# Mechanism Design for School Choice with Soft Diversity Constraints

**Haris Aziz**[1,2] , **Serge Gaspers**[1] and **Zhaohong Sun**[1,2]*
[1]UNSW Sydney, Australia
[2]Data61 CSIRO, Australia
{haziz, sergeg}@cse.unsw.edu.au, zhaohong.sun@unsw.edu.au

## Abstract

We study the controlled school choice problem where students may belong to overlapping types and schools have soft target quotas for each type. We formalize fairness concepts for the setting that extend fairness concepts considered for restricted settings without overlapping types. Our central contribution is presenting a new class of algorithms that takes into account the representations of combinations of student types. The algorithms return matchings that are non-wasteful and satisfy fairness for same types. We further prove that the algorithms are strategyproof for the students and yield a fair outcome with respect to the induced quotas for type combinations. We experimentally compare our algorithms with two existing approaches in terms of achieving diversity goals and satisfying fairness.

## 1 Introduction

Incorporating diversity constraints, transparency and fairness into systems and mechanisms are some of the prominent concerns in artificial intelligence. These concerns are also prevalent in matching markets where there has been increased attention to school choice problems that take into account affirmative action and diversity concerns when matching students to schools. One particular model of school choice with diversity constraints is *controlled school choice* [Abdulkadiroğlu and Sönmez, 2003], in which students are associated with a set of types. These types capture traits, such as being extra-talented or being from a disadvantaged group. In recent years, algorithms for matching with diversity goals have been deployed in many places including Israel [Gonczarowski *et al.*, 2019] and India [Sönmez *et al.*, 2019].

Typically, the diversity goals are achieved by setting minimum and maximum target representation of students. If diversity constraints are considered as hard bounds, there may not exist an outcome that fulfills all minimum quotas, and a fundamental tension between fairness and non-wastefulness arises [Ehlers *et al.*, 2014]. Placing hard constraints on diversity constraints may be over-constraining and may put them in head-on conflict with school priorities or other merit consideration. Kojima [2012] shows additional evidence that setting hard bounds can be counter-productive. There are challenges on the computational front as well: it is NP-hard to check whether there exists a feasible or stable matching under hard bounds [Aziz *et al.*, 2019].

Because of these issues with hard bounds, the recent literature on controlled school choice problems treats diversity constraints as *soft bounds* which are soft goals that schools attempt to achieve [Hafalir *et al.*, 2013; Ehlers *et al.*, 2014; Kurata *et al.*, 2015; Kurata *et al.*, 2017]. In particular, these quotas are used to determine which types should be given higher precedence when allocating school seats.

Most papers on controlled school choice assume that each student is associated with only one type. In reality, students may satisfy multiple types. For example, a student could be both female and aboriginal. Kurata *et al.* [2015] studied the setting where diversity constraints are soft bounds and students are allowed to have multiple types. One important issue in existing work on multiple types is the imbalance of representation for certain type combinations. For example, the existing algorithms may achieve a reasonable representation of girls as well as aboriginals but have zero representation of aboriginal girls. Similar issues have also been debated machine learning algorithms where fairness across types has become increasingly important [Kearns *et al.*, 2018].

In this paper, we study the controlled school choice problem where students may have overlapping types, and diversity constraints are viewed as soft bounds. The research question we consider is *how to design mechanisms that cater to diversity objectives while still satisfying desirable fairness, non-wastefulness and strategy-proofness properties?*

### Contributions

We propose a new fairness concept generalizing the standard one for the model in which each student has one type [Ehlers *et al.*, 2014]. Then we present a clear taxonomy of fairness and non-wastefulness concepts in the literature as shown in Figure 1 and show an impossibility result that fairness is incompatible with non-wastefulness.

We then present a novel class of algorithms *Generalized Deferred Acceptance for Type Combinations (GDA-TC)*. Unlike a previous approach DA-OT [Kurata *et al.*, 2015; Kurata *et al.*, 2017] that modifies the structure of preferences

*Contact Author

and priorities, we take an alternative route to overcome this incompatibility. A central idea is to create a set of type combinations such that each student is associated with precisely one type combination. Then we directly set target quotas for each of the relevant type combinations. We show that our GDA-TC algorithm is strategy-proof for the students and yields a non-wasteful and fair outcome for students of same types. We compare with existing algorithms and summarize the theoretical properties satisfied by each algorithm in Table 1.

Finally, we undertake the first experimental comparative analysis of school choice algorithms for students with multiple types. Our generated data uses similar features as the private data set used by Gonczarowski *et al.* [2019]. We show that our new algorithm performs well across several axes, including fairness, diversity goals as well as running time. Note that although DA-OT additionally satisfies KHIY-fairness, it performs the worst in achieving diversity goals and running time. In contrast, GDA-TC performs better than the other two algorithms in terms of consistently satisfying a reasonable relaxation of targets representations.

### Related Work

Our paper belongs to the line of work on school choice with affirmative action goals. Methodologically, our work is closest to that of Ehlers *et al.* [2014] whose objective is to deal with soft diversity constraints. Our results are more general in a strict sense because their model and algorithms do not cater to multiple types. The GDA-PMA algorithm of Gonczarowski *et al.* [2019] directly fits into our model. As the authors point out, GDA-PMA does not guarantee a fair and non-wasteful outcome. We further show that it is also not fair for same types.

There are other recent papers on school choice that concern overlapping types [Aygün and Turhan, 2016; Kurata *et al.*, 2015; Kurata *et al.*, 2017]. However, the models and algorithms considered in these papers are different from ours in one crucial respect. Students and schools are asked to express strict preferences and priorities over contracts *that also involve types*. However, in several scenarios, students may not care about which privilege type they were granted for admission as long as they obtained a school seat. They also may be averse to reveal their contract explicitly corresponding to some type. The algorithms in the papers can be tailored for our model by letting the mechanism form artificial preference over school-type pairs. We pursue this adaptation to experimentally compare the algorithms.

Kominers and Sönmez; Kominers and Sönmez [2013; 2016] studied matching with slot-specific priorities in which each slot may have different priorities from the school and the school has a predefined precedence ordering over all slots. Sönmez *et al.* [2019] studied the affirmative action system in Indian in which students may belong to multiple types. They show desirable outcomes only exist under the assumption that types form a nested structure such that for any two distinct types, either they are unrelated or one is the other's ancestor. Baswana *et al.* [2019] designed and deployed an algorithm for Indian engineering colleges. For their setting with its very particular features, they used a heuristic to deal with non-nested common quotas and their algorithm does not guarantee a fair outcome. Echenique and Yenmez [2015] consider different choice functions based on merit and diversity concerns but do not focus on algorithms or markets as a whole. Ahmed *et al.*; Dickerson *et al.* [2017; 2019] concern a different model with cardinal information.

## 2 Preliminaries

To simplify the presentation, we only focus minimum quotas only for the rest of the paper, as was the focus of Kurata *et al.* [2015]. The impossibility result in Theorem 1 carries over to maximum quotas, and our new algorithms can be easily extended to cater to maximum quotas.

An instance $I^T$ of the school choice problem with diversity constraints consists of a tuple $(S, C, q_C, T, \underline{\eta}, \mathcal{X}, \succ_S, \succ_C)$ where $S$ and $C$ denote the set of students and schools respectively. The capacity vector $q_C = (q_c)_{c\in C}$ assigns each school $c$ a capacity $q_c$. Let $T$ denote the type space. For each student $s$, $T(s) \subseteq T$ denotes the subset of types to which student $s$ belongs. If $T(s) = \emptyset$, then student $s$ does not have any privileged type. Each school $c$ imposes a minimum quota $\underline{\eta}_c^t$ on each type $t$. Let $\underline{\eta}_c = (\underline{\eta}_c^t)_{t\in T}$ denote the type-specific minimum quota vector of school $c$ and let $\underline{\eta}$ be a matrix consisting of all schools' type-specific minimum quotas.

Each contract denoted by $x = (s, c)$ consists of a student-school pair representing that student $s$ is matched to school $c$. Let $\mathcal{X} \subseteq S \times C$ denote the set of available contracts. Given any $X \subseteq \mathcal{X}$, let $X_s$ be the set of contracts involving student $s$, let $X_c$ be the set of contracts involving school $c$ and let $X_c^t$ be the set of contracts involving type $t$ and school $c$.

Each student $s$ has a strict preference ordering $\succ_s$ over $\mathcal{X}_s \cup \{\emptyset\}$ where $\emptyset$ is a null contract representing the option of being unmatched for student $s$. A contract $(s, c)$ is *acceptable* to student $s$ if $(s, c) \succ_s \emptyset$. Let $\succ_S = \{\succ_{s_1}, ..., \succ_{s_n}\}$ be the preference profile of all students $S$. Each school $c$ has a strict priority ordering $\succ_c$ over $\mathcal{X}_c \cup \{\emptyset\}$ where $\emptyset$ represents the option of leaving seats vacant for school $c$. A contract $(s, c)$ is *acceptable* to school $c$ if $(s, c) \succ_c \emptyset$. Let $\succ_C = \{\succ_{c_1}, ..., \succ_{c_m}\}$ be the priority profile of all the schools.

An outcome (or a matching) $X$ is a subset of $\mathcal{X}$. An outcome $X$ is *feasible* (under soft bounds) for $I^T$ if i) each student $s$ is matched with at most one school, i.e., $|X_s| \leq 1$, and ii) the number of students matched to each school $c$ does not exceed its capacity, i.e., $|X_c| \leq q_c$. A feasible outcome $X$ is *individually rational* if each contract $(s, c) \in X$ is acceptable to both student $s$ and school $c$. Without loss of generality, we focus on acceptable contracts.

**Definition 1** (Non-wastefulness)**.** *Given a feasible outcome $X$, student $s$ claims an empty seat of school $c$ if $(s, c) \succ_s X_s$ and $|X_c| < q_c$. A feasible outcome is non-wasteful if no student claims an empty seat.*

An *algorithm* for our problem takes an instance $I^T$ as input and outputs a set of contracts. An algorithm is *strategy-proof* for students if there exists no student who can misreport his preferences to be matched with a better school.

Next, we briefly introduce the generalized deferred acceptance (GDA) algorithm, which extends the classical deferred acceptance (DA) algorithm to the setting of matching with

**Algorithm 1** Generalized Deferred Acceptance (GDA)

**Input:** A set of contracts $X \subseteq \mathcal{X}, Ch_S, Ch_C$
**Output:** An outcome $Y \subseteq X$

1: $R \leftarrow \emptyset, Y \leftarrow X, Z \leftarrow \emptyset$ % $R$: rejected contracts
2: **while** $Y \neq Z$ **do**
3: $Y \leftarrow Ch_S(X \setminus R), Z \leftarrow Ch_C(Y), R \leftarrow R \cup (Y \setminus Z)$
4: **return** $Y$

contracts [Hatfield and Milgrom, 2005]. Given any $X \subseteq \mathcal{X}$, let $Ch_S(X)$ denote the choice function of students $S$ which selects each student's favorite contract from $X_S$. Similarly, the choice function $Ch_C(X)$ of schools $C$ selects a set of contracts from $X_C$. Note that the way to specify $Ch_C$ is not unique and different implementations of the GDA algorithm vary on how to define the choice function of schools [Kurata *et al.*, 2015; Gonczarowski *et al.*, 2019].

The GDA algorithm works in the same way as the original deferred acceptance algorithm [Roth and Sotomayor, 1990] does: each student first selects one contract involving her favorite school that has not rejected her yet; then schools choose a set of contracts among the proposals and reject others. Repeat this procedure until no more contract is rejected by any school.

## 3 Fairness under Diversity Constraints

In this section, we discuss how to define fairness under soft diversity constraints. In a seminal paper on school choice, Abdulkadiroğlu and Sönmez [2003] proposed an algorithm that eliminates justified envy among students who have the same type. Ehlers *et al.* [2014] also considered this concept, while the difference from our Definition 2 is that in our setting, each student may belong to multiple types.

**Definition 2** (Fairness for same types)**.** *Given an instance $I^T$ and a feasible outcome $X$, student $s$ has justified envy towards student $s'$ of same types if i) $(s,c) \succ_s X_s$, $(s',c) \in X$, ii) $(s,c) \succ_c (s',c)$ and iii) $T(s) = T(s')$. A feasible outcome is* fair for same types *if no student has justified envy towards any student of same types.*

In real-life, the number of distinct type combinations is relatively small. For instance, in the Indian college admission which involves 1.2 million annual applicants, the affirmative action was imposed on three backward classes (Scheduled Castes, Scheduled Tribes, Other Backward Castes) due to historical discrimination and two disadvantaged groups (female and disables) [Baswana *et al.*, 2019].

Thus fairness for same types is a meaningful concept and it is a suitable way to measure the outcomes in terms of fairness, because it concerns a huge number of students who have the same types. On the other hand, as we will show later, that s stronger fairness concept is incompatible with non-wastefulness.

### 3.1 General Impossibility

Next we propose a natural way to measure justified envy among students of different types. The idea is that student $s$ is given higher precedence over student $s'$ if student $s$ has a superset of types that are below the minimum quotas compared to student $s'$. In other words, student $s$ makes more contribution in terms of satisfying diversity goals. Formally, given a feasible outcome $X$, let $\underline{V}_c^X = \{t \in T \mid \underline{\eta}_c^t > |X_c^t|\}$ denote the set of types that are undersubscribed at school $c$.

**Definition 3** (Binary Relation $\trianglerighteq$)**.** *Given a feasible outcome $X$ and two students $s$, $s'$ with $(s,c) \notin X$ and $(s',c) \notin X$,*

- $s \trianglerighteq_c^X s' \Leftrightarrow T(s) \cap \underline{V}_c^X \supseteq T(s') \cap \underline{V}_c^X$,
- $s \triangleright_c^X s' \Leftrightarrow s \trianglerighteq_c^X s'$ *and* $s' \ntrianglerighteq_c^X s$,
- $s \sim_c^X s' \Leftrightarrow s \trianglerighteq_c^X s'$ *and* $s' \trianglerighteq_c^X s$.

Given a feasible outcome $X$, the notation $s \trianglerighteq_c^X s'$ means that student $s$ contributes at least as much as student $s'$ to school $c$. The notations $s \triangleright_c^X s'$ and $s \sim_c^X s'$ specify the strict and equivalent binary relation respectively.

Based on binary relation $\trianglerighteq$ we propose a new and extremely weak fairness concept in Definition 4.

**Definition 4** (Fairness)**.** *Given an instance $I^T$ and a feasible outcome $X$, student $s$ has* justified envy *towards student $s'$ if i) $(s,c) \succ_s X_s$, $(s',c) \in X$ and ii) for the outcome $X' = X \setminus \{(s',c)\}$, either ii-a) $s \triangleright_c^{X'} s'$ and $(s,c) \succ_c \emptyset$, or ii-b) $s' \ntriangleright_c^{X'} s$ and $(s,c) \succ_c (s',c)$ hold. An outcome is fair if no student has justified envy towards any student.*

Although there are other ways to define fairness, we show in Theorem 1 that even such weak fairness concept in Definition 4, is incompatible with non-wastefulness.

**Theorem 1.** *The set of fair and non-wasteful outcomes could be empty, even if there are only two types.*

*Proof.* We prove Theorem 1 by the following counterexample in which each contract is acceptable.

$$
\begin{aligned}
&S = \{s_1, s_2, s_3, s_4\}, C = \{c_1, c_2\}, q_{c_1} = 2, q_{c_2} = 1,\\
&T = \{t_1, t_2\}, \underline{\eta}_{c_1} = (1,1), \underline{\eta}_{c_2} = (0,0), T(s_1) = \{\emptyset\},\\
&T(s_2) = \{t_1, t_2\}, T(s_3) = \{t_1\}, T(s_4) = \{t_2\},\\
&(s_1,c_1) \succ_{s_1} (s_1,c_2), (s_2,c_2) \succ_{s_2} (s_2,c_1),\\
&(s_3,c_1) \succ_{s_3} (s_3,c_2), (s_4,c_1) \succ_{s_4} (s_4,c_2),\\
&(s_1,c_1) \succ_{c_1} (s_2,c_1) \succ_{c_1} (s_3,c_1) \succ_{c_1} (s_4,c_1),\\
&(s_1,c_2) \succ_{c_2} (s_2,c_2) \succ_{c_2} (s_3,c_2) \succ_{c_2} (s_4,c_2).
\end{aligned}
$$

It can be argued that the set of fair and non-wasteful outcomes is empty for the instance and we omit the detailed proof. ☐

### 3.2 Taxonomy of Concepts

In this subsection, we present a clear taxonomy of fairness and non-wastefulness concepts in the literature. Ehlers *et al.* [2014] proposed Definition 5 for the model where each student belongs to one type.

**Definition 5** (Fairness across types)**.** *Given an instance $I^T$ with distinct types and a feasible outcome $X$, student $s$ has justified envy towards student $s'$ of a different type if i) $(s,c) \succ_s X_s$, $(s',c) \in X$ and ii) one of the following cases holds, where $T(s) = \{t\}$ and $T(s') = \{t'\}$ with $t \neq t'$:*

*a)* $|X_c^t| < \underline{\eta}_c^t$, $|X_c^{t'}| > \underline{\eta}_c^{t'}$ *and* $(s,c) \succ_c \emptyset$;

*b)* $|X_c^t| < \underline{\eta}_c^t$, $|X_c^{t'}| \leq \underline{\eta}_c^{t'}$ *and* $(s,c) \succ_c (s',c)$;

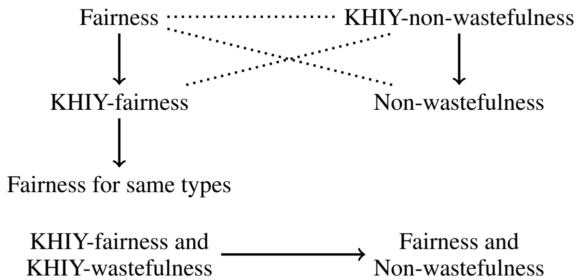


Figure 1: Relations between all concepts for school choice with overlapping types. An arrow from A to B denotes that A implies B. A dashed line between two concepts implies they are incompatible.

*c) $|X_c^t| \geq \underline{\eta}_c^t$, $|X_c^{t'}| > \underline{\eta}_c^{t'}$ and $(s,c) \succ_c (s',c)$.*

*An outcome is fair across types if no student has justified envy towards any student of different type.*

Kurata *et al.* [2015] proposed concepts of fairness and non-wastefulness for school choice with multiple types, where KHIY consists of the first letter of each of the authors.

**Definition 6** (KHIY-fairness)**.** *Given an instance $I^T$ with overlapping types and a feasible outcome $X$, student $s$ has KHIY-justified-envy towards student $s'$ if i) $(s,c) \succ_s X_s$, $(s',c) \in X$, ii) $(s,c) \succ_c (s',c)$ and iii) $\forall t \in T(s') \setminus T(s), |X_c^t| > \underline{\eta}_c^t$ or $T(s') \setminus T(s) = \emptyset$. A feasible outcome $X$ is* KHIY-fair *if no student has KHIY-justified-envy.*

**Definition 7** (KHIY-non-wastefulness)**.** *Given a feasible outcome $X$, student $s$ claims an empty seat of school $c$ if $(s,c) \succ_s X_s$ and $|X_c| < q_c$; and student $s$ claims an empty seat of school $c$ by type if $(s,c) \succ_s X_s$, and $\exists t \in T(s), |X_c^t| < \underline{\eta}_c^t$. A feasible outcome is KHIY-non-wasteful if no student claims an empty seat or an empty seat by type.*

Next, we present a clear taxonomy of all non-wastefulness and fairness concepts through Theorem 2 and Theorem 3. Due to space limitations, proofs are omitted and we depict the results of Theorem 3 in Figure 1.

**Theorem 2.** *Given an instance $I^T$ in which each student belongs to one type and a feasible outcome $X$,*

*i) $X$ is fair if and only if it is fair for same types and fair across types;*

*ii) $X$ is non-wasteful and fair if and only if it is KHIY-non-wasteful and KHIY-fair.*

**Theorem 3.** *Given an instance $I^T$ in which each student belongs to multiple types and a feasible outcome $X$,*

*i) if $X$ is KHIY-non-wasteful, then it is non-wasteful;*

*ii) if $X$ is fair, then it is KHIY-fair;*

*iii) if $X$ is KHIY-fair, then it is fair for same types;*

*iv) if $X$ is KHIY-non-wasteful and KHIY-fair, then it is non-wasteful and fair.*

Although Kurata *et al.* [2017] showed that KHIY-fairness is incompatible with KHIY-non-wastefulness, our impossibility result in Theorem 1 is stronger than the impossibility result of Kurata *et al.* [2017] in two respects. First, as shown in Theorem 3, the combination of non-wastefulness and fairness is weaker than the combination of KHIY-non-wastefulness and KHIY-fairness. Second, our incompatibility result still holds even if there are only two types, while the proof of Kurata *et al.* [2017] uses three types.

**Algorithm 2** GDA-TC

**Input:** $I^T\!=\!(S, C, q_C, T, \eta, \mathcal{X}, \succ_S, \succ_C)$
**Output:** An outcome $X \subseteq \mathcal{X}$

Create a set of type combinations $U$ from types $T$.
Determine quotas $\underline{\delta}$ for type combinations $U$.
Incorporate quotas $\underline{\delta}$ into choice function $Ch_c^{TC}$.
Run GDA with choice function $Ch_c^{TC}$.

## 4 A Class of Algorithms GDA-TC

In this section, we propose a new class of algorithms *Generalized Deferred Acceptance for Type Combinations (GDA-TC)* which yields outcomes that are non-wasteful and fair for students of same types. The general idea is to eliminate overlapping types by creating a new set $U$ corresponding to type combinations of $T$ so that each student is associated with exactly one type combination. Then we establish new quotas for type combinations $U$ and incorporate the induced quotas into the choice function $Ch_c^{TC}$ of schools. We employ the GDA algorithm with choice function $Ch_c^{TC}$ to determine the outcome. All these procedures consist of our new class of algorithms GDA-TC, as shown in Algorithm 2.

There are different ways to establish quotas for type combinations $U$ and each different method specifies one particular algorithm of GDA-TC. For instance, we can invoke linear programming to divide minimum quotas $\underline{\eta}$ for types $T$ into minimum quotas $\underline{\delta}$ for type combinations $U$. We refer to this algorithm as GDA-TC-LP that makes use of linear programming and we explain how GDA-TC-LP works in detail.

### Creating Type Combinations

Let $U \subseteq 2^T$ denote the set of type combinations over types $T$ and let $U(s)$ represent the type combination of student $s$. Note that we only consider the set of type combinations associated with students $S$, whose number is bounded by the number of students.

Let $T(u)$ represent the set of types associated with $u \in U$. For ease of exposition, the index of type combination $u$ is represented in binary where the $i$-th element is 1 if type $t_i \in T(u)$ and 0 otherwise. For a set of type combinations $R \subseteq U$, we say set $R$ *covers* type $t$ if $\forall u \in U \backslash R$, we have $t \notin T(u)$. We say set $R$ *exactly covers* type $t$ if set $R$ is the smallest set that covers type $t$. For type $t \in T$, let $U^t \subseteq U$ denote the set of type combinations that exactly covers type $t$. Note that any superset $R$ of set $U^t$ also covers type $t$.

**Example 1.** *Consider two types $T = \{t_1, t_2\}$ with four type combinations $U\!=\!\{u_{00}, u_{01}, u_{10}, u_{11}\}$. The set $\{u_{00}, u_{10}, u_{11}\}$ covers type $t_1$ and $U^{t_1}\!=\!\{u_{10}, u_{11}\}$ exactly covers type $t_1$.*

### Setting Quotas for Type Combinations

Let $\underline{\delta}_c\!=\!(\underline{\delta}_c^u)_{u\in U}$ denote a minimum target vector of school $c$ where each element $\underline{\delta}_c^u$ is the minimum target quota of type

**Algorithm 3** Choice function $Ch_c^{TC}$ of school $c$

**Input:** An instance $I^T$, quotas $\underline{\delta}$ for $U$, a set of contracts $X$
**Output:** A set of contracts $Y \subseteq X$

1: $Y \leftarrow \emptyset$
2: **for** $x = (s, c) \in X$ in descending ordering of $\succ_c$ **do**
3: **if** $|Y_c| < q_c$ **and** $|Y_c^u| < \underline{\delta}_c^u$ with $u = U(s)$ **then**
4: $Y \leftarrow Y \cup \{x\}, X \leftarrow X \setminus \{x\}$
5: **while** $|Y| < q_c$ **and** $|X_c| > 0$ **do**
6: Select $x \in X$ with highest priority based on $\succ_c$
7: $Y \leftarrow Y \cup \{x\}, X \leftarrow X \setminus \{x\}$
8: **return** $Y$

combination $u$. Let $\underline{\delta}=(\underline{\delta}_c)_{c\in C}$ be a matrix consisting of minimum target quota of each type combination for each school. Next, we explain one possible way to calculate the vector $\underline{\delta}_c$ by the following linear program LP 1.

$$\min \quad \sum\nolimits_{u\in U} \underline{\delta}_c^u \tag{1}$$

$$\sum\nolimits_{u\in U^t} \underline{\delta}_c^u \geq \underline{\eta}_c^t, \quad \forall c \in C, \forall t \in T \tag{2}$$

$$\underline{\delta}_c^u \geq 0, \quad \forall u \in U \tag{3}$$

$$\underline{\delta}_c^u \times |S^v| = \underline{\delta}_c^v \times |S^u|, \quad \forall c \in C, \forall u, v \in U \tag{4}$$

The objective of LP 1 is to minimize the sum of minimum quotas $\underline{\delta}_c^u$ of each type combination $u$ at school $c$. Inequalities (2) specify the basic requirement on how to convert quotas for types into quotas for type combinations. Given a school $c$ with minimum target vector $\underline{\eta}_c = (\underline{\eta}_c^t)_{t\in T}$, the sum of minimum target quota $\underline{\delta}_c^u$ of each type combination $u \in U^t$ that exactly covers type $t$ should be at least as large as the minimum quota $\underline{\eta}_c^t$ for type $t$. Inequalities (3) require that the quota for each type combination should be positive.

We also consider a set of inequalities (4) that takes the proportion of different type combinations into account. In words, in a school $c$, for every two type combinations $u, v \in U$, the minimum quotas $\underline{\delta}_c^u$ and $\underline{\delta}_c^v$ should be proportional to the total number of students with corresponding type combinations.

**Specifying Choice Function for Schools**

We take the minimum targets $\underline{\delta}$ for type combinations $U$ into account when defining choice function $Ch_c^{TC}$, as described in Algorithm 3. Given a set of contracts $X$, the choice function $Ch_c^U$ traverses the set of contracts $X_c$ involving school $c$ twice in accordance with the priority order of school $c$: in the first round, it selects a set of contracts without exceeding any minimum quota for type combinations and the capacity $q_c$ of school $c$; in the second round, it selects a set of contracts without exceeding the capacity only. Next, we show the key properties that the class of GDA-TC algorithms guarantees.

**Theorem 4.** *The class of GDA-TC algorithms with choice function 3 is strategy-proof for students, and yields a feasible outcome that is fair for same types and non-wasteful.*

*Proof.* Ehlers *et al.* [2014] study the model in which each student has one type and they show that GDA algorithm with their choice function is strategy-proof for students and the outcome yielded by their algorithm is fair for same types and non-wasteful. If we consider a new instance $I^U=(S, C, q_C, U, \underline{\delta}, \mathcal{X}, \succ_S, \succ_C)$ obtained from $I^T$ by replacing types $T$ with $U$ and replacing matrix $\underline{\eta}$ with $\underline{\delta}$, then we have a new instance of school choice in which each student has one type combination and our choice function becomes equivalent to the choice function of [Ehlers *et al.*, 2014]. It is easy to infer that the class of GDA-TC algorithms satisfies fairness for same types and non-wastefulness.

Suppose GDA-TC is not strategy-proof. Say student $s$ is matched with $X_s$ if he truly reports his preference $\succ_s$ and is matched with $X'_s$ if misreports his preference $\succ'_s$ with $X_s \succ_s X'_s$. Note that quotas $\underline{\delta}$ for type combinations $U$ are independent from preference profile. Then for instance $I^U$, student $s$ is matched to a better school by manipulating his preference, however, this violates the fact that GDA-TC is strategy-proof for the case each student has one type. □

If we set minimum quotas to be zero for all type combinations, then GDA-TC is equivalent to the Deferred Acceptance algorithm. By Theorem 4, we infer that DA also satisfies fairness for same types and non-wastefulness. However, DA algorithm completely ignores diversity goals.

| | GDA-TC | GDA-PMA | DA-OT |
|---|---|---|---|
| Fairness | ✗ | ✗ | ✗ |
| KHIY-fairness | ✗ | ✗ | ✓ |
| Fairness for same types | ✓ | ✗ | ✓ |
| KHIY-non-wastefulness | ✗ | ✗ | ✗ |
| Non-wastefulness | ✓ | ✓ | ✓ |
| Strategy-proofness | ✓ | ✗ | ✓ |

Table 1: Comparison of our new algorithm GDA-TC with two existing algorithms GDA-PMA and DA-OT.

## 5 Comparison with Existing Algorithms

In this section, we compare GDA-TC with two existing algorithms designed by Kurata *et al.* [2015] and Gonczarowski *et al.* [2019] for school choice with multiple types.

Kurata *et al.* [2015] proposed Deferred Acceptance Algorithm for Overlapping Types (DA-OT) under the assumption that each student is assumed to consume only one unit of some type rather than one unit of each type to which she belongs. Gonczarowski *et al.* [2019] proposed GDA-PMA algorithm to handle soft minimum quotas and multiple types. During the process of GDA-PMA, schools give higher precedence to students who have some type that has not reached the minimum quota. GDA-PMA algorithm is not strategy-proof for students and does not yield a fair outcome [Gonczarowski *et al.*, 2019]. We further show that GDA-PMA algorithm does not eliminate justified envy among students who have the same types. When each student has one type, the choice functions in GDA-TC, GDA-PMA and DA-OT are the same as the choice function defined by Ehlers *et al.* [2014].

In Table 1, we summarize the theoretical properties that are satisfied by three algorithms. Our new proposed GDA-TC algorithm satisfies more properties than GDA-PMA does. Although DA-OT additionally satisfies KHIY-fairness, it performs the worst in terms of achieving diversity goals and running time. In contrast, the experimental results show that

GDA-TC performs better than the other two algorithms in terms of consistently satisfying a reasonable relaxation of targets representations.

### Setup of Experiments

We consider a market with $|S| = 2000$, $|C| = 40$ and $q_c = 50$, which are close to the number of students, the number of schools and the average number of slots at each institution in the 'Mechinot' market [Gonczarowski *et al.*, 2019]. The number of types is $|T| \in \{2, 4, 6, 8\}$. For a given type $t$, its percentage $per(t)$ is determined by the number of students $|S^t|$ associated with type $t$ divided by the total number of students $|S|$. The percentage of each type is randomly chosen from the set $\{0.1, 0.2, 0.3, 0.4, 0.5\}$. The same minimum vector $\underline{\eta}_c$ is imposed on all schools. The minimum target $\underline{\eta}_c^t$ for type $t$ is set as $\underline{\eta}_c^t = |S^t|/|C| * \alpha$ where $\alpha$ is the *target ratio*. The target ratio $\alpha$ takes two values from $\{1.0, 1.3\}$.

The preference profile of students and the priority profile of schools are generated by *Mallows Model (MM)*. Let $\Phi$ be the set of all possible preference orders. MM is a distribution over permutations of $\Phi$ determined by two parameters, a reference order $\sigma \in \Phi$ and a dispersion parameter $\theta \in (0, 1]$ [Lu and Boutilier, 2011].

We generate 100 instances for each setting and compute the average results. Because the relative performances of the three algorithms are consistent, we present the results for target ratio $\alpha = 1.3$ with two dispersion parameters $\theta = 0.1$ and 0.9, as shown in Figure 2 and Figure 3 respectively.

### Measurement of Achieving Diversity Goals

We measure the performance of three algorithms in terms of achieving diversity goals by calculating the percentage of types that satisfy different fractional relaxation of minimum targets. For instance, suppose the minimum target for type $t$ of school $c$ is $\underline{\eta}_c^t = 20$. If an outcome assigns 16 students of type $t$ to school $c$, then it satisfies 0.8 fractional relaxation of target $\underline{\eta}_c^t$ of school $c$, but not 1.0 fractional relaxation.

Note that reaching more minimum targets of large fractional relaxation is not always desirable. This is because the total number of students is fixed: the more students of a given type are assigned to some schools, the fewer students of that type could be assigned to other schools, which may lead to serious imbalanced distribution of students across schools. This violates the motivation of diversity targets, which attempts to eliminate segmentation of students of different types.

An outcome is more *balanced* if it satisfies more percentage of reasonable fractional relaxation $(1/\alpha)$ of minimum quotas. For instance, when target ratio $\alpha = 1.3$, in an ideal outcome, there should be more type targets satisfying fractional relaxation less than $1/1.3 \approx 0.77$.

### Results of the Experiment

In Figure 2 and 3, we present the experimental results for the setting of target ratio $\alpha = 1.3$, MM model with dispersion parameter $\theta = 0.1$ and $\theta = 0.9$. The x-axis denotes the fractional relaxation of type targets at schools and the y-axis denotes the percentage of type targets whose fractional relaxation are satisfied. For instance, in the first subfigure Figure 2 of with 2 types, the orange bar at 0.6 indicates that in the outcome returned by GDA-PMA, around 75% of all types at all

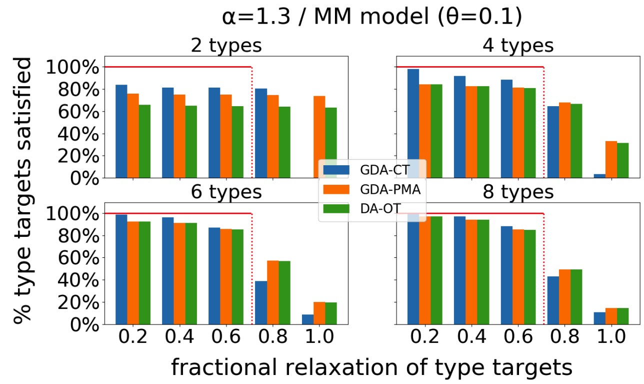


Figure 2: Experimental results of achieving diversity goals for the setting of target ratio $\alpha = 1.3$, MM model with $\theta = 0.1$.

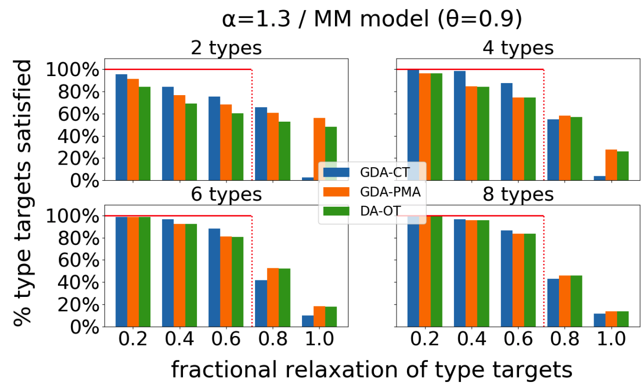


Figure 3: Experimental results of achieving diversity goals for the setting of target ratio $\alpha = 1.3$, MM model with $\theta = 0.9$.

schools satisfied a 0.6 fraction of the minimum targets on average. The red horizontal line represents the ideal outcome, and the red dotted line represents reasonable fractional relaxation $1/\alpha \approx 0.77$.

The experimental results show that GDA-TC performs better than the other two algorithms in terms of consistently satisfying a reasonable relaxation of targets representations. In contrast, GDA-PMA and DA-OT algorithms satisfy much more targets of 1.0 fractional relaxation than GDA-TC does, but at the expense of more type targets *not* satisfying 0.2 / 0.4 / 0.6 fractional relaxation of targets.

We also analysed the number of justified envy relations among students of same types in the outcomes returned by GDA-PMA. When $\alpha = 1.0$ and $\theta = 0.1$, there is a significant number of justified envy relations. We also compare three algorithms in terms of running time. GDA-TC runs the fastest, followed by GDA-PMA. DA-OT is the slowest and spends up to 30 times more time than GDA-TC does.

In conclusion, GDA-TC-LP is a suitable alternative algorithm to GDA-PMA and DA-OT. It outperforms DA-OT in terms of achieving diversity goals and returns a much more balanced outcome. It also has satisfies several important theoretical properties that GDA-PMA does not. In addition, it takes the representation of type combinations into account which was overlooked by the other two algorithms.

## References

[Abdulkadiroğlu and Sönmez, 2003] Atila Abdulkadiroğlu and Tayfun Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, 2003.

[Ahmed *et al.*, 2017] Faez Ahmed, John P. Dickerson, and Mark Fuge. Diverse weighted bipartite b-matching. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 35–41. AAAI Press, 2017.

[Aygün and Turhan, 2016] Orhan Aygün and Bertan Turhan. Dynamic reserves in matching markets: Theory and applications. 2016.

[Aziz *et al.*, 2019] Haris Aziz, Serge Gaspers, Zhaohong Sun, and Toby Walsh. From matching with diversity constraints to matching with regional quotas. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pages 377–385, 2019.

[Baswana *et al.*, 2019] Surender Baswana, Partha Pratim Chakrabarti, Sharat Chandran, Yashodhan Kanoria, and Utkarsh Patange. Centralized admissions for engineering colleges in india. In *Proceedings of the 20th ACM Conference on Economics and Computation*, pages 323–324, 2019.

[Dickerson *et al.*, 2019] John P. Dickerson, Karthik Abinav Sankararaman, Aravind Srinivasan, and Pan Xu. Balancing relevance and diversity in online bipartite matching via submodularity. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1877–1884, 2019.

[Echenique and Yenmez, 2015] Federico Echenique and M Bumin Yenmez. How to control controlled school choice. *American Economic Review*, 105(8):2679–94, August 2015.

[Ehlers *et al.*, 2014] Lars Ehlers, Isa E. Hafalir, M Bumin Yenmez, and Muhammed A. Yildirim. School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory*, 153:648–683, 2014.

[Gonczarowski *et al.*, 2019] Yannai A. Gonczarowski, Lior Kovalio, Noam Nisan, and Assaf Romm. Matching for the Israeli "Mechinot" gap year: Handling rich diversity requirements. In *Proceedings of the 20th ACM Conference on Economics and Computation*, pages 321–321, 2019.

[Hafalir *et al.*, 2013] Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. Effective affirmative action in school choice. *Theoretical Economics*, 8(2):325–363, 2013.

[Hatfield and Milgrom, 2005] John William Hatfield and Paul R. Milgrom. Matching with contracts. *American Economic Review*, 95(4):913–935, 2005.

[Kearns *et al.*, 2018] Michael Kearns, Seth Neel, Aaron Roth, and Zhiwei Steven Wu. Preventing fairness gerrymandering: Auditing and learning for subgroup fairness. In *Proceedings of the 35th International Conference on Machine Learning*, pages 2569–2577, 2018.

[Kojima, 2012] Fuhito Kojima. School choice: Impossibilities for affirmative action. *Games and Economic Behavior*, 75(2):685–693, 2012.

[Kominers and Sönmez, 2013] Scott Duke Kominers and Tayfun Sönmez. Designing for diversity in matching. In *Proceedings of the fourteenth ACM Conference on Electronic Commerce*, pages 603–604, 2013.

[Kominers and Sönmez, 2016] Scott Duke Kominers and Tayfun Sönmez. Matching with slot-specific priorities: Theory. *Theoretical Economics*, 11(2):683–710, 2016.

[Kurata *et al.*, 2015] Ryoji Kurata, Naoto Hamada, Atsushi Iwasaki, and Makoto Yokoo. Controlled school choice with soft bounds and overlapping types. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 951–957, 2015.

[Kurata *et al.*, 2017] Ryoji Kurata, Naoto Hamada, Atsushi Iwasaki, and Makoto Yokoo. Controlled school choice with soft bounds and overlapping types. *Journal of Artificial Intelligence Research*, 58:153–184, 2017.

[Lu and Boutilier, 2011] Tyler Lu and Craig Boutilier. Learning Mallows models with pairwise preferences. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 145–152, 2011.

[Roth and Sotomayor, 1990] Alvin E. Roth and Marilda A. Oliveira Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, 1990.

[Sönmez *et al.*, 2019] Tayfun Sönmez, M Bumin Yenmez, et al. Constitutional implementation of vertical and horizontal reservations in India: A unified mechanism for civil service allocation and college admissions. 2019.