Two for One & One for All: Two-Sided Manipulation in Matching Markets

Hadi Hosseini¹, Fatima Umar², Rohit Vaish³

¹Pennsylvania State University
²Rochester Institute of Technology
³Indian Institute of Technology Delhi

hadi@psu.edu, fu1476@rit.edu, rvaish@iitd.ac.in

Abstract

Strategic behavior in two-sided matching markets has been traditionally studied in a “one-sided” manipulation setting where the agent who misreports is also the intended beneficiary. Our work investigates “two-sided” manipulation of the deferred acceptance algorithm where the misreporting agent and the manipulator (or beneficiary) are on different sides. Specifically, we generalize the recently proposed accomplice manipulation model (where a man misreports on behalf of a woman) along two complementary dimensions: (a) the two for one model, with a pair of misreporting agents (man and woman) and a single beneficiary (the misreporting woman), and (b) the one for all model, with one misreporting agent (man) and a coalition of beneficiaries (all women). Our main contribution is to develop polynomial-time algorithms for finding an optimal manipulation in both settings. We show this despite the fact that an optimal one for all strategy fails to be inconspicuous, while it is unclear whether an optimal two for one strategy has this property. We also study the conditions under which stability of the resulting matching is preserved. Experimentally, we show that two-sided manipulations are more frequently available and offer better quality matches than their one-sided counterparts.

1 Introduction

The deferred acceptance algorithm [Gale and Shapley, 1962] is one of the biggest success stories of matching theory and market design. It has profoundly impacted numerous practical applications including school choice [Abdulkadiroglu et al., 2005a; Abdulkadiroglu et al., 2005b] and entry-level labor markets [Roth and Peranson, 1999] and has inspired a long line of work in economics and computer science [Gusfield and Irving, 1989; Roth and Sotomayor, 1992; Roth, 2008; Manlove, 2013], and has been a key predictor of the long-term sustainability of many real-world matching markets [Roth, 2002].

Unfortunately, any stable matching algorithm is known to be vulnerable to strategic misreporting of preferences by the agents [Roth, 1982]. For the DA algorithm, in particular, it is known that truth-telling is a dominant strategy for the proposing side—colloquially, the men—implying that any strategic behavior is confined to the proposed-to side—the women [Dubins and Freedman, 1981; Roth, 1982].

One-sided vs. two-sided manipulation. Given the strong practical appeal of DA algorithm, significant research effort has been devoted towards understanding its incentive properties. Much of this work has focused on “one-sided” manipulation wherein the agent who misreports is also the intended beneficiary; that is, the misreporting agent and the beneficiary are on the same side. For this self manipulation problem, the structural and computational questions have been extensively studied [Dubins and Freedman, 1981; Gale and Sotomayor, 1985a; Gale and Sotomayor, 1985b; Teo et al., 2001; Vaish and Garg, 2017]. By contrast, there are many real-world settings that, in essence, resemble “two-sided” manipulations where the misreporting agent and the beneficiary are on different sides. For example, in the student-proposing school choice, schools could influence the preferences of students they find “undesirable” (such as those from low-income backgrounds) by using indirect measures such as fee hike [Hatfield et al., 2016]. In ridesharing platforms, a driver may influence the preferences of certain riders by strategically moving to a farther distance. Similarly, in a gig economy, freelancers’ preferences over tasks may be affected by an employer’s restrictive requirements.

Motivated by these examples, recent works have studied the accomplice manipulation model wherein a man misreports his preferences in order to help a specific woman [Bendlin and Hosseini, 2019; Hosseini et al., 2021]. It has been shown via simulations that accomplice manipulation strategies are more frequently available than self manipulation and result in better matches for the woman. Further, an optimal misreport for the accomplice is known to be inconspicuous (i.e., the manipulated list can be derived from his true list by promoting exactly one woman), efficiently computable, and stability-preserving (i.e., the manipulated DA matching is stable with respect to the true preferences).
Towards coalitional two-sided manipulation. The aforementioned advantages of two-sided manipulation call for a deeper investigation into the topic. Our work takes a step in this direction by focusing on coalitional aspects of the two-sided manipulation problem. Our starting point is the accomplice manipulation model [Hosseini et al., 2021], which involves one misreporting agent and one beneficiary (i.e., a one for one setting). We consider two coalitional generalizations of this model: (i) The two for one model, with a coalition of two misreporting agents (a man and a woman) and a single beneficiary (the woman), and (ii) the one for all model, with one misreporting agent (man) and a coalition of beneficiaries (all women). Coalitional manipulation of the DA algorithm has received considerable theoretical interest over the years [Dubins and Freedman, 1981; Gale and Sotomayor, 1985a; Gale and Sotomayor, 1985b; Demange et al., 1987; Kobayashi and Matsui, 2010], and recently its practical relevance has also been discussed. Indeed, in college admissions in China, universities have been known to form “leagues” for conducting independent recruitment exams allowing them to jointly manipulate admission results [Shen et al., 2021]. Prior work on coalitional manipulation has focused exclusively on one-sided manipulation.

Our contributions. We study two coalitional generalizations of accomplice manipulation and make the following theoretical and experimental contributions (see Table 1):

- **Two for one**: We show that when the accomplice and the beneficiary can jointly misreport, an optimal pair manipulation strategy can be strictly better for the beneficiary than either of the optimal individual (i.e., self or accomplice) manipulation strategies (Example 1). In contrast to self and accomplice manipulation, an optimal pair manipulation may not be stability-preserving (Remark 1) and it is unclear whether it is inconspicuous. Nevertheless, we provide a polynomial-time algorithm for computing an optimal pair manipulation (Theorem 1).

- **One for all**: We observe that optimal manipulation by the accomplice for helping all women could fail to be inconspicuous (Example 2). By contrast, when helping a single woman, an optimal strategy for the accomplice is known to be inconspicuous [Hosseini et al., 2021]. Despite losing this structural benefit, we develop a polynomial-time algorithm for computing an optimal one-for-all strategy (Corollary 1), and show that such a strategy is stability-preserving (Corollary 3 in [Hosseini et al., 2022]). In fact, we show that a minimum optimal misreport (i.e., one that pushes up as few women as possible) can be efficiently computed and provide tight bounds on size of promoted set [Hosseini et al., 2022].

**Experiments**: Our simulations on uniformly random preferences show that two-sided strategies are more frequently available (Figure 2) and result in better matches than one-sided strategies [Hosseini et al., 2022].

Related Work. The literature on strategic aspects of stable matching procedures has classically focused on truncation strategies where the misreported list is a prefix of the true list [Dubins and Freedman, 1981; Roth, 1982; Roth and Rothblum, 1999]. Our work, on the other hand, focuses on permutation manipulation where the manipulated list is a reordering of the true list. Teo et al. [2001] initiated the study of permutation manipulation by a single woman and provided a polynomial-time algorithm for finding an optimal misreport. Vaish and Garg [2017] showed that an optimal strategy for the woman is inconspicuous and stability-preserving. Permutation manipulation by a group of agents has been studied for a coalition of men [Huang, 2006; Huang, 2007] and for a coalition of women [Kobayashi and Matsui, 2010; Shen et al., 2021]. In particular, Shen et al. [2021] provided an algorithm for finding a strategy for a coalition of women that is Pareto optimal among all stability-preserving strategies, and showed that such a strategy is inconspicuous. In the two-sided setting, Bendlin and Hosseini [2019] introduced the accomplice manipulation model and observed that it can be more beneficial for a woman than optimal self manipulation. Subsequently, Hosseini et al. [2021] studied with-regret and no-regret accomplice manipulation, depending on whether the accomplice’s match worsens or stays the same. They showed that an optimal no-regret manipulation is stability-preserving while its with-regret counterpart is not, and that optimal strategies under both models are inconspicuous and therefore efficiently computable. Our work will focus exclusively on no-regret strategies.

## 2 Preliminaries

**Problem instance.** An instance of the stable marriage problem [Gale and Shapley, 1962] is given by a tuple $(M, W, \succ)$, where $M$ is a set of $n$ men, $W$ is a set of $n$ women, and $\succ$ is a preference profile which specifies the preference lists of the agents. The preference list of a man $m \in M$, denoted by $\succ_m$, is a strict total order over all women in $W$. The list $\succ_w$ of a woman $w \in W$ is defined analogously. We will write $w_1 \succ_m w_2$ to denote “either $w_1 \succ_m w_2$ or $w_1 = w_2$”, and write $w_1 \prec_m w_2$ to denote the profile without the list of man $m$; thus, $\prec = \{\prec_m, \succ_m\}$.

**Stable matching.** A matching $\mu : M \cup W \rightarrow M \cup W$ such that $\mu(m) \in W$ for all $m \in M$, $\mu(w) \in M$ for all $w \in W$, and $\mu(m) = w$ if and only if $\mu(w) = m$. Given a matching $\mu$, a blocking pair with respect to the preference

---

### Table 1: Summary of previously known (top two rows) and new results (bottom two rows).

<table>
<thead>
<tr>
<th>Who misreports?</th>
<th>Who benefits?</th>
<th>Results for optimal manipulation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woman $w$</td>
<td>Woman $w$</td>
<td>Open</td>
<td>Vaish and Garg [2017]</td>
</tr>
<tr>
<td>Man $m$</td>
<td>Woman $w$</td>
<td>$\checkmark$</td>
<td>Hosseini et al. [2021]</td>
</tr>
<tr>
<td>Man $m$ and woman $w$</td>
<td>Woman $w$</td>
<td>$\checkmark$</td>
<td>Section 3</td>
</tr>
<tr>
<td>Man $m$</td>
<td>All women</td>
<td></td>
<td>Section 4</td>
</tr>
</tbody>
</table>

---

Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22)

322
profile $\succ$ is a man–woman pair $(m, w)$ who prefer each other over their assigned partners, i.e., $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. A matching is said to be stable if it does not have any blocking pair. We will write $S_\succ$ to denote the set of all stable matchings with respect to $\succ$. Note that in the worst case, the size of $S_\succ$ can be exponential in $n$ [Knuth, 1997].

For any pair of matchings $\mu, \mu'$, we will write $\mu \succeq_M \mu'$ to denote that all men weakly prefer $\mu$ over $\mu'$, i.e., $\mu(m) \succeq_M \mu'(m)$ for all $m \in M$ (analogously $\mu \succeq_W \mu'$ for women).

**Deferred acceptance algorithm.** The deferred acceptance (DA) algorithm is a well-known procedure for finding stable matchings [Gale and Shapley, 1962]. Given as input a preference profile, the algorithm alternates between a proposal phase, where each currently unmatched man proposes to his favorite woman among those who haven’t rejected him yet, and a rejection phase, where each woman tentatively accepts her favorite proposal and rejects the rest. The algorithm terminates when no further proposals can be made.

Gale and Shapley [1962] showed that given any preference profile $\succ$ as input, the matching computed by the DA algorithm, which we will denote by $DA(\succ)$, is stable. Furthermore, this matching is men-optimal as it assigns to each man his favorite partner among all stable matchings in $S_\succ$. Subsequently, it was observed that the same matching is also women-pessimal [McVitie and Wilson, 1971].

**Proposition 1** ([Gale and Shapley, 1962; McVitie and Wilson, 1971]). Let $\succ$ be a profile and let $\mu := DA(\succ)$. Then, $\mu \in S_\succ$. Furthermore, for any $\mu' \in S_\succ$, $\mu \succeq_M \mu'$ and $\mu' \succeq_W \mu$.

**Self manipulation.** Given a profile $\succ$ and the matching $\mu := DA(\succ)$, we say that woman $w$ can self-manipulate if there exists a list $\succ'_w$ (which is a permutation of $w$’s true list $\succ_w$) such that $\mu'(w) \succ_w \mu(w)$, where $\mu' := DA(\succ_w \succ'_w)$. An optimal self-manipulation $\succ'_w$ (with respect to the profile $\succ$) is one for which there is no other list $\succ''_w$ such that $\mu''(w) \succ_w \mu'(w)$, where $\mu'' := DA(\succ_w \succ''_w)$.

**Accomplish manipulation.** A different model of strategic behavior is accomplish manipulation [Bendlin and Hosseini, 2019; Hosseini et al., 2021], wherein a woman $w$, instead of misreporting herself, asks a man $m$ to misreport his preference in order to improve $w$’s match. Formally, given a profile $\succ$ and a fixed man $m$, we say that woman $w$ can manipulate via accomplish $m$ if there exists a list $\succ'_m$ for man $m$ (which is a permutation of his true list $\succ_m$) such that $\mu'(w) \succ_w \mu(w)$, where $\mu := DA(\succ_m \succ'_m)$ and $\mu' := DA(\succ_m \succ'_m)$. An optimal accomplish manipulation $\succ'_m$ (with respect to $\succ$) is one for which there is no other list $\succ''_m$ such that $\mu''(w) \succ_w \mu'(w)$, where $\mu'' := DA(\succ_m \succ''_m)$. We will call woman $w$ a ‘beneficiary’ and man $m$ an ‘accomplish’.

**No-regret assumption.** In this paper, we will consider two generalizations of accomplish manipulation: (a) the “two for one” problem with two misreporting agents (man $m$ and woman $w$) and a single beneficiary (woman $w$), and (b) the “one for all” problem with a single misreporting agent (man $m$) and a coalition of beneficiaries (all women). In both cases, we will assume no-regret manipulation for the man which means that $m$’s match does not worsen upon misreporting, i.e., $\mu(m) \succeq_m \mu(m)$.

Interestingly, for both generalizations mentioned above, the no-regret assumption implies that the accomplish’s match stays the same, i.e., $\mu(m) = \mu'(m)$. Indeed, in the one-for-all problem, it follows from the strategy-proofness of the DA algorithm for the proposing side that $\mu(m) \succeq_m \mu'(m)$. Along with the no-regret assumption, this implies $\mu(m) = \mu'(m)$.

For the two-for-one problem where both man $m$ and woman $w$ can misreport, it is known that $m$ cannot be strictly better off unless $w$ is strictly worse off [Huang, 2007, Corollary 4]. To prevent the beneficiary $w$ from being worse off, we must ensure that man $m$’s match does not improve, implying once again that $\mu(m) = \mu'(m)$.

**Inconspicuous manipulation.** A misreported list or strategy $\succ'_m$ for an accomplish $m$ is said to be inconspicuous if it can be derived from his true list $\succ_m$ by promoting at most one woman and making no other changes. Similarly, when the misreporting agent is a woman $w$, inconspicuous involves promoting at most one man in her true list $\succ_w$.

**Push up and push down operations.** For any man $m \in M$, let $\succ'_m \succ'_w$ denote the parts of $m$’s list above and below his DA partner, respectively. That is, $\succ_m := (\succ'_m, \succ'_w, m)$. We say that man $m$ pushes up a set $X \subseteq W$ if the new list is $\succ^{X := (\succ'_m \cup X, \mu(m), \succ'_w \setminus X)}$. Similarly, pushing down a set $Y \subseteq W$ results in $\succ^{Y := (\succ'_m \setminus Y, \mu(m), \succ'_w \cup Y)}$. The exact positions at which agents in $X$ (or $Y$) are placed above (or below) $\mu(m)$. Huang [2006] has shown that the DA outcome remains unchanged if each man $m$ arbitrarily permutes the part of his list above and below his DA-partner $\mu(m)$.

**Proposition 2** ([Huang, 2006]). Let $\succ$ be a profile and let $\mu := DA(\succ)$. For any man $m \in M$ with true list $\succ_m := (\succ^L_m, \mu(m), \succ^R_m)$, let $\succ := (\pi^L(\succ^L_m), \mu(m), \pi^R(\succ^R_m))$, where $\pi^L$ and $\pi^R$ are arbitrary permutations. Let $\mu' := DA(\succ_m)$. Then, $\mu' = \mu$.

### 3 Two for One: Helping a Single Woman Through Pair Manipulation

In this section, we will consider the “two for one” generalization of accomplish manipulation where the accomplish and the strategic woman can jointly misreport in order to benefit the latter. To see how such a generalization can be useful, let us start with an example showing that pair manipulation can be strictly more beneficial for the woman compared to either self or accomplish manipulation (Example 1).

Example 1 (Pair manipulation can be strictly better than accomplish or self manipulation). Consider the following preference profile where the DA outcome is underlined. The notation “$m_1 : w_5 w_3 w_4 w_2 w_1$” denotes that $m_1$’s top choice is $w_5$, second choice is $w_3$, and so on.

$\begin{align*}
m_1 : w_5 w_3 w_4 w_2 w_1 \quad w_1 : m_1 m_2 m_3 m_5 m_4 \\
m_2 : w_4 w_3 w_5 w_1 w_2 \quad w_2 : m_1 m_2 m_3 m_5 m_4 \\
m_3 : w_5 w_4 w_1 w_3 w_2 \quad w_3 : m_5 m_4 m_3 m_2 m_1 \\
m_4 : w_4 w_5 w_2 w_3 w_1 \quad w_4 : m_5 m_3 m_1 m_4 \\
m_5 : w_5 w_4 w_3 w_1 w_2 \quad w_5 : m_2 m_4 m_3 m_1 m_3
\end{align*}$

Suppose the manipulating pair is $(m_1, w_1)$. Since $m_4$ is the only man who proposes to $w_1$ during the execution of DA
algorithm on this profile, it follows that there is no beneficial self manipulation strategy for \( w_1 \).

To find an optimal accomplice manipulation for \( m_1 \), it suffices to focus on inconsistent strategies [Hosseini et al., 2021]. It is straightforward to verify that promoting any woman below \( w_3 \) in \( m_1 \)'s list does not result in a better match for \( w_1 \). In fact, none of the other men can give \( w_1 \) a better partner via no-regret manipulation.

The DA matching when \( m_1 \) and \( w_1 \) jointly misreport with \( m_1 := w_1 \succ w_5 \succ w_3 \succ w_4 \succ w_2 \) and \( w_1 \succ w_3 \succ m_1 \succ m_2 \) is marked by “*”. Note that the strategic woman \( w_1 \) is now able to match with her top choice without worsening the match of the accomplice \( m_1 \).

Since pair manipulation can be strictly more beneficial than either self or accomplice manipulation, it is natural to ask whether an optimal pair manipulation can be efficiently computed. Our main result in this section is that an optimal pair manipulation can be computed in polynomial time.

### 3.1 Computing an Optimal Joint Strategy

A natural approach for finding an optimal joint strategy is to combine (or “concatenate”) an optimal self manipulation for the woman and an optimal accomplice manipulation for the man. However, as we saw in Example 1, an optimal pair manipulation may exist despite there being no beneficial accomplice nor self manipulations. Further, Example 3 in the full version of the paper [Hosseini et al., 2022] shows that the woman’s match could actually worsen by naively combining the respective individual strategies. Thus, pair manipulation appears to be “more than just the sum of its parts”.

Another natural approach is to combine inconspicuous (but not necessarily individually optimal) strategies of the accomplice and the strategic woman. However, as we discuss in the appendix of [Hosseini et al., 2022], there are some subtleties that arise from this approach that become difficult to resolve. We leave the question of determining whether optimal pair manipulation is inconspicuous as an open problem.

Nevertheless, the idea of looking for structure in the individual strategies turns out to be useful. To see why, consider manipulating a pair \((m,w)\). Let \( \succ^m \) and \( \succ^w \) denote the respective lists of \( m \) and \( w \) under an optimal pair manipulation, and let \( \succ^* := \{ \succ^+(m,w), \succ^*, \succ^w \} \) denote the corresponding preference profile.

Since it is easier to think about single-agent misreports, let us break down the transition from the true profile \( \succ \) to the pair manipulation profile \( \succ^* \) in two steps: First, swap \( w \)'s list in \( \succ \) to obtain the intermediate profile \( \succ^w := \{ \succ^+, \succ^w \} \), and then swap \( m \)'s list in \( \succ^w \) to get \( \succ^* \); see Figure 1. We will show that this two-step approach allows us to impose additional structure on the individual strategies \( \succ^m \) and \( \succ^w \).

Let us start by analyzing woman \( w \)'s strategy \( \succ^w \). Consider the transition \( \succ \rightarrow \succ^w \) in Figure 1, where \( w \) is the only misreporting agent. Let \( R_w \) denote the set of preference lists that can be obtained from \( w \)'s true list \( \succ^w \) by moving some pair of men to the top two positions, i.e., \( R_w := \{ (m_i, m_j, \succ^w \setminus \{m_i, m_j\}) : m_i, m_j \in M \} \). In Lemma 1, we show that for an arbitrary misreport by \( w \), there exists a list in \( R_w \) that creates the same matching for all agents. Thus, it follows that \( \succ^w \in R_w \). Observe that the set \( R_w \) is of polynomial size \( O(n^2) \) and can be efficiently enumerated.

**Lemma 1.** Let \( \succ \) be a profile and let \( \succ' \) be any misreport for a fixed woman \( w \). Then, there exists a list \( \succ'' \in R_w \) that achieves the same matching, i.e., \( \mu'' = \mu' \), where \( \mu' := DA(\succ, \succ') \) and \( \mu'' := DA(\succ, \succ'') \).

Next consider the transition \( \succ \rightarrow \succ^* \) in Figure 1. For this step, man \( m \) is the only misreporting agent. We define \( \succ_m \) as the list obtained by promoting \( m \)'s original match, namely \( \mu(m) \), to the top of his original list \( \succ_m \), and define \( R_m := \{ \succ_m \} \cup \{ \succ_m^w \uparrow : \succ_m^w \neq \mu(m) \} \) as the set consisting of the list \( \succ_m \) as well as all preference lists that are obtained by individually pushing up each woman other than \( \mu(m) \) to the top position in the list \( \succ_m \). Further, we say that an arbitrary misreport \( \succ_m' \) is feasible if \( m \) matches with \( \mu(m) \) under \( \succ_m' = \{ \succ^+(m,w), \succ^m, \succ^w \} \). In Lemma 2, we show that for an arbitrary feasible misreport by the accomplice, there exists another feasible list in \( R_m \) that results in the same partner for \( w \). Thus, we can assume that \( \succ_m'' \in R_m \). Again, observe that the set \( R_m \) is of polynomial size \( O(n) \).

**Lemma 2.** Let \( \succ \) be a profile and let \( \succ_m' \) and \( \succ_m'' \) be any misreport for a fixed pair \((m,w)\) such that \( \mu'(m) = \mu''(m) \), where \( \mu := DA(\succ, \succ_m') \) and \( \mu' := DA(\succ, \succ_m'') \). Then, there exists a list \( \succ_m'' \in R_m \) such that \( \mu''(m) = \mu''(m) \) and \( \mu''(w) = \mu''(w) \), where \( \mu'' := DA(\{w, \succ^w \}) \).

Although the lists in sets \( R_m \) and \( R_w \) are not necessarily inconspicuous versions of the true lists \( \succ_m \) and \( \succ_w \), respectively, we have been able to identify nominally-sized sets of misreports \( R_m \) and \( R_w \) that are sufficient to check, leading to a simple algorithm for finding an optimal pair manipulation strategy: Enumerate the sets \( R_m \) and \( R_w \), evaluate the DA outcome for each possible \( \succ_m'' \in R_m, \succ_w'' \in R_w \), and return the strategy that gives the best match for the woman \( w \) without regret for the accomplice \( m \); see Algorithm 1.

**Theorem 1.** An optimal pair manipulation can be computed in \( O(n^2) \) time.

**Remark 1.** In Example 5 in [Hosseini et al., 2022], we show that optimal pair manipulation could fail to be stability-preserving. Thus, an unrestricted pair manipulation (i.e.,
4 One for All: Helping All Women Through a Single Accomplice

Let us now consider a different generalization of accomplice manipulation which we call “one for all” manipulation where a single accomplice (man \( m \)) misreports in order to improve the outcome for all women in \( W \). Recall that due to the no-regret assumption, the manipulated match of the accomplice \( m \) is the same as his true match. As we are interested in improving a group of agents, it will be helpful to define the notions of Pareto improvement and Pareto optimal strategies.

Pareto optimal and optimal strategies. Let \( \succ \) be the true preference profile and let \( \mu := \text{DA}(\succ) \). We say that a strategy \( \succ_m' \) of the accomplice Pareto improves another strategy \( \succ_m'' \) if \( \mu' \succeq_w \mu'' \) and \( \mu'(w) \succ_w \mu''(w) \) for some \( w \in W \), where \( \mu' := \text{DA}(\succ_m',\succ_m'') \) and \( \mu'' := \text{DA}(\succ_m',\succ_m'') \). A strategy \( \succ_m' \) is Pareto optimal if \( \mu' \succeq_w \mu \) and there is no other strategy \( \succ_m'' \) that Pareto improves \( \succ_m' \). Further, a strategy \( \succ_m' \) is optimal if \( \mu' \succeq_w \mu \) and for any other strategy \( \succ_m'' \), we have \( \mu' \succeq_w \mu'' \). Thus, given a Pareto optimal strategy, any other strategy that improves some woman must make some other woman worse off, while the outcome under an optimal strategy simply cannot be improved for any woman. Similarly, we say that a matching \( \mu' := \text{DA}(\succ_m',\succ_m'') \) is “Pareto optimal” (respectively, “optimal”) if the corresponding strategy \( \succ_m' \) is Pareto optimal (respectively, optimal). Note that an optimal strategy is also Pareto optimal. The finiteness of the strategy space implies that a Pareto optimal strategy is guaranteed to exist. Whether an optimal strategy also always exists is not immediately clear; however, if an optimal strategy exists, then the set of Pareto optimal matchings—the Pareto frontier—must be a singleton, consisting only of the optimal matching. There can be multiple optimal strategies, but all such strategies must induce the same optimal matching.

Let us now proceed to analyzing the structure of (Pareto) optimal strategies. When an accomplice manipulates on behalf of a single beneficiary woman (i.e., “one for one”), it is known that there always exists an optimal strategy that is inconspicuous [Hosseini et al., 2021]. By contrast, when an accomplice misreports on behalf of multiple women (i.e., “one for all”), an inconspicuous strategy may no longer be optimal (Example 2).

Example 2 (Inconspicuous strategy can be suboptimal). Consider the following preference profile where the DA outcome is (\[325\]).

\[
\begin{align*}
\succ_m' := & \left( w_1 \succ w_2 \succ w_3 \succ w_4 \succ w_5 \mid w_1 : m_1, m_2, m_3, m_4, m_5 \\
\succ_m'' := & \left( w_1 \succ w_2 \succ w_3 \succ w_4 \succ w_5 \mid w_1 : m_1, m_2, m_3, m_4, m_5 \\
\succ_m''' := & \left( w_1 \succ w_2 \succ w_3 \succ w_4 \succ w_5 \mid w_1 : m_1, m_2, m_3, m_4, m_5 \\
\succ_m^* := & \left( w_1 \succ w_2 \succ w_3 \succ w_4 \succ w_5 \mid w_1 : m_1, m_2, m_3, m_4, m_5 \\
\end{align*}
\]

Suppose the accomplice is \( m_1 \) and all women are beneficiaries. The DA matching after \( m_1 \) submits the optimal no-regret manipulated list \( \succ_m^* := w_2 \succ w_4 \succ w_3 \succ w_5 \succ w_1 \) is marked by “\( * \)”. Notice that \( \succ_m^* \) is derived from \( \succ_m' \) by pushing up \( w_2 \) and \( w_4 \) and therefore is not inconspicuous. The manipulation results in the women-optimal matching, where all women are matched with their top choices.

There is no inconspicuous strategy that \( m_1 \) (or any other man for that matter) can report to produce the same matching; indeed, if \( m_1 \) were to push up only \( w_2 \), then only \( w_2 \) and \( w_3 \) would improve, and if he were to push up only \( w_4 \), then only \( w_4 \) and \( w_5 \) would improve. This observation highlights the conflict between optimality and inconspicuousness when the set of beneficiaries consists of all women.

Our main result in this section is that an optimal strategy for the accomplice is guaranteed to exist and computable in polynomial time (Theorem 2). In the full version [Hosseini et al., 2022], we prove a stronger result: Among all optimal strategies, we can efficiently compute one that promotes the smallest number of women in accomplice’s list.

4.1 Computing an Optimal Strategy

Recall from Proposition 2 that even if each man \( m \) arbitrarily permutes the part of his list above and below his DA-partner \( \mu(m) \), the DA outcome remains unchanged. This result shows that any strategy of the accomplice \( m \), without loss of generality, can be expressed in terms of only push up and push down operations, where a set of women is pushed above the DA partner \( \mu(m) \), and another disjoint set of women is pushed below \( \mu(m) \). We will now provide a structural simplification: Any matching obtained by a combination of push up and push down operations in the accomplice’s list can be weakly improved for all women by the push up operation alone.

Proposition 3. Let \( \succ \) be a profile. For any fixed man \( m \) and any subsets \( X \subseteq W \) and \( Y \subseteq W \) of women who are ranked below and above \( \mu(m) \), respectively, let \( \succ^* := \{\succ_{m, X}, \succ_{m, Y}\} \) denote the profile after pushing up the set \( X \) and let \( \succ^* := \{\succ_{m, X}, \succ_{m, Y}\} \) denote the profile after pushing up \( X \) and push down \( Y \) in the true preference list \( \succ_m \) of man \( m \). Then, \( \mu' \succeq_w \mu'' \), where \( \mu' := \text{DA}(\succ') \) and \( \mu'' := \text{DA}(\succ'') \).

Having established that push up operations suffice, let us now examine which subset of women the accomplice should push up. Given a profile \( \succ \) and an accomplice \( m \), define the no-regret set \( W^{NR} := \{ w \in W : \succ_m := \{\succ_{m, w}\} \) is a no-regret profile\} as the set of all women who do not cause \( m \) to incur regret when pushed up individually, and its complement with-regret set \( W^R := W \setminus W^{NR} \).

We will first show that pushing up any subset of no-regret women does not cause regret for the accomplice (Lemma 3).

Lemma 3. Let \( \succ \) be a profile and let \( \mu := \text{DA}(\succ) \). For any subset \( Y \subseteq W^R \), let \( \succ^* := \{\succ_{m, X}, \succ_{m, Y}\} \) denote the preference profile after pushing up the set \( Y \) in the true preference list \( \succ_m \) of man \( m \), and let \( \mu^* := \text{DA}(\succ^*) \). Then, \( m \) does not incur regret under \( \succ^* \), i.e., \( \mu^*(m) = \mu(m) \).

In contrast to Lemma 3, any subset \( Y \subseteq W \) that contains at least one woman from the with-regret set (i.e., \( Y \cap W^R = \emptyset \)) causes regret for the man \( m \) (Lemma 4).

Lemma 4. Let \( w' \in W^R \) and let \( Y \subseteq W \) be such that \( w' \in Y \). Then, \( m \) incurs regret under \( \succ^* := \{\succ_{m, X}, \succ_{m, Y}\} \), i.e., \( \mu^*(m) \succ_{m} \mu(m) \), where \( \mu^*(m) := \text{DA}(\succ^*) \).
Together, Lemmas 3 and 4 imply that a push up operation is no-regret if and only if the pushed-up set is a subset of \( W^{NR} \). Thus, an optimal (or Pareto optimal) strategy should promote some subset of \( W^{NR} \). This observation, however, does not automatically provide an efficient algorithm for computing the desired strategy because brute force enumeration of subsets of \( W^{NR} \) could take exponential time. Also, in case an optimal strategy does not exist, the Pareto frontier of strategies can be exponential in size, again ruling out exhaustive search.

Our main result of this section (Theorem 2) alleviates both of the above concerns. We show that not only does an optimal strategy always exist, but also that pushing up the entire no-regret set \( W^{NR} \) achieves such an outcome.

**Theorem 2.** An optimal one-for-all strategy for the accomplice is to push up the no-regret set \( W^{NR} \) in his true list.

Our proof of Theorem 2 leverages the following known result which says that the matching resulting from a no-regret push up operation is weakly preferred by all women.

**Proposition 4** ([Hosseini et al., 2021]). Let \( \succ \) be a preference profile and let \( \mu := DA(\succ) \). For any man \( m \), let \( \succ^m := \{ \succ_{-m}, \succ^X_m \} \) and \( \mu' := DA(\succ^m) \). If \( m \) does not incur regret, then \( \mu' \in S_{\succ} \) and thus \( \mu' \succeq_W \mu \) and \( \mu \succeq_M \mu' \).

**Proof.** (of Theorem 2) From Proposition 2, we know that any accomplice manipulation can be simulated via push up and push down operations. Proposition 3 shows that any combination of push up and push down operations can be weakly improved for all women by push up only. From Lemma 4, we know that the desired push up set, say \( Y \subseteq W \), should not contain any woman from the with-regret set \( W^{WR} \). Therefore, \( Y \subseteq W^{NR} \). From Lemma 3, we know that pushing up \( Y \) satisfies no-regret assumption. If \( Y \neq W^{NR} \) (thus, \( Y \subseteq W^{WR} \)), then Proposition 4 shows that \( \succ^Y := \{ \succ_{-m}, \succ^Y_m \} \) can be weakly improved for all women by additionally pushing up the women in \( W^{WR} \setminus Y \). Thus, pushing up all women in \( W^{NR} \) gives an optimal no-regret accomplice manipulation strategy for helping all women, as desired.

Theorem 2 readily gives a polynomial-time algorithm for computing an optimal strategy (Corollary 1).

**Corollary 1.** An optimal one-for-all strategy for the accomplice can be computed in \( O(n^3) \) time.

Although pushing up the entire no-regret set \( W^{NR} \) is optimal (Theorem 2), the accomplice may want to displace as few women as possible in order to remain close to his true preference list. In the full version [Hosseini et al., 2022], we provide a polynomial-time algorithm for computing a minimum optimal strategy (i.e., one that promotes the smallest number of women). We also show that the size of the promoted set is at most \( \lceil \frac{n-1}{2}\rceil \) and that this bound is tight.

**5 Experimental Results**

Let us now experimentally compare the two-sided and one-sided models in terms of the fraction of instances where each model improves upon truthful reporting. In our experimental setup, the preferences of \( n \) men and \( n \) women are drawn uniformly at random.\(^1\) For each value of \( n \in \{4, 6, \ldots, 20\} \), we independently sample 1,000 preference profiles. For the two-for-one part, we compute the fraction of instances where some man \( m \) can jointly misreport with a fixed woman \( w \) to improve her match, and compare it with the analogous fraction where only one of \( m \) (accomplice) or \( w \) (self) can misreport; see Figure 2 (left). Similarly, for the one-for-all part, we compute the fraction of instances where some man can misreport to help all women (i.e., weakly improve all and strictly improve some compared to their true matches), and compare it with the analogous fraction where a woman helps all women; see Figure 2 (right). Figure 2 shows that two-sided strategies are more frequently available than one-sided; in roughly 2% more instances under the two-for-one setting and roughly 10% more instances under the one-for-all. In the full version [Hosseini et al., 2022], we show that two-sided manipulation outperforms one-sided in terms of the extent of improvement for the beneficiary/beneficiaries, i.e., the difference between the ranks of old and new matched partner(s).

**6 Concluding Remarks**

We studied two coalitional generalizations of two-sided manipulation of the DA algorithm. Moving from single-agent to coalitional manipulation impacted the structure of optimal strategies in the form of loss of inconspicuousness, but we showed that efficient computation can still be achieved. Going forward, it will be interesting to consider manipulation by arbitrary coalitions of men and women. Another relevant direction could be to interpret two-sided manipulation as a bribery problem [Faliszewska et al., 2009; Boehmer et al., 2021] wherein there is a cost associated with each pairwise swap in an agent’s true list. Finally, extensions of our work to more general preference models (e.g., partial orders), as well as experimental evaluation on non-uniform distributions or real-world data, will also be of interest.

**Acknowledgments**

We thank the anonymous reviewers for helpful comments, HH acknowledges support from NSF IIS grants #2144413, #2052448, and #2107173. RV acknowledges support from DST INSPIRE grant no. DST/INSPIRE/04/2020/000107.

\(^1\)The assumption about uniformly random preferences is quite common in the literature on strategic aspects of stable matchings; see, for example, [Teo et al., 2001; Kojima et al., 2013; Immorlica and Mahdian, 2015; Aziz et al., 2015; Ashlagi et al., 2017].
References


