

The Attention Market: Interpreting Online Fair Re-ranking as Manifold Optimization under Walrasian Equilibrium

Chen Xu

Gaoling School of Artificial
Intelligence
Renmin University of China
Beijing, China
xc_chen@ruc.edu.cn

Fengran Mo

University of Montreal
Montreal, Canada
fengran.mo@umontreal.ca

Wei Chu

Gaoling School of Artificial
Intelligence
Renmin University of China
Beijing, China
wchu4112@gmail.com

Jun Xu*

Gaoling School of Artificial
Intelligence
Renmin University of China
Beijing, China
junxu@ruc.edu.cn

Wenyu Hu

School of Mathematics
Renmin University of China
Beijing, China
huwenyu_123@ruc.edu.cn

Maarten de Rijke

University of Amsterdam
Amsterdam, The Netherlands
m.derijke@uva.nl

Abstract

Fair re-ranking aims to promote long-tail items and enhance diversity within groups in information retrieval. While previous research on online fairness-aware re-ranking has shown promising outcomes, our comprehensive evaluation of online fair re-ranking methods over 20 settings reveals significant performance disparities among existing methods. To uncover the root causes of these inconsistencies, we reformulate fair re-ranking within an attentional market framework governed by a Walrasian Equilibrium, where the fairness is treated as a taxation cost. This market-based formulation is then coupled with manifold optimization, demonstrating that seeking this equilibrium is equivalent to performing gradient descent on a specific ranking manifold constructed by the market. Different re-ranking settings induce distinct manifold geometries, and these intrinsic geometric differences dictate the gradient landscapes and optimization trajectories. We propose ManifoldRank, an efficient online fair re-ranking algorithm. ManifoldRank adjusts gradients to align with the ranking manifold, considering various contextual settings. On the supply side, it incorporates a gradient adjustment based on different fairness requirements, accounting for associated costs. On the demand side, it empirically predicts an additional gradient adjustment term derived from the ranking scores. By integrating these two gradient adjustments, ManifoldRank effectively balances fairness and accuracy. Experimental results across multiple datasets confirm ManifoldRank's effectiveness.

CCS Concepts

• Information systems → Information retrieval.

*Jun Xu is the corresponding author.



This work is licensed under a Creative Commons Attribution 4.0 International License.
SIGIR '26, Melbourne, VIC, Australia
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2599-9/2026/07
<https://doi.org/10.1145/3805712.3809614>

Keywords

Re-ranking, Fairness, Attention economics, Manifold

ACM Reference Format:

Chen Xu, Wei Chu, Wenyu Hu, Fengran Mo, Jun Xu, and Maarten de Rijke. 2026. The Attention Market: Interpreting Online Fair Re-ranking as Manifold Optimization under Walrasian Equilibrium. In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '26)*, July 20–24, 2026, Melbourne, VIC, Australia. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3805712.3809614>

1 Introduction

Fair re-ranking plays a crucial role in promoting long-tail items and enhancing diversity within groups in information retrieval (IR) [20, 26]. Generally, re-ranking algorithms operate under either offline or online paradigms. Offline solutions typically optimize ranking slots based on global user information, thereby achieving better performances [39, 43]. However, real-world industrial systems operate under strict latency constraints, which necessitate the use of online solutions. To achieve this, prior work often uses online gradient descent to dynamically adjust ranking scores based on historical interaction data [24, 36, 40].

While previous work has demonstrated promising results, evaluations are often restricted to limited settings, such as a narrow range of datasets or base models. To investigate the stability and robustness of online fair re-ranking approaches, we conduct a comprehensive evaluation spanning over 20 re-ranking settings, including the seven most widely used datasets and 3 distinct base ranking models. Within each setting, we tuned over 50 times across a sweep of accuracy-fairness trade-off parameters. Detailed experimental settings are provided in Section 6.1. Figure 1 details the evaluation results, with models plotted on the x-axis and performance ranking on the y-axis (where lower values indicate better performance). The box plots depict the performance distribution, including the median, quartiles, and range. These results reveal a significant lack of consistency in algorithmic performance and the sensitivity of existing online approaches to fair re-ranking to diverse environments.

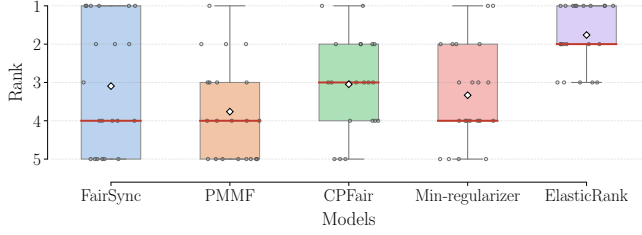


Figure 1: Performance rankings (lower is better) of fair on-line re-ranking models under multiple re-ranking settings, including seven datasets and three base models.

Table 1: Analogy between the economic market and the fair re-ranking market (attention economy).

Economic market	Fair re-ranking market
Buying products (<i>demand</i>)	Consuming attention (<i>users</i>)
Producing products (<i>supply</i>)	Producing attention (<i>providers</i>)
Binding income	Limited attention
Taxation mechanism	Fairness function
Selling price	Re-ranking score
Seeking equilibrium	Optimizing re-ranking
Market context	Manifold geometries

Theoretical analysis. Drawing from economic principles, we will apply the concept of attention economics to examine this issue. Attention economics is a theoretical framework that treats human attention as a scarce and valuable commodity that must be efficiently allocated [31]. Prior work [39, 40] has adopted this framework to model fairness as a form of economic taxation. We expand on this line of work and draw a parallel between the traditional economic market and the fair re-ranking market, see Table 1. In our formulation, user satisfaction functions as the market demand, where users seek to maximize the utility of their limited attention. Conversely, content provider fairness is modeled as a taxation mechanism, introducing costs for producing attention that the user needs. The optimal re-ranking score acts as the equilibrium price balancing these forces, a so-called Walrasian Equilibrium [3]. Crucially, we establish that seeking this equilibrium is equivalent to gradient descent on a specific re-ranking manifold. Consequently, performance inconsistency is explained by the fact that different re-ranking settings induce distinct manifold geometries and optimization trajectories. That is, we attribute the root cause of the performance inconsistency to the neglect of the underlying geometric properties of the optimization landscape in economics. See Section 4 for a detailed theoretical analysis of such a comparison.

Empirical analysis. To further investigate how different settings influence manifold geometries and optimization trajectories, we conduct an ordinal regression analysis to identify which factors of the ranking scores significantly contribute to algorithm performance inconsistency. Given the difficulty of directly modeling the manifold induced by ranking scores, we instead used statistical features of the ranking score matrix as proxy variables. Our results indicate a strong correlation between the entropy and skew of the scores and the gradient step size. Consequently, we propose to

calibrate the gradient magnitude in our online fair re-ranking algorithm by jointly considering the fairness (supply side) and statistical features of ranking scores (demand side).

Building on this observation, we propose ManifoldRank, an efficient online fair re-ranking algorithm that aligns optimization gradients with the ranking manifold by incorporating two adjustment terms. The algorithm first introduces a fairness-based gradient term on the supply side to identify the equilibrium direction. Second, it empirically estimates a gradient step adjustment derived from the statistical features of the scores to fit the underlying manifold structure from the demand side. By combining these two terms, ManifoldRank achieves both efficient and effective fair re-ranking. Extensive experiments verify the effectiveness of the ManifoldRank method across diverse settings. Our source code and all baselines are integrated into the toolkit FairDiverse [37] in GitHub <https://github.com/XuChen0427/FairDiverse>.

Main contributions. We summarize our major contributions:

- (1) We theoretically frame fair re-ranking as a Walrasian Equilibrium within an attention economy. We demonstrate that the process of seeking this market equilibrium can be viewed as a process analogous to gradient descent on a ranking manifold.
- (2) We propose an online fair re-ranking algorithm, ManifoldRank, which aligns optimization gradients with the manifold structure by incorporating two adaptive adjustment fairness terms.
- (3) Extensive empirical evaluations on seven publicly available ranking datasets demonstrate that ManifoldRank consistently surpasses state-of-the-art baselines.

2 Related Work

Fair re-ranking. Fairness is a central concern in information retrieval (IR), prompting extensive research into embedding fairness constraints within ranking systems [11, 12, 20, 33, 34]. Existing literature generally categorizes fair ranking algorithms into three pipeline stages. First, *pre-processing* methods, such as causal models [42] and probabilistic clustering [18], attempt to remove bias from input data prior to training [29]. Second, *in-processing* methods integrate constraints directly into the training process via techniques like re-weighting [15, 16] and regularization [44]. Third, *post-processing* methods (*i.e.*, re-ranking) adjust the final candidate list after the model has been trained [21, 26, 35, 36]. In this paper, we focus on fairness in the post-processing stage, as re-ranking offers a flexible and effective mechanism for ensuring fairness without re-training the underlying base ranking models.

Online fair re-ranking. Re-ranking strategies can be split into offline and online paradigms. While offline methods [9, 10, 21, 26, 39] use mixed-integer linear programming for global optimization, they are computationally expensive and impractical for real-time needs. Consequently, online algorithms are essential for scenarios requiring immediate user satisfaction [36]. MACE proposes causal embedding learning for fair streaming IR systems to improve OOD generalization [45]. These include heuristic methods like *CP-Fair* [24], which use greedy solutions to obtain the fair scores. On the other hand, gradient-based methods [6, 36, 38, 40] that use online optimization. For example, *P-MMF* [36] employ online learning techniques to optimize fairness objectives via subgradients, and *ElasticRank* [40] uses the elasticity of ranking as

the online gradient. However, current online approaches tend to be ranking-score-free; they fail to account for the geometry of the optimization manifold constructed by the ranking scores.

Multi-stakeholder fairness. Depending on the stakeholders involved [1, 2], re-ranking fairness can be categorized into user-side or provider-side. User fairness typically focuses on invariance to sensitive user attributes [19, 28, 32]. Item (or provider) fairness, however, aims to make item exposure more equitable [36, 40]. We prioritize item fairness in this work because it presents a unique challenge in online settings: unlike user fairness, item fairness requires managing cumulative exposure that accrues dynamically over a period of time, such as one day.

Economics and fair re-ranking. Fairness has long been a central subject of study in economics, particularly through the lenses of taxation and social welfare functions [25]. Economic study the fairness mainly through the demand-supply equilibrium analysis method [4]. In the context of ranking, prior work [30] treats fair ranking as a resource allocation problem, formulating it via the economic concept of Nash Social Welfare [8]. Similarly, other studies [39, 40] model fair re-ranking as a taxation process, a key mechanism for redistribution. However, these approaches typically adopt economic concepts only at a conceptual level. In contrast, in this paper we propose a rigorous mathematical comparison establishing the formal equivalence between economic theory and fair re-ranking formulations.

3 Preliminaries

In this section we define the fair re-ranking tasks and introduce our framework for modeling economic markets.

3.1 Fair re-ranking

In re-ranking tasks, let $\mathcal{U}$ denote the set of users and $\mathcal{I}$ denote the set of items. Each item $i \in \mathcal{I}$ belongs to a unique group $g \in \mathcal{G}$, and the set of items within a specific group g is represented as $\mathcal{I}_g$. When a user $u \in \mathcal{U}$ accesses the system, a candidate list $L_K(u)$ is generated with ranking scores $s_{u,i}$ for each item $i \in L_K(u)$, K is the size of the candidate list. The ranking score $s_{u,i}$ represents the relevance of item i to user u , is typically pre-computed based on ranking models, such as *DMF* [41], *SASRec* [17].

To formulate the re-ranking decision, we introduce a binary decision variable $x_{u,i} \in \{0, 1\}$. Here, $x_{u,i} = 1$ indicates that item i is displayed to user u (i.e., selected in the final ranked list), and $x_{u,i} = 0$ otherwise. Consequently, the user and item group utilities are determined by the displayed items. The item group utility v_g and user utility w_u are defined as the accumulated scores of the displayed items:

$$v_g = \sum_{u \in \mathcal{U}} \sum_{i \in L(u)} x_{u,i} \cdot s_{u,i} \cdot \mathbb{I}(i \in \mathcal{I}_g), \quad w_u = \sum_{i \in L(u)} x_{u,i} \cdot s_{u,i}, \quad (1)$$

where $\mathbb{I}(\cdot)$ is the indicator function.

The objective of the fair re-ranking problem is to maximize the overall user utility (i.e., social welfare) while ensuring that the exposure or utility distributed among different item groups satisfies fairness constraints. This problem can be formulated as a constrained optimization problem:

$$\max_{x_{u,i} \in \mathcal{X}} \sum_{u \in \mathcal{U}} w_u, \text{ such that } v_r \approx v_p, \quad \forall r, p \in \mathcal{G}, \quad (2)$$

where $\mathcal{X} = \{x_{u,i} \mid x_{u,i} = K, x_{u,i} \in \{0, 1\}\}$ represents the set of binary decision variables. The first constraint enforces that the utility difference between any pair of groups r and p remains within an acceptable range. To enforce the constraint $v_r \approx v_p$, a standard approach involves maximizing a fairness function $r'(v)$, which is formulated such that its optimum yields $v_r \approx v_p$.

Typical online fair re-ranking methods aim to learn a fair re-ranking score $f_{u,g}$ with the ranking score to achieve an objective like the following:

$$L_K(u) = \arg \max_{S \in \mathcal{I}^K} \sum_{i \in S} (s_{u,i} - \hat{f}_{u,g}), \quad (3)$$

where S is the subset of $\mathcal{I}$ with size K and g indicates that item i belongs to the group g : $i \in \mathcal{I}_g$.

3.2 Economic markets

In microeconomics, market forces are driven by the interaction between buyers (demand) and sellers (supply). Figure 2(a) provides an intuitive illustration of this concept.

1. *The demand curve (The buyer)* represents how much consumers want to buy. It follows the *Law of Demand*: when prices are lower, people buy more.

Example: Imagine buying apples. If an apple costs \$5, you might buy none. If it drops to \$0.50, you might buy 10 to make a pie. Mathematically, the quantity demanded (Q_d) is a decreasing function of price (P). The mathematical formulation is:

$$Q_d = f_d(P), \quad \frac{dQ_d}{dP} \leq 0.$$

2. *The supply curve (The seller)* represents how much producers are willing to sell. It follows the *Law of Supply*: higher prices make production more profitable, so producers will supply more product.

Example: Imagine you are an apple farmer. If apples sell for \$0.10, it is not worth the effort, so you supply 0. If they sell for \$5, you work harder to sell as many as possible. Mathematically, the quantity supplied (Q_s) is an increasing function of price:

$$Q_s = f_s(P), \quad \frac{dQ_s}{dP} \geq 0.$$

In fact, the supply curve is not determined solely by producers' production costs. In real-world economies, governments often regulate supply behavior through policy instruments such as taxes or subsidies [4]. When the government imposes a per-unit tax τ on a commodity, the producer's effective revenue from selling each unit decreases from P to $P - \tau$. As a result, producers' willingness to supply declines at every price level, which is reflected by an overall leftward shift of the supply curve.

Example: Consider the apple market again. Suppose the government imposes an environmental tax of \$0.50 per kilogram of apples in order to reduce pesticide residues. Even if the market price remains \$2.0 per kilogram, apple growers' effective net revenue is only \$1.50 yuan per kilogram. This additional cost may force some growers to exit the market or reduce their planting scale, causing the supply curve to shift leftward. The tax revenue collected by the

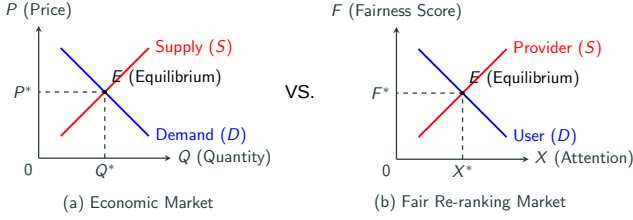


Figure 2: Parallels between (a) an economic market, where the x-axis represents production/consumption quantity and the y-axis represents selling price, and (b) a fair re-ranking market, where the x-axis represents the user attention for an item and the y-axis represents the fair re-ranking score.

government can then be used to subsidize ecological growers or low-income consumers, thereby achieving resource redistribution across groups.

3. *Market Equilibrium* occurs when the buyer's desire matches the seller's willingness. At this optimal price, there is no shortage and no surplus.

Example: At a price of \$2, the farmer brings exactly 100 apples, and customers want exactly 100 apples. If the price rises above \$2, a surplus of unsold apples forces the farmer to lower the price to clear his stock. Conversely, if the price falls below \$2, a shortage creates competition among buyers, naturally driving the price back up until it returns to the equilibrium level. Mathematically, we set demand equal to supply: $Q_d(P^*) = Q_s(P^*)$, and solve for the equilibrium point: $E = (Q^*, P^*)$.

Achieving market equilibrium represents a state of allocative efficiency where total economic surplus is maximized, corresponding to a Pareto optimal allocation of resources as described by the First Fundamental Theorem of Welfare Economics [25]. In the next section we use these basic notions to formalize the online fair re-ranking problem.

4 Theory: Attention Market for Fair Re-ranking

In this section, we reformulate the fair re-ranking task as a market for selling and consuming users' attention, while explaining why this market-based perspective is crucial for optimizing fair re-ranking. We model the fair re-ranking market by comparing it to the economic market discussed in Section 3.2.

Overall idea. The key distinction between a traditional economic market and fair re-ranking lies in redefining exchange variables. The price is interpreted as the fairness score f_g , regulating the trade-off between accuracy and fairness, while the quantity corresponds to user attention $x_{u,i}$, treated as a scarce resource allocated to items (see Figure 2(b)). We next elaborate on this formulation.

4.1 Demand side

For each user u , we can rewrite their goal to maximize their individual utility for paying attention cost (i.e., global fairness score) f_g for group $g \in \mathcal{G}$:

$$X^d = \arg \max_{x_{u,i} \in \mathcal{X}} \sum_{i \in \mathcal{I}} \underbrace{s_{u,i}}_{\text{ranking score}} x_{u,i} - \underbrace{f_g}_{\text{price of paying attention}} x_{u,i}. \quad (4)$$

The user u operates under a constrained attention budget (e.g., the capacity to view only K items) and allocates this scarce resource to maximize their cumulative utility derived from $s_{u,i}$. Within this framework, the fairness score f_g functions as the market price, representing the cost incurred when allocating attention $x_{u,i}$ to a specific item.

THEOREM 1 (LAW OF DEMAND). Let X_g^d denote the aggregate attention consumed by users. We have the following law of demand. User attention consumption is inversely related to the fairness score f_g , implying:

$$X_g^d = \sum_u \sum_{i \in \mathcal{I}_g} x_{u,i} \mathbb{I}(i \in \mathcal{I}_g), \quad \frac{\partial X_g^d}{\partial f_g} \leq 0.$$

A detailed proof is provided in Appendix A. Intuitively, a higher price f_g reduces the attractiveness of item i , leading to less or equal attention allocation $x_{u,i}$, which in turn decreases or maintains the consuming attention X_g^d for provider g .

4.2 Supply side

For the supply side, each provider g produces attentions that user needs based on the price of getting attention (i.e., fairness score):

$$X_g^s = \arg \max_{x_{u,i} \in \mathcal{X}} \sum_u \sum_{i \in \mathcal{I}_g} \underbrace{f_g}_{\text{price of getting attention}} x_{u,i} - \underbrace{r(v_g)}_{\text{taxation cost}}. \quad (5)$$

In this formulation, $r(v_g)$ represents the fairness penalty function for different group utility v_g (defined in Eq. (1)), which relates to the attention $x_{u,i}$. Please note that f_g here should be consistent with Equation (4), as the total revenue from sales must equal the total amount consumed.

Conceptually, this parallels the mechanism of taxation in economics [40], where a cost is imposed to discourage inequitable resource allocation. Next, we define the $r(v_g)$ as:

$$r(v_g) = \text{sign}(\alpha\beta) v_g^\beta \left(\sum_g v_g^\beta \right)^{\alpha-1}, \quad (6)$$

where $\alpha \in \mathbb{R}$ is the global taxation and $\beta \in \mathbb{R}$ is the local taxation rate. This definition will ensure the previous fairness function is equivalent to the taxation cost of the market, according to the following theorem:

THEOREM 2 (FAIRNESS FUNCTION AS TAXATION). The total market cost $r'(v) = \sum_g r(v_g)$ satisfies the taxation requirement [4]: The first derivative represents the marginal tax rate. This condition implies that the tax liability is strictly increasing with income: $\frac{\partial r'(v)}{\partial v_g} \geq 0$.

Once we have established this supply side, we can effectively generalize the market cost $r'(v)$ to encompass any fairness metric listed in Table 2. And we can easily obtain the Law of Supply:

THEOREM 3 (LAW OF SUPPLY). Let X_g^s be the aggregate attention allocated to provider g . We posit a law of supply such that the attention received increases with the fairness score f_g , yielding:

$$X_g^s = \sum_{i \in \mathcal{I}_g} x_{u,i} \mathbb{I}(i \in \mathcal{I}_g), \quad \frac{\partial X_g^s}{\partial f_g} \geq 0.$$

Table 2: Generalization of fairness metrics via α and β

Fairness metric $r'(v)$	α	β
Max-min fairness [36] $\min_g(v_g)$	∞	$-\infty$
p -norm [7] $(\sum_{i=1}^g v_i^p)^{1/p}$	$1/p$	p
Elastic fairness [40] $\text{sign}(1-t) (\sum_{i=1}^g \bar{v}_i^{1-t})^{(1/t)}$	$1/t$	$1-t$
α -fairness [39] $\sum_i v_i^{1-\alpha}$	1	$1-\alpha$
Proportional fairness [8] $\sum_i \log(v_i)$	1	$\beta \rightarrow 0$

A detailed proof is provided in Appendix C. Intuitively, a higher price f_g increases the likelihood of item i capturing attention due to its attractiveness. Consequently, the attention X_g^s obtained by provider g is non-decreasing.

4.3 Walrasian Equilibrium of two sides

Next, we show that optimizing the fair re-ranking process is equivalent to seeking the Walrasian Equilibrium of the previously introduced demand and supply sides, where the fairness objective is viewed as a single total taxation.

First, fair re-ranking algorithms often seek to learn an optimal re-ranking score f_g , which corresponds to the price of attention defined earlier. Next, we use the following theorem to prove that the fair re-ranking process is equivalent to the Walrasian Equilibrium of the demand and supply sides described above.

THEOREM 4 (WALRASIAN EQUILIBRIUM IN RE-RANKING). *Assuming $r'(v)$ is convex, we have:*

$$\max_{x_{u,i} \in \mathcal{X}} \sum_u w_u - r'(v) \iff \max_{x_{u,i} \in \mathcal{X}} \sum_u \sum_i (s_{u,i} - f_g^*) x_{u,i}, \quad (7)$$

where $\mathcal{X} = \{x_{u,i} \mid \sum_i x_{u,i} \leq K, x_{u,i} \in \{0, 1\}\}$ and w_u is the user utility defined in Section 3.1.

Moreover, assume that the price f_g^* is such that

$$\sum_g X_g^d(f_g^*) = \sum_g X_g^s(f_g^*) = |\mathcal{U}|K.$$

Then the allocation for demand and supply $\{X_g^d, X_g^s\}$ is Pareto optimal. That is, there does not exist another re-ranking allocation:

$$\forall u, w_u(X^d) \leq w_u(\tilde{X}^d), \quad \exists u_k, w_{u_k}(X^d) < w_{u_k}(\tilde{X}^d).$$

A detailed proof can be found in Appendix D. Intuitively, if the fair re-ranking score f_g^* (serving as the price) deviates from the optimum, market forces will correct it. The price will either rise in response to excess demand or fall in response to excess supply, driving the system toward improved accuracy and fairness.

Next, we use the following theorem to prove that the fairness learning objective is essentially equivalent to the single taxation mechanism. This result constitutes an equivalent formulation of the Second Fundamental Theorem of Welfare Economics [25]:

THEOREM 5 (ECONOMIC EXPLANATION FOR FAIRNESS). *Assuming $r'(v)$ is convex, the fairness objective is equivalent to implementing a single taxation mechanism T_g . In this context, $T_g > 0$ represents a tax levied on “rich” providers, while $T_g < 0$ means a subsidy for “poor” providers. Eventually, it will reach the Pareto optimal point.*

A detailed proof is provided in Appendix E. This theorem demonstrates that the concept of economic fairness aligns closely with the

Algorithm 1: Algorithm of ManifoldRank

Input: User set $\mathcal{U}$, item set $\mathcal{I}$, group set $\mathcal{G}$, ranking size K , taxation parameters α, β , demand parameters $a_e, a_s > 0$, user-item ranking score $s_{u,i}, \forall u \in \mathcal{U}, \forall i \in \mathcal{I}$

- 1: Set $v_g = 1, \forall g \in \mathcal{G}$
- 2: **for** $u \in \mathcal{U}$ **do**
- 3: // Gradient computation
- 4: Valid direction $c_g = \text{sign}((\alpha - 1)\beta v_g^\beta + (\beta - 1)(\sum_i v_g^\beta))$
- 5: Get normalized factor $G = \sum_j x_j^\beta$
- 6: Supply gradient: $\eta_g = cr(v_g)\beta \left(\frac{1}{v_g} + (\alpha - 1)\frac{v_g^{\beta-1}}{G} \right)$
- 7: Demand gradient $\zeta = a_e \text{Entropy}(s) - a_s \text{Skew}(s)$
- 8: // Online re-ranking
- 9: $L_K(u) = \arg \max_{S \in \mathcal{I}^K} \sum_{i \in S} (s_{u,i} - \eta_g * \zeta), i \in \mathcal{I}_g$
- 10: $v_g = v_g + \sum_{i \in L_K(u)} s_{u,i} \mathbb{I}(i \in \mathcal{I}_g), \forall g \in \mathcal{G}$
- 11: **end for**

objectives of fair re-ranking. This theorem implies that a fundamental requirement for the fairness function is to redistribute utility, effectively taxing high-resource entities to compensate those in the lower-resource brackets. This theoretical connection motivates our analysis of existing fair re-ranking algorithms.

4.4 Regarding attention market as manifold

In this section, we will use the manifold concept to describe the process for optimizing re-ranking. In advanced economic theory, rather than viewing a market merely as a system of linear constraints or a static vector in $\mathbb{R}^{|\mathcal{G}|}$, we can model the entire set of feasible market states as a smooth manifold [5].

Let the state of the economy be defined by a vector x (representing allocations, prices, or attention distributions). The structural constraints of the economy define the geometry of $\mathcal{M}$:

$$\mathcal{M} = \{f \in \mathbb{R}^{|\mathcal{G}|} \mid \sum_g X_g^d(f) = \sum_g X_g^s(f)\}.$$

Under this formulation, the “Market” is not just a point, but a geometric landscape (re-ranking scores and fairness objective) for finding the optimal price. Conducting gradient descent on such a manifold is subject to constraints from two sources: the supply side, governed by fairness parameters α and β , and the demand side, determined by the ranking score attributes $s_{u,i}$. Both factors influence the trajectory of the gradient step.

However, existing online fair re-ranking methods often overlook these constraints within the context of Demand-Supply Equilibrium, leading to the unstable performance illustrated in Figure 1.

5 Method: ManifoldRank

Building on the previously discussed theory, this section focuses on developing a method named ManifoldRank. The overall idea can be seen in the Algorithm 1.

5.1 Online re-ranking

During online re-ranking, whenever a user u enters the system, ManifoldRank will compute the fairness score:

$$L_K(u) = \arg \max_{S \in I^k} \sum_{i \in S} (s_{u,i} - \eta_g(v_g) * \zeta(v_g)), \quad (8)$$

where $i \in I_g$ and v_g is updated after returning a ranked list:

$$v_g = v_g + \sum_{i \in L_K(u)} s_{u,i} \mathbb{I}(i \in I_g), \forall g \in \mathcal{G}.$$

The fairness score is composed of $\eta_g(v_g)$ and $\zeta(v_g)$, which can be interpreted as the gradients on the supply side and demand side, respectively. They will be detailed in the next two subsections.

A comment on efficiency. ManifoldRank derives the gradient in closed-form, necessitating only fundamental arithmetic operations. This ensures high computational efficiency and makes the approach well-suited for real-time online re-ranking.

5.2 Gradient on the supply side

From the supply side formulation in Section 4.2, we can observe that the price (i.e., fairness score) is related to the cost function $r(v_g)$. Overall, the gradient can be written as:

$$\eta_g = c_g * f_g, \quad (9)$$

where c_g is the gradient direction and f_g is the gradient value.

Gradient value. First, according to the supply side formulation in Eq. (5), we can observe that the price of getting attention f_g is positively related to the first order of the cost function $r(v_g)$:

$$f_g = \frac{\partial r(v_g)}{\partial x_{u,i}} = \frac{\partial r(v_g)}{\partial v_g} \frac{\partial v_g}{\partial x_{u,i}} \propto \frac{\partial r(v_g)}{\partial v_g} = r(v_g) \gamma_g, \quad (10)$$

where $\frac{\partial v_g}{\partial x_{u,i}} > 0$ means allocating more user attention to the provider consistently enhances their utility, and γ_g can be regarded as the manifold degree of curvature:

$$\gamma_g = \beta \left[\frac{1}{v_g} + (\alpha - 1) \frac{v_g^{\beta-1}}{G} \right], G = \sum_g v_g^\beta,$$

where G can be viewed as the taxation normalization for all groups.

Gradient direction. Theorem 4 and 5 both require that the total cost function $r'(v)$ is convex. Therefore, we need to adjust the direction $c_g = \{-1, 1\}$ to make the term $\frac{\partial^2 r'(v)}{\partial v_g^2}$ be positive:

$$c_g = \text{sign}\left(\frac{\partial^2 r'(v)}{\partial v_g^2}\right) = \text{sign}\left((\alpha - 1)\beta x_i^\beta + (\beta - 1)\left(\sum_i x_i^\beta\right)\right). \quad (11)$$

5.3 Gradient on the demand side

The demand side presents a greater level of complexity. As outlined in Section 4.1, the impact on price is associated with the ranking score $s_{u,i}$, and the constraints form a simplex structure. Consequently, we opt to use the statistical information of $s_{u,i}$ for each user u : $\mathbf{s} = [s_{u,i}]_{i=1}^{|I|}$ to empirically assess the effect:

$$\zeta = a_e \text{Entropy}(\mathbf{s}) - a_s \text{Skew}(\mathbf{s}), \quad (12)$$

where $\text{Entropy}(\mathbf{s})$, $\text{Skew}(\mathbf{s})$ are the the entropy and skewness of $\mathbf{s}$ and $a_e > 0$ and $a_s > 0$ are the hyper parameters.

Intuitively, higher entropy indicates greater diversity in user preferences, meaning that implementing fairness taxation on such users will have a minimal impact on performance. Conversely, higher skewness suggests that users prefer a very limited range of items, and imposing a higher taxation in this case could significantly harm performance. The influences of other statistical information is discussed in Section 6.3.4.

Discussion. We note that the connection to manifold-based optimization is an empirical approximation rather than a strict theoretical equivalence. While the supply-side formulation is derived analytically, the demand-side term ζ is constructed using statistical properties as a practical proxy for the complex simplex-constrained structure of user preferences. This design trades theoretical completeness for tractability and efficiency in online settings. Although heuristic, our experiments across diverse datasets suggest that this approximation captures stable and transferable patterns in practice, while we leave more principled formulations as future work.

6 Experiments

We evaluate ManifoldRank using seven public ranking datasets.

6.1 Experimental settings

Dataset. Our experiments are based on seven large-scale, public datasets: **Steam** [17]: a ranking dataset for games on the Steam platform; **Five Amazon subsets** [13]: a widely used recommendation dataset, and **AliEC**¹: a large-scale e-commerce ad recommendation dataset. The detailed information are listed as following.

- **Steam** [17]: a ranking dataset for games on the Steam platform. We use the data for games played for more than 10 hours in our experiments. The publishers of games are considered item groups. It has 169,160 samples, which contain 4,444 users, 1,237 items, and 43 item groups.²

- **Five Amazon subsets** [13]: a widely used recommendation dataset that consists of a variety of product categories and includes interactions such as product ratings, and purchase histories. They all use brands as item groups. After pre-processing, they include: *Digital music domains*: 46,640 samples, with 873 users, 11,952 items, and 33 item groups; *Fashion domains*: 7,985 samples, with 629 users, 3,915 items, and 11 item groups; *Software domains*: 13,115 samples, with 583 users, 2,707 items, and 34 item groups; *Industrial and scientific domains*: 16,585 samples, with 1,078 users, 1,756 items, and 29 item groups; *Arts, crafts, and sewing domains*: 113,365 samples, with 3,690 users, 3,109 items, and 100 item groups.³

- **AliEC**: a large-scale e-commerce ad recommendation dataset. The categories of items are considered as item groups. After pre-processing it has 505,430 samples, with 21,242 users, 17,310 items, and 117 item groups.⁴

During the pre-processing step, users that have interactions with fewer than L items are excluded from the entire dataset to mitigate the issue of extreme sparsity. We set $L = 8$ for Amazon-Arts-Crafts-and-Sewing, $L = 5$ for AliEC, Steam, and Amazon-Industrial-and-Scientific, and $L = 3$ for the other three datasets. Similarly, for the

¹<https://tianchi.aliyun.com/dataset/56>

²http://cseweb.ucsd.edu/~wckang/Steam_games.json.gz.

³<http://jmcauley.ucsd.edu/data/amazon/>.

⁴<https://tianchi.aliyun.com/dataset/56>

item side, we set $L = 8$, $L = 5$, and $L = 1$, respectively. Following Xu et al. [40], we consider groups with fewer than 10 items as a single group, which we name the “infrequent group”. Following [24, 36], we sort all interactions by time and use the first 80% of the interactions as data to train the base ranking model. The remaining 20% of interactions are used as the test data for re-ranking tasks.

Evaluation. The performance of the models is evaluated on two aspects: re-ranking accuracy and fairness degree. For accuracy, following [36, 40], we use NDCG@K:

$$\text{NDCG@K} = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{\sum_{i \in L_K^F(u)} s_{u,i} / \log(\text{rank}_i^F + 1)}{\sum_{i \in L_K(u)} s_{u,i} / \log(\text{rank}_i + 1)}, \quad (13)$$

where $L_K(u_t)$ is the original ranked list and $L_K^F(u_t)$ is the fair-aware re-ranked list, and rank_i and rank_i^F are the ranking positions of the item i in $L_K(u_t)$ and $L_K^F(u_t)$, respectively.

To evaluate fairness, we use three distinct metrics: Elastic Fairness (EF) [40], the Gini Index (GINI) [10], and Max-Min Fairness (MMF) [36]. We adopt the MMF@K formulation from [10], which quantifies the aggregated ranking score of the bottom 20% “worst-off” groups, thereby measuring the system’s ability to protect the most disadvantaged entities. Higher values for EF@K and MMF@K, coupled with a lower GINI@K, indicate a more fair distribution of provider groups.

Baselines. The following representative online re-ranking models were chosen as baselines: *CPFair* [24] formulates the re-ranking task as a knapsack problem and employs a greedy strategy; *Min-regularizer* [36] introduces a regularization term that penalizes exposure gaps between target providers and worst-off groups; *P-MMF* [36] uses mirror gradient descent specifically to improve the utility of the worst-off item groups; *ElasticRank* [40], which determines re-ranking priority by calculating a fairness score based on the exposure elasticity between two item groups; and *FairSync* [38], which implements an online gradient descent framework designed to guarantee minimum exposure thresholds for every item group.

Implementation details. Our experiments are implemented using Python 3.9. All experiments are conducted on a server with Ubuntu 18.04. All our experiments are conducted based on the *FairDiverse* toolkit [37]. As for the hyperparameters in all models, the fairness parameter β, α is tuned among $[-3, 3]$. The demand parameter a_e, a_s is tuned among $[0.1, 1]$.

6.2 Experimental results

We report on the performance of ManifoldRank and the baselines. Our main experiments are conducted based on the well-performing ranking base model SASRec [17].

6.2.1 Fairness performance comparison. To ensure a rigorous evaluation, we follow the methodology established by ElasticRank. We conduct experiments to assess the trade-off between accuracy and fairness for ManifoldRank and our selected baselines. Specifically, we constrain the accuracy (measured by NDCG@K) to remain above 99% of the base model’s performance. This approach is motivated by the fact that an accuracy loss of less than 1% is typically considered negligible in production environments.

Comprehensive results are shown in Table 3 details the experimental outcomes for ManifoldRank and the baseline methods. We

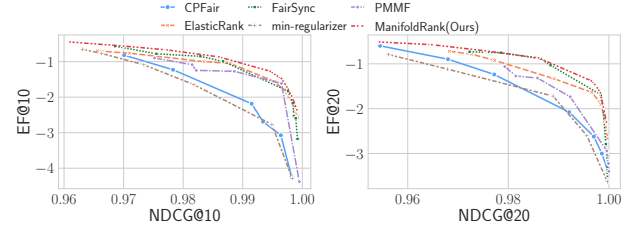


Figure 3: Pareto frontier with different size K under Amazon industrial and scientific domains on SASRec.

use datasets with different data distributions to ensure the robustness of our model. From Table 3, we can observe that ManifoldRank outperforms almost all baselines in terms of all fairness metrics on all datasets (except for 1 out of 42 settings), including EF@K, GINI@K, and MMF@K. The experimental results demonstrate that ManifoldRank achieves superior fairness performance at the same accuracy level, highlighting our effectiveness and robustness.

6.2.2 Performance on Pareto frontiers. Figure 3 illustrates the Pareto frontiers [22] between the accuracy metric NDCG@K and the fairness metric EF@K at $K = 10$ and $K = 20$. Due to space limitations, we report results on the Amazon Industrial and Scientific domains; similar patterns are consistently observed across other datasets and base models. The Pareto frontiers are obtained by systematically varying model parameters and selecting configurations that jointly optimize NDCG@K and EF@K, thereby characterizing the trade-off between ranking accuracy and item fairness. As expected, a clear trade-off emerges: improving fairness often comes at the cost of accuracy, and vice versa. Compared with baseline methods, the proposed ManifoldRank consistently achieves Pareto-superior performance, as its curves occupy the upper-right region of the frontier. This dominance indicates that, for a given NDCG@K level, ManifoldRank attains higher EF@K, and for a given EF@K level, it delivers better NDCG@K. These results demonstrate the substantial advantage of ManifoldRank over existing baselines.

6.3 Experimental analysis

In this section, we report on an experimental analysis on Amazon Industrial and Scientific domains; similar patterns are consistently observed across other datasets and base models.

6.3.1 Ablation study on different base models. To verify the generalizability of ManifoldRank, we integrate it with three distinct base ranking models to obtain the preliminary preference scores $s_{u,i}$. These include: (i) BPR [27], a pairwise collaborative filtering model; (ii) GRU4Rec [14], a recurrent neural network designed for session-based patterns; and (iii) SASRec [17], a transformer-based model that uses attention mechanisms for sequential modeling.

Table 4 presents a comparative analysis of ManifoldRank against the two strongest baselines, FairSync and ElasticRank, across various base ranking architectures. Following the experimental configuration established in Table 4, we observe that ManifoldRank consistently yields superior performance. These results validate the robustness of our framework, demonstrating that its effectiveness is independent of the underlying base model. This performance gain is largely attributed to our gradient formulation, which explicitly

Table 3: Performance comparison of ManifoldRank and the baselines on seven datasets, where “Music”, “Fashion”, “Software”, “Industrial”, and “Arts” denote five Amazon subsets. We tuned various re-ranking models to achieve accuracy performances (NDCG) exceeding 99%, and evaluated fairness across two cut-offs K to assess three fairness metrics (EF@ K , GINI@ K , MMF@ K). Bold number indicates that the improvements over the baselines are statistically significant (t-tests and p -value < 0.05).

Metric	Datasets	Steam		Music		Fashion		Software		Industrial		Arts		AliEC	
	Ranking size	K=10	K=20	K=10	K=20	K=10	K=20	K=10	K=20	K=10	K=20	K=10	K=20	K=10	K=20
EF	CPFair	-12.797	<u>-2.0865</u>	-15.884	-4.1494	-4.6663	-6.2038	-13.487	-2.0288	-2.1845	-2.0778	-0.8414	-1.1086	-21.618	-14.428
	Min-regularizer	-20.326	-4.9795	-15.907	-6.0151	-6.8100	-5.8191	-1.9039	-3.0153	-2.7667	-2.5984	-1.6493	-1.1742	-21.623	-16.551
	P-MMF	-26.329	-2.1071	<u>-2.8046</u>	-3.0857	-4.5436	-5.7765	-1.9605	-1.8360	-1.6128	-1.7352	-1.1208	-1.6035	-8.7536	-9.2907
	FairSync	<u>-3.5548</u>	-5.9235	-2.8208	<u>-2.7674</u>	-4.6886	<u>-4.4956</u>	<u>-1.0825</u>	<u>-1.2412</u>	<u>-1.2458</u>	<u>-1.1884</u>	<u>-0.6494</u>	-0.8975	-3.8817	-5.5189
	ElasticRank	-4.9473	-2.2201	-4.9056	-7.6306	<u>-4.5156</u>	-5.9379	-1.4596	-1.3225	-1.2652	-1.4249	-0.6862	<u>-0.7460</u>	<u>-3.5795</u>	<u>-3.5835</u>
	ManifoldRank (ours)	-3.3971	-1.8908	-2.3834	-2.7253	-4.2612	-4.3041	-1.0263	-0.9437	-1.0860	-1.0774	-0.6276	-0.7013	-3.2473	-3.2589
GINI	CPFair	0.6612	0.6444	0.7547	0.8590	0.7731	0.8191	0.4999	0.4928	0.5165	<u>0.5393</u>	0.2412	0.3978	0.7735	0.7689
	Min-regularizer	0.6763	0.6418	0.7655	0.8614	0.8020	0.8192	0.5638	0.4726	0.5831	0.6008	0.3028	0.3918	0.7903	<u>0.7673</u>
	P-MMF	0.6675	<u>0.6222</u>	<u>0.7537</u>	0.8471	0.7650	0.8172	<u>0.4740</u>	<u>0.4342</u>	0.5571	0.5348	0.2438	0.4694	0.8147	0.8059
	FairSync	0.6985	0.6291	0.7629	<u>0.8419</u>	0.7697	<u>0.8091</u>	0.4863	0.4805	<u>0.4863</u>	0.5475	0.2947	0.3958	<u>0.7621</u>	0.7675
	ElasticRank	0.7742	0.6377	0.8736	0.9005	0.8199	0.8490	0.5556	0.5547	0.5770	0.6162	0.2429	<u>0.3868</u>	0.7937	0.8354
	ManifoldRank (ours)	0.6608	0.6155	0.7363	0.8346	0.7495	0.7914	0.4539	0.3919	0.4608	0.5074	0.2226	0.3571	0.7065	0.7647
MMF	CPFair	0.0430	0.0840	0.0745	0.0664	0.0527	0.0358	0.1706	0.1767	0.1700	0.1618	0.3377	0.3041	0.0490	0.0487
	Min-regularizer	0.0259	0.0567	0.0945	0.0554	0.0415	0.0361	0.1374	0.1871	0.1364	0.1317	0.2982	0.2990	0.0552	0.0491
	P-MMF	0.0530	<u>0.1053</u>	0.0957	<u>0.0771</u>	<u>0.0637</u>	0.0476	0.2074	<u>0.2304</u>	0.1607	0.1815	0.3425	<u>0.3218</u>	0.0424	0.0433
	FairSync	0.0688	0.0794	<u>0.1167</u>	0.0764	0.0609	<u>0.0532</u>	<u>0.2101</u>	0.2111	0.2140	0.1830	0.3238	0.2586	<u>0.0721</u>	0.0590
	ElasticRank	0.0452	0.0915	0.0465	0.0284	0.0427	0.0319	0.1879	0.1902	0.1721	0.1533	<u>0.3577</u>	0.3124	0.0680	0.0640
	ManifoldRank (ours)	0.0754	0.1096	0.1255	0.0799	0.0706	0.0620	0.2227	0.2617	0.2190	0.2005	0.3621	0.3158	0.0774	0.0704

integrates the base model’s preference scores into the demand-side gradient calculation.

Table 4: Fairness performances for different base model BPR, GRU4Rec, SASRec. Also, we tuned various re-ranking models to achieve accuracy performances (NDCG) exceeding 99%, and evaluated fairness across two cut-offs K to assess three fairness metrics (EF@ K).

Model	BPR		GRU4Rec		SASRec	
	EF@10	EF@20	EF@10	EF@20	EF@10	EF@20
FairSync	-0.1861	-0.5712	-0.9317	-1.0366	-1.0610	-1.0326
ElasticRank	-0.0136	-0.1442	-1.1096	-1.1680	-1.0514	-1.3328
ManifoldRank	-0.0120	-0.0412	-0.7876	-0.9635	-0.9124	-0.9088

6.3.2 Ablation study on fairness parameters. We conduct an ablation analysis to investigate the influence of the fairness parameters α and β . We evaluate in EF@10 and NDCG@ K by varying each parameter individually while fixing its counterpart; see Figure 4.

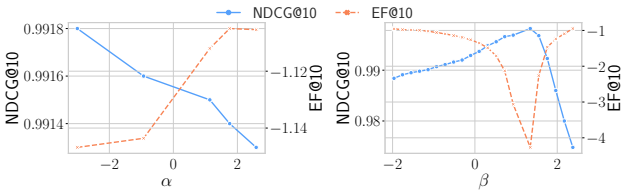


Figure 4: Ablation study on NDCG@ K and EF@ K performance w.r.t. fairness parameter α, β .

As illustrated in Figure 4 (left), an increase in α leads to a corresponding improvement in the fairness metric EF@10, albeit at the expense of a decrease in NDCG@10. This observation aligns with the theoretical analysis presented in Section 4, as α governs the global taxation degree; a larger α effectively amplifies the gradient associated with the fairness objective.

Second, Figure 4 (right) illustrates a non-monotonic trend for the local fairness parameter β . As β increases within the range $[-2, 1]$, we observe a decline in EF@10 accompanied by an improvement in NDCG@10. However, as β continues to rise from $[1, 3]$, this trend reverses: EF@10 increases while NDCG@10 decreases. This behavior aligns with the theoretical framework in Section 4; specifically, when $\beta < 1$, decreasing its value prioritizes the utility of the “worst-off” providers. Conversely, when $\beta > 1$, larger values impose a heavier penalty on “rich” providers, thereby progressively enhancing global fairness.

Finally, we observe that the model exhibits lower sensitivity to α , suggesting that this parameter requires tuning over a broader range to induce significant performance shifts. In contrast, the local fairness parameter β demonstrates a more pronounced impact on fairness metrics; this heightened sensitivity stems from its capacity to exert a larger influence on the gradient magnitude.

6.3.3 Performance at each online step. In this section, we provide a comparative analysis of the real-time accuracy and fairness trajectories between our proposed ManifoldRank and the strongest baseline, FairSync. The performance evolution is depicted in Figure 5.

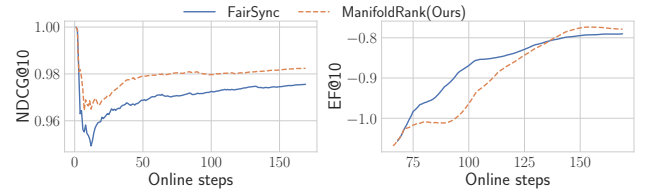


Figure 5: NDCG@ K and EF@ K performance w.r.t. online step.

As illustrated in Figure 5, both ManifoldRank and FairSync initially prioritize the fairness objective, which leads to a transient decline in NDCG. However, as the online procedure progresses, both frameworks successfully recalibrate, demonstrating a simultaneous improvement in fairness and accuracy.

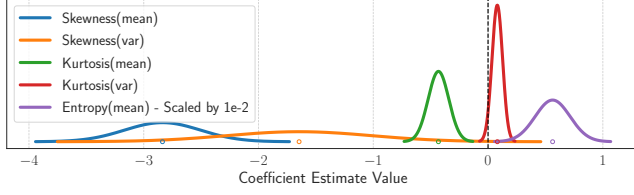


Figure 6: Influence of the statistical properties of the demand-side vector s , where “mean” and “var” in the legend denote the mean and variance across different users.

Compared to FairSync, ManifoldRank consistently maintains superior accuracy throughout the online process (left panel), reflecting high utility on the demand side. Regarding fairness (right panel), while ManifoldRank initially exhibits a more gradual improvement trend within the first 140 steps, it eventually surpasses the baseline to achieve superior long-term fairness. These results underscore ManifoldRank’s suitability for real-world applications; in large-scale deployment scenarios involving massive user bases and extended operational timelines.

6.3.4 Influence on information of demand side. To investigate the influences of different statistical properties of the demand side vector s , we tune ManifoldRank without the demand-side gradient ζ using different ranking scores, and frame a regression task in which the fairness metric $EF@K$ is treated as the target variable. Then we denote by $f_i(s)$ the i -th statistical characteristic of s , which is used as a regression feature and includes measures such as entropy, skewness, kurtosis, and related statistics: $EF@K = a_i f_i(s) + \epsilon$, where ϵ is the noise variable of regression.

Figure 6 presents the estimated coefficients a_i together with their confidence intervals, where “mean” and “var” in the legend denote the mean and variance across different users. From the figure, we observe that the mean values of Skewness and Entropy have a pronounced influence on the online re-ranking process, whereas the other factors remain close to zero, indicating limited impact. Therefore, in Section 5.3, we select Skewness and Entropy as proxy properties of the manifold.

Meanwhile, the coefficient of Skewness is negative, whereas that of Entropy is positive. Higher entropy reflects greater diversity in user preferences, so enforcing fairness has little impact on performance. In contrast, higher skewness indicates concentrated preferences, and stronger taxation in this case may substantially degrade performance.

7 Conclusion & future work

In this paper, we first formulate fair re-ranking as a Walrasian Equilibrium within an attention economy. We demonstrate that varying market contexts have a significant influence on the generalizability of online re-ranking gradients on the manifold. Building on these theoretical insights, we propose Manifold. This advanced online fair re-ranking algorithm dynamically adjusts gradients by accounting for both the supply and demand sides of the market. Extensive empirical evaluations show the effectiveness of ManifoldRank.

Future work. As to limitations, the provider-fair re-ranking framework established in this paper is rooted in the economics of attention: it regards users’ limited attention as the core scarce resource in the information retrieval market, and regulates its allocation across

different providers through a mechanism analogous to taxation. However, with the rise of large language models and retrieval-augmented generation (RAG), the underlying paradigm of information retrieval is undergoing a fundamental shift: users no longer browse ranked lists item by item, but instead directly consume a unified answer generated by a large language model after synthesizing multiple sources. This shift requires a further extension of the theoretical framework developed in this paper, and also gives rise to a new research perspective: Token Economics.

Specifically, in the traditional retrieval paradigm, the scarce resources are mainly explicitly observable ranking positions and user attention, and the binary exposure decision studied in this paper is built precisely upon this foundation. Under Token Economics, however, scarce resources exhibit a dual structure on both the input and output sides: input-side context-window tokens determine which passages can enter the model’s generation context, while output-side attribution tokens measure the contribution of different passages to the final generated content. Accordingly, the original discrete exposure variable needs to be extended into continuous attribution weights. In the future, we should focus on the fact that the boundary of providers is further decomposed from ranking lists to generated tokens, which imposes new modeling requirements on both the objects and the granularity of taxation.

Acknowledgments

This research was (partially) supported by the National Natural Science Foundation of China - 62472426, the Dutch Research Council (NWO), under project numbers 024.004.022, NWA.1389.20.183, and KICH3.LTP.20.006, and the European Union under grant agreement No. 101201510 (UNITE). Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of their respective employers, funders, and/or granting authorities.

Appendix

A Proof of Theorem 1

PROOF. Based on the condition, we have:

$$\frac{\partial X_g^d}{\partial f_g} = \mathbb{I}(i \in I_g) \cdot \frac{\partial x_{u,i}^*}{\partial f_g},$$

where $x_{u,i}^*$ is the optimal attention to item i by user u .

Now, the $x_{u,i}^*$ is decided by the $\sum_i (s_{u,i} - f_g)x_{u,i}$ subject to the budget and box constraints. This is a linear program (LP) problem and by the monotonicity of optimal solutions in LP, we have:

$$\frac{\partial x_{u,i}^*}{\partial f_g} = -\lambda_u \leq 0,$$

where $\lambda_u \geq 0$ is the dual multiplier for the budget constraint $\sum_i x_{u,i} \leq K$. Thus, $\frac{\partial x_{u,i}^*}{\partial f_g} \leq 0$, implying

$$\frac{\partial X_g^d}{\partial f_g} \leq 0.$$

□

B Proof of Theorem 2

PROOF. First, the total supply cost can be written as: $r'(v) = \sum_g r(v_g) = \text{sign}(\alpha\beta)(\sum_i v_g^\beta)^\alpha$. Let $G = \sum_g v_g^\beta(v_g)$, then we have:

$$\frac{\partial r'(v)}{\partial v_g} = \text{sign}(\alpha\beta)\alpha\beta G^{\alpha-1}v_g^{\beta-1}.$$

Since $v_g \geq 0$, we have:

$$G = \sum_g v_g^\beta \geq 0, v_g^{\beta-1} \geq 0,$$

then $\frac{\partial r'(v)}{\partial v_g} \geq 0$. $\square$

C Proof of Theorem 3

PROOF. Similar to the proof of Theorem 1, we have

$$\frac{\partial X_g^s}{\partial f_g} = \mathbb{I}(i \in I_g) \cdot \frac{\partial x_{u,i}^*}{\partial f_g},$$

And the $x_{u,i}^*$ is decided by the $\sum_i (f_g x_{u,i}) - v(v_g)$ subject to the constraints. By the monotonicity of LP, we have:

$$\frac{\partial x_{u,i}^*}{\partial f_g} = \lambda_u \geq 0,$$

where $\lambda_u \geq 0$ is the dual multiplier, implying

$$\frac{\partial X_g^s}{\partial f_g} \geq 0.$$

$\square$

D Proof of Theorem 4

PROOF. *Step 1:* First, we will prove that f_g^* is the price to make demand and supply $\{X_g^d, X_g^s\}$ Pareto optimal. Analogous to the proof strategy of the First Fundamental Theorem of Welfare Economics [25], assuming there exists $\tilde{X}^d, \tilde{X}^s$ for the user that satisfies:

$$w_u(X^d) \leq w_u(\tilde{X}^d), \forall u, \quad \exists u_k, w_{u_k}(X^d) < w_{u_k}(\tilde{X}^d).$$

Then we have the price f_g^* for the supply side aims to maximize their attention gained:

$$f_g^* \tilde{X}_g^s \leq f_g^* X_g^s, \quad \sum_g f_g^* \tilde{X}_g^s \leq \sum_g f_g^* X_g^s.$$

According to the market clear condition $\tilde{X}_g^d = \tilde{X}_g^s$, we have:

$$\sum_g f_g^* \tilde{X}_g^d \leq \sum_g f_g^* X_g^d. \quad (14)$$

While for the demand side, $\exists u_k, w_{u_k}(X^d) < w_{u_k}(\tilde{X}^d)$ must satisfy that $f_g^* X_g^d > f_g^* \tilde{X}_g^d$, otherwise it does not satisfy the maximize user's utility condition:

$$w_{u_k}(X^d) = \sum_{i \in I_g} (s_{u,i} - f_g^*) x_{u,i} < \sum_{i \in I_g} (s_{u,i} - f_g^*) \tilde{x}_{u,i} = w_{u_k}(\tilde{X}^d).$$

Such a condition will lead to $\sum_g f_g^* \tilde{X}_g^d > \sum_g f_g^* X_g^d$, which contradicts Eq. (14).

Step 2: Then we will show that the fair re-ranking process can be regarded as the Walrasian Equilibrium seeking process. Let $|\mathcal{U}|K = T$, the total social welfare function can be viewed as:

$$W = \max_{x_{u,i} \in \mathcal{X}} \sum_u w_u - \sum_g r(v_g), \quad \sum_g X_g^d = \sum_g X_g^s = T \text{ (Equilibrium)}.$$

Then we move the constraints to the objective using a vector of Lagrange multipliers f_g (i.e., the price for fair re-ranking):

$$W = \max_{x_{u,i} \in \mathcal{X}} \left(\min_{f_g^d} d(f_g^d) + \min_{f_g^s} s(f_g^s) \right) \quad (15)$$

$$\text{s.t. } d(f_g^d) = \sum_u w_u + \sum_g f_g^d (\sum_g X_g^d - T), \quad (\text{Demand function})$$

$$d(f_g^s) = \sum_g f_g^s (\sum_g X_g^s - T) - \sum_g r(v_g), \quad (\text{Supply function})$$

$$f_g^s = f_g^d. \quad (\text{Market condition})$$

Such a demand function will lead to the conclusion. $\square$

E Proof of Theorem 5

PROOF. We define the aggregate production set S as: $S = \{\sum_g X_g^s\}$. Define the aggregate “better-than” set D :

$$D = \left\{ \sum_g X_g^d \mid w_u(X_g^d) \geq w_u(X^*), \exists g, w_u(X_g^d) > w_u(X^*) \right\}.$$

Since X^* is Pareto optimal, the intersection of the interior of D and the set S must be empty: $S \cap D = \emptyset$. By the Separating Hyperplane Theorem [23], there exists a non-zero price vector $f \in \mathbb{R}^{|\mathcal{G}|}$ and a scalar k :

$$f \cdot d \geq k \geq f \cdot s \quad \forall d \in D, \forall s \in S. \quad (16)$$

Then this single price vector f supports individual optimization for getting the optimal X_g^s . And the parameters β, α will change the objective, eventually reach the optimal price f . Due to the constraints of the re-ranking setting, the total produced attention is bounded by the available slots $\sum_g X_g^s \leq |\mathcal{U}|K$.

Adjusting the price f shifts the optimal attention allocation, which is effectively implemented through a single taxation mechanism T_g . In this context, $T_g > 0$ represents a tax levied on “rich” providers, while $T_g < 0$ means a subsidy for “poor” providers:

$$\max\{0, T_g\} \leq X_g^s \leq X_g^0 - \max\{0, T_g\}, \quad \sum_g T_g = 0,$$

where X_g^0 denotes the attention produced without the fairness. $\square$

References

- [1] Himan Abdollahpour, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Pizzato. 2020. Multi-stakeholder Recommendation: Survey and Research Directions. *User Modeling and User-Adapted Interaction* 30, 1 (2020), 127–158.
- [2] Himan Abdollahpour and Robin Burke. 2019. Multi-stakeholder Recommendation and its Connection to Multi-sided Fairness. *arXiv preprint arXiv:1907.13158* (2019).
- [3] Kenneth J. Arrow and Gerard Debreu. 2024. Existence of an Equilibrium for a Competitive Economy. In *The Foundations of Price Theory Vol 5*. Routledge, 289–316.
- [4] Anthony B. Atkinson and Nicholas H. Stern. 1974. Pigou, Taxation and Public Goods. *The Review of Economic Studies* 41, 1 (1974), 119–128.
- [5] Yves Balasko. 2009. *The Equilibrium Manifold: Postmodern Developments in the Theory of General Economic Equilibrium*. MIT Press.
- [6] Santiago Balseiro, Haihao Lu, and Vahab Mirrokni. 2021. Regularized Online Allocation Problems: Fairness and Beyond. In *International Conference on Machine Learning*. PMLR, 630–639.
- [7] Tolga Bektas and Adam N. Letchford. 2020. Using ℓ_p -Norms for Fairness in Combinatorial Optimisation. *Computers & Operations Research* 120 (2020), 104975.
- [8] Dimitris Bertsimas, Vivek F. Farias, and Nikolaos Trichakis. 2011. The Price of Fairness. *Operations Research* 59, 1 (2011), 17–31.

- [9] Arpita Biswas, Gourab K. Patro, Niloy Ganguly, Krishna P. Gummadi, and Abhijnan Chakraborty. 2021. Toward Fair Recommendation in Two-sided Platforms. *ACM Transactions on the Web (TWEB)* 16, 2 (2021), 1–34.
- [10] Virginie Do, Sam Corbett-Davies, Jamal Atif, and Nicolas Usunier. 2021. Two-sided Fairness in Rankings via Lorenz Dominance. *Advances in Neural Information Processing Systems* 34 (2021), 8596–8608.
- [11] Yi Fang, Ashudeep Singh, Zhiqiang Tao, et al. 2024. Fairness in Search Systems. *Foundations and Trends in Information Retrieval* 18, 3 (2024), 262–416.
- [12] Elizabeth Gómez, David Contreras, Ludovico Boratto, and Maria Salamó. 2025. Enhancing recommender systems with provider fairness through preference distribution awareness. *International Journal of Information Management Data Insights* 5, 1 (2025), 100311.
- [13] Ruining He and Julian McAuley. 2016. Ups and downs: Modeling the Visual Evolution of Fashion Trends with One-class Collaborative Filtering. In *Proceedings of the 25th International Conference on World Wide Web*. 507–517.
- [14] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2015. Session-based Recommendations with Recurrent Neural Networks. *arXiv preprint arXiv:1511.06939* (2015).
- [15] Zhihao Hu, Yiran Xu, and Xinmei Tian. 2023. Adaptive Priority Reweighting for Generalizing Fairness Improvement. In *2023 International Joint Conference on Neural Networks (IJCNN)*. 01–08.
- [16] Meng Jiang, Keqin Bao, Jizhi Zhang, Wenjie Wang, Zhengyi Yang, Fuli Feng, and Xiangnan He. 2024. Item-side Fairness of Large Language Model-based Recommendation System. In *Proceedings of the ACM Web Conference 2024*. ACM, 4717–4726.
- [17] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive Sequential Recommendation. In *2018 IEEE international conference on data mining (ICDM)*. IEEE, 197–206.
- [18] Preethi Lahoti, Krishna P. Gummadi, and Gerhard Weikum. 2019. iFair: Learning Individually Fair Data Representations for Algorithmic Decision Making. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. IEEE, 1334–1345.
- [19] Yunqi Li, Hanxiong Chen, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang. 2021. User-oriented Fairness in Recommendation. In *Proceedings of the Web Conference*. 624–632.
- [20] Yunqi Li, Hanxiong Chen, Shuyuan Xu, Yingqiang Ge, Juntao Tan, Shuchang Liu, and Yongfeng Zhang. 2023. Fairness in Recommendation: Foundations, Methods and Applications. *ACM Transactions on Intelligent Systems and Technology* 14, 5 (2023), Article No.: 95.
- [21] Yuanna Liu, Ming Li, Mohammad Aliannejadi, and Maarten de Rijke. 2025. Repeat-bias-aware Optimization of Beyond-accuracy Metrics for Next Basket Recommendation. In *ECIR 2025: 47th European Conference on Information Retrieval (Part I)*. Springer, Berlin, Heidelberg, 214–229.
- [22] Alexander V. Lotov and Kaisa Miettinen. 2008. Visualizing the Pareto frontier. In *Multiobjective Optimization*. Springer, 213–243.
- [23] Andrew McLennan. 2002. Ordinal Efficiency and the Polyhedral Separating Hyperplane Theorem. *Journal of Economic Theory* 105, 2 (2002), 435–449.
- [24] Mohammadmehdi Naghiaei, Hossein A. Rahmani, and Yashar Deldjoo. 2022. CP-Fair: Personalized Consumer and Producer Fairness Re-ranking for Recommender Systems. *arXiv preprint arXiv:2204.08085* (2022).
- [25] Yew-Kwang Ng. 1983. *Welfare Economics*. Springer.
- [26] Gourab K. Patro, Arpita Biswas, Niloy Ganguly, Krishna P. Gummadi, and Abhijnan Chakraborty. 2020. FairRec: Two-sided Fairness for Personalized Recommendations in Two-sided Platforms. In *Proceedings of the Web Conference*. 1194–1204.
- [27] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2012. BPR: Bayesian Personalized Ranking from Implicit Feedback. *arXiv preprint arXiv:1205.2618* (2012).
- [28] Clara Rus, Maarten de Rijke, and Andrew Yates. 2023. Counterfactual Representations for Intersectional Fair Ranking in Recruitment.. In *HR@ RecSys*.
- [29] Clara Rus, Gabrielle Poerwawinata, Andrew Yates, and Maarten de Rijke. 2024. AnnoRank: A Comprehensive Web-Based Framework for Collecting Annotations and Assessing Rankings. In *CIKM 2024: 33rd ACM International Conference on Information and Knowledge Management*. ACM, 5400–5404.
- [30] Yuta Saito and Thorsten Joachims. 2022. Fair Ranking as Fair Division: Impact-Based Individual Fairness in Ranking. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1514–1524.
- [31] Herbert A. Simon. 1971. *Designing Organizations for an Information-rich World*. Johns Hopkins University Press.
- [32] Lequn Wang and Thorsten Joachims. 2021. User Fairness, Item Fairness, and Diversity for Rankings in Two-Sided Markets. In *Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval*. 23–41.
- [33] Yifan Wang, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma. 2023. A Survey on the Fairness of Recommender Systems. *ACM Transactions on Information Systems* 41, 3 (2023), 1–43.
- [34] Yao Wu, Jian Cao, and Guandong Xu. 2023. Fairness in Recommender Systems: Evaluation Approaches and Assurance Strategies. *ACM Transactions on Knowledge Discovery from Data* 18, 1 (2023), 1–37.
- [35] Yao Wu, Jian Cao, Guandong Xu, and Yudong Tan. 2021. TFROM: A Two-Sided Fairness-Aware Recommendation Model for both Customers and Providers. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1013–1022.
- [36] Chen Xu, Sirui Chen, Jun Xu, Weiran Shen, Xiao Zhang, Gang Wang, and Zhenhua Dong. 2023. P-MMF: Provider Max-min Fairness Re-ranking in Recommender System. In *Proceedings of the ACM Web Conference 2023*. 3701–3711.
- [37] Chen Xu, Zhirui Deng, Clara Rus, Xiaopeng Ye, Yuanna Liu, Jun Xu, Zhicheng Dou, Ji-Rong Wen, and Maarten de Rijke. 2025. FairDiverse: A Comprehensive Toolkit for Fairness- and Diversity-aware Information Retrieval. In *SIGIR 2025: 48th international ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 3540–3550.
- [38] Chen Xu, Jun Xu, Yiming Ding, Xiao Zhang, and Qi Qi. 2024. FairSync: Ensuring Amortized Group Exposure in Distributed Recommendation Retrieval. In *Proceedings of the ACM Web Conference 2024 (Singapore, Singapore) (WWW '24)*. 1092–1102.
- [39] Chen Xu, Xiaopeng Ye, Wenjie Wang, Liang Pang, Jun Xu, and Tat-Seng Chua. 2024. A Taxation Perspective for Fair Re-ranking. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24)*. 1494–1503.
- [40] Chen Xu, Jujia Zhao, Wenjie Wang, Liang Pang, Jun Xu, Tat-Seng Chua, and Maarten de Rijke. 2025. Understanding Accuracy-Fairness Trade-offs in Re-ranking through Elasticity in Economics. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 539–548.
- [41] Hong-Jian Xue, Xinyu Dai, Jianbing Zhang, Shujian Huang, and Jiajun Chen. 2017. Deep Matrix Factorization Models for Recommender Systems.. In *IJCAI*, Vol. 17. Melbourne, Australia, 3203–3209.
- [42] Ke Yang, Joshua R. Loftus, and Julia Stoyanovich. 2020. Causal Intersectionality for Fair Ranking. In *FORC 2021: 2nd Symposium on Foundations of Responsible Computing*. Article No. 7.
- [43] Tao Yang, Zhichao Xu, and Qingyao Ai. 2023. Vertical Allocation-based Fair Exposure Amortizing in Ranking. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region (Beijing, China)*. ACM, 234–244.
- [44] Sirui Yao and Bert Huang. 2017. Beyond Parity: Fairness Objectives for Collaborative Filtering. *Advances in Neural Information Processing Systems* 30 (2017).
- [45] Xiao Zhang, Teng Shi, Jun Xu, Zhenhua Dong, and Ji-Rong Wen. 2024. Model-agnostic causal embedding learning for counterfactually group-fair recommendation. *IEEE Transactions on Knowledge and Data Engineering* 36, 12 (2024), 8801–8813.