathcal{F} \right\},
\]

(130)

(127) can be written as

\[
R_{\theta, \alpha, p}(\hat{\ell}) - R_{\theta, N, \alpha, p}(\hat{\ell}) \leq \sup_{\psi \in \Psi} \left\{ \int \psi d(\mathbb{P} - \mathbb{P}_N) \right\}.
\]

(131)

By Assumption 3 and the definition of \( \psi_{\lambda, \ell} \) as (63), we know that every \( \psi \in \Psi \) is bounded and take value in \([0, B]\). Employing symmetrization, we have

\[
R_{\theta, \alpha, p}(\hat{\ell}) - R_{\theta, N, \alpha, p}(\hat{\ell}) \leq 2\mathfrak{R}_N(\Psi) + B \sqrt{\frac{\log(2)}{N}}
\]

(132)

with probability at least \( 1 - \frac{\eta}{2} \), where

\[
\mathfrak{R}_N(\Psi) = \mathbb{E} \left[ \sup_{\psi \in \Psi} \frac{1}{N} \sum_{i=1}^{N} \epsilon_i \psi(z) \right]
\]

(133)

is the expected Rademacher average of \( \Psi \), with i.i.d Rademacher random variables \( \{\epsilon_i\} \) which are independent of \( \{z_i\}, i \in [N] \). Moreover, from Hoeffding’s inequality, it follows that

\[
R_{\theta, N, \alpha, p}^*(\ell^*) - R_{\theta, N, \alpha, p}^*(\ell^*) \leq B \sqrt{\frac{\log(2)}{2N}}
\]

(134)

with probability at least \( 1 - \frac{\eta}{2} \). Combining (132) and (134), and apply the Lemma from Appendix, we obtain the whole theorem. □

**The Stackelberg Game Attack on Pairwise Ranking.**

We study the poisoning attack on pairwise ranking, which injects the malicious pairwise comparisons into the training set of the ranking algorithm. Meanwhile, the robust ranking algorithm could prune the outlier when leaning a consensus ranking with the noise observation. Such an adversarial interaction between two opponents can be naturally a game. One player will control the ranking algorithm, and the other player will manipulate the distribution of input data, especially the pairwise comparisons. The optimal action for each player generally depends on both players’ strategies.

We adopt positive integers to index alternatives and users. Henceforth, \( V \) always is the set \( \{1, \ldots, n\} \) and denotes a set of alternatives to be ranked. In our approach to attack pairwise ranking, we represent these candidates as vertices of a graph. \( U = \{1, \ldots, m\} \) denotes a set of voters or users. For \( i, j \in V \), we write the pairwise comparison \( i \succ j \) or \( (i, j) \) to mean that alternative \( i \) is preferred over alternative \( j \). If we hope to emphasize the preference judgment of a particular user \( u \), we will write \( i \succ^u j \) or \( (u, i, j) \).

For each user \( u \in U \), the pairwise ranking matrix of user \( u \) is a skew-symmetric matrix

\[
Y^u = \{y^u_{ij}\} \in \mathbb{R}^{n \times n}, \ i, j \in V, \ u \in U,
\]

(135)

i.e. for any ordered pair \( (i, j) \in V \times V \), we have

\[
y^u_{ij} = -y^u_{ji},
\]

(136)
Informally, $y^u_{ij}$ measures the “degree of preference” of the $i^{th}$ alternative over the $j^{th}$ alternative held by the $u^{th}$ voter. Here we focus on the “binary” case of $Y^u \in \{−1, 1\}^{n \times n}$. Here $y^u_{ij} = 1$ means there exist a particular preference judgment $(u, i, j)$ made by user $u$.

Define the weight function $w : U \times V \times V \rightarrow [0, \infty)$ as the indicator function

$$w^u_{ij} = w(u, i, j) = \begin{cases} 1, & \text{if } y^u_{ij} = 1, \\ 0, & \text{otherwise}. \end{cases}$$

With the weight function $w$, we can aggregate all users’ pairwise comparison matrices $\{Y^u\}$, $u \in U$ into a single comparison matrix $Y = \{y_{ij}\}$ with weights matrix $W^0 = \{w^0_{ij}\}$, where

$$y_{ij} = 1, \forall (i, j) \in \binom{V}{2},$$

$\binom{V}{2}$ is the set of all ordered pairs of elements of $V$, and

$$w^0_{ij} = \sum_{u \in U} w^u_{ij}, \forall (i, j) \in \binom{V}{2}.$$  

A graph structure arises naturally from ranking data as follows. Let $G = (V, E)$ be a directed graph whose vertex set is $V$, the set of candidates to be ranked. The edge set is

$$E := \left\{ e = (i, j) \mid (i, j) \in \binom{V}{2} \right\}.$$ 

We call such $G$ a pairwise comparison graph. One can further associate weights on the edges as [4]. Different from the general pairwise ranking setting, we do not prune the edges whose weights equal to 0. As a consequence, the pairwise comparison graph $G$ is a complete graph. The cardinality of the edge set is

$$|E| := N = n(n - 1).$$

The comparison between $i$ and $j$ will be labeled by different annotators and their answers to the same question could be inconsistent, i.e.,

$$y_{ij}^1 = y_{ij}^2 = 1, u_1, u_2 \in U.$$ 

To obtain the true direction between vertex $i$ and $j$, we define an estimator $\hat{y}_{ij}$ of noise label $y_{ij}$ on edge $e = (i, j)$,

$$\hat{y}_e = (z_e, \theta) + \gamma_e + \varepsilon_e, \forall e \in E,$$

where $Z = \{z_e\} \in \{-1, 0, 1\}^{N \times n}$, $e \in E$ is the incident matrix of $G$, $\Theta \in \mathbb{R}^n$ is some true scaling scores on $V$, $\varepsilon_e \sim \mathcal{N}(0, \sigma^2)$ is the Gaussian noise with zero mean and variance $\sigma$, and the outlier indicator variable $\gamma_e \in \mathbb{R}$ is assumed to have a higher magnitude than $\sigma$. Here the outliers are the aggregated edges whose directions conflict with the true ranking. In order to estimate the $N + n$ unknown parameters ($N$ for $\gamma$ and $n$ for $\theta$), we aim to minimize the discrepancy between the annotation $y$ and the prediction $Z\Theta + \gamma$, as well as holding the outlier indicator $\gamma$ sparse. It gives us the following optimization problem:

$$\min_{\theta, \gamma} \ell_{w_0}(\Theta, \gamma) + \lambda \cdot R_{w_0}(\gamma),$$

where

$$\ell_{w_0}(\theta, \gamma) = \frac{1}{2} \left\| y - Z\Theta - \gamma \right\|_{w_0}^2 = \frac{1}{2} \sum_{e \in E} w^0_{ij} (y_{ij} - \gamma_{ij} - \theta_i + \theta_j)^2,$$

$$y = \text{ver}(Y), \ w^0 = \text{ver}(W^0)$$ is the vector form of $Y$ and $W^0$, and the weighted regularization term $R_{w_0}$ is

$$R_{w_0}(\gamma) = \|\gamma\|_{1, w_0} = \sum_{e \in E} w^0_{ij} |\gamma_e|.$$ 

In this situation, the weight $w^0$ and the label $y$ would be treated as the input data of the ranking problem [143]. Moreover, we introduce the variable $\beta = (\Theta, \gamma)^\top$ to define the action space of the ranking algorithm. We rewrite (143) as

$$\min_{\beta \in B} \ell_{w_0}(\beta),$$

where

$$\ell_{w_0}(\beta) = \frac{1}{2} \left\| y - \begin{bmatrix} Z & 1 \end{bmatrix} \begin{bmatrix} \Theta \\ \gamma \end{bmatrix} \right\|_{2, w_0}^2.$$
and
\[ B_\lambda = \left\{ \beta \left| \left( 0, w_0 \right), \left( \theta, \gamma \right) \right| \leq \varepsilon(\lambda) \right\} \]  

(148)
is the feasible set of [143] and the ranker’s action space.

We model poisoning attack as a game between two players, the ranker and an attacker, where the latter wants to mislead its opponent into picking parameters to generate a difference order against the true ranking. To disguise himself, the adversary needs to coordinate a poisoned \( w \) associate with \( y \). Intuitively, the adversary could not obtain \( w \) through drastic changes, neither on each \( w_{ij} \) nor \( \sum w_{ij} \). Such limitations lead to the following constraints for adversary’s action. First, the total difference between \( w_0 \) and \( w \) would be smaller than \( b \), namely,
\[ \|w - w_0\|_1 \leq b, \ b \in \mathbb{Z}_+. \]  

(149)
Furthermore, the adversary could not alter the number of votes on any pairwise comparison \( e \in E \) obviously,
\[ \|w - w_0\|_\infty \leq l, \ l \in \mathbb{Z}_+, \ l \leq \min\{\max(w_0), b\}, \]  

(150)
and the adversary’s action space \( \mathcal{W}_{w_0} \) is
\[ \mathcal{W}_{w_0} = \left\{ w \left| w \in \mathbb{Z}_+^N, l, b \in \mathbb{Z}_+, \|w - w_0\|_1 \leq b, \|w - w_0\|_\infty \leq l, l \leq \min\{\max(w_0), b\} \right. \} \]  

(151)
The robust ranking algorithm [143] observes the poisoned training set sampling from \( G \), prunes the outlier and learns a ranking from the remaining data simultaneously. Against the robust ranking algorithm that employs the defense described above, we can formulate the attacker’s goal as the following bi-level optimization problem:
\[
\begin{align*}
\text{maximize} & \quad \ell_w(\hat{\theta}, \hat{\gamma}) + \lambda \cdot R_w(\hat{\gamma}), \\
\text{subject to} & \quad \hat{\theta}, \hat{\gamma} \in \arg \min_{\theta, \gamma} \ell_w(\theta, \gamma) + \lambda \cdot R_w(\gamma), \\
& \quad w \in \mathcal{W}_{w_0}.
\end{align*}
\]  

(152)

**Distributional Perspective**

In [14], the \( \ell_1 \) and \( \ell_\infty \) distance constraints on \( w \) correspond to the attacker only being able to find the perturbation in the neighborhood of \( w^0 \). The lower problem in [152] corresponds to the robust pairwise ranking algorithm. With input data \( \{w, y\} \), the ranker obtain the relative score \( \theta \) by minimizing the discrepancy between the annotation \( y \) and the prediction \( Z\theta + \gamma \) while keeping \( \gamma \) to be sparse. Unfortunately, the bilevel nature of [152] [5], [6]—maximizing the outer loss involves an inner minimization to find the parameters \( \theta, \gamma \)—makes it difficult to solve, even less the discrete property of [12] and [13]. Next, we will discuss the poisoning attack on pairwise ranking in a different way. Generally, we can look at the poison attack [152] from a distributional perspective. The attacker and the ranker both access the weighted comparison graph \( G \) to play a game as [152]. Actually, the non-toxic training data \( \{w_0, y\} \) are drawn according to a probability distribution \( P \)
\[ p(w_0, y) = \sum_{e \in E} p(w_{0ij}, y_{ij}). \]  

(153)

The attacker chooses a perturbation function \( \psi : \mathbb{Z}_+^N \rightarrow \mathbb{Z}_+^N \) that change the weight \( w_0 \) to \( w \). The attacker constructs the perturbation \( \psi \) with the limitation as [14]. Such a perturbation \( \psi \) induces a transition from empirical distribution \( P \) to a poisoned distribution \( Q \)
\[ q(w, y) = q(\psi(w_0), y) = \sum_{e \in E} q(w_{ij}, y_{ij}). \]  

(154)
The attacker can only alter \( b \) pairwise comparisons at most, increase or decrease the number of vote on any comparison less than \( l \), and formulate the poisoned training set \( \{w, y\} \). If the attacker selects \( Q \) in a small enough neighborhood of \( P \), namely, the “distance” between the poisoned distribution \( Q \) and the empirical distribution \( P \) would be small, the attacker could obtain a good approximation of \( P \) in the sense of such a “distance” and the poisoned sample \( \{w, y\} \) would be satisfied the constraints [12] and [13].

Let \( \phi : \mathbb{R}_+ \rightarrow \mathbb{R} \) be a convex function with \( \phi(1) = 0 \). Then the \( \phi \)-divergence between distributions \( Q \) and \( P \) defined on a space \( \mathcal{X} \) is
\[
\begin{align*}
d_\phi(Q || P) &= \int \phi \left( \frac{dQ}{dP} \right) dP \\
&= \int_{\mathcal{X}} \phi \left( \frac{q(x)}{p(x)} \right) p(x) d\mu(x),
\end{align*}
\]  

(155)
where \( \mu \) is a \( \sigma \)-finite measure with \( Q, P \ll \mu \), and \( q = \frac{dQ}{d\mu}, p = \frac{dP}{d\mu} \). Given \( \phi \) and sample \( w_0 \), we reformulate the adversary’s action space, the local neighborhood of the empirical distribution \( P \) with radius \( \rho \) as

\[
Q_P = \{ \text{distribution } Q \text{ satisfies } d_\phi(Q||P) \leq \rho \},
\]

where \( P \) is the empirical distribution of the pairwise comparisons, and \( Q \) is the toxic distribution for poisoning attack. Throughout this paper, we adopt

\[
\phi(t) = \frac{1}{2}(t - 1)^2,
\]

which gives the \( \chi^2 \)-divergence \([18, 54, 68]\). It means that \( Q \) consists of discrete distributions supported on the observation \( \{(w^0_0, y)\} \). With opportunely chosen \( \rho \), the adversary could obtain \( w \) which satisfies the neighborhood constraints as \([12] \) and \([13] \).

The possible actions of two players \( a = [\beta, w] \) constitute the joint action space \( A = B_\lambda \times Q_P \) which is assumed to be nonempty, compact, and convex. Action spaces \( A \) are parameters of the game \([152] \). Then the bi-level integer programming \([152] \) can be written as a min-max optimization problem:

\[
\sup_{q \in Q_P} \inf_{\beta \in B_\lambda} \mathbb{E}_Q[f(\beta, q(w, y))] = \sup_{d_\phi(Q||P) \leq \rho} \inf_{\|q(w)\|_1 \leq \varepsilon(\lambda)} \mathbb{E}_Q[f(\beta, q(w, y))]
\]

where

\[
\begin{align*}
q(w) \circ \beta & = \begin{bmatrix} 0 & q(W) \end{bmatrix}_\beta, \\
W & = diag(w) \text{ is a diagonal matrix. } \\
\end{align*}
\]

We define the loss function \( f_{ij}(\beta, q(w, y)) \) by

\[
\begin{align*}
f_{ij}(\beta, q(w, y)) & = q(w_{ij}) \cdot \frac{1}{2}(y_{ij} - \gamma_{ij} - \theta_i + \theta_j)^2 + \lambda \cdot q(w_{ij}) |\gamma_{ij}| \\
& = q(w_{ij}) \cdot \left[ \frac{1}{2}(y_{ij} - \gamma_{ij} - \theta_i + \theta_j)^2 + \lambda \cdot |\gamma_{ij}| \right] \\
\end{align*}
\]

and the finite sum of \( \{f_{ij}\} \)

\[
f(\beta, q(w, y)) = \sum_{\ell \in E} f_{ij}(\beta, q(w, y))
\]

With fixed \( \lambda \) and some special form of \( q \), \([157] \) can be a convex problem. We swap the order of minimization and maximization in the min-max optimization problem \([157] \) as

\[
\inf_{\beta \in B_\lambda} \sup_{q \in Q_P} \mathbb{E}_Q[f(\beta, q(w, y))] = \inf_{\beta \in B_\lambda} \sup_{q \in \mathbb{Q}} \left\{ \mathbb{E}_Q[f(\beta, q(w, y))], \text{s.t. } d_\phi(Q||P) \leq \rho \right\}.
\]

In fact, the minimization of \([157] \) and \([160] \) correspond to the residual method and the Tikhonov regularization with discrepancy principle of the LASSO. Indeed, it can be shown that the constrained minimization problem is equivalent to Tikhonov regularization, when the regularization parameter \( \lambda \) is chosen according to Morozov’s discrepancy principle \([25] \). Note that the objective function of \([160] \) is a strictly convex function with respect to its arguments, then by \([7 \text{ Theorem 4.3} \), at least one Nash equilibrium exists.}

**Optimization**

From a game-theoretic viewpoint, \([160] \) can be seen as a zero-sum game between two agents: the agent ranker (the infimum) seeks to incur the least possible loss, while the agent adversary (the supremum) seeks to obtain the worst possible objective function value – both given by \( f(\beta, w, y) \).

For the supremum part of \([160] \), the integer characteristic of \( w \) obstructs obtaining a probability density function \( q \) of the toxic distribution \( Q \). Thanks to distributionally robust optimization, we reformulate the supremum part of \([160] \) as a quadratically constrained linear maximization problem. This tractable formulation can be solved by the probability simplex projection method. Suppose the total number of pairwise comparison without toxic is

\[
M^0 = \sum_{\ell \in E} w^0_{ij},
\]

and the frequencies of each comparison are

\[
p = \frac{1}{M^0} w^0.
\]

Let the maximum toxic dosage be \( \kappa \). It suggests that the number of toxic pairwise comparisons \( M \) satisfies

\[
M = \sum_{\ell \in E} w_{ij} \leq (1 + \kappa) \cdot M^0,
\]
Furthermore, we replace the toxic weight $w$ with its frequency $q = \frac{w}{M}$. We relax the integer programming problem into a general optimization by such a variable substitution. We note

$$z_e = \frac{1}{2}(y_{ij} - \gamma_{ij}^e - \theta_i^e + \theta_j^e)^2 + \lambda \cdot |\gamma_{ij}^e|, \quad e \in E$$

and $z = [z_1, \ldots, z_N] \in \mathbb{R}^N_+$. The objective function with fixed $\beta_\lambda$, maximizing the expectation $E_Q[f(\beta_\lambda, q(w, y))]$ equals to compute the worst-case linear combination of $\{z_{ij}\}, \quad e \in E$ as

$$\text{maximize } \langle q, z \rangle, \quad \text{s.t. } d_\phi(Q\|P) \leq \rho.$$  

As $Q$ is a distribution, it requires that the combination coefficients $q$ should satisfy

$$\sum_{e \in E} q_e = 1 \quad \text{or} \quad \langle 1, q \rangle = 1.$$  

It means that the distribution of $q$ is a probability simplex. Furthermore, as $P$ and $Q$ are the discrete distributions and we choose $\phi(t) = \frac{1}{2}(t - 1)^2$ in $[1]$, the neighborhood constraint $d_\phi(Q\|P) \leq \rho$ can be transformed as

$$\frac{1}{2}\|q - p\|_2^2 \leq \rho\|p\|_2^2.$$  

Now we obtain the following quadratically constrained linear maximization problem which could be used to compute the supremum problem in (160):

$$\text{maximize } \langle q, z \rangle, \quad \text{s.t. } q \in Q_p$$  

where

$$Q_p = \left\{ q \mid \frac{1}{2}\|q - p\|_2^2 \leq \rho\|p\|_2^2, \quad \langle 1, q \rangle = 1 \right\}.$$  

We reformulate the concave optimization (167) as a minimization problem for simplicity:

$$\text{minimize } \langle q, z \rangle, \quad \text{s.t. } q \in Q_p,$$  

and take a partial dual problem of this minimization, then maximize this dual problem to find the optimal $q$.

First, we introduce the dual variable $\mu \geq 0$ for the quadratical constraint (166). Notice that the strong duality exists for (169) because the Slater condition is satisfied by

$$q = p \quad \text{and} \quad 1^\top p = 1.$$  

Performing the standard min-max swap [13], it yields the following problem

$$\text{minimize } g(\mu) = \inf_{q \in Q_p} \left\{ \frac{\mu}{2}\|q - p\|_2^2 - \mu\rho\|p\|_2^2 + q^\top z \mid q \in \mathbb{R}^N_+ \right\}.$$  

Given a collection of concave functions $\{g_q\}_{q \in Q_p}$, if it attains

$$\inf g = \inf_{q \in Q_p} g_q$$

at some $q_0 \in Q_p$, we know that $\nabla g_{q_0}$ is the super-gradient of $g$ [29, Chapter VI.4.4]. Suppose $q(\mu)$ is the unique minimizer of the right hand side of (170), the dual function $g$ will be

$$g(\mu) = \frac{\mu}{2}\|q(\mu) - p\|_2^2 - \mu\rho\|p\|_2^2 + q(\mu)^\top z$$  

and the derivative with respect to $\mu$ (keeping $q(\mu)$ fixed) is

$$g'(\mu) = \frac{1}{2}\|q(\mu) - p\|_2^2 - \rho\|p\|_2^2.$$  

As the constraints $q \geq 0$ and $q^\top 1 = 1$ require $q$ is on the probability simplex, we adopt the Euclidean projection of a vector to the probability simplex [19]. Such a projection provides an efficient solver of the infimum (169). With no loss of generality, we assume that $z$ is an increasing sequence and the mean of $z$ is zero,

$$z_1 \leq z_2 \leq \cdots \leq z_N, \quad \langle z, 1 \rangle = 0.$$  

Then we use $a_i, \sigma_i \in \mathbb{R}^N_+$, the cumulative summation of $z$ and $z^2$ as

$$a_i = \sum_{j \leq i} z_j, \quad \sigma_i = \sum_{j \leq i} z_j^2, \quad i \in [N].$$  

$$z_1 \leq z_2 \leq \cdots \leq z_N, \quad \langle z, 1 \rangle = 0.$$  

Then we use $a_i, \sigma_i \in \mathbb{R}^N_+$, the cumulative summation of $z$ and $z^2$ as

$$a_i = \sum_{j \leq i} z_j, \quad \sigma_i = \sum_{j \leq i} z_j^2, \quad i \in [N].$$
The infimum in (170) is equivalent to projecting the vector $v(\mu) \in \mathbb{R}^N$ onto the probability simplex,

$$v_i = p_i - \frac{1}{\mu}z_i, \quad i \in [N]$$  \hspace{1cm} (176)

According to [19], $q(\mu)$ has the form as $q_i(\mu) = (v_i - \eta)_+$ for some $\eta \in \mathbb{R}$, where $\eta$ is selected such that $\sum q_i(\mu) = 1$. Finding such a value $\eta$ is equivalent to finding the unique index $i$ such that

$$\sum_{j=1}^{i}(v_j - v_i) < 1 \quad \text{and} \quad \sum_{j=1}^{i+1}(v_j - v_{i+1}) > 1.$$  \hspace{1cm} (177)

If no such index exists, we set $i = n$ as the sum $\sum_{j=1}^{i}(v_j - v_i)$ is increasing in $i$ and $v_1 - v_1 = 0$. Given the index $i$,

$$\eta = p_i - \frac{1}{i} - \frac{1}{i\mu} \sum_{j=1}^{i}z_i = p_i - \frac{1}{i} - \frac{1}{i\mu}a_i$$  \hspace{1cm} (178)

satisfies $\sum(v_i - \eta)_+ = 1$ and $v_j - \eta \geq 0$ for any $j \leq i$ while $v_j - \eta \leq 0$ for $j > i$. Meanwhile, the derivative $\frac{\partial}{\partial \mu} g(\mu)$ (where $q(\mu)$ is fixed) has a explicit form

$$g'(\mu) = \frac{\partial}{\partial \mu} \left\{ \frac{\mu}{2}||q(\mu) - p||^2 - \mu \rho ||p||^2 + q^\top(\mu)z \right\}$$

$$= \frac{1}{2} ||q(\mu) - p||^2 - \rho ||p||^2$$

$$= \frac{1}{2} \sum_{j=1}^{i}(v_j - \eta - p_j)^2 + \frac{1}{2} \sum_{j=i+1}^{N}p_j^2 - \rho ||p||^2$$

$$= \frac{1}{2} \sum_{j=1}^{i}(\frac{z_j}{\mu} + \eta)^2 + \frac{1}{2} \sum_{j=i+1}^{N}p_j^2 - \rho ||p||^2$$

$$= \frac{\sigma_i}{2\mu^2} + \frac{\eta^2 i}{2} + \frac{a_i \eta}{\mu} + \sum_{j=i+1}^{N}p_j^2 - \rho ||p||^2$$  \hspace{1cm} (179)

The derivative $g'(\mu)$ only needs $O(1)$ when $a$ and $\sigma$ are known. Binary search can calculate the optimal index $i$ and $q$ efficiently, which requires $O(\log \frac{1}{\epsilon} \log N)$ to find $\mu$ with accuracy $\epsilon$. We can get $\eta$ through (178) if (177) are satisfied. The solution $q(\mu)$ is

$$q_i = \left( p_i - \frac{z_i}{\mu} - \eta \right)_+, \quad i \in [N].$$  \hspace{1cm} (180)
Specifically, the computational complexity to obtain the sorted vector \( z \) is \( \mathcal{O}(N \log N) \), and that of the estimate of the frequency \( q \) is \( \mathcal{O}(\log \frac{1}{\varepsilon} \log N) \). The overall time computational complexity is \( \mathcal{O}(N \log N + \log \frac{1}{\varepsilon} \log N) \).

At last, we describe the whole optimization of the poison attack on pairwise ranking with Algorithm 4. We summarize the complete optimization procedure of the supremum in (160) as Algorithm 5 and Algorithm 3. For the infimum part of (160), the agent HodgeRank finds \( \beta \) that minimizes the regularized loss on \( \{w, y\} \) where the hyper-parameter \( \lambda \) controls the regularization strength. We include the solving process of HodgeRank as Algorithm 5 for completeness.

**Algorithm 3: WorstCase\( (z, p, \rho, \varepsilon) \)**

**Input**: the objective function value \( z \in \mathbb{R}^N \), the frequency of true comparisons \( p \in \mathbb{R}_+^N \), parameter \( \rho \), solution accuracy \( \varepsilon \).

1. Make \( z \) have the zero mean: \( z \leftarrow z - \bar{z} \), and sort \( z \).
2. Initialize \( \mu_{\min} = 0 \), \( a_i = \sum_{j \leq i} z_j \), and \( \sigma_i = \sum_{j \leq i} z_j^2 \) for all \( i \in [N] \),
   \[
   \mu_{\max} = \mu_{\infty} = \max \left\{ \|z\|_{\infty}, \sqrt{\frac{1}{\rho} \|p\|_2^2 \|z\|_2} \right\}
   \]
3. While \( |\mu_{\max} - \mu_{\min}| > \varepsilon \mu_{\infty} \) do
4.   Set \( \mu = \frac{1}{2}(\mu_{\min} + \mu_{\max}) \), and \( (\eta, i) = \text{FindShift}(z, p, a, \mu) \),
5.   Obtain the partial derivative \( g'(\mu) \) by (179)
6.   If \( g'(\mu) > 0 \) then
7.     \( \mu_{\min} \leftarrow \mu \)
8.   Else
9.     \( \mu_{\max} \leftarrow \mu \)
10. End
11. Set \( \mu = \frac{1}{2}(\mu_{\min} + \mu_{\max}) \), and \( (\eta, i) = \text{FindShift}(z, p, a, \mu) \),
12. Output: \( q \) by (180).

**Algorithm 4: FindShift\( (z, p, a, \mu) \)**

**Input**: the sorted and zero mean vector \( z \in \mathbb{R}^N \), the frequency \( p \), the cumulative sum \( a \), and the dual variable \( \mu \).

1. Initialize \( i_{\text{low}} = 1 \) and \( i_{\text{high}} = N \),
2. If \( p_N - \frac{1}{\mu} z_N \geq 0 \) then
3.   \( \eta = 0 \), \( i = N \),
4.   Break.
5. Else
6.   While \( i_{\text{low}} \neq i_{\text{high}} \) do
7.     \( i = \frac{1}{2}(i_{\text{low}} + i_{\text{high}}) \),
8.     \( a_{\text{left}} = \sum_{j=1}^{i-1} (v_j - v_i) = \frac{1}{\mu}(i z_i - a_i) \),
9.     \( a_{\text{right}} = \sum_{j=i+1}^{N} (v_j - v_{i+1}) = \frac{1}{\mu}((i+1) z_{i+1} - a_{i+1}) \),
10. If \( a_{\text{right}} \geq 1 \wedge a_{\text{left}} < 1 \) then
11.   \( \eta = p_i - \frac{1}{i} - \frac{1}{\mu} a_i \),
12.   Break.
13. Else if \( a_{\text{left}} \geq 1 \) then
14.   \( i_{\text{high}} = i - 1 \)
15. Else
16.   \( i_{\text{low}} = i + 1 \)
17. End
18. End
19. Output: \( i = i_{\text{low}}, \eta = p_i - \frac{1}{i} - \frac{1}{\mu} a_i \).
Algorithm 5: HodgeRank\((w, A, y)\)

**Input**: the weight \(w\), the comparison matrix \(A\) and the corresponding label \(y\).

1. Calculate the relative ranking score \(\hat{\theta}\)

   \[
   \hat{\theta} = (X^T X + \delta I)^{-1} X^T \sqrt{W} y
   \]

   where \(X = \sqrt{W} A, \sqrt{W} = \text{diag}(\sqrt{w})\).

**Output**: the corresponding ranking parameter \(\hat{\theta}\).