Fairness and Efficiency Trade-off in Two-Sided Matching

Sung-Ho Cho
Kyushu University
Fukuoka, Japan
cho@agent.inf.kyushu-u.ac.jp

Kei Kimura
Kyushu University
Fukuoka, Japan
kkimura@inf.kyushu-u.ac.jp

Kiki Liu
Kyushu University
Fukuoka, Japan
kiki@agent.inf.kyushu-u.ac.jp

Kwei-guu Liu
Kyushu University
Fukuoka, Japan
liu@agent.inf.kyushu-u.ac.jp

Zhengjie Liu
Kyushu University
Fukuoka, Japan
zhengjie@agent.inf.kyushu-u.ac.jp

Zhaohong Sun
Kyushu University
Fukuoka, Japan
zhaohong.sun@inf.kyushu-u.ac.jp

Kentaro Yahiro
Kyushu University
Fukuoka, Japan
yahiro@agent.inf.kyushu-u.ac.jp

Makoto Yokoo
Kyushu University
Fukuoka, Japan
yokoo@inf.kyushu-u.ac.jp

ABSTRACT

The theory of two-sided matching has been extensively developed and applied to many real-life application domains. As the theory has been applied to increasingly diverse types of environments, researchers and practitioners have encountered various forms of distributional constraints. As a mechanism can handle a more general class of constraints, we can assign students more flexibly to colleges to increase students' welfare. However, it turns out that there exists a trade-off between students' welfare (efficiency) and fairness (which means no student has justified envy). Furthermore, this trade-off becomes sharper as the class of constraints becomes more general. The first contribution of this paper is to clarify the boundary on whether a strategyproof and fair mechanism can satisfy certain efficiency properties for each class of constraints. Our second contribution is to establish a weaker fairness requirement called *envy-freeness up to k peers* (EF-k), which is inspired by a similar concept used in the fair division of indivisible items. EF-k guarantees that each student has justified envy towards at most k students. By varying k, EF-k can represent different levels of fairness. We investigate theoretical properties associated with EF-k. Furthermore, we develop two contrasting strategyproof mechanisms that work for general hereditary constraints, i.e., one mechanism can guarantee a strong efficiency requirement, while the other can guarantee EF-k for any fixed k. We evaluate the performance of these mechanisms through computer simulation.

KEYWORDS

two-sided matching; strategyproof mechanism; mechanism design

ACM Reference Format:
Sung-Ho Cho, Kei Kimura, Kiki Liu, Kwei-guu Liu, Zhengjie Liu, Zhaohong Sun, Kentaro Yahiro, and Makoto Yokoo. 2024. Fairness and Efficiency Trade-off in Two-Sided Matching. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 9 pages.

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

1 INTRODUCTION

The theory of two-sided matching has been developed and has been applied to many real-life application domains (see Roth and Sotomayor [31] for a comprehensive survey in this literature). It has attracted considerable attention from AI researchers [2, 4, 16, 17, 20, 32, 33]. As the theory has been applied to increasingly diverse types of environments, researchers and practitioners have encountered various forms of distributional constraints (see Aziz et al. [3] for a comprehensive survey on various distributional constraints). There exist three representative classes of constraints. First, the standard model of two-sided matching considers only the maximum quota of each individual college [31], which we call maximum quotas constraints.[1]

More general classes of constraints are hereditary constraints [1, 14, 19] and hereditary M$^\natural$-convex set constraints [22]. An M$^\natural$-convex set is a discrete counterpart of a convex set in a continuous domain. Hereditary constraints require that if a matching between students and colleges is feasible, then any matching that places weakly fewer students at each college is also feasible.

As a mechanism can handle a more general class of constraints, we can incorporate more complex constraints required for real-life application domains. Also, we obtain more flexibility in assigning students to colleges. As a result, we can expect that students' welfare can be increased in the obtained matching. Furthermore, maximum quotas constraints can be considered to be too *restrictive*. In a real-life situation, it is common that some flexibility exists in determining the capacity of each college, i.e., we can increase the maximum quota of a college if it turns out to be very popular (say, by assigning additional resources). Such flexibility can be modeled naturally using a more general class of constraints.

In this paper, we focus our attention on strategyproof mechanisms, which guarantee that students have no incentive to misreport their preference over colleges. From a theoretical standpoint,

[1] Although our paper is described in the context of a student-college matching problem, the obtained result is applicable to matching problems in general.

if we are interested in a property achieved in dominant strategies, strategyproof mechanisms can be exclusively considered without any loss of generality, as supported by the well-known revelation principle [13]. This principle states that if a certain property is satisfied in a dominant strategy equilibrium using a mechanism, it can also be achieved through a strategyproof mechanism. Strategyproof mechanisms are not only theoretically significant but also practically beneficial, as students do not need to speculate about the actions of others to achieve desirable outcomes; they only need to report their preferences truthfully.

Most existing works in two-sided matching require that the obtained matching must be fair, i.e., no student has justified envy. However, just requiring fairness is not sufficient since the matching that no student is assigned to any college is fair; we should achieve some requirement on students' welfare (which is referred to as *efficiency* in economics) in conjunction with fairness. In the standard maximum quotas model, the renowned Deferred Acceptance mechanism (DA) [12] can achieve an efficiency property called *nonwastefulness* in conjunction with fairness. A matching satisfying fairness and nonwastefulness together is called *stable*.

However, when some distributional constraints are imposed, there exists a trade-off between fairness and efficiency/students' welfare. In particular, Cho et al. [8] show that no strategyproof mechanism satisfies fairness and a weaker efficiency property called weak nonwastefulness under hereditary constraints.

The first goal of this paper is to clarify the tight boundaries on whether a strategyproof and fair mechanism can satisfy certain efficiency properties for each class of constraints (see Table 1 in Section 4). In particular, we show that under hereditary constraints, no strategyproof mechanism can simultaneously satisfy fairness and a very weak efficiency requirement called *no vacant college property*.

This impossibility result illustrates a dilemma: we are expanding/generalizing the classes of constraints in the hope that we can improve students' welfare. However, if we require strict fairness, we cannot guarantee a very weak requirement of students' welfare under general hereditary constraints. Given this dilemma, our next goal is to establish a weaker fairness requirement. In this paper, we propose a novel concept called *envy-freeness up to k peers* (EF-k). This concept is inspired by a criterion called envy-freeness up to k items, which is commonly used in the fair division of indivisible items [6]. EF-k guarantees that each student has justified envy towards at most k students. By varying k, EF-k can represent different levels of fairness. On one hand, EF-0 is equivalent to standard fairness. On the other hand, any matching satisfies EF-$(n-1)$, where n is the number of students. To the best of our knowledge, this paper is the first to address the relaxed notion of fairness in two-sided, many-to-one matching.

We show that there exists a case that no matching is nonwasteful and EF-k for any $k < n-1$, and checking whether a nonwasteful and EF-k matching exists or not is NP-complete. Then, we develop two contrasting strategyproof mechanisms that work for general hereditary constraints. One is based on the Serial Dictatorship mechanism (SD) [14], which utilizes an optimal master-list (where students are assigned in its order) that minimize k based on colleges' preferences, such that the obtained matching is guaranteed to satisfy EF-k. Although $k = n-1$ holds in the worst case, we experimentally show that k tends to be much smaller when colleges' preferences are similar. The other one is based on the Sample and Deferred Acceptance mechanism (SDA) [24], which is developed for a special case of hereditary constraints called student-project-resource matching-allocation problem. This mechanism satisfies EF-k for any given $0 \le k < n-1$. We extend SDA such that the obtained matching satisfies no vacant college property under a mild assumption. We experimentally show that this mechanism can significantly improve students' welfare compared to a fair (EF-0) mechanism even when k is very small.

Due to the space limitation, some of the proofs are omitted, which can be found in the full version [7].

2 MODEL

A matching market under distributional constraints is given by $I = (S, C, X, \succ_S, \succ_C, f)$. The meaning of each element is as follows.

- $S = \{s_1, \dots, s_n\}$ is a finite set of students. Let N denote $\{1, 2, \dots, n\}$.
- $C = \{c_1, \dots, c_m\}$ is a finite set of colleges. Let M denote $\{1, 2, \dots, m\}$.
- $X \subseteq S \times C$ is a finite set of contracts. Contract $x = (s, c) \in X$ represents the matching between student s and college c.
- For any $Y \subseteq X$, let $Y_s := \{(s, c) \in Y \mid c \in C\}$ and $Y_c := \{(s, c) \in Y \mid s \in S\}$ denote the sets of contracts in Y that involve s and c, respectively.
- $\succ_S = (\succ_{s_1}, \dots, \succ_{s_n})$ is a profile of the students' preferences. For each student s, $\succ_s$ represents the preference of s over $X_s \cup \{(s, \emptyset)\}$, where $(s, \emptyset)$ represents an outcome such that s is unmatched. We assume $\succ_s$ is strict for each s. We say contract (s, c) is *acceptable* for s if $(s, c) \succ_s (s, \emptyset)$ holds. We sometimes use notations like $c \succ_s c'$ instead of $(s, c) \succ_s (s, c')$.
- $\succ_C = (\succ_{c_1}, \dots, \succ_{c_m})$ is a profile of the colleges' preferences. For each college c, $\succ_c$ represents the preference of c over $X_c \cup \{(\emptyset, c)\}$, where $(\emptyset, c)$ represents an outcome such that c is unmatched. We assume $\succ_c$ is strict for each c. We say contract (s, c) is *acceptable* for c if $(s, c) \succ_c (\emptyset, c)$ holds. We sometimes write $s \succ_c s'$ instead of $(s, c) \succ_c (s', c)$.
- $f : \mathbf{Z}_+^m \to \{-\infty, 0\}$ is a function that represents distributional constraints, where m is the number of colleges and $\mathbf{Z}_+^m$ is the set of vectors of m non-negative integers. For f, we call a family of vectors $F = \{\nu \in \mathbf{Z}_+^m \mid f(\nu) = 0\}$ *induced vectors* of f.

We assume each contract x in X_c is acceptable for c. This is without loss of generality because if some contract is unacceptable for a college, we can assume it is not included in X.

We say $Y \subseteq X$ is a *matching*, if for each $s \in S$, either (i) $Y_s = \{x\}$ and x is acceptable for s, or (ii) $Y_s = \emptyset$ holds.

For two m-element vectors $\nu, \nu' \in \mathbf{Z}_+^m$, we say $\nu \le \nu'$ if for all $i \in M$, $\nu_i \le \nu'_i$ holds. We say $\nu < \nu'$ if $\nu \le \nu'$ and $\nu \ne \nu'$ hold. Also, let $|\nu|$ denote the L_1 norm of ν, i.e., $|\nu| = \sum_{i \in M} \nu_i$.

Definition 2.1 (feasibility with distributional constraints). Let ν be a vector of m non-negative integers. We say ν is *feasible* in f if $f(\nu) = 0$. For $Y \subseteq X$, let us define $\nu(Y)$ as $(|Y_{c_1}|, |Y_{c_2}|, \dots, |Y_{c_m}|)$. We say Y is *feasible* (in f) if $\nu(Y)$ is feasible in f.

We assume F is bounded, i.e., $|F|$ is finite. This is without loss of generality because we can assume each college c_i can accept at most $|X_{c_i}|$ students, i.e., $f(v) = -\infty$ holds when $\exists i \in M, v_i > |X_{c_i}|$.

Let us first introduce a very general class of constraints called *hereditary* constraints. Intuitively, heredity means that if Y is feasible in f, then any subset $Y' \subset Y$ is also feasible in f. Let e_i denote an m-element unit vector, where its i-th element is 1 and all other elements are 0. Let e_0 denote an m-element zero vector $(0, \dots, 0)$.

Definition 2.2 (heredity). We say a family of m-element vectors $F \subseteq \mathbf{Z}_+^m$ is *hereditary* if $e_0 \in F$ and for all $v, v' \in \mathbf{Z}_+^m$, if $v > v'$ and $v \in F$, then $v' \in F$ holds. We say f is *hereditary* if its induced vectors are hereditary.

Kojima et al. [22] show that when f is hereditary, and its induced vectors satisfy one additional condition called *$M^\natural$-convexity*, there exists a general mechanism called Generalized Deferred Acceptance mechanism (GDA), which satisfies several desirable properties.[2]

Let us formally define an $\mathrm{M}^\natural$-convex set.

Definition 2.3 ($M^\natural$-convex set). We say a family of vectors $F \subseteq \mathbf{Z}_+^m$ forms an *$M^\natural$-convex set*, if for all $v, v' \in F$, for all i such that $v_i > v'_i$, there exists $j \in \{0\} \cup \{k \in M \mid v_k < v'_k\}$ such that $v - e_i + e_j \in F$ and $v' + e_i - e_j \in F$ hold. We say f satisfies $\mathrm{M}^\natural$-convexity if its induced vectors form an $\mathrm{M}^\natural$-convex set.

An $\mathrm{M}^\natural$-convex set can be considered as a discrete counterpart of a convex set in a continuous domain. Intuitively, Definition 2.3 means that for two feasible vectors v and v', there exists another feasible vector, which is one step closer starting from v toward v', and vice versa. An $\mathrm{M}^\natural$-convex set has been studied extensively in discrete convex analysis, a branch of discrete mathematics. Recent advances in discrete convex analysis have found many applications in economics (see the survey paper by Murota [27]). Note that heredity and $\mathrm{M}^\natural$-convexity are independent properties.

Kojima et al. [22] show that various real-life distributional constraints can be represented as a hereditary $\mathrm{M}^\natural$-convex set. The list of applications includes matching markets with regional maximum quotas [18], individual/regional minimum quotas [11, 14], diversity requirements in school choice [10, 23], distance constraints [22], and so on. However, $\mathrm{M}^\natural$-convexity can be easily violated by introducing some additional constraints.

Let us introduce the most basic model where only distributional constraints are colleges' maximum quotas.

Definition 2.4 (maximum quotas). We say a family of vectors $F \subseteq \mathbf{Z}_+^m$ is given as colleges' maximum quotas, when for each college $c_i \in C$, its maximum quota q_{c_i} is given, and $v \in F$ iff $\forall i \in M$, $v_i \leq q_{c_i}$ holds. We say f is given as colleges' maximum quotas if its induced vectors are given as colleges' maximum quotas.

If f is given as colleges' maximum quotas, then f is a hereditary $\mathrm{M}^\natural$-convex set, but not vice versa.

With a slight abuse of notation, for two sets of contracts Y and Y', we denote $Y_s \succ_s Y'_s$ if either (i) $Y_s = \{x\}$, $Y'_s = \{x'\}$, and $x \succ_s x'$ for some $x, x' \in X_s$, or (ii) $Y_s = \{x\}$ for some $x \in X_s$ that is acceptable for s and $Y'_s = \emptyset$. Furthermore, we denote $Y_s \succeq_s Y'_s$ if either $Y_s \succ_s Y'_s$ or $Y_s = Y'_s$. Also, we use notations like $x \succ_s Y_s$ or $c \succ_s Y_s$, where x is a contract, Y is a matching, and c is a college.

Let us introduce several desirable properties of a matching and a mechanism. We say a mechanism satisfies property A if the mechanism produces a matching that satisfies property A in every possible matching market.

First, we define fairness.

Definition 2.5 (fairness). In matching Y, student s *has justified envy* toward another student s' if $(s, c) \in X$ is acceptable for s, $c \succ_s Y_s$, $(s', c) \in Y$, and $s \succ_c s'$ hold. We say matching Y is *fair* if no student has justified envy.

Fairness implies that if student s is not assigned to college c (although she hopes to be assigned), then c prefers all students assigned to it over s.

Next, we define a series of properties on students' welfare (efficiency).

Definition 2.6 (Pareto efficiency). Matching Y is Pareto dominated by another matching Y' if $\forall s \in S$, $Y'_s \succeq_s Y_s$, and $\exists s \in S$, $Y'_s \succ_s Y_s$ hold. Feasible matching Y is Pareto efficient if no other feasible matching Pareto dominates it.

Definition 2.7 (nonwastefulness). In matching Y, student s *claims an empty seat* of college c if (s, c) is acceptable for s, $c \succ_s Y_s$, and $(Y \setminus Y_s) \cup \{(s, c)\}$ is feasible. We say feasible matching Y is *nonwasteful* if no student claims an empty seat.

Intuitively, nonwastefulness means that we cannot improve the matching of one student without affecting other students.

When additional distributional constraints (besides colleges' maximum quotas) are imposed, fairness and nonwastefulness become incompatible in general. One way to address the incompatibility is weakening the requirement of nonwastefulness. Aziz et al. [1] introduce a weaker efficiency concept called *cut-off nonwastefulness*.

Definition 2.8 (cut-off nonwastefulness). Feasible matching Y is *cut-off nonwasteful* if student s claims an empty seat of college c, then there exists another student s' such that $c \succ_{s'} Y_{s'}$, $s' \succ_c s$, and $(Y \setminus Y_{s'}) \cup \{(s', c)\}$ is infeasible.

Intuitively, we consider the claim of student s to move her to college c from her current match is not considered legitimate if by doing so, another student s' would have justified envy toward s. Aziz et al. [1] show that a fair and cut-off nonwasteful matching always exists under hereditary constraints. This result carries over to less general hereditary and $\mathrm{M}^\natural$-convex set constraints, as well as weaker efficiency requirements described below. Note that the existence of a fair and cut-off nonwasteful matching does not guarantee the existence of a strategyproof mechanism for obtaining it, as shown in Section 4.

Kamada and Kojima [19] propose another weaker version of the nonwastefulness concept, which we refer to as *weak nonwastefulness.*

Definition 2.9 (weak nonwastefulness). In matching Y, student s *strongly claims an empty seat* of c if (s, c) is acceptable for s, $c \succ_s Y_s$, and $Y \cup \{(s, c)\}$ is feasible. We say feasible matching Y is *weakly nonwasteful* if no student strongly claims an empty seat.

[2] To be more precise, Kojima et al. [22] show that to apply their framework, it is necessary that the family of feasible matchings forms a matroid. When distributional constraints are defined on $\nu(Y)$ rather than on contracts Y, the fact that the family of feasible contracts forms a matroid corresponds to the fact that (i) the family of feasible vectors forms an $\mathrm{M}^\natural$-convex set, and (ii) it is hereditary [28].

Let us define two more weaker efficiency properties.

Definition 2.10 (no vacant college). We say feasible matching Y satisfies *no vacant college* property if student s claims an empty seat of college c, then $Y_s \neq \emptyset$ or $Y_c \neq \emptyset$ holds.

Definition 2.11 (no empty matching). In matching Y, student s very strongly claims an empty seat of college c, when $Y = \emptyset$, $(s, c) \in X$, $c \succ_s \emptyset$, and $\{(s, c)\}$ is feasible. Feasible matching Y satisfies no empty matching property if no student very strongly claims an empty seat of any college.

Note that this series of efficiency properties becomes monotonically weaker in this order as long as distributional constraints are hereditary. More specifically, Pareto efficiency implies nonwastefulness, but not vice versa, nonwastefulness implies cut-off nonwastefulness, but not vice versa, and so on.

Next, we introduce strategyproofness.

Definition 2.12 (strategyproofness). We say a mechanism is *strategyproof* if no student ever has any incentive to misreport her preference no matter what the other students report. More specifically, let Y denote the matching obtained when s declares her true preference $\succ_s$, and Y' denote the matching obtained when s declare something else, then $Y_s \succeq_s Y'_s$ holds.

Here, we consider strategic manipulations only by students. It is well-known that even in the most basic model of one-to-one matching [12], satisfying strategyproofness (as well as basic fairness and efficiency requirements) for both sides is impossible [29]. One rationale for ignoring the college side would be that the preference of a college must be presented in an objective way and cannot be skewed arbitrarily.

3 EXISTING MECHANISM

In this section, we briefly introduce existing mechanisms, which are strategyproof for a given class of constraints. First, let us introduce Generalized Deferred Acceptance mechanism (GDA), which works under hereditary M$^\natural$-convex set constraints [15]. As its name shows, it is a generalized version of the Deferred Acceptance mechanism [12]. To define GDA, we first introduce *choice functions* of students and colleges.

Definition 3.1 (students' choice function). For each student s, her *choice function* Ch_s specifies her most preferred contract within each $Y \subseteq X$, i.e., $Ch_s(Y) = \{x\}$, where x is the most preferred acceptable contract in Y_s if one exists, and $Ch_s(Y) = \emptyset$ if no such contract exists. Then, the choice function of all students is defined as $Ch_S(Y) := \bigcup_{s \in S} Ch_s(Y_s)$.

Definition 3.2 (colleges' choice function). We assume each contract $(s, c) \in X$ is associated with its unique strictly positive weight $w((s, c))$. We assume these weights respect each college's preference $\succ_c$, i.e., if $(s, c) \succ_c (s', c)$, then $w((s, c)) > w((s', c))$ holds. For $Y \subseteq X$, let $w(Y)$ denote $\sum_{x \in Y} w(x)$. Then, the choice function of all colleges is defined as $Ch_C(Y) := \arg\max_{Y' \subseteq Y} f(v(Y')) + w(Y')$.

As long as f induces a hereditary M$^\natural$-convex set, a unique subset Y' exists that maximizes the above formula. Furthermore, such a subset can be efficiently computed in the following greedy way. Let Y' denote the set of chosen contracts, which is initially $\emptyset$. Then, sort Y in the decreasing order of their weights. Then, choose contract x from Y one by one and add it to Y', as long as $Y' \cup \{x\}$ is feasible.

Mechanism 1 Generalized Deferred Acceptance (GDA)

Input: X, Ch_S, Ch_C
Output: matching Y

1: $Re \leftarrow \emptyset$.
2: Each student s offers her most preferred contract (s, c) which has not been rejected before (i.e., $(s, c) \notin Re$). If no remaining contract is acceptable for s, s does not make any offer. Let Y be the set of contracts offered (i.e., $Y = Ch_S(X \setminus Re)$).
3: Colleges tentatively accept $Z = Ch_C(Y)$ and reject other contracts in Y (i.e., $Y \setminus Z$).
4: If all the contracts in Y are tentatively accepted at 3, then let Y be the final matching and terminate the mechanism. Otherwise, $Re \leftarrow Re \cup (Y \setminus Z)$, and go to 2.

Using Ch_S and Ch_C, GDA is defined as Mechanism 1. Note that we describe the mechanism using terms like "student s offers" to make the description more intuitive. In reality, GDA is a direct-revelation mechanism, where the mechanism first collects the preference of each student, and the mechanism chooses a contract on behalf of each student.

Kojima et al. [22] show that when f induces a hereditary M$^\natural$-convex set, GDA is strategyproof, the obtained matching Y satisfies a property called Hatfield-Milgrom stability (HM-stability), and Y is the student-optimal matching within all HM-stable matchings (i.e., all students weakly prefer Y over any other HM-stable matching).

Definition 3.3 (HM-stability). Matching Y is HM-stable if $Y = Ch_S(Y) = Ch_C(Y)$, and there exists no contract $x \in X \setminus Y$, such that $x \in Ch_S(Y \cup \{x\})$ and $x \in Ch_C(Y \cup \{x\})$ hold.

Intuitively, HM-stability means there exists no contract in $X \setminus Y$ that is mutually preferred by students and colleges. Note that HM-stability implies fairness. If student s has justified envy in matching Y, there exists $(s, c) \in X \setminus Y$, $(s', c) \in Y$, s.t. $(s, c) \succ_s Y_s$ and $w((s, c)) > w((s', c))$ holds. Then, $(s, c) \in Ch_S(Y \cup \{(s, c)\})$ and $(s, c) \in Ch_C(Y \cup \{(s, c)\})$ hold, i.e., Y is not HM-stable.

Next, we introduce two mechanisms that work for hereditary constraints. The Serial Dictatorship (SD) mechanism [14] is parameterized by an exogenous serial order over the students called a master-list. We denote the fact that s is placed in a higher/earlier position than student s' in master-list L as $s \succ_L s'$. Students are assigned sequentially according to the master-list. In our context with constraints, student s is assigned to her most preferred college c, where c considers her acceptable (i.e., $(s, c) \in X$ holds) and assigning s to c does not cause any constraint violation. More specifically, assume the obtained matching for students placed higher than s in L is Y. Then, s can be assigned to c when $f(v(Y \cup \{(s, c)\})) = 0$ holds. SD is strategyproof and achieves Pareto efficiency.

The Artificial Cap Deferred Acceptance mechanism (ACDA) is defined as follows. First, we choose one vector v^* s.t. $f(v^*) = 0$, and there exists no $v' > v^*$ where $f(v') = 0$, i.e., a maximal feasible vector. Note that v^* must be chosen independently from students' preferences $\succ_S$ to guarantee strategyproofness. Then, we apply standard DA, where maximum quota q_{c_i} for each college c_i is given as v^*_i. Intuitively, in ACDA, the set of feasible vectors F is artificially

Table 1: Existence of fair and strategyproof mechanism (✓ means such a mechanism exists, ✗ means such a mechanism does not exist, and ✖ means even without strategyproofness, a matching that satisfies fairness and the efficiency property may not exist. A red mark represents a new result obtained in this paper)

	maximum quotas	hereditary & M$^\natural$-convex set	hereditary
Pareto efficiency	✖ [29]	✖	✖
nonwastefulness	✓ [30]	✖ [19]	✖
cut-off nonwastefulness	✓	✗ [Thm 4.1]	✗
weak nonwastefulness	✓	✓ [21]	✗ [8]
no vacant college	✓	✓	✗ [Thm 4.2]
no empty matching	✓	✓	✓ [Thm 4.3]

reduced to a hyper-rectangle, where v is feasible iff $v \le v^*$. ACDA is strategyproof and fair, assuming v^* is chosen independently from students' preferences.

4 EXISTENCE OF FAIR AND STRATEGYPROOF MECHANISM

In this section, we examine whether a fair and strategyproof mechanism exists under a given class of distributional constraints in conjunction with some efficiency property. The classes of constraints we consider are: maximum quotas constraints, hereditary and M$^\natural$-convex set constraints, and hereditary constraints.

First, we list known results.

- For maximum quotas constraints, fairness, nonwastefulness, and strategyproofness are compatible, i.e., the standard DA satisfies these properties [30]. On the other hand, fairness and Pareto efficiency are incompatible, i.e., even without strategyproofness, a matching that satisfies Pareto efficiency and fairness may not exist [29].
- For hereditary and M$^\natural$-convex set constraints, fairness, weak nonwastefulness, and strategyproofness are compatible, i.e., Generalized DA satisfies these properties [21]. On the other hand, fairness and nonwastefulness are incompatible [19].
- For hereditary constraints, fairness, weak nonwastefulness, and strategyproofness are incompatible [8]

Given these known results, the remaining open questions are as follows.

(1) Under hereditary and M$^\natural$-convex set constraints, does a strategyproof, fair, and cut-off nonwasteful mechanism exist?
(2) Under hereditary constraints, can a strategyproof and fair mechanism satisfy any property weaker than weak nonwastefulness?

For question (1), we obtain a negative answer, as shown in Theorem 4.1. For question (2), we obtain a stronger result than Cho et al. [8], i.e., Theorem 4.2 shows that no mechanism simultaneously satisfies strategyproofness, fairness, and no vacant college property. Then, we show a simple mechanism that satisfies strategyproofness, fairness, and no empty matching property (Theorem 4.3). In summary, we obtain tight boundaries (at least in the granularity of efficiency properties we consider in this paper) on whether a strategyproof and fair mechanism can satisfy certain efficiency properties for each class of constraints (Table 1).

Theorem 4.1. *No mechanism can simultaneously satisfy fairness, strategyproofness, and cut-off nonwastefulness under hereditary M$^\natural$-convex set constraints.*

Theorem 4.2. *No mechanism can simultaneously satisfy fairness, strategyproofness, and no vacant college property under hereditary constraints.*

Next, we show that there exists a mechanism that satisfies fairness, strategyproofness, and no empty matching property under hereditary constraints. This mechanism utilizes GDA. More specifically, for given f, which is hereditary, we construct a set of vectors F' such that $\forall v \in F'$, $f(v) = 0$ holds (i.e., F' is a subset of vectors induced by f), and F' is a hereditary M$^\natural$-convex set. Then, we apply GDA by using f' (where $f'(v) = 0$ iff $v \in F'$) instead of f. F' is constructed as follows. We initialize $F' \leftarrow \{e_0\}$. Then, for each $i \in M$, if $f(e_i) = 0$, we add e_i to F'. Clearly, F' is an M$^\natural$-convex set; it contains only e_0 and e_i ($i \in M$).

Theorem 4.3. *Under hereditary constraints, GDA using f' is fair, strategyproof, and satisfies no empty matching property.*

Proof. For the obtained matching Y by GDA, $f'(v(Y)) = 0$ holds. Then, by way of constructing F', $f(v(Y)) = 0$ holds, i.e., Y is feasible. Since f' induces a hereditary M$^\natural$-convex set, GDA is strategyproof and fair [22]. Also, as long as there exists $(s, c) \in X$ such that $c \succ_s \emptyset$ and $f'(v(\{(s, c)\})) = 0$ hold, $Y \neq \emptyset$ holds. This is because if $Y = \emptyset$, then $(s, c) \in Ch_S(Y \cup \{(s, c)\})$ and $(s, c) \in Ch_C(Y \cup \{(s, c)\})$ hold, which violates the fact that GDA obtains an HM-stable matching. □

5 NEW FAIRNESS CONCEPT: ENVY-FREE UP TO k PEERS (EF-k)

In this section, we introduce a weaker fairness concept called envy-free up to k peers (EF-k). For matching Y and student s, let $Ev(Y, s)$ denote $\{s' \mid s' \in S, s$ has justified envy toward s' in $Y\}$.

Definition 5.1 (Envy-free up to k peers). Matching Y is envy-free up to k peers (EF-k) if $\forall s \in S$, $|Ev(Y, s)| \le k$ holds.

EF-0 is equivalent to fairness. Any matching is EF-$(n-1)$, where $n = |S|$.

There are other ways to relax fairness than EF-k. One straightforward way is to minimize the total number of justified envies. However, this criterion can be *unfair* among students, e.g., one student has many envies while others have only a few. Our definition of EF-k is more egalitarian; it minimizes the envies of the worst student. Other egalitarian criteria are also possible. For example, instead of counting the number of students to whom each student has envy, we can count the colleges at which each student has envy. Also, we can count the number of students by whom each student is envied. Which concept is socially acceptable is difficult to tell. This work is a first step that brings up new research directions in two-sided matching, i.e., how to relax the fairness concept in a socially acceptable way.

We use the following example to show that nonwastefulness and EF-k are incompatible for any $k < n-1$ under hereditary M$^\natural$-convex set constraints.

Example 5.2. There are n students and n colleges. For each student s_i, her preference is: $c_{i+1} \succ_{s_i} c_{i+2} \succ_{s_i} \ldots \succ_{s_i} c_n \succ_{s_i} c_1 \succ_{s_i} \ldots \succ_{s_i} c_i$. For each college c_i, its preference is: $s_i \succ_{c_i} s_{i+1} \succ_{c_i} \ldots \succ_{c_i} s_n \succ_{c_i} s_1 \succ_{c_i} \ldots \succ_{c_i} s_{i-1}$. In short, for each student s_i, her most preferred college c_{i+1} considers her as the least preferred student, and her least preferred college c_i considers her as the most preferred student. Distributional constraints f is defined as: $f(\nu) = 0$ iff $\forall i \in M$, $|\nu_i| \leq 1$ and $\sum_{i \in M} |\nu_i| \leq n-1$ hold, i.e., each college can accept at most one student, and the total number of students accepted to all colleges is at most $n-1$. Clearly, f induces a hereditary M$^\natural$-convex set.

Theorem 5.3. *Under hereditary M$^\natural$-convex set constraints, there exists a case that no matching is nonwasteful and EF-k for any $k < n-1$.*

Proof. Consider the setting in Example 5.2. The total number of accepted students is at most $n-1$. Also, due to nonwastefulness, exactly one student is unassigned to any college. By symmetry, without loss of generality, let us assume s_1 is unassigned. Then, there exists exactly one vacant college, i.e., a college to which no student is assigned. The vacant college must be c_2, since if c_i ($i \neq 2$) is vacant, student s_{i-1} claims an empty seat of c_i. Also, s_n must be assigned to c_1. Otherwise, she is assigned to c_i where $3 \leq i \leq n$; she claims an empty seat of c_2. Then, s_{n-1} must be assigned to c_n. Otherwise, she is assigned to c_i where $3 \leq i \leq n-1$; she claims an empty seat of c_2. By repeating a similar argument, we obtain that each student s_i ($i \neq 1$) is assigned to her most preferred college c_{i+1}. Then, s_1 has justified envy toward $s_2, \ldots, s_n$. Thus, $|Ev(Y, s_1)| = n-1$ holds. □

Given Theorem 5.3, a natural question is the complexity of checking the existence of a nonwasteful and EF-k matching (for $k < n-1$). Let us assume f can be computed in a constant time.

Theorem 5.4. *Checking whether an EF-k ($k < n-1$) and nonwasteful matching exists or not is NP-complete, even when distributional constraints form a hereditary M$^\natural$-convex set.*

6 NEW MECHANISMS

In this section, we introduce two contrasting strategyproof mechanisms that work for general hereditary constraints. The first one (called SD*) satisfies the strongest efficiency property, i.e., Pareto efficiency, while it cannot guarantee EF-k for any fixed $k < n-1$. The second one (called SDA with reserved quotas) satisfies EF-k for any fixed $k < n-1$, while it can only guarantee a rather weak efficiency property. In the next section, we experimentally show that SD* can guarantee EF-k where k is much smaller than $n-1$ when colleges' preferences are similar. Furthermore, we experimentally show that SDA with reserved quotas can significantly improve students' welfare compared to a fair (EF-0) mechanism even when k is very small.

6.1 Pareto efficient mechanism

First, we develop a strategyproof and Pareto efficient mechanism based on SD. For master-list L, a pair of students (s, s'), and college c, we say c disagrees with L for (s, s') if $s' \succ_L s$ and $s \succ_c s' \succ_c \emptyset$ holds. Otherwise, we say c agrees with L for (s, s'). In short, c disagrees with L for (s, s'), when s' is ranked higher than s in L, both s and s' are acceptable for college c, and c prefers s over s'. Assume we use SD based on L. Then, in obtained matching Y, if c disagrees with L for (s, s'), s has a chance to have justified envy toward s' in c, since s' is chosen before s and can be allocated to c, while s might not be allocated to c. On the other hand, if c agrees with L for (s, s'), then s never has justified envy toward s' in c. This is because, the fact that c agrees with L for (s, s') means: (i) s is ranked higher than s' in L, (ii) s is ranked lower than s' in c, or (iii) either s or s' is unacceptable for c. In each of the above three cases, s cannot have justified envy toward s' in c.

Let $d(L, s)$ denote $|\{s' \mid s' \in S \setminus \{s\}, c \in C, c$ disagrees with L for $(s, s')\}|$, i.e., $d(L, s)$ counts the number of students such that for some college c, a disagreement related to s occurs.

The following theorem holds.

Theorem 6.1. *Assume for master-list L, $\forall s \in S$, $d(L, s) \leq k$ holds. Then, SD using L is EF-k.*

Theorem 6.1 means that if we can choose a good master-list L, such that $\max_{s \in S} d(L, s)$ is small, e.g. at most k, the obtained matching is guaranteed to be EF-k. Note that this guarantee holds independently from the actual distributional constraints and students' preferences; k can be computed using colleges' preference profile $\succ_C$ only. Thus, for given students' preference $\succ_S$, the obtained matching can be EF-k' for k' that is much smaller than k guaranteed by Theorem 6.1; see the experimental results that clarify this in the next section.

Let us examine the problem of finding an optimal master-list (in terms of minimizing $\max_{s \in S} d(L, s)$) for given colleges' preference profile $\succ_C$.

Theorem 6.2. *For given colleges' preference profile $\succ_C$, computing master-list L, which minimizes $\max_{s \in S} d(L, s)$ can be done in polynomial time.*

Let us call SD mechanism using optimal L as SD*. SD* is strategyproof and Pareto efficient. When we apply SD* to the matching instance presented in Example 5.2, the above algorithm returns L with $\max_{s \in S} d(L, s) = n-1$ and the obtained matching cannot be EF-k for any $k < n-1$. In the next section, we show that SD* can be EF-k for smaller k when colleges' preferences are similar.

Let us examine situations where SD* can be used in practice. Assume there exists an authority who decides a matching based on colleges'/students' preferences. The authority is allowed to override colleges' preferences to some extent in order to improve students' welfare. More specifically, the authority can use its own ordering among students to decide the matching, where the ordering is chosen such that it is as close as possible to colleges' preferences. Our SD* is based on this idea, which uses ordering L that minimizes $k = \max_{s \in S} d(L, s)$. The obtained matching is guaranteed to be EF-k. There can be alternative minimization criteria for choosing L, e.g., minimizing the sum of Kendall tau distances (the number of pairwise disagreements). However, this optimization problem is computationally hard [5] and can be *unfair* among students.

6.2 EF-k mechanism

Next, we develop a strategyproof and EF-k mechanism for any given $k \leq n-1$. First, let us define the standard Sample and Deferred Acceptance (SDA) mechanism. This mechanism is developed by Liu et al. [24] for a special case for hereditary constraints where the maximum quota of each college is determined by allocating indivisible resources to each college. The basic idea of SDA is to combine SD and ACDA. One major limitation of ACDA is that we need to determine the maximal feasible vector ν^* (which determines the maximum quota of each college) independently from students' preferences. As a result, the maximum quotas of popular colleges can be low, while those of unpopular colleges can be high. In the standard SDA, first, we choose a subset of students $S' \subseteq S$, where $|S'| = k$. We call S' *sampled* students, and $S \setminus S'$ *regular* students. We assign sampled students using SD. Assume the obtained matching for sampled students is Y'. Then, we choose a maximal feasible vector ν^* based on the preferences of sampled students. Liu et al. [24] present several alternative ways to choose ν^*. In this paper, as described later, we apply a simulation-based method using copies of sampled students, which is shown to be most effective in [24]. Then, we apply ACDA for regular students, where maximum quota q_{c_i} for each college c_i is given as $\nu^*_i - |Y'_{c_i}|$.

The standard SDA is strategyproof. It is also EF-k, since for each sampled student s, she has justified envy only toward another sampled student assigned before her. Thus, $|Ev(Y,s)| \leq k-1$ holds. Also, since DA is fair, for regular student s, she has justified envy only toward sampled students. Thus, $|Ev(Y,s)| \leq k$ holds.

However, if the preferences of sampled students are completely different from the preferences of regular students, obtained ν^* can be bad for regular students. As a result, even no vacant college property is not satisfied. We can assume SDA satisfies no empty matching property. No empty matching property is violated only in an exceptional case where all sampled students assume all colleges unacceptable. In such a case, we can choose additional sampled students until at least one student is assigned to some college.

We propose a generalized version of SDA, such that no vacant college property is satisfied under a mild assumption. The basic idea is that, since there exists a chance that the preferences of sampled students are completely different from those of regular students, we reserve some seats for each college even if the college seems unpopular based on the preferences of sampled students. Let $\widehat{\nu} = (\widehat{\nu}_1, \ldots, \widehat{\nu}_m)$ be *reserved quotas*, where $\widehat{\nu}_i \geq 0$ for each $i \in M$, and $f(\widehat{\nu}) = 0$ holds. The goal of the reserved quotas $\widehat{\nu}$ is to guarantee that each college c_i is guaranteed to accept at least $\widehat{\nu}_i$ students, as long as enough students hope to be assigned to c_i, even if c_i seems unpopular among sampled students.

For two m-element vectors ν and ν', let $\nu \vee \nu'$ denote the element-wise maximum, i.e., $\nu \vee \nu' = (\max(\nu_1, \nu'_1), \ldots, \max(\nu_m, \nu'_m))$.

First, let us define SD with reserved quotas $\widehat{\nu}$. As standard SD, we assign students sequentially based on master-list L. Let Y denote the assignment obtained so far. The current student can be assigned to c_i, as long as $f((\nu(Y) + e_i) \vee \widehat{\nu}) = 0$ holds. In short, the current student s can be assigned to c_i, if c_i can still accept one more student, assuming each college c_j will be assigned at least $\widehat{\nu}_j$ students.

Then, SDA with reserved quotas $\widehat{\nu}$ is defined as follows. Choose k sampled students (the remaining students are regular students). They are assigned by SD with reserved quotas $\widehat{\nu}$. Let Y' denote the matching for sampled students. Then, obtain a matching Y'', by further assigning multiple virtual students, each of which is a copy of sampled students by SD with reserved quotas, until no more student can be assigned. More specifically, let us assume sampled students are $s_1, \ldots, s_k$. We create virtual students $s_{i,1}, s_{i,2}, \ldots$, which are copies of each sampled student s_i. Then, after sampled students are assigned. We assign these virtual students in a round-robin order, i.e., $s_{1,1}, s_{2,1}, \ldots, s_{k,1}, s_{1,2}, s_{2,2}, \ldots, s_{k,2}, s_{1,3}, s_{2,3}, \ldots, s_{k,3}, \ldots$. Note that this procedure is just for choosing appropriate ν^*; in reality, these virtual students are not allocated to any college. Then, we choose maximal feasible vector ν^* such that $\nu^* \geq \nu(Y'') \vee \widehat{\nu}$ holds. For each college c_i, we set its maximum quota q_{c_i} as $\nu^*_i - |Y'_{c_i}|$, and run ACDA for regular students.

Theorem 6.3. *Assume for $\widehat{\nu}$, $f(\widehat{\nu}) = 0$ holds, and $\forall i \in M$, such that $f(e_i) = 0$ holds, $\widehat{\nu}_i \geq 1$ also holds. Then, SDA with reserved quotas $\widehat{\nu}$ and k-sampled students is strategyproof, EF-k, and satisfies no vacant college property.*

Let us examine situations where SDA can be used in practice. Assume there exist k *distinguished* students, e.g., they have excellent achievements in sports / volunteer works, etc., they are from financially difficult families / minority groups, or even chosen by lottery. If giving them priority in college administration is socially acceptable, we can use these *distinguished* students as sampled students in SDA. Then, the outcome is guaranteed to be EF-k.

7 EXPERIMENTAL EVALUATION

First, we show the level of k that SD* can be guaranteed by using an optimal master-list. We set the number of students n to 200 and the number of colleges m to 20. We generate the preference of each college c using the Mallows model [9, 25, 26]; college preference $\succ_c$ is drawn with probability: $\Pr(\succ_c) = \frac{\exp(-\phi_C \cdot \delta(\succ_c, \succ_{\widehat{c}}))}{\sum_{\succ'_c} \exp(-\phi_c \cdot \delta(\succ'_c, \succ_{\widehat{c}}))}$. Here $\phi_C \in \mathbf{R}_+$ denotes the spread parameter for colleges, $\succ_{\widehat{c}}$ is a central preference uniformly randomly chosen from all possible preferences, and $\delta(\succ_c, \succ_{\widehat{c}})$ represents the Kendall tau distance, which is the number of pairwise inversions between $\succ_c$ and $\succ_{\widehat{c}}$. Intuitively, colleges' preferences are distributed around a central preference with spread parameter ϕ_C. When $\phi_C = 0$, the Mallows model becomes identical to the uniform distribution, while increasing ϕ_C leads to convergence towards a constant distribution, yielding $\succ_{\widehat{c}}$. Initially, each $\succ_c$ does not include $\emptyset$. We insert $\emptyset$ at the position $\lfloor \rho \cdot n \rfloor$ (where $0 < \rho < 1$).

Figure 1 shows the guaranteed k when using an optimal master-list by varying the spread parameter ϕ_C and ρ. Each data point is an average of 10 instances. We also show the result when the master-list is randomly chosen. We can see that when the spread parameter becomes larger (colleges' preferences become more similar), SD* can guarantee EF-k for smaller k. For example, k becomes less than 5% of n when ϕ_C is 0.6. We can see ρ has almost no effect on SD*, while it significantly affects randomly selected master-lists.

Next, we apply SD* to each matching market and measure the obtained level of k that SD* achieves. We consider the following distributional constraints [24]. There exists a set of indivisible resources

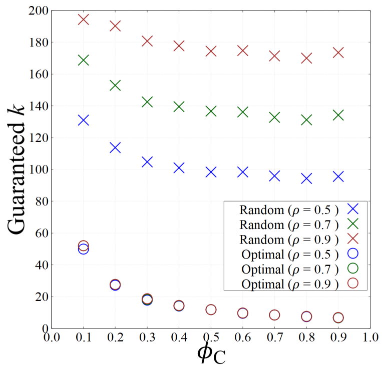

Figure 1: Guaranteed k for optimal/random master-list

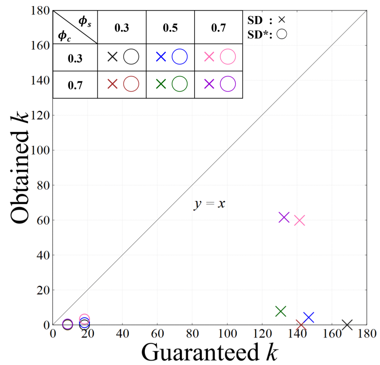

Figure 2: Comparison between obtained/guaranteed k for SD*/SD

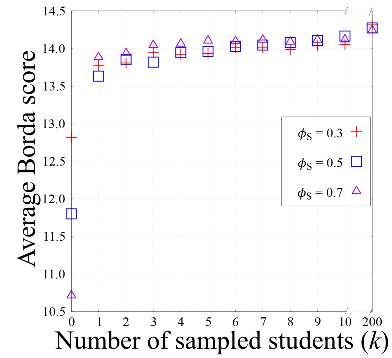

Figure 3: Average Borda score for SDA with reserved quotas

$R = \{r_1, \ldots, r_{|R|}\}$. Each resource r has its capacity $q_r \in \mathbb{N}_{>0}$. For each resource r, its college compatibility list T_r is defined; resource r can be allocated to exactly one college in $T_r \subseteq C$. Mapping μ denotes one possible allocation of resources to colleges, i.e., $\mu : R \to C$ maps each resource r to a college $\mu(r) \in T_r$. For given allocation μ, the maximum quota of college c is given as $q_\mu(c) = \sum_{r:\mu(r)=c} q_r$, i.e., the maximum quota of each college is endogenously determined as the sum of the capacities of allocated resources. We assume $f(v) = 0$ if there exists μ s.t. $v_i \leq q_\mu(c_i)$ holds for all $i \in M$. Each market has $|R| = 100$ resources. For each resource r, we generate T_r such that each college c is included in T_r with probability 0.3. There are 40, 20, and 40 resources with capacity 1, 2, and 3, respectively; thus the total capacity of colleges is equal to n. We generate each student's preference in a similar way as a college's preference, i.e., we utilize the Mallows model with spread parameter ϕ_S. We do not apply ρ for students; each student considers all colleges acceptable.

Figure 2 shows the average of 10 instances. The x-axis shows the guaranteed k and the y-axis shows the actually obtained k. We set colleges' spread parameter ϕ_C to 0.3 and 0.7, and students' spread parameter ϕ_S to 0.3, 0.5, and 0.7. ρ is set to 0.7. By definition, each data point must be located in the lower-right half. The result shows the actually obtained k is much smaller than the guaranteed k. In particular, for SD*, it is between 0 and 4. For SD, we can see that when ϕ_S becomes larger, the competition among students becomes more intense. As a result, more students tend to have justified envy.

Next, we evaluate SDA with reserved quotas. By varying k, it can be identical to ACDA (when $k = 0$) and SD* (when $k = n$), assuming we use the same master-list as SD* and the same reserved quotas. Figure 3 shows the average Borda score of the students varying k and the students' spread parameter ϕ_S. If a student is assigned to her i-th choice college, her Borda score is $m - i + 1$. We fix the colleges' spread parameter ϕ_C to 0.7 and ρ to 0.7. We set reserved quotas $\widehat{v}$ to $(1, 1, \ldots, 1)$. Each data point represents an average of 10 instances. In this setting, SDA with n sampled students (which is identical to SD*) guarantees EF-k for $k = 9$ in average. The average Borda score significantly improves as k increases from the case where $k = 0$. Note that increasing the average Borda score by one is significant; each student must be assigned to a strictly better college. The difference between $k = 0$ (where SDA is identical to ACDA) and $k = 1$ becomes larger when ϕ_S becomes larger, i.e., when students' preferences are similar. We can see that SDA achieves a high degree of fairness and efficiency with a few sampled students.

In summary, SD* is much fairer than SD with a randomly selected master-list, and can attain EF-k' for k' that is much smaller than k guaranteed by Theorem 6.1. Also, SDA with reserved quotas is much more efficient than ACDA, and attains very good fairness at the expense of a little efficiency compared to SD*.

8 CONCLUSIONS AND FUTURE WORKS

When distributional constraints are imposed in two-sided matching, there exists a trade-off between fairness and efficiency. We clarified the tight boundaries on whether a strategyproof and fair mechanism can satisfy certain efficiency properties for each class of constraints. We also established a new fairness requirement called EF-k. We examined theoretical properties related to EF-k, and developed two contrasting strategyproof mechanisms that work for general hereditary constraints. We evaluated the performance of these mechanisms via computer simulation. We believe EF-k is significant since it brings up many new research topics in constrained matching literature; there remain many open questions related to EF-k. For example, can any strategyproof mechanism guarantee EF-k for some fixed k in conjunction with some efficiency property (which is stronger than no vacant college property, e.g., weak nonwastefulness)? Furthermore, there exists another mechanism called Adaptive DA [14] that works for any hereditary constraints. Comparing this mechanism with our newly proposed mechanisms is our immediate future work.

ACKNOWLEDGMENTS

We would like to thank anonymous reviewers for their valuable comments. This work was partially supported by JST ERATO Grant Number JPMJER2301, and JSPS KAKENHI Grant Numbers JP21H04979 and JP20H00609, Japan.

REFERENCES

[1] Haris Aziz, Anton Baychkov, and Peter Biró. 2022. Cutoff stability under distributional constraints with an application to summer internship matching. *Mathematical Programming* (2022). https://doi.org/10.1007/s10107-022-01917-1

[2] Haris Aziz, Péter Biró, Tamás Fleiner, Serge Gaspers, Ronald de Haan, Nicholas Mattei, and Baharak Rastegari. 2022. Stable matching with uncertain pairwise preferences. *Theoretical Computer Science* 909 (2022), 1–11. https://doi.org/10.1016/j.tcs.2022.01.028

[3] Haris Aziz, Peter Biró, and Makoto Yokoo. 2022. Matching Market Design with Constraints. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI-2022)*. 12308–12316. https://doi.org/10.1609/aaai.v36i11.21495

[4] Haris Aziz, Serge Gaspers, Zhaohong Sun, and Toby 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 (AAMAS-2019)*. 377–385.

[5] J. Bartholdi, C. A. Tovey, and M.A. Trick. 1989. Voting Schemes for which It Can Be Difficult to Tell Who Won the Election. *Social Choice and Welfare* 6 (1989), 157–165. https://doi.org/10.1007/BF00303169

[6] Eric Budish. 2011. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy* 119, 6 (2011), 1061–1103. https://doi.org/10.1086/664613

[7] Sung-Ho Cho, Kei Kimura, Kiki Liu, Kwei-guu Liu, Zhengjie Liu, Zhaohong Sun, Kentaro Yahiro, and Makoto Yokoo. 2024. Fairness and efficiency trade-off in two-sided matching. *arXiv: 2402.01084* (2024).

[8] Sung-Ho Cho, Miyuki Koshimura, Pinaki Mandal, Kentaro Yahiro, and Makoto Yokoo. 2022. Impossibility of weakly stable and strategy-proof mechanism. *Economics Letters* 217 (2022), 110675. https://doi.org/10.1016/j.econlet.2022.110675

[9] Joanna Drummond and Craig Boutilier. 2013. Elicitation and Approximately Stable Matching with Partial Preferences. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI-2013)*. 97–105.

[10] Lars Ehlers, Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. 2014. School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds. *Journal of Economic Theory* 153 (2014), 648–683. https://doi.org/10.1016/j.jet.2014.03.004

[11] Daniel Fragiadakis, Atsushi Iwasaki, Peter Troyan, Suguru Ueda, and Makoto Yokoo. 2015. Strategyproof Matching with Minimum Quotas. *ACM Transactions on Economics and Computation* 4, 1 (2015), 6:1–6:40. https://doi.org/10.1145/2841226

[12] David Gale and Lloyd Stowell Shapley. 1962. College Admissions and the Stability of Marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15. https://doi.org/10.2307/2312726

[13] Allan Gibbard. 1973. Manipulation of Voting Schemes: A General Result. *Econometrica* 41, 4 (1973), 587–601. https://doi.org/10.2307/1914083

[14] Masahiro Goto, Fuhito Kojima, Ryoji Kurata, Akihisa Tamura, and Makoto Yokoo. 2017. Designing Matching Mechanisms under General Distributional Constraints. *American Economic Journal: Microeconomics* 9, 2 (2017), 226–262. https://doi.org/10.1257/mic.20160124

[15] John William Hatfield and Paul R. Milgrom. 2005. Matching with Contracts. *American Economic Review* 95, 4 (2005), 913–935. https://doi.org/10.1257/0002828054825466

[16] Hadi Hosseini, Kate Larson, and Robin Cohen. 2015. On Manipulablity of Random Serial Dictatorship in Sequential Matching with Dynamic Preferences. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI-2015)*. 4168–4169. https://doi.org/10.1609/aaai.v29i1.9744

[17] Anisse Ismaili, Naoto Hamada, Yuzhe Zhang, Takamasa Suzuki, and Makoto Yokoo. 2019. Weighted Matching Markets with Budget Constraints. *Journal of Artificial Intelligence Research* 65 (2019), 393–421. https://doi.org/10.1613/jair.1.11582

[18] Yuichiro Kamada and Fuhito Kojima. 2015. Efficient Matching under Distributional Constraints: Theory and Applications. *American Economic Review* 105, 1 (2015), 67–99. https://doi.org/10.1257/aer.20101552

[19] Yuichiro Kamada and Fuhito Kojima. 2017. Stability Concepts in Matching under Distributional Constraints. *Journal of Economic Theory* 168 (2017), 107–142. https://doi.org/10.1016/j.jet.2016.12.006

[20] Yasushi Kawase and Atsushi Iwasaki. 2017. Near-Feasible Stable Matchings with Budget Constraints. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-2017)*. 242–248. https://doi.org/10.24963/ijcai.2017/35

[21] Kei Kimura, Kwei-guu Liu, Zhaohong Sun, Kentaro Yahiro, and Makoto Yokoo. 2023. Multi-Stage Generalized Deferred Acceptance Mechanism: Strategyproof Mechanism for Handling General Hereditary Constraints. *arXiv: 2309.10968* (2023).

[22] Fuhito Kojima, Akihisa Tamura, and Makoto Yokoo. 2018. Designing matching mechanisms under constraints: An approach from discrete convex analysis. *Journal of Economic Theory* 176 (2018), 803–833. https://doi.org/10.1016/j.jet.2018.05.004

[23] Ryoji Kurata, Naoto Hamada, Atsushi Iwasaki, and Makoto Yokoo. 2017. Controlled School Choice with Soft Bounds and Overlapping Types. *Journal of Artificial Intelligence Research* 58 (2017), 153–184. https://doi.org/10.1613/jair.5297

[24] Kwei-guu Liu, Kentaro Yahiro, and Makoto Yokoo. 2023. Strategyproof Mechanism for Two-Sided Matching with Resource Allocation. *Artificial Intelligence* 316 (2023), 103855. https://doi.org/10.1016/j.artint.2023.103855

[25] Tyler Lu and Craig Boutilier. 2014. Effective Sampling and Learning for Mallows Models with Pairwise-Preference Data. *Journal of Machine Learning Research* 15 (2014), 3963–4009.

[26] Colin L Mallows. 1957. Non-null ranking models. I. *Biometrika* 44, 1-2 (1957), 114–130. https://doi.org/10.2307/2333244

[27] Kazuo Murota. 2016. Discrete Convex Analysis: A Tool for Economics and Game Theory. *Journal of Mechanism and Institution Design* 1 (2016), 151–273. https://doi.org/10.22574/jmid.2016.12.005

[28] Kazuo Murota and Akiyoshi Shioura. 1999. M-convex Function on Generalized Polymatroid. *Mathematics of Operations Research* 24, 1 (1999), 95–105. https://doi.org/10.1287/moor.24.1.95

[29] Alvin E. Roth. 1982. The Economics of Matching: Stability and Incentives. *Mathematics of Operations Research* 7, 4 (1982), 617–628. https://doi.org/10.1287/moor.7.4.617

[30] Alvin E. Roth. 1985. The College Admissions Problem is not Equivalent to the Marriage Problem. *Journal of Economic Theory* 36, 2 (1985), 277–288. https://doi.org/10.1016/0022-0531(85)90106-1

[31] Alvin E. Roth and Marilda A. Oliveira Sotomayor. 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis (Econometric Society Monographs)*. Cambridge University Press. https://doi.org/10.1017/CCOL052139015X

[32] Takamasa Suzuki, Akihisa Tamura, Kentaro Yahiro, Makoto Yokoo, and Yuzhe Zhang. 2023. Strategyproof Allocation Mechanisms with Endowments and M-convex Distributional Constraints. *Artificial Intelligence* 315 (2023), 103825. https://doi.org/10.1016/j.artint.2022.103825

[33] Kentaro Yahiro, Yuzhe Zhang, Nathanaël Barrot, and Makoto Yokoo. 2020. Strategyproof and Fair Matching Mechanism for Ratio Constraints. *Autonomous Agents and Multi-Agent Systems* 34 (2020), 1–29. https://doi.org/10.1007/s10458-020-09448-9