Multi-Rank Smart Reserves

HARIS AZIZ, UNSW Sydney and Data61 CSIRO, Australia
ZHAOHONG SUN, UNSW Sydney and Data61 CSIRO, Australia

We study the school choice problem where each school has flexible multi-ranked diversity goals, and each student may belong to multiple overlapping types, and consumes only one of the positions reserved for their types. We propose a novel choice function and show that it is the unique rule that satisfies three fundamental properties: maximal diversity, non-wastefulness, and justified envy-freeness. We provide a fast polynomial-time algorithm for our choice function that is based on the Dulmage Mendelsohn Decomposition Theorem as well as new insights into the combinatorial structure of constrained rank maximal matchings. Even for the case of minimum and maximum quotas for types (that capture two ranks), ours is the first known polynomial-time approach to compute an optimally diverse choice outcome. Finally, we prove that the choice function we design for schools, satisfies substitutability and hence can be directly embedded in the generalized deferred acceptance algorithm to achieve strategyproofness and stability. Our algorithms and results have immediate policy implications and directly apply to a variety of scenarios, such as where hiring positions or scarce medical resources need to be allocated while taking into account diversity concerns or ethical principles.

CCS Concepts: • **Theory of computation** → **Algorithmic game theory and mechanism design**; *Solution concepts in game theory*; • **Applied computing** → *Economics.*

Additional Key Words and Phrases: market design; matching; diversity; affirmative action; school choice

ACM Reference Format:
Haris Aziz and Zhaohong Sun. 2021. Multi-Rank Smart Reserves. In *Proceedings of the 22nd ACM Conference on Economics and Computation (EC '21), July 18–23, 2021, Budapest, Hungary.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3382129

EC '21, July 18–23, 2021, Budapest, Hungary

ACM ISBN 978-1-4503-8554-1/21/07.
https://doi.org/10.1145/3382129

1 INTRODUCTION

Diversity goals are prevalent in many scenarios including student-intake, hiring of employees, public housing, and rationing of scarce medical resources. While classical centralized matching algorithms have received widespread success, they do not directly apply to two-sided matching under diversity concerns such as those pertaining to proportionality goals or reaching minimum quotas for types. In this paper, we present a compelling solution to selection and matching problems under diversity concerns.

Our two-sided matching problem under diversity concerns is couched in the language of school choice matching with diversity goals. In such scenarios, individuals such as students are to be matched to resources such as school seats. Each individual may belong to multiple types such as being from disadvantaged groups. The goal is to match the individuals to the resources in a way that respects the preferences of the individuals, as well as the priorities and the diversity goals of the schools.

The diversity goals are typically achieved by setting minimum and maximum target representation of students of various types. Existing approaches either do not optimally achieve the diversity goals that are considered, or do so for restricted classes of diversity goals (such as those pertaining to a single level of reserves, that we will explain later). In view of the limitations of existing approaches towards diverse matchings, there is a need for a more general model of controlled school choice that can capture flexible diversity goals, and meaningful methods that achieve these diversity goals while allocating resources to individuals.

In this paper, we propose a flexible framework for specifying a rich class of diversity goals. For the framework, we design a novel school choice function and a fast combinatorial algorithm for the function that draws connections from graph theory and theoretical computer science. We characterize the choice function as the only one satisfying three fundamental properties: maximal diversity, non-wastefulness, and justified envy-freeness. It also satisfies important consistency and incentive properties that make it suitable to be embedded in the broader problem of matching students to multiple schools (with their own priorities and diversity goals).

Our central contribution is the design of a choice function and a corresponding algorithm satisfying desirable properties in a two-sided matching problem with diversity goals. In addition to the school choice problems, our study has several other applications. When there is one school, the problem we study is equivalent to finding a diverse committee of candidates under a mutually agreed ranking over the candidates. If school seats are replaced by other scarce and sought-after resources such as spots for vaccine treatment, our model and results apply as well to them.

Below, we discuss our contributions in more detail.

- Our first contribution is designing a general model of matching under diversity goals that captures various diversity objectives including proportionality or egalitarian concerns. The key idea behind our approach is to partition the reserved seats for various types into different ranks. We are interested in identifying subsets of students that give rise to matchings capturing a ‘desirable’ distribution of the reserved seats. We characterize our diversity objective via size-constrained rank maximal matchings of an underlying ‘ranked reservation graph’. Our flexible diversity approach not only captures upper quota constraints, it also captures other diversity goals such as target proportions of types.
- We design a new choice function for each school that takes into account the priority ordering of the school over students as well as its diversity goals. The choice function is based on a key link between flexible diversity goals and the concept of rank maximality (that is well-studied in theoretical computer science). We demonstrate the desirability of the choice function by showing that it is characterized by three fundamental properties: maximal diversity,

non-wastefulness, and justified envy-freeness. We also show that maximal diversity cannot be achieved by a natural adaptation of the horizontal envelope choice rule of Sönmez and Yenmez [39] to the case of multi-ranks.

- We provide a fast polynomial-time algorithm for our choice function based on the Dulmage Mendelsohn Decomposition and combinatorial insights into constrained rank-maximal matchings that are required to match a certain subset of vertices. Our algorithms are the first known polynomial-time approaches to compute an optimally diverse choice outcome when there are two levels of ranks that capture minimum and maximum quotas for the number of admitted students of given types.
- We show that our new choice function generalizes existing choice functions (that handle more restricted forms of diversity). We prove via Berge's lemma that maximal diversity implies the utilization of the maximum number of reserved seats. Our choice function also incentivizes applicants to report all of their types.
- We prove that our choice function satisfies substitutability and can hence be directly embedded in the generalized deferred acceptance algorithm to achieve strategyproofness and stability.

Since our algorithm is the unique rule satisfying fundamental axiomatic properties, we make a compelling case for its adoption in many important matching problems where diversity is a central concern.

2 RELATED WORK

The interest in affirmative action concerns has a long history in matching and market design. The topic of school choice with diversity constraints has been referred to as *controlled school choice* [21]. In one of the seminal works on school choice, racial and gender balance concerns were already alluded to by Abdulkadiroğlu and Sönmez [2]. In another early work, Abdulkadiroğlu [1] focussed on school choice with affirmative action in detail and imposed hard capacities or quotas on the number of admitted students of given types.

Typically, in controlled school choice, each school imposes a maximum quota and a minimum quota on each type [23, 30, 31]. Ehlers et al. [21] considered controlled school choice with lower and upper quotas for the types when students have exactly one type. They explored the implications of treating diversity quotas as hard bounds and soft bounds. For the case of soft quotas, they proposed a choice function of schools via dynamic priorities in which the students who belong to a type that is undersubscribed, are prioritized. Their dynamic priorities approach has inspired several followup works for more complex models. In a seminal work on choice functions for affirmative action, Echenique and Yenmez [20] examined the structure of choice functions that satisfy substitutability. They also assumed that each student has at most one type.

In reality, each student may satisfy multiple 'overlapping' types. Overlapping types have been considered in recent papers and deployed applications in the past few years including those in Brazil, Chile, Israel, and India (see, e.g., [6, 13, 17, 22, 33]). When individuals have multiple types, there are two natural conventions, namely *one-for-all* and *one-for-one*, on how many reserved seats an individual takes up (to be consistent with the majority of text) or how reserved positions are accounted for [39]. Under the one-for-all convention, an individual takes the reserved seats of all types she satisfies [9, 10, 22]. For example, an aboriginal girl could take up one seat reserved for girls and one seat reserved for aboriginals. The one-for-all convention has been employed in algorithms for the Israeli Mechinot matching system [22]. Optimally meeting diversity requirements is NP-hard under the one-for-all convention [11, 15]. Under the one-for-one convention, each individual takes up one of the reserved seats of only one of the types they belong to. This convention has the 'more widespread interpretation' [39]. Following our previous example, an Aboriginal girl could either take up a seat reserved for a girls or a seat reserved for Aboriginal persons but not both. In this sense, the convention incorporates a stronger sense of affirmative action because it leaves open the utilization of the second seat.

The one-for-one convention has been considered in the case where students either have strict preferences over reserved seats of different types [6, 33], or the indifferences are broken through fixed tie-breaking [13, 17], or priorities over the combination of students and types are dynamically updated [21]. In all of the above cases, the decision on which type a student should use, is made in a greedy or sequential manner and hence this greedy approach may not maximally satisfy diversity goals when students have overlapping types. A similar approach has been used by Kominers and Sönmez [31, 32], Aygün and Bó [4], and Aygun and Turhan [7], Aygün and Turhan [8]. In contrast, our framework captures new diversity goals and our methodology optimally achieves these goals via a so called 'smart reserves approach.'

Sönmez and Yenmez [39] were the first to pioneer the approach of using '*smart reserves*' in the context of controlled school choice, where the actual decision of which type a student is going to use is used more flexibly. This allows allows to maximally achieve the utilization of reserved seats for types. The underlying idea has also been applied to allocating medical resources to patients who satisfy various categories (Section 4, [36]). Although the smart reserve approach of Sönmez and Yenmez [39] is compelling, their rule and framework have some limitations. In particular, their work is limited to a single rank of reserves in which types whose reserves are not filled can be viewed as undersubscribed. Their approach can be viewed as softly respecting minimum quotas but is unable to capture a standard requirement in many affirmative action models, where both a minimum and a maximum quota is present for each type. In school choice models where each type has a minimum and maximum quota, types that have not reached their minimum quota have a higher priority than the types who have not reached their maximum quota. Consequently, the model of Sönmez and Yenmez [39] does not generalize the model of Ehlers et al. [21], who considered one type per student and allowed upper quotas in addition to lower quotas for types. One rank of reserves is also unable to capture a wide umbrella of diversity goals such as enforcing ratios among types or expressing interleaving goals such as follows: "only care about the types t_3 and t_4 once the reserved seats of t_1 and t_2 are used up."

The literature on school choice with diversity concerns is naturally divided across several axes: (1) individual students having a single type versus students having multiple types; (2) two types versus more than two types (3) hard diversity constraints versus soft diversity constraints; (4) students types accounted under the one-for-all convention versus one-to-one convention; (5) methods based on dynamic priority versus a smart reserve approach. Table 1 classifies some of the main works on controlled school choice across these axes. The distinction between the one-for-all and one-to-one conventions was first explicitly made by Sönmez and Yenmez [39]. Smart reserves approaches are meaningful when students may have multiple types and the one-to-one convention is assumed.

We discuss some approaches using the one-for-all convention. Gonczarowski et al. [22] study the Israeli "Mechinot" gap-year problem in which each student has multiple types and schools impose soft minimum quotas and hard maximum quotas. They propose a greedy choice function for schools where students are chosen one by one based on priority orderings as long as the student has some type that helps to meet the diversity goals. Their model is different from ours as it assumes the one-for-all convention. In their setting, it is impossible to achieve fairness and strategyproofness properties simultaneously. Aziz et al. [10] proposed a strategyproof mechanism that treats students with the same combinations of types in a fair way. Baswana et al. [13] designed and deployed an algorithm for Indian engineering colleges. They used a heuristic to deal with non-nested common quotas and their algorithm does not guarantee a fair outcome. Aziz [9] considers the one-for-all convention for diversity and proposed an algorithm for a choice rule that uses minimum quotas to

	Types per student		Convention		Bounds		Diversity Levels			Algorithmic Approach	
	1	any	many-1	1-1	Hard	Soft	1	≤ 2	any	Dynamic Priorities	Smart Reserves
Huang [24], Biró et al. [15]		✓	✓		✓			✓			
Ehlers et al. [21]	✓			✓		✓		✓		✓	
Echenique and Yenmez [20]	✓			✓	✓	✓					
Abdulkadiroğlu [1]	✓			✓	✓		✓				
Kominers and Sönmez [31]	✓			✓		✓	✓				
Kurata et al. [33], Correa et al. [17], Aygün and Bó [4]		✓		✓		✓		✓		✓	
Aygun and Turhan [7], Aygün and Turhan [8]		✓		✓		✓		✓		✓	
Aygün and Bó [4]		✓		✓		✓		✓		✓	
Gonczarowski et al. [22], Aziz et al. [10]		✓	✓			✓		✓		✓	
Aziz [9], Aziz and Sun (2020)		✓	✓			✓			✓	✓	
Sönmez and Yenmez [39], Pathak et al. [36]		✓		✓		✓	✓				✓
Our paper		✓		✓		✓			✓		✓

Table 1. Literature on school choice with affirmative action

specify diversity goals. Bredereck et al. [16] examined the complexity of multiwinner voting with diversity constraints under the one-for-all convention.

There has also been some recent work on matchings with constraints on the ratios of types [35]. The paper discusses how setting lower quotas on types may not be sufficient to achieve target ratios especially when schools seats are not fully used. The paper assumes that each student/doctor has exactly one type. Our diversity goals are more flexible (capturing other objectives such as minimum and maximum quotas), and our algorithmic solution and axiomatic focus is different as well.

Sonmez and Yenmez [38] focus on vertical and horizontal reservations that are distinctive to affirmative action in Indian college admissions. Horizontal reservations are equivalent to a single rank of reserves in our framework, whilst vertical Vertical reservations are treated as set-aside seats for students who are not selected on 'open merit'. Dur et al. [19] studied a class of mechanisms in which reserves seats are processed in a sequential manner. They highlighted the impact of the order in which reserves are processed.

Our model also bears some similarities with the hospital-resident matching problem with regional constraints [26, 27], in which students are viewed as doctors, school seats are viewed as hospitals and schools are viewed as regions. However, in the hospital-resident matching problem, the distributional constraints are imposed on the number of doctors matched to different regions , rather than on the proportional composition of types of doctors. Aziz et al. [11] focus on the connection between matching with diversity constraints and matching with regional quotas where diversity constraints are set as hard bounds. However, they do not focus on the mechanism design of choice function of schools that maximize diversity goals.

Benabbou et al. [14] study the Singapore public housing program where ethnic quotas are imposed to each block: the percentage of each ethnic group cannot exceed a certain degree. Each agent has preferences over apartments within a block, whilst in our model, each student has preferences over schools rather than school seats within the school. Furthermore, they focus on the tradeoff between diversity and social welfare, while our attention is drawn to the maximization of diversity goals for each school. Ahmadi et al. [3] and Dickerson et al. [18] consider an optimisation-based approach to diverse team formation but do not take preferences and priorities into account.

3 SCHOOL CHOICE WITH MULTI-RANK DIVERSITY GOALS

An instance I of the school choice problem with soft diversity goals consists of a tuple $(S, C, q_C, T, \succ_S, \succ_C, \eta)$ 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. We denote by T the set of types. We overload the term to also capture the information about the types of each student. For each student s, let $T(s) \subseteq T$ denote the subset of types to which student s belongs. If $T(s) = \emptyset$, then student s does not have any privileged type. The vector $\eta = (\eta_1, \ldots, \eta_{|C|})$ specifies the diversity goals of each school c. We will discuss diversity goals in detail later.

Let $\mathcal{X} = S \times C$ denote the set of possible student-school pairs. We also refer to these pairs as contracts. Given any $X \subseteq \mathcal{X}$, let X_s be the set of student-school pairs involving student s and let X_c be the set of student-school pairs involving school c.

Each student s has a strict preference ordering $\succ_s$ over $C \cup \{\emptyset\}$ where $\emptyset$ is a null school representing the option of being unmatched for student s. A school c is *acceptable* to student s if $c \succ_s \emptyset$ holds. 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 $S \cup \{\emptyset\}$ where $\emptyset$ represents the option of leaving seats vacant for school c. A student s is *acceptable* to school c if $s \succ_c \emptyset$ holds. Let $\succ_C = (\succ_{c_1}, \ldots, \succ_{c_m})$ be the priority profile of all the schools. The priority ordering of the school could be based on the entrance exam scores, or in the case of automated hiring, on some objective measure that captures the suitability of the applicants. Our results continue to hold if we assume weak preferences and priorities: we can artificially break the ties with fixed tie-breaking to induce strict orders.

An outcome (or a matching) X is a subset of $\mathcal{X}$. An outcome X is *feasible* (under soft bounds) for instance I 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 and feasible matchings.

Except for Section 9, we will focus on how a single school makes decisions. A *choice function* of a school takes as input a given set of students $S' \subseteq S$ and selects a set of students $S^* \subseteq S'$ as output. In our setup, the choice function Ch_c of a school c will take into account its priority order $\succ_c$ and its diversity goals η_c to make the choice.

Diversity Goals

In most of the literature on school choice with diversity constraints, schools typically impose minimum q_t^{min} and maximum quota q_t^{max} on each type t. We can interpret this type of diversity goal as using two ranks of quotas: for a given type t, rank 1 corresponds to the interval $[0, q_t^{min})$ and rank 2 corresponds to the interval $[q_t^{min}, q_t^{max})$. An implicit third rank is for types whose maximum quotas have already been met.

We can interpret this type of diversity goal as using two ranks of quotas: for a given type t, rank 1 corresponds to the case where the number of matched students of type t is no larger than the minimum quota, rank 2 corresponds to the case above the minimum quota and no larger than the maximum quota.[1]

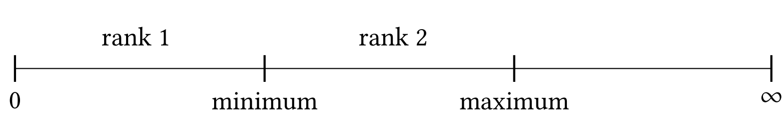

Fig. 1. An interpretation of minimum and maximum quotas.

Here, ranks are used to measure the importance of diversity goals [12]: the smaller the rank is, the more important the type is. For instance, consider two types t_1 and t_2 whose numbers of matched students at school c fall into rank i and j respectively. If $i < j$, then type i is more important than type j to school c in terms of satisfying diversity goals. If $i = j$, then both types are tied.

[1] An implicit third rank is for types whose maximum quotas have already been met.

In our model, we allow each school c to impose multiple ranks of quotas on each type. The parameter η_c specifies for each type t and rank j, the reserved quota $\eta^j_{c,t}$. We will denote by r the maximum number of ranks. A student can only take up a reserved seat of a type that she satisfies.[2] If some reserved seats remain unallocated, then in order to ensure maximal utilization of resources, any eligible student can take up a vacant seat. When a set of students are viewed as taking up reserved seats for the types they satisfy, the seats of earlier ranks are taken up before seats of later ranks are considered.[3]

Example 3.1 (Example of a problem instance). Consider the setting in which there are four students $S = \{s_1, s_2, s_3, s_4\}$ and two schools $C = \{c_1, c_2\}$. The type profile of the students is $T(s_1) = \{t_1, t_2\}$, $T(s_2) = \{t_1\}$, $T(s_3) = \emptyset$, $T(s_4) = \{t_3\}$. The capacity of c_1 is $q_{c_1} = 3$ and the capacity of c_2 is $q_{c_2} = 1$. School c_1 has diversity goals specified as follows: $\eta^1_{c_1,t_1} = 1$, $\eta^1_{c_1,t_2} = 1$, $\eta^1_{c_1,t_3} = 0$, $\eta^2_{c_1,t_3} = 1$. On the other hand, school c_2 has no diversity goals. The preferences of the students are unanimous with c_1 preferred over c_2. The priorities of both schools are unanimous: $s_1 \succ_c s_2 \succ_c s_3 \succ_c s_4$ for $c \in C$.

The interpretation of the diversity goals of school c_1 is as follows: school c_1 wants to match as many students to slots of rank 1 as possible. One of the rank 1 slot is reserved for a student of type t_1 and one is reserved for a student of type t_2. Conditional on optimising the number of students who can be matched to rank 1 slots, the school c_1 wants to next match some student of type t_3. Another interpretation of the diversity goals is in the form of setting quotas. The minimum and maximum quota for t_1 and t_2 is 1 whereas the minimum quota for t_3 is 0 and the maximum quota for t_3 is 1. ⋄

Below we further explain how our framework can capture diversity goals that the method of using a single layer of reserves used by Sönmez and Yenmez [39] cannot. Suppose that t_1 has a soft lower quota of 1 and a soft upper quota of 2. On the other hand t_2 has a soft lower quota of 0 and a soft upper quota of 3. Then it is not clear how many reserves to make for t_1 and t_2 as the diversity goals are inherently two-layered.

Before we formally specify how the reserved seats information is used to define diversity goals, we give some intution. Informally speaking, the diversity goals are achieved as follows. We want as many rank 1 seats to be filled by students who satisfy the corresponding types and conditional on that we want as many rank 2 seats filled, and so on. In the next section, we will give a formal description of how the reserved quotas $\eta^j_{c,t}$ give rise to the diversity goals discussed above. We will use the notion of a *ranked reservation graph.* This mathematical object is central to formally specifying diversity goals as well as other axiomatic properties.

4 A GRAPH-THEORETIC VIEW OF DIVERSITY

Given a set of students S' and a school c with reserved quotas η_c, a corresponding *ranked reservation graph* $G = (S' \cup V, E, \eta_c)$ is a bipartite graph whose vertices are partitioned into a set of students S' and a set of reserved seats V. Each reserved seat $v^j_{t,i} \in V$ has a rank j, a type t and an index i. For each rank j and each type t, we create $\eta^j_{c,t}$ reserved seats in G. The edge set E is specified as follows. There is an edge between a student s and a reserved seat $v^j_{t,i}$ if student s belongs to type t, i.e., $t \in T(s)$. Each edge $(s, v^j_{t,i})$ has a *rank* j corresponding to the rank j of the reserved seat $v^j_{t,i}$. We refer to all edges with rank j as j-ranked edges. The ranks of the edges lead to a natural partition of the edges: $E = E_1 \cup E_1 \cup \cdots \cup E_r$ where E_j denotes the set of j-th ranked edges. Note that the ranked reservation graph is a generalization of the reservation graph used by Sönmez

[2]This assumption is referred to as 'compliance with eligibility requirements' in the literature [36].

[3]This order of filling up seats is consistent with existing approaches in which minimum quotas are reached first before targetting the maximum quotas.

and Yenmez [39] in which all edges only have one rank. A ranked reservation graph G is a special ranked bipartite graph in which all edges incident to the same reserved seat $v^j_{t,i}$ have the same rank j.

Example 4.1 (Example of a ranked reservation graph). Consider the problem instance in Example 3.1. We construct the corresponding reservation graph as shown in Figure 2.

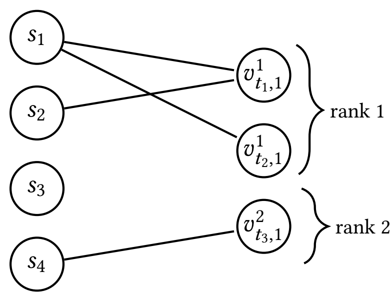

Fig. 2. The ranked reservation graph of school c_1 in Example 3.1.

Before we proceed further, we formally specify important terms and concepts from matching theory.

4.1 Matching Theory Preliminaries

Given a graph G, a matching M in G is a set of pairwise non-adjacent edges such that no two edges have common vertices. Given a matching M, an **alternating path** is a path that begins with an unmatched vertex and whose edges belong alternately to the matching and not to the matching. An **augmenting path** is an alternating path that starts from and ends on unmatched vertices.

Let P denote an alternating path with respect to matching M. Then $M \oplus P = (M \setminus P) \cup (P \setminus M)$ denotes the symmetric difference of the two set of edges, which is a new matching where the edges from $P \setminus M$ are matched while the edges from $M \cap P$ are not matched.

Consider a ranked bipartite graph $G = (A \cup B, E)$ in which each edge is assigned a rank. Suppose the edge set E is partitioned into r disjoint sets, i.e. $E = E_1 \cup E_2 \cup \cdots \cup E_r$ where E_i represents the set of edges of rank i. The **signature** $\rho(M) = \langle x_1, x_2, ..., x_r \rangle$ of a matching M in G is a tuple of integers where each element x_i represents the number of edges of rank i in M.

For a ranked bipartite graph, we compare the signatures of matchings in a lexicographical manner. A matching M' with $\rho(M') = \langle x_1, \cdots, x_r \rangle$ is **strictly better** than another matching M'' with $\rho(M'') = \langle y_1, \cdots, y_r \rangle$, if there exists an index $1 \leq k \leq r$ s.t. for $1 \leq i < k$, $x_i = y_i$ and $x_k > y_k$. A matching M' is **weakly better** than another matching M'' if M'' does not provide strictly better diversity than M'. Let $M' \succ_{lex} M''$ denote that M' is strictly better than M'' and let $M' \succsim_{lex} M''$ denote that M' is weakly better than M''.

A matching M in a ranked bipartite graph G is **rank-maximal** if M is weakly better than any other matching M' in G. A rank maximal matching can be computed in polynomial time [25, 34].

4.2 Maximally Diverse Matchings of a Reservation Graph

For a ranked reservation graph, a matching M' provides **strictly better diversity** than M'' if $M' \succ_{lex} M''$; and a matching M' provides **weakly better diversity** than M'' if $M' \succsim_{lex} M''$. A matching in a ranked reservation graph G is *maximally diverse* if it provides weakly better diversity than all other matchings of G. Alternatively, a matching in a ranked reservation graph G is *maximally diverse* if it is rank-maximal.

In this paper, one of our main goals is to formalize a way for a school to select a set of students while keeping diversity goals in mind. In the next section, we show how the concept of diverse matchings leads to a natural definition of a diverse set of students. Informally speaking, we will focus on selecting those sets of students that allow for diverse matchings in the underlying ranked reservation graph.

5 DESIRABLE PROPERTIES OF A CHOICE FUNCTION

A choice function Ch_c of school c takes as input an $S' \subseteq S$, q_c, η_c, and $\succ_c$, and selects a set of students $S^* \subseteq S'$. We will assume that the choice function satisfies feasibility requirements: $|Ch_c(S', q_c, \eta_c, \succ_c)| \leq q_c$.

Next, we extend the idea of comparing the signatures of two matchings in the reservation graph to that of comparing the diversity satisfaction of two subsets of students. The intuition is to check which subset of students leads to better utilization of the reserved seats in the ranked reservation graph.

Given a school c and two subsets of students $S_1, S_2 \subseteq S$ with $|S_1|, |S_2| \leq q_c$, we say that S_1 provides *strictly (weakly) better diversity* than S_2, if there exists a matching M in the ranked reservation graph $G = (S \cup V, E, \eta_c)$ such that

(1) the matched students in M are a subset of S_1;
(2) M provides strictly (weakly) better diversity than any matching M' of G in which the matched students in M' are a subset of S_2.

The weakly better diversity relation is a total order on the set of sets of students. We use it to define a property called maximal diversity of a choice function.

Definition 5.1 (Maximal diversity of a choice function). A choice function Ch_c satisfies *maximal diversity* if for each instance $(S', q_c, \eta_c, \succ_c)$, it selects a feasible set of students $Ch_c(S', q_c, \eta_c, \succ_c)$ such that $Ch_c(S', q_c, \eta_c, \succ_c)$ provides weakly better diversity than all feasible subsets $S'' \subseteq S'$.

Note that multiple ranks provides us with the ability to describe more complicated diversity goals. For instance, only using minimum and maximum quotas cannot capture the idea of "proportionality", a common diversity goal.

Example 5.2 (Proportional single rank reserves may not achieve proportionality). Suppose each student is associated with one of three types t_1, t_2, t_3, and the percentages of students of each type are 30%, 30%, 40%, respectively. Consider one school c with capacity 100 which imposes minimum quotas 30, 30 and 40 on type t_1, t_2, t_3 respectively. Note that the numbers of students of each type who apply to school c do not match their percentages in the population. Say 15, 60, 60 students of type t_1, t_2, t_3 respectively apply for the school c, and the priority ordering of school c is consistent with students' types, say school c prefers students of type t_1, to students of type t_2, to students of type t_3.

Then one possible matching for school c with respect to minimum quotas could be 15, 45, 40, where each integer denotes the number of students of type t_1, t_2, t_3 respectively. However, another matching with 15, 37, 48 students of each type matched, seems better for achieving a diversity balance under the proportionality constraints, in which the numbers of matched students of type t_2 and t_3 are proportional in the ratio of 3:4.

Non-wastefulness stipulates that either the school capacity should be saturated or all valid applicants must be accepted.[4]

Definition 5.3 (Non-wastefulness). A choice function satisfies non-wastefulness if a student is rejected only if the school capacity is reached.

Justified envy-freeness requires that a student with a higher priority cannot replace another student with a lower priority without compromising on the best achievable diversity goals.

Definition 5.4 (Justified envy-freeness). For an instance $(S' \subseteq S, q_c, \eta_c, \succ_c)$, if a school c selects $Ch_c(S') = S^* \subseteq S'$, we say that a student $i \in S' \setminus S^*$ has justified envy towards another student $j \in S^*$ if $i \succ_c j$ and $S^* \cup \{i\} \setminus \{j\}$ provides weakly better diversity than S^*. The choice function Ch_c of a school c satisfies *justified envy-freeness*, if for each instance $(S', q_c, \eta_c, \succ_c)$, it selects a set of students $Ch_c(S') = S^* \subseteq S'$ s.t. there does not exist any student $i \notin S^*$ who has justified envy towards any student $j \in S^*$.

Note that whereas justified envy-freeness is referred to as a property of matchings in school choice, the definition above is a property of a choice function of an individual school. If an individual school does not satisfy the property in Definition 5.4, then the justified envy-freeness of matchings is also not satisfied. Note that a choice function Ch_c of school c that completely ignores diversity goals and selects the q_c highest priority students satisfies justified envy-freeness.

Next, we introduce a property called *Maximal Reserves Utilization* that concerns whether the number of used reserved seats is maximal in the corresponding ranked reservation graph.

Given a school c and two subsets of students $S_1, S_2 \subseteq S$ with $|S_1|, |S_2| \leq q_c$, we say that S_1 *provides strictly (weakly) better reserves utilization* than S_2, if there exists a matching M in the ranked reservation graph $G = (S \cup V, E)$ such that

(1) the matched students in M are a subset of S_1;
(2) the size of M is strictly (weakly) larger than any matching M' of G in which the matched students in M' are a subset of S_2.

Definition 5.5 (Maximal Reserves Utilization of a choice function). A choice function Ch_c satisfies *maximal reserves utilization* if for each instance $(S', q_c, \eta_c, \succ_c)$, it selects a feasible set of students $Ch_c(S', q_c, \eta_c, \succ_c)$ such that $Ch_c(S', q_c, \eta_c, \succ_c)$ provides weakly better reserves utilization than any feasible subsets $S'' \subseteq S'$.

Finally, we consider another property of choice functions. An algorithm is *privilege monotonic* if no student has an incentive to report only a strict subset of their true types. Specifically, there exists no instance under which a student is selected by a choice function when it reports all her types but is not selected when she reports a strict subset of her types. The term has been used previously (see e.g. Aygün and Bó [4]).

6 A NEW CHOICE FUNCTION

Next, we design a new choice function for schools and show that it satisfies the compelling properties discussed above. The choice function relies on computing size-constrained *rank maximal* matchings of the corresponding reservation graph.

The choice function works as follows. We first compute the corresponding reservation graph G that includes the ranks of the edges according to their ranks. We then compute the signature of the rank-maximal matching in G constrained to the school's quota. At this point we do not make a

[4] It can be assumed without loss of generality that we only consider those applicants that meet minimal acceptance requirements.

decision on which individuals are to be chosen. We simply require that the students selected should give rise to the same signature. We go down the priority list of students and check whether a given student s can be matched along with the previously selected students so that we still can get the same signature from some matching. If this is possible, we select s. Otherwise, we do not select s. The process continues until we have exhausted the priority list and we have a set of students who can all be matched in a rank-maximal matching of G. At this point, if the school's quota is not exceeded, we again go through students from the start of the priority list and add unselected students if the quota is still not exceeded. The algorithm is specified formally as Algorithm 1.

Input: $S' \subseteq S$, q_c, η_c, $\succ_c$.
Output: A set of students $S^* \subseteq S'$

Selected students $S^* \leftarrow \emptyset$
Construct the corresponding ranked reservation graph $G = (S' \cup V, E, \eta_c)$.
for student $s \notin S^*$ down the list in $\succ_c$ **do**
    if there exists a matching in G of size at most q_c that satisfies the following two conditions
    (1) it is rank maximal among all matchings in G of size at most q_c
    (2) it matches all students in $S^* \cup \{s\}$
    then
        Add s to S^*
for student s down the list in $\succ_c$ **do**
    if $|S^*| < q_c$ and $s \notin S^*$ **then**
        Add s to S^*
return S^*

Algorithm 1. Choice function Ch_c of school c

In the following example, we illustrate how Algorithm 1 works.

Example 6.1. Consider the problem instance in Example 3.1. Suppose all the four students apply to school c_1. Then the school c_1 first selects s_1, then s_2, and then s_4. The corresponding maximal diversity matching is shown in the figure below. It is the unique rank maximal matching of the reservation graph. Hence s_1, s_2 and s_4 are selected.

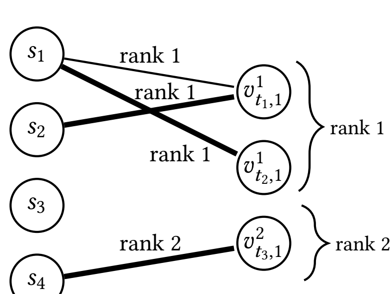

Fig. 3. The maximal diversity matching for ranked reservation graph of school c_1 in Example 3.1 is illustrated by thick bold edges.

7 AXIOMATIC PROPERTIES OF THE NEW CHOICE FUNCTION

In this section, we analyze the fundamental properties satisfied by our choice function. We show that a choice function satisfies maximal diversity, justified envy-freeness, and non-wastefulness if and only if it selects the same set of students as the choice function in Algorithm 1 does. The following lemmata show that our choice function satisfies three important axioms.

LEMMA 7.1. *The choice function in Algorithm 1 satisfies non-wastefulness.*

LEMMA 7.2. *The choice function in Algorithm 1 satisfies maximal diversity.*

LEMMA 7.3. *The choice function in Algorithm 1 satisfies justified envy-freeness.*

We have established that our choice functions satisfies maximal diversity, non-wastefulness, and justified envy-freeness. Interestingly, the converse holds as well.

LEMMA 7.4. *If a choice rule satisfies maximal diversity, non-wastefulness, and justified envy-freeness, then it is equivalent to the choice function in Algorithm 1.*

Based on the lemmata above, we obtain the following characterization of our choice function.

THEOREM 7.5 (CHARACTERIZATION OF THE CHOICE FUNCTION). *A choice rule satisfies maximal diversity, justified envy-freeness, and non-wastefulness if and only if it is the choice function in Algorithm 1.*

Next, we discuss other axiomatic properties of our choice function. Maximal diversity puts special focus on the initial layers before focusing on the latter layers. In view of this focus, a question arises whether a maximally diverse matching uses the maximum number of reserved seats. In the next lemma, we prove that a matching that satisfies maximal diversity uses the maximum number of reserved seats by students who are eligible for the reserved seats. The statement contrasts with the fact that in general bipartite graphs, a rank maximal matching of a bipartite graph need not be a maximum size matching.

LEMMA 7.6. *A matching that satisfies maximal diversity uses the maximum number of reserved seats.*

The lemma above gives us the following.

THEOREM 7.7. *The choice function in Algorithm 1 satisfies maximal reserves utilization.*

The result also shows that our choice function generalizes the horizontal envelope choice rule of Sönmez and Yenmez [39] in another way by utilizing the maximum number of reserved seats.

Next, we prove that Algorithm 1 incentivizes the students to not hide their types.

THEOREM 7.8. *Algorithm 1 is privilege monotonic.*

8 COMBINATORIAL INSIGHTS AND A FAST ALGORITHM

In this section, we take a closer look at the choice function defined in Algorithm 1. Our first observation is that the outcome of Algorithm 1 can be computed in polynomial time via a reduction to maximum weight matching. We denote $\sum_{j=1}^{m_c} \sum_{t \in T} \eta_{c,t}^{j}$ by B.

THEOREM 8.1. *An outcome of Algorithm 1 can be computed in polynomial time* $O(r|V|^3 \cdot |S|)$ *where* $|V| = (|S| + min(B, q_c))$ *and* r *is the total number of levels.*

Next, we present a tailored algorithm that is faster. Apart from its computational advantage, the design of the algorithm is based on further combinatorial insights into the structure of our solutions. Our algorithm does not directly call existing fast algorithms for rank maximal matchings but carefully relies on some of the combinatorial insights underpinning previous algorithms. The running time of the algorithm is $O(rm\sqrt{n} + nm)$ where r, n and m denotes the number of levels, the number of nodes, the number of edges in the reservation graph.

8.1 A detour to the Algorithm of Irving et al. [25]

In order to set up the groundwork for our fast algorithm, we first present the classical algorithm of Irving et al. [25] for computing a rank maximal matching of a given ranked bipartite graph. It is based on the Dulmage Mendelsohn Decomposition which we describe below.

Let M denote a maximum matching for a bipartite graph G. Then we can partition all vertices into three categories:

- $\mathcal{E}$: vertices which are reachable via even length alternating paths from a free vertex with respect to M.
- $\mathcal{O}$: vertices which are reachable via odd length alternating paths from a free vertex with respect to M.
- $\mathcal{U}$: vertices that do not belong to $\mathcal{E}$ or $\mathcal{O}$.

Theorem 8.2 (Irving et al. [25]). *The Dulmage Mendelsohn Decomposition consists of the following statement:*

- *$\mathcal{E}, \mathcal{O}, \mathcal{U}$ are invariant among all maximum matchings.*
- *G does not contain an $\mathcal{E}\mathcal{U}$ / $\mathcal{E}\mathcal{E}$ edge.*
- *No maximum matching contains an $\mathcal{O}\mathcal{O}$ / $\mathcal{O}\mathcal{U}$ edge.*
- *Every vertex in $\mathcal{O}$ is matched (to some vertex in $\mathcal{E}$) and every vertex in $\mathcal{U}$ is matched (to some vertex in $\mathcal{U}$).*
- *The cardinality of any maximum matching is $|\mathcal{O}| + |\mathcal{U}|/2$.*

The following Algorithm 2 is currently the fastest algorithm for computing a rank maximal matching [25] based on the Dulmage Mendelsohn Decomposition. The main idea in Algorithm 2 is to convert the problem of computing a rank maximal matching into a problem of computing a maximum matching in a carefully chosen subgraph, in which all edges that never belong to any rank maximal matching are deleted, as shown in Theorem 8.3.

Input: a ranked bipartite graph $G = (S \cup V, E_1 \cup \cdots \cup E_L)$
Output: a rank maximal matching M_L

$G'_1 = G_1$ where G_1 denotes the induced subgraph of G with edges of rank 1 only and M_1 is an arbitrary maximum matching in G'_1.
for $i = 1$ to $L - 1$ **do**
    Partition all nodes into $\mathcal{E}_i, \mathcal{O}_i, \mathcal{U}_i$ in subgraph G'_i;
    Delete all edges incident to a node in $\mathcal{O}_i, \mathcal{U}_i$ from E_j for all $j > i$;
    Delete all edges $\mathcal{O}_i\mathcal{O}_i$ and $\mathcal{O}_i\mathcal{U}_i$ in G'_i;
    Obtain G'_{i+1} by adding edges E_{i+1} to G'_i;
    Compute a maximum matching M_{i+1} in G'_{i+1} by finding all augmenting paths w.r.t M_i.
return matching M_L

Algorithm 2. Algorithm of Irving et al. [25] for computing a rank maximal matching

Theorem 8.3 (Irving et al. [25]). *For every $1 \leq i \leq L$, we have:*

- *Every rank maximal matching in G_i has all of its edges in G'_i.*

- *M_i is a rank maximal matching for G_i.*

Note that $G_i = (S \cup V, E_1 \cup \cdots \cup E_i)$ represents a subgraph of G with only edges of rank 1 to i and G'_i represents a pruned subgraph by removing some edges from G'_i that could not be part of any rank maximal matching. In Theorem 8.3, the matching M_i is obtained by augmenting M_{i-1}. Note that not every maximum matching M'_i in subgraph graph G'_i is a rank maximal matching.

Theorem 8.4 (Irving et al. [25]). *A rank-maximal matching can be computed in time $O(rm\sqrt{n})$ where r, n and m denotes the number of ranks, the number of nodes, the number of edges in the ranked bipartite graph.*

Although the algorithm of Irving et al. [25] provides an efficient way to compute a rank maximal matching and hence a diverse matching, it cannot be directly used to design a fast algorithm for our choice function. There are at least two aspects that need to be simultaneously handled: (1) our problem is subject to a size constraint on the matchings due to school capacity (2) we need to match target subsets of students with respect to school priority. Next we address these issues by designing a new algorithm that takes inspiration from the combinatorial insights of Irving et al. [25].

8.2 Modified Rank Maximal Matching

Since we take school capacity q_c into consideration, it is possible that the size of a rank maximal matching M in G is larger than school capacity. Given the school capacity q_c and a rank maximal matching M^* in G with the signature $\rho(M^*) = \langle x_1, \cdots, x_r \rangle$, let k denote the rank s.t. $\sum_{i=1}^{k} x_i \leq q_c < \sum_{i=1}^{k+1} x_i$. The **maximal clipped signature** of G with respect to school capacity q_c is denoted as $\hat{\rho}(G, q_c) = \langle y_1, \cdots, y_k, \cdots, y_r \rangle$ where

$$y_i = \begin{cases} x_i, & \text{if } i \leq k-1, \\ q_c - \sum_{i=1}^{k-1} x_i & \text{if } i = k, \\ 0 & \text{if } i > k. \end{cases}$$

We reduce the element y_k to be $q_c - \sum_{i=1}^{k-1} x_i$ and set each element y_i to be 0 for $i > k$. It is safe to remove all edges of rank $k+1$ or above from graph G, since they cannot be matched in any rank maximal matching due to school capacity.

Note that a matching with maximal clipped signature is rank maximal among all matchings in G of size at most q_c. By the definition of maximal diversity, a matching M is maximally diverse if it has the maximal clipped signature $\hat{\rho}(G, q_c)$.

The following Algorithm 3 is a modified version of Algorithm 2 that computes a matching with maximal clipped signature with respect to school capacity. Next, we give some intuition about Algorithm 3.

Algorithm 3 yields two objects: a **pruned reservation graph** G'_k and a rank maximal matching M_k in G'_k. Here k denotes some integer in the range $[1, r]$, depending on school capacity q_c.

Note that in subgraph G'_k, all edges of G that do not belong to any rank maximal matching are removed, while every rank maximal matching in G_k has all of its edges in G'_k by Theorem 8.3.

If the size of the returned matching M_k in G'_k is larger than q_c, then we can just remove $|M_k| - q_c$ edges of rank k from M_k to obtain a new matching M'_k of size q_c, which is still rank maximal among all matchings of size at most q_c in G'_k. However, we do not immediately make decisions about which edges of rank k should be deleted. We need to take school priorities into account to decide which students should be selected, as discussed in the next section.

Input: a ranked reservation graph $G = (S \cup V, E_1 \cup \cdots \cup E_r)$, q_c
Output: a pruned reservation graph G'_k and a rank maximal matching M_k in G'_k

$G'_1 = G_1$ where G_1 denotes the induced subgraph of G with edges of rank 1 only.
Compute a maximum matching M_1 in subgraph G'_1.
for $i = 1$ to $r - 1$ **do**
    Partition all nodes into $\mathcal{E}_i, \mathcal{O}_i, \mathcal{U}_i$ w.r.t M_i in G'_i.
    Delete all edges incident to some node in $\mathcal{O}_i, \mathcal{U}_i$ from E_j for all $j > i$.
    Delete all edges $\mathcal{O}_i\mathcal{O}_i$ and $\mathcal{O}_i\mathcal{U}_i$ in G'_i.
    if $|M_i| \geq q_c$ **then**
        $k \leftarrow i$
        return G'_k and M_k
    else
        Obtain G'_{i+1} by adding edges E_{i+1} to G'_i.
        Compute a maximum matching M_{i+1} in G'_{i+1} by finding augmenting paths w.r.t. M_i.
$k \leftarrow r$
return G'_k and M_k

Algorithm 3. Computing a modified rank maximal matching

8.3 A Fast Algorithm for Our Choice Function

Next, we present a fast solution (Algorithm 4) that implements the choice function in Algorithm 1. There are two main differences from the previous solution in Theorem 8.1. First, we do not compute a maximum weight matching, which is computational inefficient. Second, we do not compute a new rank maximal matching containing a certain set of students S^* each time. Instead, we make full use of an existing rank maximal matching by computing alternating and augmenting paths.

THEOREM 8.5. *Algorithm 1 and Algorithm 4 return the same set of students.*

THEOREM 8.6. *Algorithm 4 runs in time $O(rm\sqrt{n} + nm)$ where r, n and m denotes the number of levels, the number of nodes, the number of edges in the pruned reservation graph.*

PROOF. Algorithm 4 first computes a rank maximal matching M in the pruned reservation graph G'_r by Algorithm 3, which takes time $O(rm\sqrt{n})$ by Theorem 8.4. Then we arbitrarily remove $|M| - q_c$ edges of rank r from M if $|M| > q_c$ in constant time. In the first for loop, for each student, we check whether there exists an alternating path or an augmenting path in time $O(m)$. Thus the total running time of the first for loop is $O(nm)$. The second for loop checks whether we can add some students without exceeding school capacity in time $O(n)$. The whole running time of Algorithm 4 is $O(rm\sqrt{n} + nm)$. □

Next, we present an illustrating example for Algorithm 4.

Example 8.7. Consider the following setting in which there are four students $S = \{s_1, s_2, s_3, s_4\}$ and one school $C = \{c\}$. The priority ordering of school c is $s_1 \succ_c s_2 \succ_c s_3 \succ_c s_4$ and all students consider school c acceptable. The type profile of the students is $T(s_1) = \{t_1, t_2\}$, $T(s_2) = \{t_1\}$, $T(s_3) = \{t_3, t_4\}$, $T(s_4) = \{t_4\}$. The school capacity of c is $q_c = 3$ and it has diversity goals specified as follows: $\eta^1_{c,t_2} = \eta^1_{c,t_3} = 0$, $\eta^1_{c,t_1} = \eta^2_{c,t_2} = \eta^2_{c,t_3} = \eta^1_{c,t_4} = 1$. The corresponding reservation graph is depicted in Figure 4 where the superscript j of a seat $v^j_{t,i}$ represents its rank.

The rank maximal matching yielded by Algorithm 3 is $\{(s_1, v^2_{t_2,1}), (s_2, v^1_{t_1,1}), (s_3, v^2_{t_3,1}), (s_4, v^1_{t_4,1})\}$. Since the school capacity is 3, we randomly delete one edge of rank 2, say $(s_1, v^2_{t_2,1})$.

Input: $S' \subseteq S, q_c, \succ_c, \eta_c$.
Output: A set of students $S^* \subseteq S'$

Construct a ranked reservation graph G.
Compute a pruned reservation graph G'_k as well as a rank maximal matching M in G'_k by running Algorithm 3 on input G, q_c.
if $|M| > q_c$ **then**
    Remove $|M| - q_c$ edges of rank k from M arbitrarily.
$S^* \leftarrow \emptyset$
for each student s in descending ordering of $\succ_c$ **do**
    if s is matched in M **then**
        $S^* \leftarrow S^* \cup \{s\}$
    else if there exists an *alternating path* P w.r.t M that starts from s and ends at $s' \in S \setminus S^*$ **then**
        $S^* \leftarrow S^* \cup \{s\}$
        $M \leftarrow M \oplus P$
    else if there exist i) an *augmenting path* P w.r.t M that starts from s and ends at some free seat v of rank k and ii) some student $s' \in S \setminus S^*$ who is matched to some seat $v^k_{t,i}$ of rank k in M **then**
        $S^* \leftarrow S^* \cup \{s\}$
        $M \leftarrow M \oplus P$
        $M \setminus \{(s', v^k_{t,i})\}$ % Remove student s' from the matching M
for each student s in descending ordering of $\succ_c$ **do**
    if $|S^*| < q_c$ and $s \notin S^*$ **then**
        $S^* \leftarrow S^* \cup \{s\}$
return S^*

Algorithm 4. A Fast Algorithm for the Choice Function

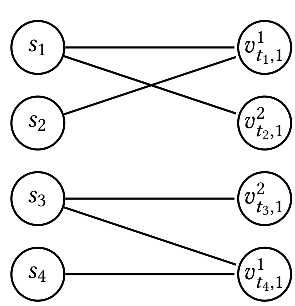

Fig. 4. The corresponding reservation graph of school c in Example 8.7.

Next, we go through students one by one based on school priority and check whether there exists a matching of size 3 with maximal clipped signature $\langle 2, 1\rangle$. Initially set S^* to be empty.

- For student s_1, there exists an alternating path $s_1, v^1_{t_1,1}, s_2$. Thus we add s_1 to S^* and update M to be $\{(s_1, v^1_{t_1,1}), (s_3, v^2_{t_3,1}), (s_4, v^1_{t_4,1})\}$.
- For student s_2, there exist an augmenting path $s_2, v^1_{t_1,1}, s_1, v^2_{t_2,1}$ and another student $s_3 \notin S^*$ who is matched to some seat of rank 2. Thus we add s_2 to S^* and update M to be $\{(s_1, v^2_{t_2,1}), (s_2, v^1_{t_1,1}), (s_4, v^1_{t_4,1})\}$.

- For student s_3, there exists an alternating path $s_3, v^1_{t_4,1}, s_4$. Thus we add s_3 to S^* and update M to be $\{(s_1, v^2_{t_2,1}), (s_2, v^1_{t_1,1}), (s_3, v^1_{t_4,1})\}$.
- For student s_4, we cannot find either i) an alternating path starting from s_4 and ending at some $s' \notin S^*$ or ii) an augmenting path starting from s_4 and another student $s' \notin S^*$ who is matched to some seat of rank 2. Thus we cannot add s_4 to S^*.

9 CONTROLLED SCHOOL CHOICE

In two-sided matching, the most important algorithm is the deferred acceptance algorithm.

Input: $\succ_S, Ch_C$
Output: A matching X

while some student has not been rejected from all the schools **do**
    All the unmatched students apply to their most preferred schools that have not rejected them.
    For each school $c \in C$, let S_c be the set of students who are matched to c or who now apply to c. School c selects students $Ch_c(S_c)$ and rejects the rest.
return Matching X that represents the current matches.

Algorithm 5. Student Proposing Deferred Acceptance (SPDA)

The Student Proposing Deferred Acceptance (SPDA) algorithm works in the same way as the original deferred acceptance algorithm [37]: 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. This procedure is repeated until no more contract is rejected by any school.

The cornerstone result in the theory of matching markets is that the student proposing deferred acceptance algorithm is strategyproof and returns a matching which satisfies a natural notion of stability. We say that an algorithm or rule is *strategyproof* if there exists no student who can misreport her preferences to get a better outcome. Stability is defined as follows. Here we use a natural notion of choice function of a given student Ch_s as taking as input a set of schools and returning the most preferred school according to $\succ_s$.

Definition 9.1 (Stability). A matching M is stable if it is

(1) $M(i) \in Ch_i(M(i))$ (individually rational)
(2) there exists no student s and school c such that $(s, c) \succ_s (s, M(s))$ and $s \in Ch_c(M(c) \cup \{s\})$. (No blocking pair)

The results rely on the assumption that the choice function satisfies certain axioms we define below. We first formally define these axioms and then show that they are satisfied by our choice function.

Definition 9.2 (Substitutability [28]). A choice rule Ch satisfies substitutability if for every $S' \subseteq S$, $i \in Ch(S')$ and $j \neq i \implies i \in Ch(S' \setminus \{j\})$.

Definition 9.3 (Law of aggregate demand (LAD)). A choice rule Ch satisfies the *law of aggregate demand (LAD)* if for $S'' \subseteq S' \subseteq S$, if $|Ch(S'')| \leq |Ch(S')|$.

Definition 9.4 (Irrelevance of rejected individuals [5]). A choice rule Ch satisfies the *irrelevance of rejected individuals* condition if for every $S' \subseteq S$, $i \in S \setminus Ch(S') \implies Ch(S' \setminus \{i\}) = Ch(S')$.

Lemma 9.5. *The choice function in Algorithm 1 satisfies substitutability, law of aggregate demand, and the irrelevance of rejected individuals condition.*

Theorem 9.6. *If the choice functions of the schools are of the class defined in Algorithm 1, then SPDA is strategyproof and returns a stable outcome.*

Proof. We have already shown that the choice function defined in Algorithm 1 satisfy the conditions of substitutability, Law of aggregate demand, and irrelevance of rejected individuals. It follows from existing results (see, e.g., Aygün and Sönmez [5]) that GDA with choice function of schools that satisfy these three properties is strategyproof and results in a stable outcome. □

10 DISCUSSION

We consider a natural and general model of diversity goals in which the schools first want to fill in reserves for the first rank and then subsequent ranks. For this model, we designed a choice function that achieves diversity maximally. The choice function also serves as a useful decision-tool for committee or set selection problems (see e.g., [29]). In the latter problems, a set of candidates are selected after aggregating the preferences of the voters. If we have already aggregated the preferences to derive common (priority) ranking over the candidates, then our devised choice function can be used to additionally cater for diversity.

In our framework, a student can use a reserved seat if she satisfies the corresponding type of the seat. Our framework and all of our technical results immediately extend to scenarios where each reserved seat has a corresponding criterion based on types and a student can use a reserved seat if she satisfied the criterion. For example a criterion could indicate that 'any student of any type can use the seat.' This captures the seats that are under termed as 'open category seats.' Another criterion could be that the user 'satisfies type t_1 and t_2 but not t_3.' Note that it can be the case that one meta type is more constrained than another meta type. For example there could be some seats reserved for people with disabilities and separate seats reserved for people with a special disability.

Finally, we mention that the general approach for our choice function (Algorithm 1) can also be applied to other settings in which the priority list of applicants needs to be processed subject to various constraints. The approach deals with the applicants in decreasing order of priority and only adds an applicant if doing so will maintain the possibility of getting a feasible set of applicants.

ACKNOWLEDGMENTS

The authors thank Acelya Altuntas, Alex Lam, Barton Lee, Matthew Olckers, and Jeremy Vollens for their valuable comments. They also thank the anonymous reviewers of ACM EC 2021.

REFERENCES

[1] A. Abdulkadiroğlu. 2005. College admissions with affirmative action. *International Journal of Game Theory* 33, 4 (2005), 535–549.

[2] A. Abdulkadiroğlu and T. Sönmez. 2003. School Choice: A Mechanism Design Approach. *American Economic Review* 93, 3 (2003), 729–747.

[3] S. Ahmadi, F. Ahmed, J. P. Dickerson, M. Fuge, and S. Khuller. 2020. An Algorithm for Multi-Attribute Diverse Matching. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*. 3–9.

[4] O. Aygün and I. Bó. 2020. College admission with multidimensional privileges: The Brazilian affirmative action case. *American Economic Journal: Microeconomics* (2020).

[5] O. Aygün and T. Sönmez. 2013. Matching with contracts: Comment. *American Economic Review* 103, 5 (2013), 2050–51.

[6] O. Aygün and B. Turhan. 2016. Dynamic reserves in matching markets: Theory and applications. (2016).

[7] O. Aygun and B. Turhan. 2020. *Designing Direct Matching Mechanism for India with Comprehensive Affirmative Action.* Papers 2004.13264. arXiv.org.

[8] O. Aygün and B. Turhan. 2020. Dynamic reserves in matching markets: Theory and applications. *Journal of Economic Theory* 188 (2020).

[9] H. Aziz. 2019. A Rule for Committee Selection with Soft Diversity Constraints. *Group Decision and Negotiation* (2019), 1—8.
[10] H. Aziz, S. Gaspers, and Z. Sun. 2020. Mechanism Design for School Choice with Soft Diversity Constraints. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI).*
[11] H. Aziz, S. Gaspers, Z. Sun, and T. Walsh. 2019. From Matching with Diversity Constraints to Matching with Regional Quotas. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems.* International Foundation for Autonomous Agents and Multiagent Systems, 377–385.
[12] H. Aziz, S. Gaspers, Z. Sun, and M. Yokoo. 2020. Multiple Levels of Importance in Matching with Distributional Constraints. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS), Extended Abstract.* 1759–1761.
[13] S. Baswana, P. P. Chakrabarti, S. Chandran, Y. Kanoria, and U. Patange. 2019. Centralized Admissions for Engineering Colleges in India. In *Proceedings of the 20th ACM Conference on Economics and Computation.* 323–324.
[14] N. Benabbou, M. Chakraborty, X. Ho, J. Sliwinski, and Y. Zick. 2018. Diversity constraints in public housing allocation. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS).* 973–981.
[15] P. Biró, T. Fleiner, R. W. Irving, and D. F. Manlove. 2010. The College Admissions problem with lower and common quotas. *Theoretical Computer Science* 411, 34 (2010), 3136 – 3153. https://doi.org/10.1016/j.tcs.2010.05.005
[16] R. Bredereck, P. Faliszewski, A. Igarashi, M. Lackner, and P. Skowron. 2018. Multiwinner elections with diversity constraints. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI).* AAAI Press.
[17] J. R. Correa, R. Epstein, J. Escobar, I. Rios, B. Bahamondes, C. Bonet, N. Epstein, N. Aramayo, M. Castillo, A. Cristi, and B. Epstein. 2019. School Choice in Chile. In *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019.* 325–343. https://doi.org/10.1145/3328526.3329580
[18] J. P. Dickerson, K. A. Sankararaman, A. Srinivasan, and P. Xu. 2019. Balancing relevance and diversity in online bipartite matching via submodularity. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 1877–1884.
[19] U. Dur, P. A. Pathak, and T. Sönmez. 2020. Explicit vs. statistical targeting in affirmative action: Theory and evidence from Chicago's exam schools. *Journal of Economic Theory* 187, C (2020).
[20] F. Echenique and M. B. Yenmez. 2015. How to Control Controlled School Choice. *American Economic Review* 105, 8 (August 2015), 2679–94.
[21] L. Ehlers, I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim. 2014. School choice with controlled choice constraints: Hard bounds versus soft bounds. *Journal of Economic Theory* 153 (2014), 648–683.
[22] Y. A. Gonczarowski, N. Nisan, L. Kovalio, and A. Romm. 2019. Matching for the Israeli "Mechinot" Gap Year: Handling Rich Diversity Requirements. In *Proceedings of the 20th ACM Conference on Economics and Computation.* 321–321.
[23] I. E. Hafalir, M. B. Yenmez, and M.A. Yildirim. 2013. Effective affirmative action in school choice. *Theoretical Economics* 8, 2 (2013), 325–363.
[24] C. Huang. 2010. Classified Stable Matching. In *Proceedings of the Twenty-first Annual ACM-SIAM Symposium on Discrete Algorithms.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1235–1253. http://dl.acm.org/citation.cfm?id=1873601.1873700
[25] R. W. Irving, T. Kavitha, K. Mehlhorn, D. Michail, and K. E. Paluch. 2006. Rank-maximal matchings. *ACM Transactions on Algorithms (TALG)* 2, 4 (2006), 602–610.
[26] Y. Kamada and F. Kojima. 2015. Efficient matching under distributional constraints: Theory and applications. *The American Economic Review* 105, 1 (2015), 67–99.
[27] Y. Kamada and F. Kojima. 2017. Recent Developments in Matching with Constraints. *The American Economic Review* 107, 5 (2017), 200–204.
[28] A. S. Kelso and V. P. Crawford. 1982. Job matching, coalition formation, and gross substitutes. *Econometrica* 50 (1982), 1483—1504.
[29] D. M. Kilgour and E. Marshall. 2012. Approval balloting for fixed-size committees. In *Electoral Systems.* Springer, Chapter 12, 305–326.
[30] F. Kojima. 2012. School choice: Impossibilities for affirmative action. *Games and Economic Behavior* 75 (2012), 685–693.
[31] S. D. Kominers and T. Sönmez. 2013. Designing for diversity in matching. In *EC.* 603–604.
[32] S. D. Kominers and T. Sönmez. 2016. Matching with slot-specific priorities: Theory. *Theoretical Economics* 11, 2 (2016), 683–710.
[33] R. Kurata, N. Hamada, A. Iwasaki, and M. Yokoo. 2017. Controlled school choice with soft bounds and overlapping types. *Journal of Artificial Intelligence Research* 58 (2017), 153–184.
[34] D. F. Manlove. 2013. *Algorithmics of Matching Under Preferences.* World Scientific.
[35] T. Nguyen and R. Vohra. 2019. Stable Matching with Proportionality Constraints. *Operations Research* 67, 6 (2019), 1503–1519.
[36] P. A. Pathak, T. Sönmez, M. U. Ünver, and M. B. Yenmez. 2020. *Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Health Care Rationing.* Boston College Working Papers in Economics

1015. Boston College Department of Economics. https://ideas.repec.org/p/boc/bocoec/1015.html
[37] A. E. Roth and M. A. O. Sotomayor. 1992. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis.* Cambridge University Press.
[38] T. Sonmez and M. B. Yenmez. 2019. Affirmative Action in India via Vertical and Horizontal Reservations. (2019).
[39] T. Sönmez and M. B. Yenmez. 2020. Affirmative Action with Overlapping Reserves. (2020). http://fmwww.bc.edu/EC-P/wp990.pdf Manuscript.