Individual Fairness under Group Fairness Constraints in Bipartite Matching - One Framework to Approximate Them All

Atasi Panda\(^1\), Anand Louis\(^1\) and Prajakta Nimbhorkar\(^2\)

\(^1\)Indian Institute of Science
\(^2\)Chennai Mathematical Institute

{atasipanda, anandl}@iisc.ac.in, prajakta.nimbhorkar@cmi.ac.in

Abstract

We study the probabilistic assignment of items to platforms that satisfies both group and individual fairness constraints. Each item belongs to specific groups and has a preference ordering over platforms. Each platform enforces group fairness by limiting the number of items per group that can be assigned to it. There could be multiple optimal solutions that satisfy the group fairness constraints, but this alone ignores item preferences. Our approach explores a ‘best of both worlds fairness’ solution to get a randomized matching, which is ex-ante individually fair and ex-post group-fair. Thus, we seek a ‘probabilistic individually fair’ distribution over ‘group-fair’ matchings where each item has a ‘high’ probability of matching to one of its top choices. This distribution is also ex-ante group-fair. Users can customize fairness constraints to suit their requirements. Our first result is a polynomial-time algorithm that computes a distribution over ‘group-fair’ matchings such that the individual fairness constraints are approximately satisfied and the expected size of a matching is close to OPT. We empirically test this on real-world datasets. We present two additional polynomial-time bi-criteria approximation algorithms that users can choose from to balance group fairness and individual fairness trade-offs.

For disjoint groups, we provide an exact polynomial-time algorithm adaptable to additional lower ‘group fairness’ bounds. Extending our model, we encompass ‘maxmin group fairness,’ amplifying under-represented groups, and ‘mindom group fairness,’ reducing the representation of dominant groups.

1 Introduction

Matching is a foundational concept in theoretical computer science, well-studied over several years. Maximum bipartite matching finds applications in real-world scenarios, such as ad-auctions [Mehta et al., 2007; Mehta, 2013], resource allocation [Halabian et al., 2012], scheduling [McKeown et al., 1996], school choice [Abdulkadiroglu and Sönmez, 2003; Kurata et al., 2017], and healthcare rationing [Aziz and Brandl, 2021; Ganesh et al., 2023]. In this paper, we refer to the two partitions of the underlying bipartite graph as items and platforms. A matching is an allocation of items to platforms, allowing multiple assignments for each item and platform. Real-world items often have diverse attributes, leading to their categorization into different groups. To ensure equitable representation among these groups, it is natural to enforce group fairness constraints [Definition 3.1], which limit the number of items per group assigned to a platform by enforcing upper bounds. Also, one can specify lower bounds on the number of items from each group that need to be assigned to a platform, so as to ensure a minimum representation from each group among the items matched to a platform. The above constraints thus achieve Restricted Dominance introduced in [Bera et al., 2019], which asserts that the representation from any group on any platform does not exceed a user-specified cap, and Minority Protection [Bera et al., 2019], which asserts that the representation from any group, among the items matched to any platform is at least a user-specified bound [Definition 3.2].

Both the definitions of group fairness are well-motivated by various applications like school choice, formation of committees in an organization, or teams to work on projects. For instance, in school choice, group-fairness constraints can promote diversity among students assigned to each school based on attributes like ethnicity and socioeconomic background, as observed in practical implementations [Cowen Institute, 2011]. Similarly, in project teams, group fairness constraints ensure the inclusion of experts from all required fields. Both definitions of group fairness are motivated by the Disparate Impact doctrine [Feldman et al., 2015], which broadly posits addressing unintentional bias, which leads to widely different outcomes for different groups.

However, since items have preferences over platforms, a matching meeting group fairness constraints alone may not be fair to individual items. The exclusive use of group fairness constraints can lead to sub-optimal outcomes for individuals. Furthermore, deterministic algorithms for matching assign top choices to some individuals while assigning less preferred choices to others. This necessitates the introduction of individual fairness constraints. In this paper, we consider probabilistic individual fairness constraints, first introduced in robust clustering [Anegg et al., 2020; Harris et al., 2019]. Instead of a single matching, the goal is to generate a distribution on group-fair matchings such that, in a matching sampled from the said distribution, the proba-
bility of each item being matched to one of its top choices is within the user-specified bounds [Definition 3.3]. Thus, this approach, known as the best of both worlds fairness approach in literature, aims to compute an outcome with both ex-ante and ex-post fairness guarantees.

In this paper, the central objective is to design efficient algorithms that compute an ex-ante probabilistic individually fair distribution over deterministic group-fair matchings.

1.1 An Overview of Our Results and Techniques

Our approach revolves around formulating various notions of individual and group fairness using linear programming (LP). The key idea is to represent the LP’s optimal solution as a convex combination of integer group-fair matchings, enabling the satisfaction of probabilistic individual fairness constraints through sampling. However, depending on the structure of the groups, it may not be possible to express the LP optimum as an exact convex combination of integral group-fair matchings. Nonetheless, our algorithms express an approximate LP optimum as a convex combination of integer matchings.

Our technique leads to a unified framework for different group fairness notions beyond fixed upper and lower bound constraints on the number of items from each group that can be matched to a platform. Two such notions, referred to as maxmin group fairness and mindom group fairness in this paper, are discussed in a later section (see Section 3). In maxmin group fairness, the goal is to maximize the minimum number of items that get matched to any platform from any one group. In mindom group fairness, the goal is to minimize the maximum number of items that get matched to any platform from any one group. Informally, both these notions aim to get a matching with nearly equal representation from all groups. (See Section 3 Definitions 3.9 and 3.10 for formal definitions).

In a similar spirit, for individual fairness, one can aim to provide the strongest possible guarantee simultaneously to all individuals in terms of the probability of being matched. We refer to this as maxmin individual fairness [Definition 3.8].

2 Related Work

Several allocation problems like resource allocation [Halabian et al., 2012], kidney exchange programs [Farnadi et al., 2021], school choice [Abdulkadiroglu and Sönmez, 2003], candidate selection [Bei et al., 2020], summer internship programs [Aziz et al., 2020], and matching residents to hospitals [Goko et al., 2022] are modeled as matching problems. [Manlove, 2013] extensively examines preference-based matching in the stable marriage and roommate problems, hospitals/residents matching, and the house allocation problem. Since the people/items to be matched may belong to different groups, bipartite matchings under various notions of group fairness have been studied and their significance has been emphasized in literature [Celis et al., 2017; Luss, 1999; Devanur et al., 2013; Costello et al., 2016; Segal-Halevi and Suksompong, 2019; Kay et al., 2015; Bolukbasi et al., 2016]. [Aziz et al., 2022] survey the developments in the field of matching with constraints, including those based on regions, diversity, multi-dimensional capacities, and matroids. The fairness constraints are captured by upper and lower bounds [Huang, 2010; Gonczarowski et al., 2019], justified envy-freeness [Abdulkadiroglu and Sönmez, 2003], or in terms of proportion of the final matching size [Bei et al., 2020]. Historically, discriminated groups in India are protected with vertical reservations implemented as set-asides, and other disadvantaged groups are protected with horizontal reservations implemented as minimum guarantees(lower bounds) [Sönmez and Yenmez, 2022].

In some applications, the items could belong to multiple groups as well. [Sankar et al., 2021] present a polynomial-time algorithm with an approximation ratio of \( \frac{1}{\Delta+1} \) where each item belongs to at most \( \Delta \) laminar families of groups per platform, and [Nasre et al., 2019] show the NP-hardness of the problem without a laminar structure. While both papers focus primarily on group-fairness upper bounds, [Louis et al., 2023] focus on proportional diversity constraints with an emphasis on lower bounds in the general context. However, group fairness constraints alone do not account for individual preferences. Our work aims to introduce individual fairness considerations into the problem and explore both upper and lower bounds for specific scenarios.

The notions of maxmin individual fairness, maxmin group fairness, and mindom group fairness are motivated by existing literature. Maxmin individual fairness, originally termed as the "distributional maxmin fairness" framework in [García-Soriano and Bonchi, 2020], was further explored in group-fair ranking problems by [García-Soriano and Bonchi, 2021]. Their distribution is only over maximum matchings, and we extend this idea to a distribution over maximum group-fair matchings and a stronger notion of individual fairness. Maxmin group fairness is a natural extension of Maxmin fairness, initially introduced as a network design objective by [Bertsekas and Gallager, 2021](Section 6.5.2) and extensively studied in various areas of networking [Radunovic and Le Boudec, 2007; Hahne, 1991]. Mindom group fairness has been studied in network load distribution [Georgiadis et al., 2001], transmission cost sharing [Abhyankar et al., 2007], and other network applications [Radunovic and Le Boudec, 2007]. This concept of probabilistic individual fairness also has applications in fair-ranking [García-Soriano and Bonchi, 2021; Gorantla et al., 2023] and graph-cut problems [Dinitz et al., 2022].

[Singh et al., 2021] study individual fairness in ranking under uncertainty, extending fairness definitions by explicitly modeling incomplete information. Their approach mirrors [Racke, 2008] but assumes a posterior distribution over candidates’ merits. Similarly, we assume that the individual and group fairness parameters are given. They express a distribution \( \pi \) over rankings as a bistochastic matrix, where each entry denotes the probability of a candidate’s position under \( \pi \). They use an LP to maximize utility while enforcing fairness constraints and ensure marginal probabilities form a doubly stochastic matrix. The optimal solution is then decomposed as a distribution over rankings using the Birkhoff-von Neumann algorithm [Birkhoff, 1946]. Though we use a similar approach for Theorems 3.5, 3.4 and 3.7, we have both group and individual fairness constraints. Our marginal probabilities do not form a doubly stochastic matrix, so we cannot use the Birkhoff-von Neumann decomposition [Birkhoff, 1946].
directly, and hence need a different approach for Theorem 3.4. Similar to our group and individual fairness constraints (Definitions 3.2 and 3.3), [Gorantla et al., 2023] address a related problem in fair ranking, particularly addressing laminar set structures using techniques akin to the Birkhoff-von Neumann decomposition.

Fairness constraints with bounds on the number of items with each attribute are also studied in ranking and multi-winner voting [Celis et al., 2018b; Celis et al., 2018a]. Among other notions of fairness, [Sühk et al., 2019] propose a fairness notion for ride-hailing platforms that distributes fairness over time, ensuring benefits proportional to drivers’ platform engagement duration. Kletti et al. [Kletti et al., 2022] present an algorithm for optimizing rankings to maximize consumer utility while minimizing producer-side individual exposure unfairness. [García-Soriano and Bonchi, 2021] explore maximin fair distributions in general search problems with group fairness constraints, while [Esmaili et al., 2022] examine Rawlsian fairness (maximin fairness) in online bipartite matching, considering both group and individual fairness. [Esmaili et al., 2022] simultaneously address two-sided fairness but treat group and individual fairness separately. In contrast, we handle both individual and group fairness on the item side within a single bipartite matching instance, which has not been explored in the existing literature to the best of our knowledge.

Our solution fits into the best of both worlds (BoBW) fairness paradigm, which is gaining attention in the fair allocation of indivisible items [Aziz et al., 2023; Babaioff et al., 2022; Freeman et al., 2020; Aziz, 2020]. In literature, popular target fairness properties have been envy-freeness, envy-freeness up to one item [Freeman et al., 2020; Aziz, 2020], proportionality, and proportionality up to one item [Aziz et al., 2023; Hoefer et al., 2023; H.V. and Nimubhokar, 2024]. Other than these, [Babaioff et al., 2022] study truncated proportional share. [Freeman et al., 2020] showed that ex-ante envy-freeness (EF) and ex-post envy-freeness up to one item (EF1) BoBW outcomes are achievable for any allocation problem instance. [Aziz et al., 2023] studied BoBW outcomes based on envy-based fairness in allocating indivisible items to agents with additive valuations and weighted entitlements. [Babaioff et al., 2022] approach BoBW fairness from a fair-share guarantee perspective. While our target fairness properties are group fairness and probabilistic individual fairness, our technique, in essence, resembles that of [Aziz, 2020], where a randomized EF allocation is first generated and then decomposed as the convex combination of EF1 deterministic allocations.

Fairness constraints, with various notions of fairness, have been considered in preference-based matchings, e.g., for kidney-exchange [Farnadi et al., 2021], for rank-maximality and popularity [Nasre et al., 2019], stability [Huang, 2010], stability under matroid constraints [Feiiner and Kamiyama, 2016], and in various settings of two-sided matching markets [Beyhaghi and Éva Tardos, 2021], [Patro et al., 2020], [Huang et al., 2016].

3 Preliminaries

Our problem: The input instance consists of a bipartite graph denoted as \(G = (A \cup P, E)\). Here \(A\) denotes the set of items and \(P\) is the set of platforms. There is an edge, \((a, p) \in E\) if \(a\) can be assigned to \(p\). The items are grouped into possibly non-disjoint subsets \(A_1, A_2, \ldots, A_\chi\) for an integer \(\chi \geq 1\) such that \(\bigcup_{h \in [\chi]} A_h = A\). Here \(\chi\) denotes the total number of groups. Let \(|A| = n, |P| = m\), \(\Delta\) denote the maximum number of distinct groups to which any item belongs, and \(N(v)\) denote the neighborhood of any node \(v \in A \cup P\). Each item \(a \in A\) has a preference list \(R_a\), which contains a ranking of platforms, and let \(R_{a,k}\) denote the set of top \(k\) preferred platforms of \(a\).

We define the group fairness and individual fairness notions below, these constraints are also part of the input.

**Definition 3.1 (Group fairness).** Each platform, \(p\), has upper bounds, \(u_{p,h}\), for all \(h \in [\chi]\) denoting the maximum number of items from group \(h\) that can be assigned to \(p\). These are referred to as group fairness constraints in this paper. \(E_{p,h}\) is the set of edges \(((a, p) : a \in A_h)\). A matching \(M \subseteq E\) is said to be group-fair if and only if

\[|E_{p,h} \cap M| \leq u_{p,h} \forall p \in P, h \in [\chi].\]  

(1)

This notion of group fairness is also known as Restricted Dominance, introduced in [Bera et al., 2019].

**Definition 3.2 (Strong group fairness).** Along with upper bounds, each platform, \(p\), has lower bounds, \(l_{p,h}\), for all \(h \in [\chi]\) denoting the minimum number of items from group \(h\) that should be assigned to \(p\). A matching \(M \subseteq E\) is said to be strong-group-fair if and only if

\[l_{p,h} \leq |E_{p,h} \cap M| \leq u_{p,h} \forall p \in P, h \in [\chi].\]  

(2)

This notion of group fairness encompasses Minority Protection, also introduced in [Bera et al., 2019], along with Restricted Dominance.

**Definition 3.3 (Probabilistic individual fairness).** In addition to the group fairness constraints, the input also contains individual fairness parameters, \(L_{a,k}, U_{a,k} \in [0, 1]\) for each item \(a\) and \(k \in [m]\). A distribution \(D\) on matchings in \(G\) is probabilistically individually fair if and only if \(\forall a \in A, k \in [m]\)

\[L_{a,k} \leq \Pr_{M \sim D} \exists p \in R_{a,k} \text{ s.t. } (a, p) \in M \leq U_{a,k}\]  

(3)

It is easy to see how Equation (3) can capture the requirement that items are matched to a high-ranking platform in their preference list with high probability and a low-ranking platform in their preference list with low probability. Our model allows users to set individual fairness constraints based on their requirements.

**Objective:** Let \(I = (G, A_1 \cdots A_\chi, \vec{l}, \vec{u}, \vec{L}, \vec{U})\) denote an instance of our problem. Our objective is to calculate a probabilistic individually fair distribution over a set of group-fair matchings, aiming to maximize the expected matching size when a matching is sampled from this distribution.

Note that our model provides a generic framework that accommodates various fairness settings, elaborated in Section 3.2.

3.1 Results

We provide four different algorithms under different settings to compute a distribution over matchings. The support of the distribution is of size polynomial in the size of the instance, and is in fact of the same size as the number of iterations in the algorithms. Below, we list some known hardness results.
Known Hardness Results

Even without individual fairness constraints, finding a maximum size group-fair matching [Definition 3.1] is NP-hard [Nasre et al., 2019]. Additionally, when there is a single platform and each item appears in at most $\Delta$ classes, the group fairness problem with only upper bounds is NP-hard to approximate within a factor of $O(\log \frac{\Delta}{\epsilon})$ [Sankar et al., 2021].

When an item can belong to multiple groups, determining if a feasible solution exists for group fairness constraints even with lower bounds alone is NP-hard [Louis et al., 2023], making the computation of a strong group-fair matching [Definition 3.2] NP-hard.

Algorithmic Results

Our first contribution is an algorithm that computes a distribution over group-fair matchings such that the individual fairness constraints are approximately satisfied and the expected size of a matching, at the cost of an additive violation in feasibility if no such distribution exists.

**Theorem 3.5.** (Informal version of [Panda et al., 2024])

Our next algorithm improves the multiplicative factor for making the computation of a strong group-fair matching [Definition 3.1] is NP-hard, even with lower bounds alone [Louis et al., 2023]. Additionally, when there is a single platform and each item appears in at most $\Delta$ classes, the group fairness problem with only upper bounds is NP-hard to approximate within a factor of $O(\log \frac{\Delta}{\epsilon})$ [Sankar et al., 2021].

Given an instance of our problem where each item belongs to exactly one group, there is a polynomial-time algorithm that either computes a probabilistic individually fair distribution over a set of strong group-fair matchings or reports infeasibility if no such distribution exists.

The proofs of all the Theorems can be found in the full version [Panda et al., 2024].

### 3.2 Extension to Other Fairness Notions

Our results can be extended to accommodate other fairness notions mentioned below.

**Definition 3.8 (Maxmin individual fairness).** Let $D[a]$ denote $Pr_{M \sim D}[\exists p \in P \text{ s.t. } (a, p) \in M]$. A distribution, $D$, over matchings is Maxmin individually fair if for all distributions $F$ over matchings and all $a \in A$.

$$F[a] > D[a] \implies \exists a' \in A \text{ s.t. } D[a'] > F[a']$$

We refer to the goal of maximizing the representation of the worst-off groups as maxmin group fairness, defined below.

**Definition 3.9 (Maxmin group fairness).** Let $X_{h, p}$ denote the total number of items matched under a feasible matching, $M \subseteq E$, from group $h$ to platform $p$. The matching $M$ is said to be maxmin group-fair if, for any other feasible matching $M'$, if $\exists p \in P, h \in [\chi]$ such that $X_{h, p} > X_{h', p}$, then there is some $p' \in P, h' \in [\chi]$ with $X_{h, p} \geq X_{h', p}$ and $X_{h, p} > X_{h', p'}$. Here at least $p' \neq p$, or $h' \neq h$.

In maxmin group fairness, defined below, the goal is to minimize the representation of the most dominant groups. This is a dual to maxmin group fairness.

**Definition 3.10 (Mindom group fairness).** Let $X_{h, p}$ denote the total number of items matched under a feasible matching, $M \subseteq E$, from group $h$ to platform $p$. $M$ is said to be mindom group-fair if, for any other feasible matching $M'$, if $\exists p \in P, h \in [\chi]$ such that $X_{h, p} < X_{h, p'}$, then there is some $p' \in P, h' \in [\chi]$ with $X_{h, p} < X_{h', p'}$ and $X_{h, p} > X_{h', p'}$. Here at least $p' \neq p$, or $h' \neq h$.

### Extension of Results

**Theorem 3.11.** Given a bipartite graph with disjoint groups and a lower bound on the expected matching size, our framework and the polynomial-time algorithm from Theorem 3.7 can be extended to compute the following:

1. A probabilistic individually fair distribution over a set of maxmin or mindom group-fair matchings, with probabilistic individual fairness constraints.
2. A maxmin individually fair distribution over strong group-fair matchings.

**Theorem 3.12.** Given a bipartite graph and a lower bound on the expected matching size, say $lb$, our framework and the polynomial-time algorithm from Theorems 3.4, 3.5 and 3.6 can be extended to compute the following:
<table>
<thead>
<tr>
<th>Size-approximation</th>
<th>( \frac{1}{f_e} (\OPT + \epsilon) )</th>
<th>( \OPT )</th>
<th>( \OPT )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Fairness Violation</td>
<td>None</td>
<td>( \frac{1}{g} )-additive</td>
<td>( \frac{1}{g} )-additive</td>
</tr>
<tr>
<td>Individual Fairness Violation</td>
<td>( \frac{1}{f_e} )-multiplicative, ( \frac{1}{f_e} )-additive</td>
<td>( \frac{1}{g} )-multiplicative</td>
<td>( \frac{1}{g} )-multiplicative</td>
</tr>
</tbody>
</table>

Table 1: Comparison of Approximation Algorithms.

1. A distribution over mindom group-fair or group-fair matchings, ensuring an expected matching size of at least \( \frac{1}{f_e} (lb + \epsilon) \), with \( f_e \) and the violation of probabilistic individual fairness or maxmin individual fairness as in Theorem 3.4.

2. A distribution over mindom group-fair or group-fair matchings, guaranteeing an expected matching size of at least \( \frac{lb}{2g} \), and a violation of probabilistic individual constraints or maxmin individual fairness by at most \( \frac{1}{2g} \). The mindom group-fairness or group-fairness is violated by an additive factor of at most \( \Delta \).

3. A distribution over matchings, guaranteeing an expected matching size of at least \( \frac{lb}{2} \), and a violation of probabilistic individual constraints or maxmin individual fairness by at most \( \frac{1}{g} \). The proof of Theorems 3.11 and 3.12, and details of how to extend our results to these settings are in the appendix of the full version [Panda et al., 2024].

## 4 \( O(\Delta \log n) \) Bicriteria Approximation Algorithm

In this section, we have an instance of a bipartite graph \( G = (A \cup P, E) \), where any arbitrary item, \( a \in A \), can belong to at most \( \Delta \) distinct groups. Our objective is to compute a probabilistic individually fair distribution over group fair matchings, so as to maximize the expected size of any matching sampled from this distribution. We design a polynomial-time algorithm that provides an approximation factor dependent on \( O(\Delta) \) and prove our main result, Theorem 3.4, formally stated below.

**Theorem 4.1 (Formal version of Theorem 3.4).** Given any \( \epsilon > 0 \), and an instance of our problem where each item can belong to at most \( \Delta \) groups, there is a polynomial-time algorithm that computes a distribution \( D \) over a set of group-fair matchings such that the expected size is at least \( \frac{1}{f_e} (\OPT + \epsilon) \) where \( f_e = O(\Delta \log(n/\epsilon)) \) and \( n \) is the total number of items. Given the individual fairness parameters, \( L_{a,k}, U_{a,k} \in [0, 1] \), for each item \( a \in A \) and subset \( R_{a,k} \forall k \in [n] \),

\[
\frac{1}{f_e} (L_{a,k} - \epsilon) \leq \Pr \left[ \exists p \in R_{a,k} \text{ s.t. } (a, p) \in M \right] \leq \frac{1}{f_e} (U_{a,k} + \epsilon).
\]

Note that if we set \( \epsilon = \min_{a \in A, k \in [n]} \frac{L_{a,k}}{2} \) in Theorem 4.1, then

\[
\forall a \in A, k \in [n],
\]

\[
\frac{L_{a,k}}{2f_e} \leq \frac{1}{f_e} (L_{a,k} - \epsilon) \leq \Pr \left[ \exists p \in R_{a,k} \text{ s.t. } (a, p) \in M \right] \leq \frac{1}{f_e} (U_{a,k} + \epsilon) \leq \frac{3U_{a,k}}{2f_e}.
\]

Therefore, we only get a multiplicative violation of individual fairness for \( \epsilon = \min_{a \in A, k \in [n]} \frac{L_{a,k}}{2} \).

### 4.1 Model Formulation

We begin by formulating a Linear Programming (LP) model for our problem, specifically tailored to address Theorem 4.1. We design an extensive LP formulation, applicable to all our theorems, including Theorem 3.7, which integrates additional lower group fairness constraints, is detailed in Section 5 of the full version [Panda et al., 2024]. Since items assumed to be indivisible, we assume that the group fairness bounds are integers.

**LP 4.2.**

\[
\max \sum_{(a, p) \in E} x_{ap}
\]

subject to

\[
L_{a,k} \leq \sum_{p \in R_{a,k}} x_{ap}, \quad \forall a \in A, \forall k \in [n]
\]

\[
\sum_{a \in A_h} x_{ap} \leq u_{p,h}, \quad \forall p \in P, \forall h \in [\chi]
\]

\[
0 \leq x_{ap} \leq 1
\]

where \( a \in A, p \in P \).

LP 4.2 is a relaxation of the Integer Linear Programming (ILP) formulation of our problem with group and individual fairness constraints. In the ILP version, \( x_{ap} = 1 \) iff the edge \((a, p)\) is picked in the matching and 0 otherwise. Constraints 5 and 6 capture the individual fairness and group fairness requirements, respectively.

Before delving into the algorithm and the proof sketch of Theorem 4.1, it is essential to note that LP 4.2 may become infeasible if the fairness constraints are inconsistent. This is possible due to the individual fairness constraints. To address this, one could introduce a variable to calculate the smallest multiplicative relaxation of the fairness constraints required to ensure the feasibility of LP 4.2. This method is detailed in the section titled ‘Dealing with infeasibility’ in the full version [Panda et al., 2024].
13

11

D

whether such a convex combination exists. Therefore, we

12

M

value of

10

sists of tuples, and each tuple consists of a group-fair matching

negative values in

M

Clearly,

P

M

Greedily find a Maximal Matching \( M^{(i)} \) in \( G^{(i)} \) such that constraints (6) are not violated.

\[
\alpha^{(i)} \leftarrow \min_{(a,p) \in M^{(i)}} \{ x_{ap}^{(i-1)} \}
\]

\[
\begin{align*}
\text{sum} & \leftarrow \text{sum} + \alpha^{(i)}, \\
\mathcal{D} & \leftarrow \mathcal{D} \cup \{ M^{(i)}, \alpha^{(i)} \}
\end{align*}
\]

for \( D^{(i)} \in \mathcal{D} \) do

\[
D^{(i)} \leftarrow (M^{(i)}, \alpha^{(i)} / \text{sum})
\]

if \( \mathcal{D} == \phi \) then

Return ‘Infeasible’

Return \( \mathcal{D} \)

\[ \text{Algorithm 1: } O(\Delta \log n)\text{-BicriteriaApprox}(\mathcal{I} = (G, A_1 \cdots A_k, \bar{u}, \bar{L}, \bar{U}), \epsilon) \]

\[ \text{Input : } \mathcal{I}, \epsilon \]

\[ \text{Output : } \text{Distribution over matchings satisfying the guarantees in Theorem 4.1.} \]

1 Solve LP 4.2 on \( G \) with the parameters in the input instance, \( \mathcal{I} \), and store the result in \( x \).

2 \( i \leftarrow 0, \alpha^{(0)} \leftarrow 0, G^{(0)} \leftarrow G, x^{(0)} \leftarrow x, \text{sum} \leftarrow 0, \mathcal{D} \leftarrow \phi \)

3 while \( \|x^{(i)}\|_1 \geq \epsilon \) do

4 \( i \leftarrow i + 1, G^{(i)} \leftarrow G^{(i-1)} - \{ (a, p) \mid x_{ap}^{(i-1)} = 0 \} \)

5 Greedily find a Maximal Matching \( M^{(i)} \) in \( G^{(i)} \) such that constraints (6) are not violated.

6 \( \alpha^{(i)} \leftarrow \min_{(a,p) \in M^{(i)}} \{ x_{ap}^{(i-1)} \} \)

7 \( \text{sum} \leftarrow \text{sum} + \alpha^{(i)}, \mathcal{D} \leftarrow \mathcal{D} \cup \{ M^{(i)}, \alpha^{(i)} \} \)

8 \( x^{(i)} \leftarrow x^{(i-1)} - \alpha^{(i)} \cdot M^{(i)} \)

9 for \( D^{(i)} \in \mathcal{D} \) do

10 \( D^{(i)} \leftarrow (M^{(i)}, \alpha^{(i)} / \text{sum}) \)

11 if \( \mathcal{D} == \phi \) then

12 Return ‘Infeasible’

13 Return \( \mathcal{D} \)

4.2 Algorithm

First, we describe our algorithm (Algorithm 1) and its intuition. The key idea is to express a feasible solution, \( x \), of LP 4.2 as a convex combination of integer group-fair matchings. This approach allows us to satisfy our probabilistic individual fairness constraints by sampling from the probability distribution corresponding to this convex combination. If the groups are not disjoint, as is the case in our problem, then it is not known whether such a convex combination exists. Therefore, we show that \( \frac{x - x^f}{f^1} \) can be written as a convex combination of integer group-fair matchings, where \( \|x^{(i)}\|_1 < \epsilon \) for some \( \epsilon > 0 \) and \( f = O(\Delta \log n) \). Algorithm 1 computes such a convex combination.

The algorithm begins with a feasible solution of LP 4.2 solved on the input instance \( \mathcal{I} \), denoted by variable \( x \). At round \( i \) of the while loop (step 3), \( G^{(i)} \) denotes the state of the input graph after the \( i \)-th iteration. It is a graph where edges with a corresponding zero value in \( x^{(i-1)} \) are discarded in Step 4, \( M^{(i)} \) represents a group-maximal matching computed on \( G^{(i)} \) in step 5, and \( x^{(i)} \) denotes the state of \( x \) after \( i \) rounds. It is the residue after a scaled down \( M^{(i)} \) is “deduced” from \( x^{(i-1)} \) in step 8. \( \alpha^{(i)} \), denoting the minimum non-zero value associated with any edge in \( x^{(i-1)} \) (step 6), is used to scale down \( M^{(i)} \) before “deducting” it from \( x^{(i-1)} \) to ensure non-negative values in \( x^{(i)} \). The algorithm terminates when the value of \( \|x^{(i)}\|_1 \) goes below \( \epsilon \). \( \mathcal{D} \) returned by Algorithm 1 consists of tuples, and each tuple consists of a group-fair matching because \( M^{(i)} \) is group-fair and its corresponding coefficient, \( \alpha^{(i)} / \text{sum} \). If the loop terminates after \( k \) rounds, \( \text{sum} = \sum_{i=1}^{k} \alpha^{(i)} \). Clearly, \( \sum_{i=1}^{k} \alpha^{(i)} / \text{sum} = 1 \), therefore, \( \mathcal{D} \) is a distribution over group-fair matchings.

One key intuition behind Algorithm 1 is that in every iteration, we start with a solution, \( x^{(i-1)} \), that satisfies group-fairness constraints (Equation (6)), which allows us to greedily compute a group-fair matching \( M^{(i)} \) in the support of \( x^{(i-1)} \) (step 5 of Algorithm 1). This ensures that step 5 always returns a non-empty group-fair matching as long as \( x^{(i-1)} \) has non-zero entries.

Next, we provide a proof overview of Theorem 4.1 using Algorithm 1. The proof of Theorem 4.1 is detailed in the full version [Panda et al., 2024].

4.3 Proof Overview

The proof of Theorem 4.1 is based on a careful analysis of our simple (and fast) greedy algorithm (Algorithm 1). We first construct an LP for our problem, concentrating solely on group-fairness constraints [Definition 3.1], excluding individual fairness constraints. This LP, (LP 4.3), and its dual (LP 4.4), facilitate our analysis. This choice is made because, instead of grounding our analysis on LP 4.2, it suffices to focus on LP 4.3, as clarified in Observation 4.5.

LP 4.3.

\[
\begin{align*}
\max & \sum_{(a,p) \in E} x_{ap} \\
\text{such that} & \sum_{a \in A_k} x_{ap} \leq u_{p,h}, & \forall h \in [\chi], \forall p \in P \\
& 0 \leq x_{ap} \leq 1, & \forall(a, p) \in E
\end{align*}
\]

LP 4.4.

\[
\begin{align*}
\min & \sum_{p \in P} \sum_{h \in [\chi]} u_{p,h}w_{p,h} + \sum_{(a,p) \in E} y_{ap} \\
\text{such that} & 1 \leq \sum_{h:a \in A_h} w_{p,h} + y_{ap}, & \forall(a, p) \in E
\end{align*}
\]

Observation 4.5. Any feasible solution of LP 4.2 is also a feasible solution of LP 4.3.

We show that the size of the matching \( M^{(i)} \), in the \( i \)-th round of Algorithm 1, is at least \( \|x^{(i-1)}\|_1 / \Delta + 1 \) using dual fitting analysis technique [Williamson and Shmoys, 2011; Vazirani, 2013; Jain et al., 2003] in the full version [Panda et al., 2024][Lemma 4.7]. We update the LP solution to \( x^{(i)} \) by “removing” \( \alpha^{(i)} \) \( M^{(i)} \) from \( x^{(i-1)} \) (step 8 of Algorithm 1). \( \alpha^{(i)} \) is the largest possible value such that the remaining LP solution is still a feasible solution of LP 4.3 after step 8. Therefore, if a “large” mass of the LP solution remains in the \( i \)-th iteration, i.e., \( \|x^{(i)}\|_1 \) is large, then we make “large” progress in the current iteration. This can essentially be used to show that \( \sum_{i=1}^{\delta} \alpha^{(i)} \) is bounded by \( f_\epsilon = 2(\Delta + 1)(\log(n/\epsilon) + 1) \) when \( \|x^{(i)}\|_1 < \epsilon \) [Panda et al., 2024][Lemma 4.11]. Here, \( k \) is the total number of iterations by Algorithm 1. Finally, setting \( \hat{\epsilon} = \frac{x - x^f}{f^1}, t = f_\epsilon, \) and \( \delta = \frac{\epsilon}{f^1} \) in Lemma 4.6, proves the approximation guarantee on probabilistic individual fairness given by Theorem 4.1. Lemma 4.6 is stated below.
Lemma 4.6. Let us consider a set of tuples, \( D = \{(M^{(i)}, \beta^{(i)})\}_{i \in [k]} \) where \( M^{(i)} \) is an integer matching and \( \beta^{(i)} \) is a scalar, \( \forall i \in [k] \), where \( k \in \mathbb{Z} \). Let \( \hat{x} = \sum_{i=1}^{k} \beta^{(i)} M^{(i)} \) such that \( \sum_{i=1}^{k} \beta^{(i)} = 1 \), and \( \|\hat{x} - \frac{x}{t}\|_1 \leq \delta \) where \( x \) is any feasible solution of LP 4.2, \( \delta \in [0, 1) \), and \( t \geq 1 \). The probability that an item, \( a \in A \), is matched to a platform \( p \in R_{a,k}, \forall k \in [n] \), in a matching sampled from the support of \( D \) is

\[
\frac{L_{a,k}}{t} - \delta \leq \Pr_{M \sim D}[p \in R_{a,k} \text{ s.t. } (a, p) \in M] \leq \frac{U_{a,k}}{t} + \delta.
\]

The proof of Lemma 4.6 can be found in the full version [Panda et al., 2024]. We use Lemma 4.6 to prove the individual fairness guarantees provided not just in Theorem 4.1 but also in the rest of the theorems.

5 Experiments

In this section, we apply our main algorithm 1 from Theorem 4.1, on two real-world datasets. The runtime bottleneck of our primary solution (Algorithm 1) is the execution time of LP 4.3. LP 4.3 has a polynomial number of variables and constraints, and Algorithm 1 solves it exactly once. Therefore, this solution is scalable with practical LP solvers. In our experiments on standard datasets, the algorithm performs much better than the 2(\( \Delta + 1 \))(log(\( n/e \)) + 1) approximation guarantee provided by Theorem 4.1. Here \( n \) is the total number of items, \( \Delta \) is the maximum number of groups an item can belong to, and \( e > 0 \) is a small value. There are no comparison experiments since there are no benchmarks for solving this exact problem. We use experiments to validate and demonstrate the practical efficiency of Algorithm 1.

5.1 Datasets

Employee Access Data[EA]: This is the University of Melbourne’s dataset on grant applications collected from 2004-2008 and published on Kaggle. We use the entire dataset of size 8,707 for our experiments. In our model, the applicants and the grants correspond to items and platforms, respectively, and each grant application represents an edge. We group the applicants based on their research fields. The same applicant can apply to different grants under different research fields represented as RFCD code. Therefore, each item could have edges to multiple platforms and belong to more than one group.

Grant Application Data[GA]: This is the University of Melbourne’s dataset on grant applications collected from 1995-2008 and published on Kaggle. We use the entire dataset of size 8,707 for our experiments. In our model, the applicants and the grants correspond to items and platforms, respectively, and each grant application represents an edge. We group the applicants based on their research fields. The same applicant can apply to different grants under different research fields represented as RFCD code. Therefore, each item could have edges to multiple platforms and belong to more than one group.

5.2 Experimental Setup and Results

We implement our algorithm in Python 3.7 using the libraries NumPy, scipy, and Pandas. All the experiments were run using Google colab notebook on a virtual machine with Intel(R) Xeon(R) CPU @ 2.20GHz and 13GB RAM. Both the datasets for our experiments are taken from Kaggle. We run our experiments on one complete dataset and three different sample sizes on another dataset. The sample size denotes the total number of rows present in the unprocessed sample. The total number of edges can differ from the sample size after data cleaning like removing null values and dropping duplicate edges if any. For group fairness bounds, we set all the upper bounds for each platform group pair to \( \frac{n}{mg} \), where \( n \) is the number of items, \( m \) is the number of platforms, \( g \) is the number of groups, and \( k = \lceil \frac{mg}{20} \rceil \). \( \frac{n}{mg} \) is a uniform distribution of the items amongst group-platform pairs, however, if \( \frac{n}{mg} < 1 \), step 5 in Algorithm 1 would always result in an empty matching. Therefore, to avoid that, the factor \( k \) is added. For individual fairness constraints, we first choose a random permutation of the platforms to create a ranking and then add constraints such that an item should have \( \frac{\epsilon}{n} \) chance of being matched to a platform in the top \( r \) positions in the ranking. For all the runs, \( \epsilon = 0.0001 \).

We use the solution to LP 4.2 to upper bound OPT. We denote it by \( UB \), and SOL denotes the expected size of the solution given by Algorithm 1 on different samples. Let ‘approx’ = 2(\( \Delta + 1 \))(log(\( n/e \)) + 1). In Table 2, we compare the actual approximation ratio, \( UB/SOL \), with the theoretical one, ‘approx’. As seen in Table 2, in our experiments on standard real datasets, the algorithm performs much better than the worst-case theoretical guarantee of Algorithm 1. We repeatedly apply our approximation algorithm from Theorem 4.1 on multiple datasets sampled from the Employee Access dataset under the same experimental setup except for the \( \epsilon \)-value which is now set to 0.001. We see that the algorithm continues to perform much better than the approximation guarantee of 2(\( \Delta + 1 \))(log(\( n/e \)) + 1) shown in Section 4. The results can be seen in the full version [Panda et al., 2024].

Acknowledgements

We thank Shankar Ram for the initial discussions. AP was supported in part by the Empower Programme of the Kotak IISc AI-ML Centre and the Walmart Center for Tech Excellence at IISc (CSR Grant WMTG-23-0001). AL was supported in part by SERB Award CRG/2023/002896, Pratiksha Trust

---

1https://www.kaggle.com/datasets/lucamassaron/amazon-employee-access-challenge

3https://www.kaggle.com/competitions/unimehl/data
Young Investigator Award and the Walmart Center for Tech Excellence at IISc (CSR Grant WMGT-23-0001). PN was supported by SERB Grant CRG/2019/004757.

**Contribution Statement**
AL and PN contributed equally to this paper.

**References**


[Panda et al., 2024] Atasi Panda, Anand Louis, and Prajakta Nimbhorkar. Individual fairness under varied notions of group fairness in bipartite matching - one framework to approximate them all, 2024.


