

# Coalitional Permutation Manipulations in the Gale-Shapley Algorithm\*

Yuan Deng  
Duke University  
ericdy@cs.duke.edu

Weiran Shen  
IIIS, Tsinghua University  
emersonswr@gmail.com

Pingzhong Tang  
IIIS, Tsinghua University  
kenshinping@gmail.com

## ABSTRACT

In this paper, we consider permutation manipulations by any subset of women in the Gale-Shapley algorithm. This paper is motivated by the college admissions process in China. Our results also answer an open problem on what can be achieved by permutation manipulations. We present an efficient algorithm to find a strategy profile such that the induced matching is stable and Pareto-optimal while the strategy profile itself is inconspicuous. Surprisingly, we show that such a strategy profile actually forms a Nash equilibrium of the manipulation game.

In the end, we show that it is NP-complete to find a manipulation that is strictly better for all members of the coalition. This result demonstrates a sharp contrast between weakly better-off outcomes and strictly better-off outcomes.

## KEYWORDS

Coalition Manipulation; Permutation Manipulation; Gale-Shapley Algorithm

### ACM Reference Format:

Yuan Deng, Weiran Shen, and Pingzhong Tang. 2018. Coalitional Permutation Manipulations in the Gale-Shapley Algorithm. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018*, IFAAMAS, 9 pages.

## 1 INTRODUCTION

Stability has been a central concept in economic design, ever since the seminal work by Gale and Shapley [9]. Intensive research has been done over the years. A variety of applications of this problem have also been developed, ranging from college admissions and school choice [2, 3, 9] to centralized kidney exchange programs [4, 24, 30, 31] to hospitals-residents matchings [19, 20, 29] to recently proposed water right trading [25, 35].

In the standard stable matching model, there is a set of men and a set of women. Each agent has a preference list over a subset of the opposite sex. A matching between men and women is stable if no pair of agents prefer to match with each other than their designated partners. Gale and Shapley [9] put forward an algorithm, aka. the Gale-Shapley algorithm, that computes a stable matching in  $O(n^2)$  time. The algorithm (men-proposing version) proceeds in multiple rounds. At each round, each man proposes to his favorite woman that has not rejected him yet; and each woman keeps her favorite

proposal, if any, and rejects all others. The algorithm iterates until no further proposal can be made.

The algorithm enjoys many desirable properties. It is well-known that the matching returned by the algorithm is preferred by every man to any other stable matching, hence called the M-optimal (for men-optimal) matching. It is also known that all stable matchings form a lattice defined by such a preference relation and the M-optimal matching is the greatest element in the lattice [21]. Furthermore, men and women have strictly opposite preferences over two stable matchings: every man prefers stable matching  $\mu_1$  to stable matching  $\mu_2$  if and only if every woman prefers  $\mu_2$  to  $\mu_1$ . As a result, the M-optimal matching is the W-pessimal (for women-pessimal) matching [26]. The smallest element in the lattice, the W-optimal (M-pessimal) matching, can be obtained by swapping the roles of men and women.

## 1.1 Motivations

This work is motivated by the college admission process in China [7], where the stable matching model is adopted. The admissions process consists of two phases: the examination phase and the application phase. In the examination phase, all students are required to take the National College Entrance Examination (NCEE, aka. the National Higher Education Entrance Examination), which is held nation-wide annually. Millions of students take the NCEE every year, and the number peaked at 10.5 millions in the year of 2008. The NCEE contains a series of exams on different subjects. After the examination, each student receives a total score which is the sum of the scores of the subjects. The total score uniquely determines an ordering of all students, which is also the preference ordering adopted by all colleges and universities. In the application phase, each student who takes the NCEE is required to submit an ordered list of about 4 to 6 intended colleges or universities. In the end, the Ministry of Education settles the applications using the student-proposing version of the Gale-Shapley algorithm.

However, a major concern of the Gale-Shapley algorithm is its non-truthfulness. While it is known that the algorithm is group strategy-proof<sup>1</sup> for all men [8], it is not truthful for women. In fact, Roth [28] shows that there is no stable matching algorithm that is strategy-proof for all agents.

Such an undesirable property gives rise to the so-called “manipulation” problem for the women. In China’s college admissions process, besides the NCEE, some top universities are also allowed to conduct *independent recruitment* exams. These universities promise

\*This paper is supported in part by the National Natural Science Foundation of China Grant 61561146398, a China Youth 1000-talent Program and an Alibaba Innovative Research Program.

*Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018)*, M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden. © 2018 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

<sup>1</sup>Precisely, *group strategy-proof* means no coalition manipulation can make all men in the coalition strictly better off, in this context. If considering the case where no man is worse off and at least one man is strictly better off, the Gale-Shapley algorithm is not group strategy-proof [15].

to the students who perform well in these exams that, when applying to these universities, a certain amount of extra scores will be added to their NCEE total scores. In other words, such independent recruitment exams give the universities the ability to manipulate the admissions result by changing the ordering of the students in their preference ordering.

Starting from 2010, several leagues of such universities emerged, with the two most influential leagues represented by China’s two major universities, the Tsinghua University (the Tsinghua league) and the Peking University (the Peking league). Each league contains universities of similar types and tiers. Thus universities of the same league attract about the same set of students, and they conduct the independent recruitment programs together [1]. The benefits of such leagues are obvious: (1) the costs of organizing such exams are greatly reduced since they are shared by the universities; (2) the students only need to participate in one such exam instead of many. However, such leagues are widely conjectured to be beneficial to universities inside the leagues when it comes to the quality of finally admitted students, since they can cooperatively manipulate the admissions result to benefit them all. Besides cooperations, the universities in the same league are also faced with the problem of competition because they share a similar set of candidate students. In 2012, two top universities (Fudan University and Nankai University) quit the Peking league, both claiming that they were not able to recruit their desired students. Such leagues were urged to dissolve in 2015 by the Ministry of Education for the belief that it is unfair for universities that are not in any of the leagues.

## 1.2 Results

We study the problem where a coalition of women (universities) can manipulate the Gale-Shapley algorithm. Most existing works consider the general case where women can report any preference list (potentially incomplete) without ties. In contrast, we focus on the setting where all women must report a complete list, which indicates that women can only permute their true preferences. This type of manipulation comes directly from the independent recruitment programs in China, where the universities can only permute the ordering of the students by adding scores to some of them, but are not allowed to remove any student from the lists.

We model the coalition manipulation problem as a game among the members of the coalition (called the manipulation game hereafter). We first show that a coalition of women could get worse off if they perform their optimal single-agent manipulation separately (see Table 1 for details). This result confirms that there are conflicts between different universities in the same league so that they need to find a way to manipulate jointly to achieve a better outcome.

We present an efficient algorithm to find a strategy profile such that (1) the induced matching is stable with respect to the true preference, (2) the induced matching is Pareto-optimal among all stable matchings that can be achieved by coalitional permutation manipulations, and (3) the strategy profile is inconspicuous, where inconspicuous manipulations are those in which each woman of the coalition only moves one man to a higher rank (Algorithm 1 and Algorithm 2). Surprisingly, we show that such a strategy profile actually forms a Nash equilibrium of the manipulation game

(Theorem 5.2). Therefore, the strategy profile found by our algorithm captures both the cooperation and the competition among the universities in the same league. This result implies that it is computationally easy to find a “profitable” manipulation that is weakly better off and Pareto-optimal for all members of the coalition, supporting the wide conjecture that such leagues of universities can benefit from forming coalitions.

All these results confirm the belief of the Ministry of Education that such leagues of universities are unfair for other universities. In the end, we show that it is NP-complete to find a manipulation that is strictly better off for all members of the coalition (Theorem 6.1). This result demonstrates a sharp contrast between weakly better-off outcomes and strictly better-off outcomes: if a manipulation is costly so that every manipulator must be strictly better off to ensure nonnegative payoff, a coalition manipulation is unlikely to happen due to computational burdens.

Our results also give answers to the open problem raised by Gusfield and Irving [14] on what can be produced by permutation manipulations (see also [23] and [32] for more of the problem).

## 1.3 Related Works

There is a large body of literature that focuses on finding manipulations for women when fixing men’s preferences in the Gale-Shapley algorithm. Gale and Sotomayor [10] show that it is possible for all women to strategically truncate their preference lists so that each of them is matched with their partner in the  $W$ -optimal matching, and Teo et al. [33] provide a polynomial time algorithm to find the optimal single-agent truncation manipulation.

Teo et al. [33] study permutation manipulations, where a woman can report any permutation of her true preference list. Their work is motivated by the primary student assignment process in Singapore. They give an efficient algorithm to compute the best manipulation for a single manipulator. Vaish and Garg [34] shows that the resulting matching from optimal singleton permutation manipulation is stable with respect to true preference lists and there exists an inconspicuous singleton manipulation which is optimal. Gupta et al. [12] extends the algorithm from Teo et al. [33] to the so-called  $P$ -stable (stable with respect to preferences  $P$ ) Nash equilibrium setting. Aziz et al. [6] also study permutation manipulations in a many-to-one setting, but focus on a single manipulator with quota more than one. Pini et al. [27] create a stable matching mechanism and show that for a single agent, it is computationally hard to manipulate the matching result. All the results, except for the last, do not apply to cases where a coalition of women jointly manipulate.

## 2 PRELIMINARIES

We consider a stable matching model with a set of men  $M$  and a set of women  $W$ , where only complete and strict preference lists are allowed.<sup>2</sup> The preference list of a man  $m$ , denoted by  $P(m)$ , in a preference profile  $P$ , is a strict total order  $>_m^P$  over the set of women  $W$ . Let  $w_1 >_m^P w_2$  denote that  $m$  prefers  $w_1$  to  $w_2$  in profile  $P$ . Similarly, the preference list  $P(w)$  of a woman  $w$  is a strict total

<sup>2</sup>We consider the case where men also report complete preference lists for simplicity. Our result can be generalized to the case where men may report incomplete preference lists.

order over  $M$ . For simplicity, we sometimes use  $\succ_w$  to denote the true preference list when it is clear from the context.

A matching is a function  $\mu : M \cup W \mapsto M \cup W$ . We write  $\mu(m) = w$  if a man  $m$  is matched to a woman  $w$ . Similarly,  $\mu(w) = m$  if  $w$  is matched to  $m$ . Also,  $\mu(m) = w$  if and only if  $\mu(w) = m$ . We will also write  $\mu(m) = m$  (or  $\mu(w) = w$ ) if  $m$  (or  $w$ ) is not matched. For two matchings  $\mu_1$  and  $\mu_2$ , if for all  $w \in W$ ,  $\mu_1(w) \succeq_w \mu_2(w)$ , we say  $\mu_1 \succeq_W \mu_2$ . If in a matching  $\mu$ , a man  $m$  and a woman  $w$  are not matched together, yet prefer each other to their partners in  $\mu$ , then  $(m, w)$  is called a *blocking pair*. A matching is *stable* if and only if it contains no blocking pair.

The Gale-Shapley algorithm is not truthful for women [8]. Given a set of women manipulators, the algorithm can be thought of as a game (henceforth, the *manipulation game*), between them.

*Definition 2.1 (Manipulation game).* Given a true preference profile  $P$ , a manipulation game is a tuple  $(L, \mathcal{A})$ , where:

- (1)  $L \subseteq W$  is the set of manipulators;
- (2)  $\mathcal{A} = \prod_{i \in L} A_i$  is the set of all possible reported preference profiles.

The outcome of the manipulation game (also called induced matching in this paper) is the matching resulted from the Gale-Shapley algorithm with respect to the reported preference profiles. A manipulator's preference in this game is her true preference in  $P$ .

Motivated by the NCEE in China, we focus on the setting where all women must report a complete list of men, which indicates that women can only permute their true preferences in the manipulation.

*Definition 2.2 (Permutation manipulation).* Let  $\circlearrowleft$  be the set of strict total orders over  $M$ . In permutation manipulations,  $A_i = \circlearrowleft$ ,  $\forall i \in L$ .

Let  $P(M) = (P(m) : m \in M)$  be the preference profile of all men. Similarly, denote the preference profiles for all women, all manipulators and all non-manipulators by  $P(W)$ ,  $P(L)$  and  $P(N)$ , respectively, where  $N = W \setminus L$  is the set of non-manipulators. Thus the overall preference profile is  $P = (P(M), P(N), P(L))$ . Denote by  $S(P(M), P(W))$  the set of all stable matchings under profile  $(P(M), P(W))$ . Let  $S_A(P(M), P(W)) \subseteq S(P(M), P(W))$  be the set of all stable matchings that can be achieved by a coalition manipulation of  $L$ . We sometimes write  $S_A$  for short when  $(P(M), P(W))$  is clear from the context. We define Pareto-optimality within the set  $S_A$ .

*Definition 2.3 (Pareto-optimal matching).* A matching  $\mu$  is Pareto-optimal if  $\mu \in S_A$  and there is no  $\mu' \in S_A$  such that all manipulators are weakly better off and at least one is strictly better off.

We say a strategy profile  $P(L)$  of a manipulation game is Pareto-optimal if its induced matching is Pareto-optimal. In a manipulation game, the solution concept we are interested in is Nash equilibrium.

*Definition 2.4 (Nash equilibrium).* A preference profile  $P(L) = \bigcup_{l \in L} P(l)$  of a manipulation game is a Nash equilibrium if  $\forall l \in L$ ,  $l$  cannot get a strictly better partner with respect to the true preference list by reporting any other preference list.

Our algorithm is enabled by two special structures, the *rotation* [13] and the *suitor graph* [23].

## 2.1 Rotations

The concept of rotations was first introduced by Irving [17] when solving the stable roommate problem, which is a natural generalization of the stable marriage problem.

In the Gale-Shapley algorithm, if a woman  $w_i$  rejects a man  $m_j$ , then  $w_i$  must have a better partner than  $m_j$  in the  $W$ -pessimal matching. Thus in any stable matching,  $w_i$  cannot be matched with any man ranked below  $m_j$  in  $w_i$ 's list. As a result, we can safely remove all impossible partners from each man or woman's preference list after each iteration of the algorithm. We call each man or woman's preference list after the removal a *reduced list*, and the set of all reduced lists a *reduced table*.

*Definition 2.5 (Rotation).* A rotation is a sequence of men  $R = (m_1, m_2, \dots, m_r)$ , where the first woman in  $m_{i+1}$ 's reduced list is the second in  $m_i$ 's reduced list ( $i+1$  is taken modulo  $r$ ).

Note that rotations are known as *improvement cycles* in some literature and is useful in converting the  $M$ -optimal matching to the  $W$ -optimal matching [5, 11, 16].

We also use  $R = (\mathcal{M}, \mathcal{W}, \mathcal{W}')$  to represent a rotation, where  $\mathcal{M}$  is the sequence of men and  $\mathcal{W}$  and  $\mathcal{W}'$  are the sequences of the first and the second women in  $\mathcal{M}$ 's reduced lists. Since  $\mathcal{W}_{i+1} = \mathcal{W}'_i$  by definition of rotations, we write  $\mathcal{W}^r = \mathcal{W}'$ , where  $\mathcal{W}^r$  is the sequence  $\mathcal{W}$  with each woman shifted left by one position. We may sometimes use  $m_i$  and  $w_i$  to mean the  $i$ -th agent in  $\mathcal{M}$  and  $\mathcal{W}$  when the order is important.

After the termination of the Gale-Shapley algorithm, one can still change the matching by eliminating rotations. The elimination of a rotation  $R$  is to force each woman  $w_i$  in  $\mathcal{W}$  to reject her current proposer  $m_i$  and let  $m_i$  propose to  $w_{i+1}$ . It is clear that after the elimination, each woman still holds a proposal, i.e. there is still a matching between men and women. More importantly, it can be shown that the matching is stable with respect to the true preference. We say a rotation  $R = (\mathcal{M}, \mathcal{W}, \mathcal{W}')$  moves  $m_i$  from  $w_i$  to  $w_{i+1}$  and moves  $w_i$  from  $m_i$  to  $m_{i-1}$  since after eliminating the rotation, the corresponding matching matches  $m_i$  and  $w_{i+1}$ . It is known that each stable matching corresponds to a set of rotations, and there exists an order of elimination that produces the matching, which we do not discuss in detail here, but refer readers to [14].

## 2.2 Suitor Graph

Suitor graph is another important structure for our analysis. It is proposed by Kobayashi and Matsui [23] when considering the problem that given a preference profile for all truthful agents  $P(M)$  and  $P(N)$ , is there a profile  $P(L)$  for the manipulators such that the  $M$ -optimal matching of the combined preference profile is a certain matching  $\mu$ ? The detailed definition of suitor graph is as follows:

*Definition 2.6 (Suitor graph; Kobayashi and Matsui [23]).* Given a matching  $\mu$ , a preference profile for all men  $P(M)$  and a preference profile for all non-manipulators  $P(N)$ , the corresponding suitor graph  $G(P(M), P(N), \mu)$  is a directed graph  $(V, E)$ , which can be constructed using the following steps:

- (1)  $V = M \cup W \cup \{s\}$ , where  $s$  is a virtual vertex;
- (2)  $\forall w \in W$ , add edges  $(w, \mu(w))$  and  $(\mu(w), w)$ , and let  $\delta(w) = \{m \mid w \succ_m \mu(m)\}$ ;
- (3)  $\forall w \in L$  and for each  $m$  in  $\delta(w)$ , add edges  $(m, w)$ ;

- (4)  $\forall w \in N$ , if  $\delta(w)$  is nonempty, add the edge  $(m, w)$ , where  $m$  is  $w$ 's favorite in  $\delta(w)$ ;
- (5)  $\forall w \in W$ , if  $\delta(w) = \emptyset$ , add an edge  $(s, w)$  to the graph;

Kobayashi and Matsui [23] also give a characterization of the existence of such profiles and an  $O(n^2)$  time algorithm that can be found directly from their constructive proof.

**THEOREM 2.7 (KOBAYASHI AND MATSUI [23]).** *Given a matching  $\mu$ , a preference profile with  $P(M)$  for all men and  $P(N)$  for all non-manipulators, there exists a profile for the manipulators  $P(L)$  such that  $\mu$  is the  $M$ -optimal stable matching for the total preference profile  $(P(M), P(N), P(L))$ , if and only if for every vertex  $v$  in the corresponding suitor graph  $G(P(M), P(N), \mu)$ , there exists a directed path from  $s$  to  $v$  ( $s$  is the virtual vertex in the graph). Moreover, if such a  $P(L)$  exists, it can be constructed in  $O(n^2)$ .*

### 3 PARETO-OPTIMAL STRATEGY PROFILES

We analyze the manipulation problem in the independent recruitment program of China's universities. In fact, this is also an open problem raised by Gusfield and Irving [14] on what can be achieved by permutation manipulations. Formally, we have the following results in this section.

**THEOREM 3.1.** *There exists a polynomial time algorithm (Algorithm 1) that, given any complete preference profile  $P$  and any set of manipulators  $L \subseteq W$  as input, computes a strategy profile  $P'(L)$  such that when  $L$  reports  $P'(L)$ , the induced matching  $\mu'$  is Pareto-optimal.<sup>3</sup>*

Moreover, our algorithm provides an algorithmic characterization of Pareto-optimal optimal matchings.

**THEOREM 3.2.** *A matching is Pareto-optimal if and only if it is an induced matching of a strategy profile found by Algorithm 1.*

#### 3.1 Conflicts between Manipulators

Before we develop our algorithms, we first show an example to demonstrate that a coalition of women could get worse off if they perform their optimal single-agent manipulation separately.

$m_1$	$w_1$	$w_4$	$w_2$	$w_3$	$w_1$	$m_3$	$m_2$	$m_1$	$m_4$
$m_2$	$w_1$	$w_3$	$w_2$	$w_4$	$w_2$	$m_1$	$m_4$	$m_3$	$m_2$
$m_3$	$w_2$	$w_3$	$w_1$	$w_4$	$w_3$	$m_2$	$m_3$	$m_1$	$m_4$
$m_4$	$w_2$	$w_4$	$w_1$	$w_3$	$w_4$	$m_4$	$m_1$	$m_3$	$m_2$

(a) Men's preference lists.      (b) Women's preference lists.

**Table 1: Example of non-cooperativeness**

Consider the preference lists in Table 1. The  $M$ -optimal matching is  $\{(m_1, w_4), (m_2, w_1), (m_3, w_3), (m_4, w_2)\}$ . Suppose  $L = \{w_1, w_2\}$  and consider individual manipulations by  $w_1$  and  $w_2$ .

- (1)  $w_1$  exchanges  $m_1$  and  $m_2$  and get  $\{(m_1, w_4), (m_2, w_3), (m_3, w_1), (m_4, w_2)\}$ ;
- (2)  $w_2$  exchanges  $m_3$  and  $m_4$  and get  $\{(m_1, w_2), (m_2, w_1), (m_3, w_3), (m_4, w_4)\}$ ;

<sup>3</sup>It is weakly better off for all manipulators to follow the strategy  $P'(L)$  rather than  $P$ , since  $\mu'$  is stable under  $P$ , which is preferred by each manipulator to the  $W$ -pessimal matching under  $P$ .

In both cases,  $w_1$  and  $w_2$  can manipulate to get their  $W$ -optimal partner and these manipulations are their optimal single-agent manipulation. However, if they jointly perform their optimal single-agent manipulations, the induced matching is  $(m_1, w_1), (m_2, w_3), (m_3, w_2), (m_4, w_4)$ . It is surprising that they both get worse off than the matching corresponding to their true preference lists.

This example shows a sharp contrast between permutation manipulations and general manipulations, where removing men from the preference lists is allowed. In general manipulations, women can jointly perform their optimal single-agent manipulations to be matched with their  $W$ -optimal partner [10, 33].

#### 3.2 Our Algorithm

To develop our algorithm, we extensively use two structures, rotations [17] and suitor graphs [22], introduced in Section 2.1 and 2.2 respectively. We further develop several new structures such as maximal rotations and principle sets to derive connections between suitor graphs and permutation manipulations.

Notice that eliminating more rotations results in weakly better matchings for all women. Thus, the manipulators' objective is to eliminate as many rotations as possible by permuting their preference lists. Since there is no direct rotation elimination in the Gale-Shapley algorithm, we try to figure out what kind of rotations can be eliminated, i.e., after eliminating these rotations, the corresponding matching is in  $S_A$ .

We first analyze the structure of the sets of rotations. Rotations are not always exposed in a reduced table. Some rotations become exposed only after other rotations are eliminated. Thus, we define the *precedence relation* between rotations and based on that, we incorporate notions from [18] (closed set, maximal rotations), and introduce the concept *principle sets* to analyze the problem.

The high-level idea behind our algorithm is that, with our theoretical analysis, we can actually reduce the search space from the set of all closed sets to the set of all principle sets, which enables our algorithm to run in polynomial time.

**Definition 3.3 (Precedence).** A rotation  $R_1 = (\mathcal{M}_1, \mathcal{W}_1, \mathcal{W}_1^r)$  is said to explicitly precede another  $R_2 = (\mathcal{M}_2, \mathcal{W}_2, \mathcal{W}_2^r)$  if  $R_1$  and  $R_2$  share a common man  $m$  such that  $R_1$  moves  $m$  from some woman to  $w$  and  $R_2$  moves  $m$  from  $w$  to some other woman. Let the relation precede be the transitive closure of the explicit precedence relation, denoted by  $<$ . Also,  $R_1 \sim R_2$  if neither  $R_1 < R_2$  nor  $R_2 < R_1$ .

**Definition 3.4 (Closed set).** A set of rotations  $\mathcal{R}$  is closed if for each  $R \in \mathcal{R}$ , any rotation  $R'$  with  $R' < R$  is also in  $\mathcal{R}$ . A closed set  $C$  is minimal in a family of closed sets  $\mathcal{C}$ , if there is no other closed set in  $\mathcal{C}$  that is a subset of  $C$ . Moreover, define  $CloSet(\mathcal{R})$  to be the minimal closed set that contains  $\mathcal{R}$ .

**Definition 3.5 (Maximal rotation & Principle set).** Given a closed set of rotations  $\mathcal{R}$ ,  $R$  is a *maximal rotation* of  $\mathcal{R}$  if no rotation  $R' \in \mathcal{R}$  satisfies  $R < R'$ . Let  $Max(\mathcal{R})$  be the set of all the maximal rotations in  $\mathcal{R}$ . Furthermore,  $\mathcal{R}$  is a *principle set* if  $Max(\mathcal{R})$  contains only one rotation. We will slightly abuse notations and write  $CloSet(R)$  to mean the principle set  $CloSet(\mathcal{R})$  if  $Max(\mathcal{R}) = \{R\}$ .

Henceforth,  $R_1$  precedes  $R_2$  if  $R_2$  can only be exposed after  $R_1$  is eliminated. A rotation  $R$  can only be exposed after all rotations preceding  $R$  are eliminated. Thus only closed sets can be validly

eliminated. Also, a closed set of rotations  $\mathcal{R}$  is uniquely determined by  $\text{Max}(\mathcal{R})$ . Therefore, given a closed set  $\mathcal{R}$ , the corresponding matching after eliminating rotations in  $\mathcal{R}$  is determined by  $\text{Max}(\mathcal{R})$ .

The following theorem shows that closed sets of rotations are all that we need to consider.

**THEOREM 3.6** (IRVING AND LEATHER [18]). *Let  $S$  be the set of all stable matchings for a given preference profile, there is a one-to-one correspondence between  $S$  and the family of all closed sets.*

Therefore, we need to understand the changes made to the suitor graph when a rotation  $R$  is eliminated. We keep track of every proposal made by men in  $R$  and modify the graph accordingly. We first assume that the virtual vertex  $s$  is comparable with each man and for every  $w \in W$  and every  $m \in M$ ,  $m \succ_w s$ . When eliminating a rotation, we follow the steps below to modify the graph:

- (1) Let all women in  $R$  reject their current partner, i.e., delete the edge  $(w_i, m_i)$  involved in  $R$  for each  $i$ ;
- (2) Arbitrarily choose a man  $m_i$  in  $R$  who does not have an incoming edge from a woman and let him propose to the next woman  $w$  in his preference list:
  - (a) If  $w$  is a manipulator, add an edge from  $m_i$  to  $w$  and delete edge  $(s, w)$  if it exists;
  - (b) If  $w$  is not a manipulator, then compare  $m_i$  with the two men (one is possibly  $s$ ) in  $V' = \{v \mid (v, w) \in E\}$ . If  $m_i$  is not the worst choice, add an edge from  $m_i$  to  $w$  and delete the worst edge, and we say  $w$  is *overtaken* by  $m_i$ ;
  - (c) If  $w$  accepts  $m_i$ , add an edge from  $w$  to  $m_i$ ;
- (3) Repeat step 2 until all men in  $R$  are accepted.

Let  $G$  and  $G'$  be the suitor graphs corresponding to the reduced tables before and after the elimination of  $R$ . It is easy to check that after modifying  $G$  using the steps defined above, the resulting suitor graph is exactly  $G'$ . From Theorem 2.7, the most important property of the suitor graph is the existence of a path from  $s$  to any other vertex. Therefore, we focus on the change of *strongly connected components* and their connectivity in the suitor graph before and after the elimination of rotations.

**Definition 3.7.** A sub-graph  $G'$  is strongly connected if for any two vertices  $u, v$  in  $G'$ , there is a path from  $u$  to  $v$  in  $G'$ . A strongly connected component is a maximal strongly connected sub-graph.

The following lemma gives some connectivity properties of the suitor graph after eliminating a rotation.

**LEMMA 3.8.** *After eliminating a rotation  $R$ , (1) all agents in  $R$  are in the same strongly connected component; (2) vertices that are formerly reachable from a vertex in  $R$  remain reachable from  $R$ ; (3) vertices that are overtaken during the elimination of  $R$  are reachable from  $R$ .*

To prove Lemma 3.8, we first show the following claim.

**CLAIM 1.** *For each man  $m_i$  in  $R$ , in the procedure of eliminating the rotation  $R$ ,  $w_{i+1}$  (the subscript is taken modulo  $r$ ) is the first woman to accept him, and each woman in  $R$  accepts only one proposal during the procedure.*

**PROOF.** According to the definition of rotations,  $w_{i+1}$  is the second in  $m_i$ 's reduced list. If there are other women between  $w_i$  and  $w_{i+1}$  in  $m_i$ 's preference list, they are absent from the reduced list because these women already hold proposals from better men.

Henceforth, even though  $m_i$  proposes to these women, they reject him. But  $m_i$  is in  $w_{i+1}$ 's reduced list since  $w_{i+1}$  is in  $m_i$ 's. Therefore,  $m_i$  is a better choice for  $w_{i+1}$  and  $w_{i+1}$  accepts him.

After the elimination, each man  $m_i$  in  $R$  proposes to  $w_{i+1}$  and each man is accepted only once. Also each woman  $w_{i+1}$  holds a new proposal from  $m_i$  and thus accepts at least once. The conclusion is immediate since the total number of accepted men is equal to the total number of women who accept a new partner.  $\square$

**PROOF OF LEMMA 3.8.** For each  $m_i$  in  $R$ ,  $R$  moves  $m_i$  from  $w_i$  to  $w_{i+1}$ . As a result, there exists an edge from  $w_{i+1}$  to  $m_i$ . We now prove that each  $m_i$  has an outgoing edge pointing to  $w_i$ , and all agents in  $R$  then form a cycle, and thus in the same strongly connected component. Before the elimination,  $w_i$  is the partner of  $m_i$ , so there is an edge from  $m_i$  to  $w_i$ . If  $w_i$  is a manipulator, the edge  $(m_i, w_i)$  is not removed during the elimination according to the steps described above. If  $w_i$  is not a manipulator, then only two incoming edges are remained after the elimination and these edges are from the best two men among those who propose to her. According to Claim 1, only one man, namely  $m_{i-1}$ , is accepted. Thus,  $m_{i-1}$  is the best suitor of  $w_i$ . We claim that  $m_i$  is the second best and the edge from  $m_i$  is still in the suitor graph. Otherwise, suppose  $m'$  is a better choice than  $m_i$  to  $w_i$ . Then  $m'$  is also in  $R$ . We let  $m'$  propose first, and  $w_i$  accepts  $m'$ , which makes  $w_i$  accepts at least twice. A contradiction.

Since each woman can be reached from her partner before the elimination, it is without loss of generality to assume that a vertex  $v$  can be reached from a man  $m$  in  $R$  through a path  $p$ . Let  $u$  be the last vertex in  $p$  such that  $u$  is in  $R$  or is overtaken by a vertex in  $R$ . If  $u$  is in  $R$ , then after the elimination,  $m$  can reach  $u$  since they are in the same strongly connected component. If  $u$  is overtaken by some vertex  $m'$ , then during the elimination, an edge  $(m', u)$  is added to the graph. Thus,  $m$  can reach  $u$  through  $m'$ . Henceforth, in any case,  $u$  is reachable. Since in  $p$  the vertices between  $u$  and  $v$  are neither in  $R$  nor overtaken by some vertex in  $R$ , the path from  $u$  to  $v$  remains in the modified graph. Therefore  $v$  is reachable from  $m$  and also from any vertex in  $R$  for they are in the same strongly connected component after the elimination.  $\square$

With Lemma 3.8, we do not need to worry about vertices that are reachable from vertices in  $R$ , for they will remain reachable after the elimination. Also, vertices that are overtaken and the other vertices reachable from overtaken vertices can be reached from vertices in  $R$  after the elimination.

In fact, every vertex is reachable from  $s$  in the initial suitor graph. Therefore, if a vertex becomes unreachable from  $s$  after eliminating a rotation, there must exist some edge that is deleted during the elimination, which only happens when some woman is overtaken if she is a non-manipulator. The next lemma extends Lemma 3.8 to a closed set of rotations.

**LEMMA 3.9.** *After eliminating a closed set of rotations  $\mathcal{R}$ , each  $v$  in  $\mathcal{R}$  is reachable from at least one vertex in  $\text{Max}(\mathcal{R})$ , i.e., there exists a path to  $v$  from a vertex in  $\text{Max}(\mathcal{R})$ .*

**PROOF.** We eliminate the rotations in  $\mathcal{R}$  one by one and generate a sequence of rotations  $q = (R_1, R_2, \dots, R_n)$ .  $R_i$  is the  $i$ -th rotation to eliminate. After eliminating  $R_n$ , all rotations in  $\mathcal{R}$  are eliminated.

Denote  $q_i = \bigcup_{j=1}^i R_j$ . For each  $i$ ,  $q_i$  is a closed set. We call  $i$  the sequence number of  $q_i$  and we prove by induction on the sequence number that after eliminating  $q_i$ , all vertices in  $q_i$  can be reached from a vertex in  $Max(q_i)$ . For  $i = 1$ ,  $q_1 = \{R_1\}$ , the case is trivial from Lemma 3.8. Assume the statement is true for  $i = k$ , then for  $i = k+1$ , we only eliminate one more rotation  $R_{k+1}$  than in the case with  $i = k$ .  $R_{k+1}$  is in  $Max(q_{k+1})$  otherwise there exists another rotation  $R'$  in  $q_k$  such that  $R_{k+1} < R'$  and then  $q_k$  is not a closed set. Let  $D = Max(q_k) \setminus Max(q_{k+1})$ . Rotations in  $D$  are no longer maximal rotations because  $R_{k+1}$  is eliminated, which indicates that rotations in  $D$  explicitly precede  $R_{k+1}$ . Henceforth, every rotation  $R$  in  $D$  has a common agent with  $R_{k+1}$  and each vertex  $u$  reachable from  $R$  is reachable from that common agent. According to Lemma 3.8,  $u$  can be reached from  $R_{k+1}$ . For each vertex  $u'$  that is not reachable from rotations in  $D$ , it must be reachable from another rotation  $R'$  in  $Max(q_k)$  through path  $p$  and  $R'$  is still in  $Max(q_{k+1})$ . If  $p$  is still in the suitor graph, then we are done. Otherwise, some vertices in  $p$  must be in  $R_{k+1}$  or overtaken by a man in  $R_{k+1}$ . Let  $z$  be the last vertex in  $p$  such that  $z$  is in  $R_{k+1}$  or overtaken.  $z$  can be reached from  $R_{k+1}$  and the path from  $z$  to  $u'$  is not affected by the elimination. Therefore,  $u'$  is reachable from  $R_{k+1}$ .  $\square$

If  $R_2$  explicitly precedes  $R_1$ , then they must contain a common man. Therefore, after eliminating  $R_1$ , vertices in  $R_1$  can reach any vertex that is previously reachable from  $R_2$ . The analysis goes on recursively until some rotation has no predecessors.

Given a closed set of rotations  $\mathcal{R}$ , we say  $\mathcal{R}$  can be eliminated for simplicity, if the corresponding matching after eliminating rotations in  $\mathcal{R}$  is in  $S_A$ . The following lemma provides us a simpler way to check whether a closed set of rotations can be eliminated.

**LEMMA 3.10.** *A closed set of rotations  $\mathcal{R}$  can be eliminated if and only if after eliminating  $\mathcal{R}$ , every vertex in  $Max(\mathcal{R})$  can be reached from  $s$ .*

**PROOF.** If a closed set of rotations  $\mathcal{R}$  can be eliminated, then every vertex is reachable after  $\mathcal{R}$  is eliminated. As a result, any member of  $Max(\mathcal{R})$  is reachable.

If after eliminating  $\mathcal{R}$ , any member of  $Max(\mathcal{R})$  can be reached from  $s$ , then we need to show that all other vertices are also reachable from  $s$ . We split all vertices into two parts. Let  $V$  denote the set of all the vertices that can be reached from members of  $Max(\mathcal{R})$ . If a vertex  $v$  is in  $V$ , then  $v$  is reachable from  $s$  through  $Max(\mathcal{R})$ . If  $v$  is not in  $V$ , then in the initial graph, there is a path  $p$  from  $s$  to  $v$ . We claim that no vertex in path  $p$  is either in any of the rotations in  $\mathcal{R}$  or overtaken when eliminating a rotation. Otherwise, according to Lemma 3.9,  $v$  is reachable from  $Max(\mathcal{R})$ . Thus, the path  $p$  is still in the graph after eliminating all the rotations in  $\mathcal{R}$ .  $\square$

However, we still cannot afford to enumerate all possible closed sets of rotations, whose number is exponential with respect to the number of women.

**THEOREM 3.11.** *Given a closed set of rotations  $\mathcal{R}$ , if  $\mathcal{R}$  can be eliminated, then there exists a rotation  $R \in \mathcal{R}$  such that  $CloSet(R)$  can be eliminated.*

In order to prove Theorem 3.11, we first show the following claim about the maximal rotations of a closed set that can be eliminated.

**CLAIM 2.** *If a closed set  $\mathcal{R}$  can be eliminated, then every rotation in  $Max(\mathcal{R})$  must contain a manipulator.*

**PROOF.** Assume there exists a rotation  $R \in Max(\mathcal{R})$  such that  $R$  contains no manipulator. We can change the order of elimination to make  $R$  the last to eliminate. We prove that after eliminating  $R$ , any vertex in  $R$  is not reachable from  $s$ . From the proof of Lemma 3.8, we know that all vertices in  $R$  form a cycle after eliminating  $R$ . Each man in  $R$  has only one incoming edge from his current partner who is also in  $R$ . Each woman has two incoming edges, one from her partner in  $R$  and another from her former partner which is also in  $R$ . Thus, every vertex in  $R$  has no incoming edges from outside the cycle and thus is not reachable from  $s$ .  $\square$

**PROOF OF THEOREM 3.11.** Let  $V$  be the set of all vertices in  $\mathcal{R}$ . After eliminating  $\mathcal{R}$ , we arbitrarily choose a vertex  $v$  in  $V$ . In the corresponding suitor graph, there is a path  $p = (v_0 = s, v_1, v_2, \dots, v_n = v)$  from  $s$  to  $v$  since  $\mathcal{R}$  can be eliminated. Let  $u$  be the first vertex in  $p$  such that  $u$  is in  $V$ .  $u$  is obviously not  $v_1$ , or otherwise the edge  $(s, u)$  will be deleted. Moreover,  $u$  must be in  $L$ , since any non-manipulator can only be reached from a node in  $V$  if she is overtaken during the elimination. Assume  $u = v_l$  and  $l > 1$ . Then the sub-path  $p' = (v_0, v_1, \dots, v_l = u)$  is not affected (no vertices in  $V$  or overtaken) during the elimination. Henceforth,  $p'$  is in the original suitor graph before eliminating  $\mathcal{R}$ . Now we consider the set  $\mathcal{R}' = \{R \in \mathcal{R} | u \in R\}$ . For any  $R$  in  $\mathcal{R}'$ , if we eliminate  $CloSet(R)$ , the sub-path is also not affected. Therefore  $CloSet(R)$  can be eliminated according to Lemma 3.10.  $\square$

The above theorem reduces the search space from *the set of closed sets* to *the set of principles sets*. We are ready to design Algorithm 1 to compute a Pareto-optimal strategy profile. For any iteration of Algorithm 1, the matching at the beginning of each iteration is in  $S_A$ . Therefore, according to Theorem 3.11, if a closed set of rotations  $\mathcal{R}$  can be eliminated, we can always find a principle set  $\mathcal{P}^*$  contained in  $\mathcal{R}$  such that  $\mathcal{P}^*$  can be eliminated. Since the number of principle sets equals the number of rotations, which is polynomial and can be efficiently computed [13], given a matching in  $S_A$ , we figure out an efficient way to find a weakly better matching in  $S_A$ . Using this method as a sub-routine, we are able to design an algorithm to find a Pareto-optimal strategy profile for permutation manipulations.

---

**Algorithm 1:** Find a Pareto-optimal strategy profile

---

Find the set of all rotations  $\mathcal{R}$  and all principle sets  
 $\mathcal{P} = \{CloSet(R) | R \in \mathcal{R}\};$

**while** True **do**

    Construct  $\mathcal{C} = \{\mathcal{P} \in \mathcal{P} | \mathcal{P} \text{ can be eliminated}\};$

**if**  $\mathcal{C} = \emptyset$  **then**

        Construct  $P(L)$  for  $L$  and return;

**else**

        Arbitrarily choose a principle set  $\mathcal{P}^* \in \mathcal{C}$  and  
        eliminate  $\mathcal{P}^*$ ;

---

**3.2.1 Correctness of Algorithm 1.** To prove the correctness of the algorithm. We begin with the following lemma.

LEMMA 3.12. *Given a set of manipulators  $L \in W$ , and the true preference profile  $P = (P(M), P(W))$ . Let  $\mu$  be any matching in  $S_A$  and  $\mathcal{R}$  be the corresponding closed set of rotations. Then there exists a preference profile  $P_\mu(L)$  for  $L$  such that  $\mu$  is the  $M$ -optimal stable matching of the preference profile  $P_\mu = (P(M), P(N), P_\mu(L))$ , and the reduced table of  $P$  after eliminating  $\mathcal{R}$  is exactly the reduced table of  $P_\mu$  before eliminating any rotation.*

PROOF. Since  $\mu$  is in  $S_A$ , there exists  $P' = (P(M), P(N), P'(L))$  such that  $\mu$  is the induced matching for  $P'$ . For each  $w \in L$ , we modify  $P'(w)$  as follows:

- (1) delete all men  $m$  such that  $m \succ_w^P \mu(w)$ ;
- (2) reinsert them at the beginning according to their order in  $w$ 's true preference list;
- (3) move  $\mu(w)$  to the position right after all men  $m$  such that  $m \succ_w^P \mu(w)$ ;

Denote the modified preference profile by  $P'_\mu$ . In fact,  $P'_\mu$  is the  $P_\mu$  we are looking for.

We first prove that  $\mu$  is the  $M$ -optimal matching under  $P'_\mu$ . After the first two steps of modifications, the  $M$ -optimal matching is still  $\mu$ , since for each  $w$ , we only change the position of men ranked higher than  $\mu(w)$  in her true preference list, who must have not proposed to  $w$  under  $P'$ , and thus do not change the output of the Gale-Shapley algorithm. Otherwise, if a man  $m$  with  $m \succ_w^P \mu(w)$  has proposed to  $w$ , then we must have  $w \succ_m^{P'} \mu(m)$ , which is equivalent to  $w \succ_m^P \mu(m)$ . Thus  $(m, w)$  forms a blocking pair in  $\mu$  under the true preference profile  $P$ , contradicting to the stability of  $\mu$  under  $P$ . In the third step, we move  $\mu(w)$  to the position right after all men ranked higher than  $\mu(w)$  in the true preference list  $P(w)$ . Consider all the men  $m'$  with  $m' \succ_w^{P'} \mu(w)$  but  $\mu(w) \succ_w^{P'} m'$ .  $m'$  must have not proposed to  $w$  under  $P'$ , or otherwise  $\mu(w)$  cannot be the partner of  $w$ . Therefore, the positions of the men in  $P'_\mu$  do not affect the output of the Gale-Shapley algorithm.

Let  $T_{P_\mu}$  be the reduced table of  $P$  after eliminating  $\mathcal{R}$  and  $T_{P'_\mu}$  be the reduced tables of  $P'_\mu$ . We already know that for each woman, her partners in the two reduced tables are the same, which is  $\mu(w)$ . In fact, a change of reduced table happens if and only if a woman accepts a proposal from a man  $m$  and removes everyone ranked below  $m$  in her preference list. So in the reduced list of each woman, no man is ranked below her current partner. Therefore, to prove that  $T_{P_\mu}$  is the same as  $T_{P'_\mu}$ , it suffices to show that for each woman,  $P$  and  $P'_\mu$  are the same after removing all men ranked below her current partner, which is clear from the construction of  $P'_\mu$ .  $\square$

From Theorem 3.6, we know that only closed sets need to be considered. Although Claim 2 has already ruled out all closed sets that have a maximal rotation containing only non-manipulators, there are still exponentially many possibilities. However, Theorem 3.11 shows that every closed set that can be eliminated contains a principle set, which can also be eliminated. A natural idea is to iteratively grow the closed set by adding principle sets. The above lemma shows that after each iteration, we can construct a problem that has the current matching as its initial matching, and contains rotations that are not yet eliminated. If we find a principle set that can be eliminated in the constructed problem, it can also be eliminated in the original problem.

3.2.2 *Complexity of Algorithm 1.* To analyze the time complexity of Algorithm 1, we define a graph describing the precedence relation between rotations.

Definition 3.13 (*Precedence graph*). Given a set of rotations  $\mathcal{R}$ , let  $D$  be a directed acyclic graph, where the vertices in  $D$  are exactly  $\mathcal{R}$ , and there is an edge  $(R_1, R_2)$  in  $D$  if  $R_1 < R_2$ . Moreover, let  $H$  be the transitive reduction of  $D$  defined above, and  $H_r$  be the graph  $H$  with all edges reversed.

Note that  $H$  is exactly the directed version Hasse diagram of the precedence relation between rotations. For a rotation  $R$ ,  $CloSet(R)$  is the set of vertices that can be reached from  $R$  through a directed path in  $H_r$ . We split the algorithm into the initialization part and iteration part, and assume  $|M| = |W| = n$ .

In the initialization part, we first compute the initial matching using the Gale-Shapley algorithm, which can be computed in  $O(n^2)$  time. Next we find all rotations with respect to preference profile  $P$  and also find all the principle sets. These two operations depend on the graph  $H_r$ . However, the graph  $H$  is the transitive reduction of  $D$ , and the construction of  $H$  is somewhat complex. Gusfield [13] discusses how to find all rotations, whose number is  $O(n^2)$ , in  $O(n^2)$  time. Instead of constructing  $H$ , Gusfield considered a subgraph  $H'$  of  $D$ , whose transitive closure is identical to  $D$ . Moreover,  $H'$  can be constructed in  $O(n^2)$  time. We will not discuss how to construct  $H'$  in detail but only apply Gusfield's results here. Then for each rotation  $R$ , we only need to search  $H'$  to find  $CloSet(R)$ , which takes  $O(n^2)$  time. Thus, we finish the initialization step in  $O(n^4)$  time since there are  $O(n^2)$  rotations altogether.

The iteration part is the bottleneck of the algorithm. At least one rotation is eliminated for each iteration, and thus  $O(n^2)$  iterations are needed. Inside each iteration, we need to construct the set  $\mathcal{C}$ . There are  $O(n^2)$  principle sets and to determine whether a principle set can be eliminated, we need to simulate the Gale-Shapley algorithm and modify the suitor graph accordingly. After the modification, we traverse the suitor graph to see if each vertex is reachable. Both of the two operations takes  $O(n^2)$  time. Thus, the construction of  $\mathcal{C}$  takes  $O(n^4)$  time. In the *If-Else* statement, if we find a principle set that can be eliminated, we eliminate the principle set and modify the suitor graph in  $O(n^2)$ . Otherwise, we traverse the suitor graph to construct the preference profile for  $L$  according to Theorem 2.7. Thus, the *If-Else* statement takes  $O(n^2)$  time and totally, the time complexity is  $O(n^6)$ .

### 3.3 Algorithmic Characterization

Notice that at each iteration, the algorithm has multiple principle sets to select from. To prove our characterization result in Theorem 3.2, we have already shown that for each Pareto-optimal matching  $\mu$ , there exists a way to select the principle sets to eliminate in each iteration such that the induced matching from the output of Algorithm 1 is  $\mu$ .

## 4 INCONSPICUOUSNESS

In fact, if a stable matching with respect to the true preference lists can be obtained by permutation manipulations, the manipulators can also obtain this matching by an inconspicuous manipulation.

*Definition 4.1 (Inconspicuous Strategy Profile).* A strategy profile is inconspicuous if each manipulator permutes their preference lists by moving only one man to a higher rank.

For convenience, we introduce a new notation  $Pro(w)$  for each  $w \in W$ . A proposal list  $Pro(w)$  of a woman is a list of all men who have proposed to her in the Gale-Shapley algorithm, and the orderings of its entries are the same as her stated preference list. A reduced proposal list contains the top two entries (first entry if only one entry exists) of  $Pro(w)$ , denoted by  $Pro_r(w)$ . Clearly, each woman  $w$  is matched to the first man of  $Pro_r(w)$ .

**THEOREM 4.2.** *For any stable matching with respect to the true preference lists that can be obtained by permutation manipulations, there exists a preference profile for the manipulators, in which each manipulator only needs to move at most one man to some higher ranking, that yields the same matching.*

Theorem 4.2 suggests that for each woman  $w$ , let  $m_1$  and  $m_2$  be the two men in  $Pro_r(w)$ <sup>4</sup>, then  $w$  can modify her true preference list by moving  $m_2$  to the place right after  $m_1$  to generate the same induced matching (see Algorithm 2 for details).

---

**Algorithm 2:** Find a Pareto-optimal and inconspicuous preference profile

---

Use Algorithm 1 to compute a strategy profile  $P'(L)$  for  $L$ ;

Compute  $Pro_r(w)$  for each  $w \in L$  with respect to  $P'(L)$ ;

**for**  $w$  **in**  $L$  **do**

Modify the true preference list  $P(w)$  by moving the second man in  $Pro_r(w)$  to the position right after the first man in  $Pro_r(w)$ ;

**return** the modified preference profile  $P$ ;

---

## 5 INCENTIVE PROPERTIES

Although we only have been focusing on constructing Pareto-optimal strategy profiles, a Pareto-optimal strategy profile, which is also inconspicuous, actually forms a Nash equilibrium.

**LEMMA 5.1.** *Suppose there is only one manipulator  $w$ . Then the best matching  $\mu'$  that  $w$  can obtain via permutation manipulation is stable with respect to the true preference  $P$ .*

**PROOF.** Let  $P'$  be the preference profile corresponding to  $\mu'$ . Assume on the contrary that  $\mu'$  is not stable with respect to  $P$ . Then there must be a blocking pair. However, any pair  $(m, w')$  with  $w' \neq w$  cannot block  $\mu'$  under  $P$ , since they have the same preferences in both  $P$  and  $P'$ . It follows that the woman in the blocking pair must be  $w$ . Let  $(m, w)$  be the blocking pair. We move  $m$  to the top of  $P'(w)$ . If we run the Gale-Shapley algorithm with the new preference profile,  $m$  will still propose to  $w$  and will finally be matched to  $w$  since  $m$  is now the favorite man of  $w$ . But  $m \succ_w \mu'(w)$ , which contradicts to the fact that  $\mu'$  is the best matching that  $w$  can obtain.  $\square$

**THEOREM 5.2.** *A strategy profile, that is Pareto-optimal and inconspicuous, forms a Nash equilibrium.*

<sup>4</sup>If woman  $w$  only receives one proposal, she cannot implement any manipulation.

**PROOF.** Denote by  $P$  and  $\mu$  the true preference profile of agents and the corresponding matching. Let  $P_1$  be the preference profile returned by Algorithm 2 given  $P$ , and  $\mu_1$  be the corresponding matching. It is clear that for each  $w \in L$ , Algorithm 2 only changes the order of the men ranked strictly lower than  $\mu_1(w)$ . For the sake of contradiction, assume there exists a manipulator  $w' \in L$  such that  $w'$  can get a strictly better partner  $m$  ( $m \succ_{w'}^P \mu_1(w')$ ) by misreporting a different preference list. Let  $P_2$  and  $\mu_2$  be the preference profile after misreporting and the corresponding matching. Without loss of generality, we assume that  $m$  is the best partner (according to both  $P$  and  $P_1$ ) that  $w'$  can obtain. Then we know from Lemma 5.1 that,  $\mu_2$  is stable with respect to  $P_1$ , and thus for each  $w \in W$ , we have that  $\mu_2(w) \geq_{P_1}^w \mu_1(w)$ . It follows that  $\mu_2(w) \geq_w^P \mu_1(w)$ , since Algorithm 2 does not change the order of the men who are ranked higher than  $\mu_1(w)$ . It follows that  $\mu_2$  is also stable with respect to  $P$ , and  $\mu_2$  Pareto-dominates  $\mu_1$ . However,  $\mu_2$  is not found by Algorithm 2. A contradiction.  $\square$

Since all Pareto-optimal strategy profiles can be turned into an inconspicuous one by Algorithm 2, we have the following corollary.

**COROLLARY 5.3.** *All Pareto-optimal matchings can be induced by a Nash equilibrium.*

Therefore, Pareto-optimal matchings exactly address both the cooperation and the competition among the women in the coalition.

## 6 STRICTLY BETTER-OFF OUTCOMES

The above results show that the Gale-Shapley algorithm is vulnerable to coalition manipulation. However, if a manipulation is costly such that every manipulator needs to be strictly better off after the manipulation to preserve individual rationality, we show a hardness result:

**THEOREM 6.1.** *It is NP-complete to find a strategy profile, the induced matching of which is strictly better off for all manipulators.*

Therefore, when the manipulation is costly, a manipulation coalition is unlikely to form and the Gale-Shapley algorithm is immune to coalition manipulations. According to Theorem 6.1, one immediate corollary is that the number of Pareto-optimal matchings cannot be polynomial in the number of men and women. For otherwise, we can enumerate all such matchings by Algorithm 1 to develop a polynomial time algorithm. Last but not least, we show that the problem to compute the number of Pareto-optimal matchings, which are strictly better off for all manipulators, is #P-complete.

**THEOREM 6.2.** *It is #P-complete to compute the number of Pareto-optimal matchings, which are strictly better off for all manipulators.*

## 7 CONCLUSION

Motivated by a real life phenomenon risen in recent years in the college admissions in China, we consider manipulations by subsets of women in the Gale-Shapley algorithm. We show that a Nash equilibrium with Pareto-optimal matching can be efficiently computed in general. These results confirm that the leagues of universities can benefit from forming coalitions. On the contrary, we show that it is NP-complete to find a strictly better off matching for all the manipulators, implying that Gale-Shapley algorithm is immune from permutation manipulations when the manipulations are costly.

## REFERENCES

- [1] 2011. China's top universities start joint entrance exams. [http://english.pku.edu.cn/News\\_Events/News/Outlook/7740.htm](http://english.pku.edu.cn/News_Events/News/Outlook/7740.htm). (Feb. 21, 2011). Accessed: 2017-07-08.
- [2] Atila Abdulkadiroglu, Parag A Pathak, and Alvin E Roth. 2005. The new york city high school match. *American Economic Review* (2005), 364–367.
- [3] Atila Abdulkadiroglu and Tayfun Sönmez. 2003. School choice: A mechanism design approach. *The American Economic Review* 93, 3 (2003), 729–747.
- [4] David J Abraham, Avrim Blum, and Tuomas Sandholm. 2007. Clearing algorithms for barter exchange markets: Enabling nationwide kidney exchanges. In *Proceedings of the 8th ACM conference on Electronic commerce*. ACM, 295–304.
- [5] Itai Ashlagi, Yashodhan Kanoria, and Jacob Leshno. 2013. Unbalanced Random Matching Markets. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce (EC '13)*. ACM, New York, NY, USA, 27–28. <https://doi.org/10.1145/2482540.2482590>
- [6] Haris Aziz, Hans Georg Seedig, and Jana Karina von Wedel. 2015. On the Susceptibility of the Deferred Acceptance Algorithm. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 939–947.
- [7] Yan Chen and Onur Kesten. 2017. Chinese college admissions and school choice reforms: A theoretical analysis. *Journal of Political Economy* 125, 1 (2017), 99–139.
- [8] Lester E Dubins and David A Freedman. 1981. Machiavelli and the Gale-Shapley algorithm. *American mathematical monthly* (1981), 485–494.
- [9] D. Gale and L. S. Shapley. 1962. College Admissions and the Stability of Marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15.
- [10] David Gale and Marilda Sotomayor. 1985. Ms. Machiavelli and the stable matching problem. *Amer. Math. Monthly* (1985), 261–268.
- [11] Yannai A. Gonczarowski. 2014. Manipulation of Stable Matchings Using Minimal Blacklists. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation (EC '14)*. ACM, New York, NY, USA, 449–449. <https://doi.org/10.1145/2600057.2602840>
- [12] Sushmita Gupta, Kazuo Iwama, and Shuichi Miyazaki. 2015. Stable Nash Equilibria in the Gale-Shapley Matching Game. *arXiv preprint arXiv:1509.04344* (2015).
- [13] Dan Gusfield. 1987. Three fast algorithms for four problems in stable marriage. *SIAM J. Comput.* 16, 1 (1987), 111–128.
- [14] Dan Gusfield and Robert W. Irving. 1989. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, Cambridge, MA, USA.
- [15] Chien-Chung Huang. 2006. Cheating by men in the Gale-Shapley stable matching algorithm. In *Algorithms-ESA 2006*. Springer, 418–431.
- [16] Nicole Immorlica and Mohammad Mahdian. 2005. Marriage, honesty, and stability. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 53–62.
- [17] Robert W Irving. 1985. An efficient algorithm for the “stable roommates” problem. *Journal of Algorithms* 6, 4 (1985), 577–595.
- [18] Robert W Irving and Paul Leather. 1986. The complexity of counting stable marriages. *SIAM J. Comput.* 15, 3 (1986), 655–667.
- [19] Robert W Irving and David F Manlove. 2009. Finding large stable matchings. *Journal of Experimental Algorithmics (JEA)* 14 (2009), 2.
- [20] Robert W Irving, David F Manlove, and Sandy Scott. 2000. The hospitals/residents problem with ties. In *Algorithm Theory-SWAT 2000*. Springer, 259–271.
- [21] Donald E. Knuth. 1976. *Marriages Stables*. Les Presses de l'Université de Montreal.
- [22] Hirotatsu Kobayashi and Tomomi Matsui. 2009. Successful manipulation in stable marriage model with complete preference lists. *IEICE TRANSACTIONS on Information and Systems* 92, 2 (2009), 116–119.
- [23] Hirotatsu Kobayashi and Tomomi Matsui. 2010. Cheating strategies for the Gale-Shapley algorithm with complete preference lists. *Algorithmica* 58, 1 (2010), 151–169.
- [24] Yicheng Liu, Pingzhong Tang, and Wenyi Fang. 2014. Internally Stable Matchings and Exchanges. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada*. 1433–1439.
- [25] Yicheng Liu, Pingzhong Tang, Tingting Xu, and Hang Zheng. 2016. Optimizing Trading Assignments in Water Right Markets. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA*. 551–557.
- [26] David G McVitie and Leslie B Wilson. 1971. The stable marriage problem. *Commun. ACM* 14, 7 (1971), 486–490.
- [27] Maria Silvia Pini, Francesca Rossi, K Brent Venable, and Toby Walsh. 2009. Manipulation and gender neutrality in stable marriage procedures. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 665–672.
- [28] Alvin E Roth. 1982. The economics of matching: Stability and incentives. *Mathematics of operations research* 7, 4 (1982), 617–628.
- [29] Alvin E Roth. 1996. The national residency matching program as a labor market. *Journal of the American Medical Association* 275, 13 (1996), 1054–1056.
- [30] Alvin E Roth, Tayfun Sönmez, and M Utku Ünver. 2004. Kidney Exchange. *The Quarterly journal of economics* 119, 2 (2004), 457–488.
- [31] Alvin E Roth, Tayfun Sönmez, and M Utku Ünver. 2005. A kidney exchange clearinghouse in New England. *American Economic Review* (2005), 376–380.
- [32] Noriyoshi Sukegawa and Yoshitsugu Yamamoto. 2012. Preference profiles determining the proposals in the Gale-Shapley algorithm for stable matching problems. *Japan journal of industrial and applied mathematics* 29, 3 (2012), 547–560.
- [33] Chung-Piaw Teo, Jay Sethuraman, and Wee-Peng Tan. 2001. Gale-shapley stable marriage problem revisited: Strategic issues and applications. *Management Science* 47, 9 (2001), 1252–1267.
- [34] Rohit Vaish and Dinesh Garg. 2017. Manipulating Gale-Shapley Algorithm: Preserving Stability and Remaining Inconspicuous. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*. 437–443. <https://doi.org/10.24963/ijcai.2017/62>
- [35] Wei Zhan, Zhiyuan Li, Pingzhong Tang, Yicheng Liu, and Tingting Xu. 2017. Stability of generalized two-sided markets with transaction thresholds. In *Proceedings of the AAMAS*.