# Sampling Ex-Post Group-Fair Rankings

**Sruthi Gorantla**[1], **Amit Deshpande**[2] and **Anand Louis**[1]
[1]Indian Institute of Science, Bengaluru
[2]Microsoft Research, Bengaluru
gorantlas@iisc.ac.in, amitdesh@microsoft.com, anandl@iisc.ac.in

## Abstract

Randomized rankings have been of recent interest to achieve ex-ante fairer exposure and better robustness than deterministic rankings. We propose a set of natural axioms for randomized group-fair rankings and prove that there exists a unique distribution $\mathcal{D}$ that satisfies our axioms and is supported only over ex-post group-fair rankings, i.e., rankings that satisfy given lower and upper bounds on group-wise representation in the top-$k$ ranks. Our problem formulation works even when there is implicit bias, incomplete relevance information, or only ordinal ranking is available instead of relevance scores or utility values.

We propose two algorithms to sample a random group-fair ranking from the distribution $\mathcal{D}$ mentioned above. Our first dynamic programming-based algorithm samples ex-post group-fair rankings uniformly at random in time $O(k^2\ell)$, where $\ell$ is the number of groups. Our second random walk-based algorithm samples ex-post group-fair rankings from a distribution $\delta$-close to $\mathcal{D}$ in total variation distance and has expected running time $O^*(k^2\ell^2)$[1], when there is a sufficient gap between the given upper and lower bounds on the group-wise representation. The former does exact sampling, but the latter runs significantly faster on real-world data sets for larger values of $k$. We give empirical evidence that our algorithms compare favorably against recent baselines for fairness and ranking utility on real-world data sets.

## 1 Introduction

Ranking individuals using algorithms has become ubiquitous in many applications such as college admissions [Baswana *et al.*, 2019], recruitment [Geyik *et al.*, 2019], among others. In many such scenarios, individuals who belong to certain demographic groups based on race, gender, age, etc., face discrimination due to human and historical biases [Uhlmann and Cohen, 2005; Okonofua and Eberhardt, 2015; Hassani, 2021]. Algorithms learning from biased data exacerbate representational harms for certain groups in the top ranks[2,3], leading to loss of opportunities. One way to mitigate representational harms is by imposing explicit representation-based fairness constraints that the ranking output by the algorithm must contain a certain minimum and maximum number of candidates from each group [Geyik *et al.*, 2019; Celis *et al.*, 2020b]. Many fair processes such as the Rooney rule [Collins, 2007], the $4/5$-th rule [Bobko and Roth, 2004], and fairness metrics such as demographic parity impose representation-based constraints.

A large body of work has proposed deterministic post-processing of rankings to satisfy representation-based group-fairness constraints [Celis *et al.*, 2018b; Geyik *et al.*, 2019; Wu *et al.*, 2018; Zehlike *et al.*, 2017; Gorantla *et al.*, 2021]. These methods essentially merge the group-wise ranked lists to create a common ranking that satisfies representation-based constraints. However, these methods contain two critical flaws. First, deterministic rankings cannot create opportunities for *all* the groups at the top, especially when the number of groups is large. Second, observations of merit are often noisy in the real world [Okonofua and Eberhardt, 2015] and could contain *implicit bias* towards protected groups [Uhlmann and Cohen, 2005]. Hence, inter-group comparisons of merit while merging the group-wise rankings could lead to a loss of opportunities for certain groups. For example, suppose multiple companies intend to hire for a limited number of similar open positions and use the same recruitment system to rank a common candidate pool for job interviews. In a deterministic top-$k$ ranking based on biased merit scores, every company would see the same ranking, where the protected groups could be systematically ranked lower. Hence, equal representation at the top-$k$ may not translate into equal opportunities.

We consider randomized group-fair rankings as a way to create opportunity for every group in the top ranks (or, more generally, any rank). We assume that we are given only the ordinal rankings of items within each group (i.e., intra-group ordering without any scores) and no comparison of

[1] $O^*$ suppresses logarithmic and error terms.

[2] Jeffrey Dastin, *Amazon scraps secret AI recruiting tool that showed bias against women*

[3] Bogen and Rieke, *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias*

items across different groups (i.e., inter-group comparisons). This assumption is strong but circumvents implicit bias and allows us to consider group-fair rankings even under incomplete or biased data about pairwise comparisons. Our randomized ranking algorithms output ex-post group-fair rankings, i.e., every sampled ranking output is group-fair and satisfies representation-based fairness constraints as a stronger ex-post (or actual) guarantee instead of a weaker ex-ante (or expected) guarantee.

The rest of the paper is organized as follows: In Section 2, we survey related work and summarize its limitations for our problem. In Section 3 we describe our axiomatic approach to define a distribution over ex-post group-fair rankings (Axioms 3.1-3.3) and show that there is a unique distribution that satisfies our axioms (Theorem 3.4). The same distribution satisfies sufficient representation of every group in any consecutive ranks in the top-$k$ (Corollary 3.6), a natural characteristic derived from our axioms. In Section 4, we give an efficient dynamic programming-based algorithm (Algorithm 1) and a random walk-based algorithm (Algorithm 2) to sample an ex-post group-fair ranking from the above distribution. We also extend our algorithms to handle representation-based constraints on the top prefixes (Section 4.3). In Section 5, we empirically validate our theoretical and algorithmic guarantees on real-world datasets. Finally, Section 6 contains some limitations of our work and open problems. All the proofs and additional experimental results are included in the appendix.

## 2 Related Work

Representation-based fairness constraints for ranking $\ell$ groups in top $k$ ranks are typically given by numbers $L_j$ and $U_j$, for each group $j \in [\ell]$, representing the lower and upper bound on the group representation respectively. They capture several fairness notions; setting $L_j = U_j = \frac{k}{\ell}, \forall j \in [\ell]$ ensures equal representation for all groups (see section 5 of [Zehlike *et al.*, 2022b]), whereas $L_j = U_j = p_j \cdot k, \forall j \in [\ell]$, ensures proportional representation[4], where $p_j$ is the proportion of the group $j$ in the population [Gorantla *et al.*, 2021; Gao and Shah, 2020; Geyik *et al.*, 2019]. These constraints have also been studied in other problems such as fair subset selection [Stoyanovich *et al.*, 2018], fair matching [Goto *et al.*, 2016], and fair clustering [Chierichetti *et al.*, 2017].

Previous work has tried to formalize the general principles of fair ranking as treating similar items consistently, maintaining a sufficient presence of items from minority groups, and proportional representation from every group [Castillo, 2019]. There has been work on quantifying fairness requirements [Yang and Stoyanovich, 2017; Geyik *et al.*, 2019; Beutel *et al.*, 2019; Narasimhan *et al.*, 2020; Kuhlman *et al.*, 2019], which has predominantly proposed deterministic algorithms for group-fair ranking [Yang and Stoyanovich, 2017; Geyik *et al.*, 2019; Gorantla *et al.*, 2021; Celis *et al.*, 2018b; Zehlike *et al.*, 2022b].

Our recruitment example in the previous section shows the inadequacy of deterministic ranking to improve opportunities. Recent works have also observed this and proposed randomized ranking algorithms to achieve equality or proportionality of expected exposure [Diaz *et al.*, 2020; Singh and Joachims, 2018; Biega *et al.*, 2018; Memarrast *et al.*, 2021; Kletti *et al.*, 2022]. All of them require utilities or scores of the items to be ranked and hence, are susceptible to implicit bias or incomplete information about the true utilities. Moreover, they do not give ex-post guarantees on the representation of each group, which can be a legal or necessary requirement if the exposure cannot be computed efficiently and reliably [Heuss *et al.*, 2022].

Another recent workaround is to model the uncertainty in merit. Assuming access to the true merit distribution, Singh *et al.* (2021) give a randomized ranking algorithm for a notion of individual fairness. On the other hand, Celis *et al.* (2020b) try to model the systematic bias. Under strong distributional assumptions, they show that representation constraints are sufficient to achieve a fair ranking. However, their assumptions may not hold in the real world as unconscious human biases are unlikely to be systematic [Uhlmann and Cohen, 2005; Okonofua and Eberhardt, 2015].

Most aligned to our work is a heuristic randomized ranking algorithm – *fair $\epsilon$-greedy* – proposed by [Gao and Shah, 2020]. Similar to our setup, they eschew comparing items across different groups. However, their algorithm does not come with any theoretical guarantees, and it does not always sample ex-post group-fair rankings. Moreover, it works only when no upper-bound constraints exist on the group-wise representation. To the best of our knowledge, our work is the first to propose a distribution over ex-post group-fair rankings, using lower and upper bounds on the group-wise representations, and to give provably correct and efficient sampling algorithms for it.

We note here that previous work has also used randomization in ranking, recommendations, and summarization of ranked results to achieve other benefits such as controlling polarization [Celis *et al.*, 2019], mitigating data bias [Celis *et al.*, 2020a], and promoting diversity [Celis *et al.*, 2018a].

## 3 Group Fairness in Ranking

Given a set $N := [n]$ of items, a *top-$k$ ranking* is a selection of $k < n$ items followed by the assignment of each rank in $[k]$ to exactly one of the selected items. We use index $i$ to refer to a rank, index $j$ to refer to a group, and $a$ to refer to elements in the set $N$. Let $a, a' \in N$ be two different items such that the item $a$ is assigned to rank $i$ and item $a'$ is assigned to rank $i'$. Whenever $i < i'$ we say that item $a$ is *ranked lower* than item $a'$. Going by the convention, we assume that being ranked at lower ranks gives items better visibility [Gorantla *et al.*, 2021]. Throughout the paper, we refer to a top-$k$ ranking by just ranking. The set $N$ can be partitioned into $\ell$ disjoint groups of items depending on a sensitive attribute. A group-fair ranking is any ranking that satisfies a set of group fairness constraints. Our fairness constraints are *representation constraints*; lower and upper bounds, $L_j, U_j \in [k]$ respectively, on the number of top $k$ ranks assigned to group $j$, for each group $j \in [\ell]$. Throughout the paper, we assume that we are given a ranking of the items within the same group for all

[4]It may not always be possible to satisfy equal or proportional representation constraints exactly. In that case, the algorithms need to say that the instance is infeasible.

groups. We call these rankings *in-group rankings*. We now take an axiomatic approach to characterize a random group-fair ranking.

## 3.1 Random Group-Fair Ranking

The three axioms we state below are natural consistency and fairness requirements for distribution over all the rankings.

**Axiom 3.1** (In-group consistency)**.** *For any ranking sampled from the distribution, for all items $a, a'$ belonging to the same group $j \in [\ell]$, item $a$ is ranked lower than item $a'$ if and only if item $a$ is ranked lower than item $a'$ in the in-group ranking of group $j$.*

Since the intra-group merit comparisons are reliable, their in-group ranking must remain consistent, which is what the axiom asks for. Many post-processing algorithms for group-fairness ranking satisfy this axiom [Gorantla *et al.*, 2021; Celis *et al.*, 2018b; Zehlike *et al.*, 2022a; Zehlike *et al.*, 2017]. Once we assign the ranks to groups, Axiom 3.1 determines the corresponding items to be placed there consistent with in-group ranking. Hence, for the next axioms, we look at the group assignments instead of rankings. A *group assignment* assigns each rank in the top $k$ ranking to exactly one of the $\ell$ groups. Let $Y_i$ be a random variable representing the group $i$th rank is assigned to. Therefore $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_k)$ is a random vector representing a group assignment. Let $y = (y_1, y_2, \ldots, y_k)$ represent an instance of a group assignment. A *group-fair assignment* is a group assignment that satisfies the representation constraints. Therefore the set of group-fair assignments is $\left\{y \in [\ell]^k : L_j \leqslant \sum_{i\in[k]} \mathbb{I}[y_i = j] \leqslant U_j, \forall j \in [\ell]\right\}$, where $\mathbb{I}[\cdot]$ is an indicator function. The ranking can then be obtained by assigning the items within the same group, according to their in-group ranking, to the ranks assigned to the group. We use $Y_0$ to represent a dummy group assignment of length 0 for notational convenience when no group assignment is made to any group (e.g. in Axiom 3.3).

Let $X_j$ be a random variable representing the number of ranks assigned to group $j$ in a group assignment for all $j \in [\ell]$. Therefore $\boldsymbol{X} = (X_1, X_2, \ldots, X_\ell)$ represents a random vector for a *group representation*. Let $x = (x_1, x_2, \ldots, x_\ell)$ represent an instance of a group representation. Then the set of group-fair representations is $\left\{x \in \mathbb{Z}^\ell_{\geqslant 0} : \sum_{j\in[\ell]} x_j = k \text{ and } L_j \leqslant x_j \leqslant U_j, \forall j \in [\ell]\right\}$.

Since the inter-group comparisons are unreliable, any feasible group-fair representation is equally likely to be the best. That is, the distribution should be maximally non-committal distribution over the group-fair representations, which is nothing but a uniform distribution over all feasible group-fair representations. This is captured by our next axiom as follows,

**Axiom 3.2** (Representation Fairness)**.** *All the non-group-fair representations should be sampled with probability zero, and all the group-fair representations should be sampled uniformly at random.*

**Remark.** Any distribution for top $k$ ranking that satisfies Axiom 3.2 is *ex-post* group fair since the support of the distribution consists only of rankings that satisfy representation constraints. This is important when the fairness constraints are legal or strict requirements.

Many distributions over rankings could satisfy Axiom 3.1 and Axiom 3.2. Consider a distribution that samples a group representation $x$ uniformly at random. Let $x_1 \in [L_1, U_1]$ be the representation corresponding to group 1. Let us assume that this distribution always assigns ranks $k - x_1 + 1$ to $k$ to group 1. Due to in-group consistency, the best $x_1$ items in group 1 get assigned to these ranks. However, always being at the bottom of the ranking is not fair to group 1, since it gets low visibility. Therefore, we introduce a third axiom that asks for fairness in the second step of ranking – assigning the top $k$ ranks to the groups in a *rank-aware* manner.

**Axiom 3.3** (Ranking Fairness)**.** *For any two groups $j, j' \in [\ell]$, for all $i \in \{0, \ldots, k-2\}$, conditioned on the top $i$ ranks and a group representation $x$, the $(i+1)$-th and the $(i+2)$-th ranks are assigned to $j$ and $j'$ interchangeably with equal probability. That is, $\forall j, j' \in [\ell], \forall i \in \{0, \ldots, k-2\}$,*

$$\begin{aligned}\Pr\left[Y_{i+1} = j, Y_{i+2} = j' \mid Y_0, Y_1, \ldots, Y_i, \boldsymbol{X}\right] \\ = \Pr\left[Y_{i+1} = j', Y_{i+2} = j \mid Y_0, Y_1, \ldots, Y_i, \boldsymbol{X}\right].\end{aligned}$$

Let $\mathcal{U}$ represent a uniform distribution. In the result below, we prove that there exists a unique distribution over the rankings that satisfies all three axioms.

**Theorem 3.4.** *Let $\mathcal{D}$ be a distribution from which a ranking is sampled as follows,*

1. *Sample an $x$ from,*
$$\mathcal{U}\left\{x \in \mathbb{Z}^\ell_{\geqslant 0} : \sum_{j\in[\ell]} x_j = k \wedge L_j \leqslant x_j \leqslant U_j, \forall j \in [\ell]\right\}.$$
2. *Sample a $y$, given $x$, from*
$$\mathcal{U}\left\{y \in [\ell]^k : \sum_{i\in k} \mathbb{I}[y_i = j] = x_j, \forall j \in [\ell]\right\}.$$
3. *Rank the items within the same group in the order consistent with their in-group ranking in the ranks assigned to the groups in the group assignment $y$.*

*Then $\mathcal{D}$ is the unique distribution that satisfies all three axioms.*

We also have the following additional characteristic of the distribution in Theorem 3.4. It guarantees that every rank in a randomly sampled group assignment is assigned to group $j$ with probability at least $\frac{L_j}{k}$ and at most $\frac{U_j}{k}$. Hence, every rank gets a sufficient representation of each group. Note that no deterministic group-fair ranking can achieve this.

Let $\mathcal{D}_\delta$ be a distribution that differs from $\mathcal{D}$ as follows: $\boldsymbol{X}$ is sampled from a distribution $\delta$-close to a uniform distribution in Step 1 of $\mathcal{D}$, in the total-variation distance, $\boldsymbol{Y}|x$ is sampled as in Step 2 of $\mathcal{D}$. The items are also assigned as in Step 3 of $\mathcal{D}$. Then it is easy to show that $\mathcal{D}_\delta$ is $\delta$-close to $\mathcal{D}$ in total-variation distance. We then prove the following theorem and its corollary.

**Theorem 3.5.** *For any $\delta > 0$ and group assignment $\boldsymbol{Y}$ sampled from $\mathcal{D}_\delta$, for every group $j \in [\ell]$ and for every rank $i \in [k]$, $\frac{L_j}{k} \leqslant \Pr_{\mathcal{D}_\delta}\left[Y_i = j\right] \leqslant \frac{U_j}{k}$.*

**Corollary 3.6.** *For any $\delta > 0$, let $i, i' \in [k]$ be such that $i \leqslant i'$ and let $Z^j_{i,i'}$ be a random variable representing the number of ranks assigned to group $j$ in ranks $i$ to $i'$, for a ranking sampled from $\mathcal{D}_\delta$. Then, for every group $j \in [\ell]$ and for every rank $i \in [k]$, $\left(\frac{i'-i+1}{k}\right) \cdot L_j \leqslant \mathbb{E}_{\mathcal{D}_\delta}\left[Z^j_{i,i'}\right] \leqslant \left(\frac{i'-i+1}{k}\right) \cdot U_j$.*

Two comments are in order. First, fixing $i = 0$ in Corollary 3.6 gives us that every prefix of the ranking sampled from $\mathcal{D}_\delta$ will have sufficient representation from the groups, in expectation. Such fairness requirements are consistent with those studied in the ranking literature [Celis *et al.*, 2018b]. Second, let $k' := i' - i$ for some $i, i' \in [k]$ such that $i \leqslant i'$. Then Corollary 3.6 also gives us that any consecutive $k'$ ranks of the ranking sampled from $\mathcal{D}_\delta$ also satisfy representation constraints. Such fairness requirements are consistent with those studied in [Gorantla *et al.*, 2021].

In the ranking, one might ask for different representation requirements for different prefixes of the ranking. We extend our algorithms to handle prefix fairness constraints in the next section (see Section 4.3).

**Time taken to sample from $\mathcal{D}$ (in Theorem 3.4).** Given a group representation $x$, one can find a uniform random $y$, for a given $x$, by sampling a random binary string of length at most $\log k!$. Since $x$ is fixed, this gives us a uniform random sample of $y$ conditioned on $x$. This takes time $\mathcal{O}(k \log k)$. This sampling takes care of Step 2. Step 3 simply takes $\mathcal{O}(k)$ time, given in-group rankings of all the groups. The main challenge is to provide an efficient algorithm to perform Step 1. Therefore, in the next section, we focus on sampling a uniform random group-fair representation in Step 1.

# 4 Sampling a Uniform Random Group-Fair Representation

We first note that each group-fair representation corresponds to a unique integral point in the convex polytope $K$ defined below,

$$K = \left\{ x \in \mathbb{R}^\ell \,\middle|\, \sum_{j \in [\ell]} x_j = k,\ L_j \leqslant x_j \leqslant U_j, \forall j \in [\ell] \right\}. \quad (1)$$

Therefore, sampling a uniform random group-fair representation is equivalent to sampling an integral or a lattice point uniformly at random from the convex set $K$.

## 4.1 Dynamic Programming for Exact Sampling

In this section, we give a dynamic programming-based algorithm (see Algorithm 1) for uniform random sampling of integer points from the polytope $K$. Each entry $D[k', i], \forall k' = \{0, 1, \ldots, k\}$ and $\forall i \in \{0, 1, \ldots, \ell\}$ in Algorithm 1 corresponds to the number of integer points in $K_{i,k'} = \left\{ x \in \mathbb{R}^i \mid \sum_{h \in [i]} x_h = k', L_h \leqslant x_h \leqslant U_h, \forall h \in [i] \right\}$. That is, the DP table keeps track of the number of feasible solutions that sum to $k'$ with the first $i$ groups. Therefore, $D[k, \ell]$ contains all feasible integer points of $K_{\ell,k}$, which is nothing but $K$ defined in Equation (1). The reader should note that the entry $D[0, i] = 1$ if assigning 0 to the first $i$ groups is feasible with respect to the fairness constraints and 0 otherwise. However, the entry $D[k', 0]$ is always 0 for $k' > 0$, since we can

**Algorithm 1** Sampling a uniform random group-fair representation

**Require:** Fairness constraints $L_j, U_j$ for all the groups $j \in [\ell]$, a number $k \in \mathbb{Z}_{\geqslant 0}$

1: **Initialize:** Set $D[k', i] := 0, \forall k' = \{0, 1, \ldots, k\}$ and $\forall i \in \{0, 1, \ldots, \ell\}$, and $D[0, 0] := 1$
// Counting
2: **for** $k' = 0$ to $k$ **do**
3:   **for** $i = 1$ to $\ell$ **do**
4:     $D[k', i] = \sum_{L_i \leqslant x_i \leqslant U_i} D[k' - x_i, i - 1]$
5:   **end for**
6: **end for**
// Sampling
7: Set $k' := k$ and $i := \ell$
8: **while** $i \neq 0$ **do**
9:   Sample $x_i$ from the categorical distribution on $L_i, L_{i+1}, \ldots, U_i$ with corresponding probabilities $\frac{D[k'-L_i, i-1]}{D[k', i]}, \frac{D[k'-L_{i+1}, i-1]}{D[k', i]}, \ldots, \frac{D[k'-U_i, i-1]}{D[k', i]}$
  // by convention $D[t, i] = 0$ for $t < 0$.
10:   Update $k' := k' - x_i$ and $i := i - 1$
11: **end while**

not construct a ranking of non-zero length without assigning the ranks to any of the groups. In Step 1, we initialize all the entries of the DP to 0 except for the entry $D[0, 0]$, which is set to 1. Steps 2 to 6 then count the number of feasible solutions for $D[k', i]$ by recursively summing over all feasible values for $x_i$. We note that this DP is similar to the DP, given by Štefankovič *et al.* (2012), for counting $0/1$ knapsack solutions where the feasible values of an item are 0 (including it) and 1 (not including it). Now, let us assume that we have sampled the value of $x_i$ for all $i+1, i+2, \ldots, \ell$ for some $0 < i < \ell$, and let $k' := k - x_{i+1} - x_{i+2} - \cdots - x_\ell$. Then for any $x_i \in [L_i, U_i]$ the probability that we sample $x_i$ is given by the number of feasible solutions after fixing $x_i$, divided by the total number of solutions for $x_i \in [L_i, U_i]$, which is nothing but $\frac{D[k'-x_i, i-1]}{D[k', i]}$ (see Step 9). Therefore, expanding the probability of sampling a feasible solution $(x_1, x_2, \ldots, x_\ell)$ gives us a telescoping product that evaluates to $1/D[k, \ell]$. Hence, we have the following theorem, whose proof appears in Appendix B.

**Theorem 4.1.** *Algorithm 1 samples a uniform random group-fair representation in time $\mathcal{O}(k^2 \ell)$.*

## 4.2 Approximate Uniform Sampling

Our second algorithm outputs an integral point from $K$, defined in Equation (1), from a density that is close to the uniform distribution over the set of integral points in $K$, with respect to the total variation distance (see Algorithm 2).

There is a long line of work on polynomial-time algorithms to sample a point approximately uniformly from a given convex polytope or a convex body [Dyer *et al.*, 1991; Lovász and Vempala, 2006; Cousins and Vempala, 2018]. We use the algorithm by [Cousins and Vempala, 2018] as SAMPLING-ORACLE in Algorithm 2. We get an algorithm

**Algorithm 2** Sampling an approximately uniform random group-fair representation

**Require:** Fairness constraints $L_j, U_j, \forall j \in [\ell], k \in \mathbb{Z}^+, \delta$.

1: $H := \left\{x \in \mathbb{R}^\ell \;\middle|\; \sum_{j\in[\ell]} x_j = k\right\}$.
2: $P := \left\{x \in \mathbb{R}^\ell \;\middle|\; L_j \leqslant x_j \leqslant U_j, \forall j \in [\ell]\right\}$.
3: $\Delta := \min \left\{ \left\lfloor \frac{k-(\sum_{j\in[\ell]} L_j)}{\ell} \right\rfloor, \left\lfloor \frac{(\sum_{j\in[\ell]} U_j)-k}{\ell} \right\rfloor, \min_{j\in[\ell]} \left\lfloor \frac{U_j - L_j}{2} \right\rfloor \right\}$.
4: $x_j^* := L_j + \Delta, \forall j \in [\ell]$
5: **for** $j := 1, 2, \ldots, \ell$ **do**
6:     **if** $\sum_{j'\in[\ell]} x_{j'}^* < k$ **then**
7:         $x_j^* := \min\left\{k - \sum_{j'\neq j} x_{j'}^*,\; U_j - \Delta\right\}$.
8:     **end if**
9: **end for**
10: $K' := K - x^*$.
11: $z :=$ SAMPLING-ORACLE$\left(\left(1 + \frac{\sqrt{\ell}}{\Delta}\right) K', \delta\right)$.
12: **if** $j \in \left[\left|\sum_j \lfloor z_j \rfloor\right|\right]$, $x_j := \lceil z_j \rceil$; **else** $x_j := \lfloor z_j \rfloor$.
13: **if** $x \in K'$, return $x + x^*$; **else** reject $x$, go to Step 11.

with expected running time $\mathcal{O}^*(k^2\ell^2)$ to sample a close to uniform random group-fair representation (Theorem 4.2).

**Theorem 4.2.** *Let $L_j, U_j \in \mathbb{Z}_{\geqslant 0}, \forall j \in [\ell]$ be the fairness constraints and $k \in \mathbb{Z}_{\geqslant 0}$ be the size of the ranking. Let $\Delta$ be as defined in Algorithm 2. Then for any non-negative number $\delta < e^{-2\frac{\ell\sqrt{\ell}}{\Delta}}$, Algorithm 2 samples a random point from a density that is within total variation distance $\delta$ from the uniform distribution on the integral points in $K$ by making $1/\left(e^{-2\frac{\ell\sqrt{\ell}}{\Delta}} - \delta\right)$ calls to the oracle in expectation. When $\delta$ is a non-negative constant, such that $\delta < e^{-2}$ and $\Delta = \Omega\left(\ell^{1.5}\right)$, Algorithm 2 calls the oracle only a constant number of times in expectation, and each oracle call takes time $\mathcal{O}^*\left(k^2\ell^2\right)$.*

**Overview of Algorithm 2 and the proof of Theorem 4.2**

Let $H, P$, and $\Delta$ be as defined in Steps 1, 2 and 3 respectively. Clearly, $K = H \cap P$. We first find an integral center in $x^* \in H \cap P$ (Steps 4 to 7) such that there is a ball of radius $\Delta$ in $P$ (see Lemma B.1) and translate the origin to this point $x^*$ (Step 10). This ensures that there exists a bijection between the set of integral points in the translated polytope $K'$ and the original polytope $K$ (see proof of Theorem 4.2). We now sample a rational point $z$ uniformly at random from the expanded polytope $\left(1 + \frac{\sqrt{\ell}}{\Delta}\right) K'$, using SAMPLING-ORACLE (Step 11). We then round the point $z$ to an integer point on $H'$ (Step 12). We prove that our deterministic rounding algorithm ensures that the set of points in the expanded polytope that get rounded to an integral point on $H'$ is contained inside a cube of side length 2 around this point (Lemma B.2) and that this cube is fully contained in this expanded polytope (Lemma B.3). Lemma B.5 gives us that for any two integral points $x$ and $x'$, there is a bijection between the set of points that get rounded to these points. Therefore, every integral point is sampled from a distribution close to uniform, given the SAMPLING-ORACLE samples any rational point in the expanded polytope from a distribution $\delta = 0.1$ close to uniform. If the rounded point belongs to $K'$, we accept, else we reject and go to Step 11. We then lower bound the probability of acceptance. The algorithm is run until a point is accepted. Hence, the expected running time polynomial is inversely proportional to the probability of acceptance, which is exponential in $\ell$ in expectation. However, if $\Delta = \Omega\left(\ell^{1.5}\right)$ and $\delta < e^{-2}$ the probability of acceptance is at least a constant. Note that the value of $R^2$ in Theorem B.1 for the polytope $K'$ is $k^2$. Therefore, the algorithm by Cousins and Vempala (2018) gives a rational point from $K'$, from a distribution close to uniform, in time $\mathcal{O}^*\left(k^2\ell^2\right)$. Therefore each oracle call in Theorem 4.2 takes time $\mathcal{O}^*(k^2\ell^2)$ when we use Theorem B.1 as the SAMPLING-ORACLE.

**Comparison with Kannan and Vempala (1997).** Algorithm 2 is inspired by the algorithm by Kannan and Vempala (1997) to sample integral points from a convex polytope from a distribution close to uniform. On a high level, their algorithm on polytope $K'$ works slightly differently when compared to our algorithm on $K'$. They first expand $K'$ by $\mathcal{O}\left(\sqrt{\ell \log \ell}\right)$, sample a rational point from a distribution close to uniform over this expanded polytope (similar to Step 11), and use a probabilistic rounding method to round it to an integral point. Their algorithm requires that a ball of radius $\Omega\left(\ell^{1.5}\sqrt{\log \ell}\right)$ lies entirely inside $K'$. We expand the polytope $K'$ by $\mathcal{O}(\sqrt{\ell})$, sample a rational point from this polytope from a distribution close to uniform, and then deterministically round it to an integral point. Our algorithm only requires that a ball of radius $\Omega\left(\ell^{1.5}\right)$ lies inside $P - x^*$ with center on $H - x^*$, where $P$, $H$ and $x^*$ are as defined in Algorithm 2. As a result, we get an expected polynomial time algorithm for a larger set of fairness constraints. We also note here that the analysis of the success probability of Algorithm 2 is the same as that of the algorithm by Kannan and Vempala (1997).

In Appendix B.1, we give our third exact sampling algorithm, which runs faster than our DP for small values of $\ell$.

### 4.3 Prefix Fairness Constraints

Prefix fairness constraints are represented by the numbers $L_{ij}$ and $U_{ij}$, for all $j \in [\ell]$ and $i \in M$, where $M \subseteq [k]$, which give lower and upper bounds on the representation of group $j$ in the top $i$ ranks, i.e., the top $i$ prefix of the top $k$ ranking. When $M = \{k\}$, this gives us the setup we started with. This model has been first studied by [Celis *et al.*, 2018b; Yang and Stoyanovich, 2017], who give a deterministic algorithm to get a utility-maximizing ranking while satisfying prefix fairness constraints. To overcome the limitations of a deterministic ranking, we propose to use our algorithm[5] as a heuristic to output randomized rankings under prefix fairness constraints. It inductively finds a group-fair assignment in *blocks* between two ranks (ranks $i + 1$ to $i'$ such that $i, i' \in M$ and $i'' \notin M, \forall i'' \in [i + 1, i' - 1]$), as follows:

[5]We use Algorithm 2 as it is faster than Algorithm 1 in practice.

1. Let us assume we have a random sample of the top $i$ ranking for some $i \in M$. Let $w_{ij}$ be the counts of groups $j \in [\ell]$ in this top-$i$ ranking.
2. Let $i' \in M$ be the smallest rank larger than $i$ in the set $M$. Use Algorithm 2 to find an integer solution $x_j^{(i')}$ in $K = \{x \in \mathbb{R}^\ell | \sum_{j\in[\ell]} x_j = i' - i + 1,\ \max\{0, L_{i'j} - w_{ij}\} \leqslant x_j \leqslant \min\{i' - i + 1, U_{i'j} - w_{ij}\}, \forall j \in [\ell]\}$ to get a group-fair representation for ranks $i + 1$ to $i'$.
3. Find a uniform random permutation of $x_j^{(i')}$ similar to Step 2 in Theorem 3.4 to get a group-fair assignment for ranks $i + 1$ to $i'$, and go to Step 1 with $i = i'$.

We call this algorithm 'Prefix Random Walk'.

# 5 Experimental Results

In this section, we run our algorithms on various real-world datasets. We implement Algorithm 2 (called 'Random walk' in plots) using the tool called *PolytopeSampler*[6] to sample a point from a distribution close to uniform, on the convex rational polytope $K$. This tool implements a constrained Riemannian Hamiltonian Monte Carlo for sampling from high dimensional distributions on polytopes [Kook *et al.*, 2022].

**Datasets.** We evaluate our results on the German Credit Risk dataset comprising credit risk scoring of 1000 adult German residents [Dua and Graff, 2017], along with their demographic information (e.g., gender, age, etc.). We use the Schufa scores of these individuals to get the in-group rankings; grouping based on $age < 25$ (see Figure 1) similar to [Zehlike *et al.*, 2017; Gorantla *et al.*, 2021; Castillo, 2019], who observed that Schufa scores are biased against the adults of $age < 25$. Their representation in the top 100 ranks is $10\%$ even though their true representation in the whole dataset is $15\%$. We also evaluate our algorithm on the IIT-JEE 2009 dataset, also used in Celis *et al.* (2020b). The dataset consists of the student test scores of the joint entrance examination (JEE) conducted for undergrad admissions at the Indian Institutes of Technology (IITs). Information about the students includes gender details (25% women and 75% men)[7]. Students' test scores give score-based in-group rankings. We evaluate our algorithm with *female* as the protected group, as they are consistently underrepresented ($0.04\%$ in top 100 [Celis *et al.*, 2020b]), in a score-based ranking on the entire dataset, despite $25\%$ female representation in the dataset.

**Baselines.** $(i)$ We compare our experimental results with *fair $\epsilon$-greedy* [Gao and Shah, 2020], which is a greedy algorithm with $\epsilon$ as a parameter (explained in detail in Section 5.1). To the best of our knowledge, this algorithm is the closest state-of-the-art baseline to our setting, as it does not rely on comparing the scores of two candidates from different groups. $(ii)$ We also compare our results with a recent deterministic re-ranking algorithm (GDL21) given by Gorantla *et al.* (2021), which achieves the best balance of both group fairness and underranking of individual items compared to their original ranks in top-$k$.

**Plots.** We plot our results for the protected groups in each dataset (see Figures 1 and 2). We use the representation constraints $L_j = \lceil (p_j^* - \eta)k \rceil$ and $U_j = \lfloor (p_j^* + \eta)k \rfloor$ for group $j$ where $p_j^*$ is the total fraction of items from group $j$ in the dataset and $\eta = 0.1$. For "prefix random walk" we put constraints at every $b$ ranks, i.e., $L_{ij} = \left\lceil \left(p_j^* - \frac{\eta b}{\max\{b, k-i\}}\right) k \right\rceil$ and $U_{ij} = \left\lfloor \left(p_j^* + \frac{\eta b}{\max\{b, k-i\}}\right) k \right\rfloor$ with $i \in \{b, 2b, \ldots\}$. We use $k = 100$ and $b = 50$ in the experiments. With these, the representation constraints are stronger in the top 50 ranks than in the top 100 ranks. The "representation" (on the y-axis) plot shows the fraction of ranks assigned to the protected group in the top $i$ ranks (on the x-axis). We sample 1000 rankings for randomized algorithms and output the mean and standard deviation. The dashed red line is the true representation of the protected group in the dataset, which we call $p^*$, dropping the subscript. The "fraction of rankings" plot for randomized ranking algorithms represents the fraction of 1000 rankings that assign rank $i$ to the protected group. For completeness, we plot the results for the ranking utility metric, normalized discounted cumulative gain, defined as nDCG@$i$ $= \left(\sum_{i'\in[i]} \frac{(2^{\hat{s}_{i'}}-1)}{\log_2(i'+1)}\right) / \left(\sum_{i'\in[i]} \frac{(2^{s_{i'}}-1)}{\log_2(i'+1)}\right)$, where $\hat{s}_{i'}$ and $s_{i'}$ are the scores of the items assigned to rank $i'$ in the group-fair and the score-based ranking, respectively.

## 5.1 Observations

The rankings sampled by our algorithms have the following property: for any rank $i$, rank $i$ is assigned to the protected group in a sufficient fraction of rankings (see plots with "fraction of rankings" on the y-axis). This experimentally validates our Theorem 3.5. Moreover, this fraction is stable across the ranks. Whereas fair $\epsilon$-greedy fluctuates a lot, which can be explained as follows. For each rank $k' = 1$ to $k$, with $\epsilon$ probability, it assigns a group uniformly at random, and with $1 - \epsilon$ probability, it assigns group $G_1$ := (age$\geqslant$ 25) if the number of ranks assigned to $G_1$ is less than $(\frac{L_1 k'}{k})$ in the top $k'$ ranks, and to $G_2$ := (age $<$ 25) otherwise. Consider Figure 1 top row where $L_1 = 80$, $L_2 = 10$, and $k = 100$, and the plot on the right shows the fraction of rankings (y-axis) assigning rank $i$ (x-axis) to $G_1$. Note that if $\epsilon = 0$ (no randomization), this algorithm gives a deterministic ranking where the first four ranks are assigned to $G_2$, and the fifth to $G_1$, and this pattern repeats after every five ranks. Hence, there would be a peak in the plot at ranks $k' = 5, 10, 15, 20, \ldots$. Now, when $\epsilon = 0.3$, fair-$\epsilon$-greedy introduces randomness in group assignment at each rank and, as a result, smoothens out the peaks as $k'$ increases, which is exactly what is observed. Therefore, the first four ranks will have very low representation, even in expectation. Similarly, the ranks 6 to 9. Clearly, fair-$\epsilon$-greedy does not satisfy fairness for any $k' < k$ consecutive ranks. But our algorithm satisfies this property, as is also confirmed by Corollary 3.6.

Our algorithms satisfy representation constraints for the protected group in the top $k'$ ranks in expectation (see plots with "representation" on the y-axis). Fair $\epsilon$-greedy overcompensates for representing the protected group. The deterministic algorithm GDL21 achieves very high nDCG but very low representation for smaller values of $k'$, although all run with

[6] PolytopeSamplerMatlab (License: GNU GPL v3.0)

[7] Only binary gender information was annotated in the dataset.

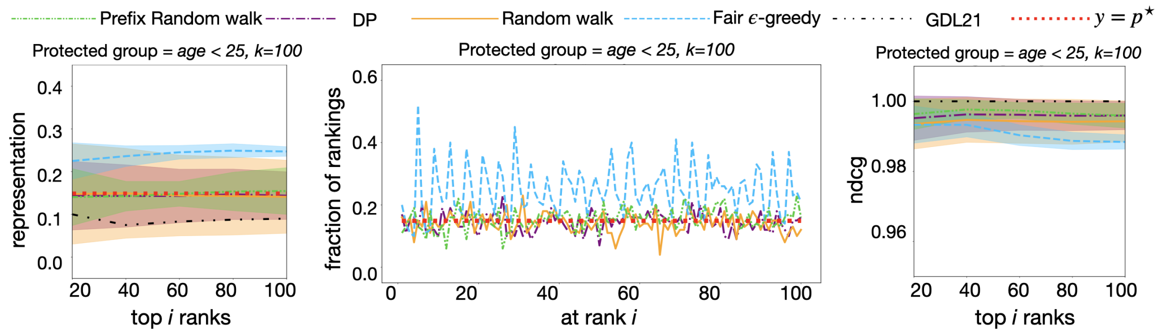


Figure 1: Results on the German Credit Risk dataset with *age* $< 25$ as the protected group. For Fair $\epsilon$-greedy we use $\epsilon = 0.3$ (see Figures 5 and 7 for other values of $\epsilon$). In the plots, the means of the DP and the random walk algorithms are almost coinciding.

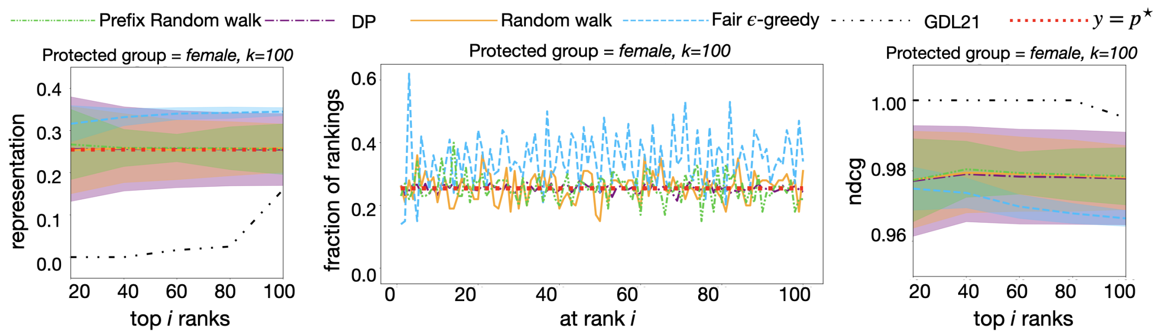


Figure 2: Results on the JEE 2009 dataset with *gender* as the protected group. For Fair $\epsilon$-greedy we use $\epsilon = 0.3$ (see Figures 6 and 8 for other values of $\epsilon$). In the plots, the means of the DP and the random walk algorithms are almost coinciding.

similar representation constraints. This is because the deterministic algorithm uses comparisons based on the scores, hence putting most of the protected group items in higher ranks (towards $k$). With a larger value of $\epsilon$, fair $\epsilon$-greedy gets a much higher "representation" of the protected group than necessary (see Figures 7 and 8), whereas, with a smaller value of $\epsilon$, it fluctuates a lot in the "fraction of rankings" (see see Figures 5 and 6). Our "Prefix Random Walk" algorithm is run with stronger fairness requirements than "Random Walk" and "DP" in the top 50 ranks, which can be observed by its smaller deviation from the line $y = p^*$ in the left-most plots. The random walk runs very fast, even for a large number of groups (Figure 3). We also run experiments on the JEE 2009 dataset with birth category defining 5 groups (see Appendix C). The experiments were run on a Quad-Core Intel Core i5 processor consisting of 4 cores, with a clock speed of 2.3 GHz and DRAM of 8GB. Implementation of our algorithms and the baselines has been made available for reproducibility[8].

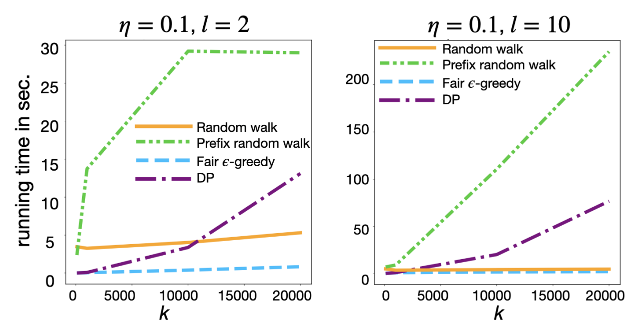


Figure 3: Average running time in seconds of the algorithms, over 5 runs, to sample a ranking. For prefix random walk, we add prefix constraints with $b = 50, 200, 400, 400$ for $k = 100, 1000, 10000, 20000$ respectively.

## 6 Conclusion

We take an axiomatic approach to define randomized group-fair rankings and show that it leads to a unique distribution over all feasible rankings that satisfy lower and upper bounds on the group-wise representation in the top ranks. We propose practical and efficient algorithms to exactly and approximately sample a random group-fair ranking from this distribution. Our approach requires merging a given set of ranked lists, one for each group, and can help circumvent implicit bias or incomplete comparison data across groups.

The natural open problem is to extend our method to work even for noisy, uncertain inputs about rankings within each group. Even though our heuristic algorithm does output ex-post group-fair rankings under prefix constraints, it is important to investigate the possibility of polynomial-time algorithms to sample from the distribution that satisfies natural extensions of our axioms for prefix group-fair rankings.

[8] github.com/sruthigorantla/SamplingExPostGroupFairRankings.

## Ethical Statement

A limitation of our work as a post-processing method is that it cannot fix all sources of bias, e.g., bias in data collection and labeling. Randomized rankings can be risky and opaque in high-risk, one-time ranking applications. Our guarantees for group fairness may not necessarily reflect the right fairness metrics for all downstream applications for reasons including biased, noisy, incomplete data and legal or ethical considerations in quantifying the eventual adverse impact on individuals and groups.

## Acknowledgements

SG was supported by a Google PhD Fellowship. The proof of Theorem 4.1 was suggested to us by Santosh Vempala. The authors would like to thank him for allowing us to include it in our paper. AL was supported in part by the SERB Award ECR/2017/003296 and a Pratiksha Trust Young Investigator Award. AL is also grateful to Microsoft Research for supporting this collaboration.

## Contribution Statement

AD and AL have made equal contributions to this work.

## References

[Baswana *et al.*, 2019] Surender Baswana, Partha Pratim Chakrabarti, Sharat Chandran, Yashodhan Kanoria, and Utkarsh Patange. Centralized admissions for engineering colleges in india. *Interfaces*, 49(5):338–354, 2019.

[Beutel *et al.*, 2019] Alex Beutel, Jilin Chen, Tulsee Doshi, Hai Qian, Li Wei, Yi Wu, Lukasz Heldt, Zhe Zhao, Lichan Hong, Ed H. Chi, and Cristos Goodrow. Fairness in recommendation ranking through pairwise comparisons. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '19, page 2212–2220, New York, NY, USA, 2019. Association for Computing Machinery.

[Biega *et al.*, 2018] Asia J. Biega, Krishna P. Gummadi, and Gerhard Weikum. Equity of attention: Amortizing individual fairness in rankings. In *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '18, page 405–414, New York, NY, USA, 2018. Association for Computing Machinery.

[Bobko and Roth, 2004] P. Bobko and P.L. Roth. The four-fifths rule for assessing adverse impact: An arithmetic, intuitive, and logical analysis of the rule and implications for future research and practice. *Research in Personnel and Human Resources Management*, 23, 2004.

[Castillo, 2019] Carlos Castillo. Fairness and transparency in ranking. *SIGIR Forum*, 52(2):64–71, jan 2019.

[Celis *et al.*, 2018a] L. Elisa Celis, Vijay Keswani, Damian Straszak, Amit Deshpande, Tarun Kathuria, and Nisheeth K. Vishnoi. Fair and diverse dpp-based data summarization. In *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 715–724. PMLR, 2018.

[Celis *et al.*, 2018b] L. Elisa Celis, Damian Straszak, and Nisheeth K. Vishnoi. Ranking with fairness constraints. In *ICALP*, 2018.

[Celis *et al.*, 2019] L. Elisa Celis, Sayash Kapoor, Farnood Salehi, and Nisheeth Vishnoi. Controlling polarization in personalization: An algorithmic framework. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 160–169. Association for Computing Machinery, 2019.

[Celis *et al.*, 2020a] L. Elisa Celis, Vijay Keswani, and Nisheeth K. Vishnoi. Data preprocessing to mitigate bias: A maximum entropy based approach. In *ICML*, 2020.

[Celis *et al.*, 2020b] L. Elisa Celis, Anay Mehrotra, and Nisheeth K. Vishnoi. Interventions for ranking in the presence of implicit bias. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, page 369–380. Association for Computing Machinery, 2020.

[Chierichetti *et al.*, 2017] Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[Collins, 2007] Brian W. Collins. Tackling unconscious bias in hiring practices: The plight of the rooney rule. *New York University Law Review*, 82, 2007.

[Cousins and Vempala, 2018] Benjamin R. Cousins and Santosh S. Vempala. Gaussian cooling and $\mathcal{O}^*(n^3)$ algorithms for volume and gaussian volume. *SIAM J. Comput.*, 47:1237–1273, 2018.

[Diaz *et al.*, 2020] Fernando Diaz, Bhaskar Mitra, Michael D. Ekstrand, Asia J. Biega, and Ben Carterette. *Evaluating Stochastic Rankings with Expected Exposure*, page 275–284. 2020.

[Dua and Graff, 2017] Dheeru Dua and Casey Graff. UCI machine learning repository. *University of California, Irvine, School of Information and Computer Sciences*, 2017.

[Dyer *et al.*, 1991] Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. ACM*, 38(1):1–17, jan 1991.

[Gao and Shah, 2020] Ruoyuan Gao and Chirag Shah. Toward creating a fairer ranking in search engine results. *Information Processing and Management*, 57(1):102138, 2020.

[Geyik *et al.*, 2019] Sahin Cem Geyik, Stuart Ambler, and Krishnaram Kenthapadi. Fairness-aware ranking in search and recommendation systems with application to linkedin talent search. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '19, page 2221–2231, New York, NY, USA, 2019. Association for Computing Machinery.

[Gorantla *et al.*, 2021] Sruthi Gorantla, Amit Deshpande, and Anand Louis. On the problem of underranking in group-fair ranking. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 3777–3787. PMLR, 18–24 Jul 2021.

[Goto *et al.*, 2016] Masahiro Goto, Atsushi Iwasaki, Yujiro Kawasaki, Ryoji Kurata, Yosuke Yasuda, and Makoto Yokoo. Strategyproof matching with regional minimum and maximum quotas. *Artificial Intelligence*, 235:40–57, 2016.

[Hassani, 2021] Bertrand K. Hassani. Societal bias reinforcement through machine learning: a credit scoring perspective. *AI and Ethics*, 1:239–247, 2021.

[Heuss *et al.*, 2022] Maria Heuss, Fatemeh Sarvi, and Maarten de Rijke. Fairness of exposure in light of incomplete exposure estimation. SIGIR '22, page 759–769, 2022.

[Kannan and Vempala, 1997] Ravi Kannan and Santosh Vempala. Sampling lattice points. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97, page 696–700, New York, NY, USA, 1997. Association for Computing Machinery.

[Kletti *et al.*, 2022] Till Kletti, Jean-Michel Renders, and Patrick Loiseau. Introducing the expohedron for efficient pareto-optimal fairness-utility amortizations in repeated rankings. WSDM '22. Association for Computing Machinery, 2022.

[Kook *et al.*, 2022] Yunbum Kook, Yin Tat Lee, Ruoqi Shen, and Santosh S. Vempala. Sampling with riemannian hamiltonian monte carlo in a constrained space. In *NeurIPS*, 2022.

[Kuhlman *et al.*, 2019] Caitlin Kuhlman, MaryAnn Van-Valkenburg, and Elke Rundensteiner. Fare: Diagnostics for fair ranking using pairwise error metrics. In *The World Wide Web Conference*, WWW '19, page 2936–2942, New York, NY, USA, 2019. Association for Computing Machinery.

[Lovász and Vempala, 2006] László Lovász and Santosh Vempala. Hit-and-run from a corner. *SIAM J. Comput.*, 35(4):985–1005, apr 2006.

[Memarrast *et al.*, 2021] Omid Memarrast, Ashkan Rezaei, Rizal Fathony, and Brian D. Ziebart. Fairness for robust learning to rank. *CoRR*, abs/2112.06288, 2021.

[Narasimhan *et al.*, 2020] Harikrishna Narasimhan, Andy Cotter, Maya Gupta, and Serena Lutong Wang. Pairwise fairness for ranking and regression. In *33rd AAAI Conference on Artificial Intelligence*, 2020.

[Okonofua and Eberhardt, 2015] Jason A. Okonofua and Jennifer L. Eberhardt. Two strikes: Race and the disciplining of young students. *Psychological Science*, 26(5):617–624, 2015.

[Singh and Joachims, 2018] Ashudeep Singh and Thorsten Joachims. Fairness of exposure in rankings. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 2219–2228, 2018.

[Singh *et al.*, 2021] Ashudeep Singh, David Kempe, and Thorsten Joachims. Fairness in ranking under uncertainty. In *NeurIPS 2021, December 6-14, 2021*, pages 11896–11908, 2021.

[Stoyanovich *et al.*, 2018] Julia Stoyanovich, Ke Yang, and HV Jagadish. Online set selection with fairness and diversity constraints. Advances in Database Technology - EDBT, pages 241–252, 2018.

[Uhlmann and Cohen, 2005] Eric Luis Uhlmann and Geoffrey L. Cohen. Constructed criteria: Redefining merit to justify discrimination. *Psychological Science*, 16(6):474–480, 2005.

[Štefankovič *et al.*, 2012] Daniel Štefankovič, Santosh Vempala, and Eric Vigoda. A deterministic polynomial-time approximation scheme for counting knapsack solutions. *SIAM Journal on Computing*, 41(2):356–366, 2012.

[Wu *et al.*, 2018] Yongkai Wu, Lu Zhang, and Xintao Wu. On discrimination discovery and removal in ranked data using causal graph. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '18, page 2536–2544, New York, NY, USA, 2018. Association for Computing Machinery.

[Yang and Stoyanovich, 2017] Ke Yang and Julia Stoyanovich. Measuring fairness in ranked outputs. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, SSDBM '17, New York, NY, USA, 2017. Association for Computing Machinery.

[Zehlike *et al.*, 2017] Meike Zehlike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed, and Ricardo Baeza-Yates. Fa*ir: A fair top-k ranking algorithm. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, CIKM '17, page 1569–1578, New York, NY, USA, 2017. Association for Computing Machinery.

[Zehlike *et al.*, 2022a] Meike Zehlike, Tom Sühr, Ricardo Baeza-Yates, Francesco Bonchi, Carlos Castillo, and Sara Hajian. Fair top-k ranking with multiple protected groups. *Information Processing and Management*, 59(1):102707, 2022.

[Zehlike *et al.*, 2022b] Meike Zehlike, Ke Yang, and Julia Stoyanovich. Fairness in ranking, part i: Score-based ranking. *ACM Comput. Surv.*, apr 2022. Just Accepted.