Approximation Algorithms for Stable Marriage Problems

A Dissertation
Presented to the Graduate School of Informatics
Kyoto University
in Candidacy for the Degree of
Doctor of Philosophy

by
Hiroki Yanagisawa
Kyoto University, Japan

Supervisor: Prof. Kazuo Iwama

Abstract

Approximation Algorithms for Stable Marriage Problems

Hiroki Yanagisawa

Kyoto University, Japan

2007

The stable marriage problem is a classical matching problem introduced by Gale and Shapley. An instance of the stable marriage problem consists of men and women, where each person totally orders all members of the opposite sex. A matching is *stable* if there is no pair that prefer each other to their current partners. The problem is to find a stable matching for a given instance. It is known for any instance, there exists a solution, and there is a polynomial time algorithm (Gale-Shapley algorithm) to find one.

In Chapter 3, we consider the quality of solutions for the stable marriage problem. The matching obtained by Gale-Shapley algorithm is man-optimal, that is, the matching is preferable for men but unpreferable for women, (or, if we exchange the role of men and women, the resulting matching is woman-optimal). The sex-equal stable marriage problem posed by Gusfield and Irving asks to find a stable matching "fair" for both genders, namely it asks to find a stable matching with the property that the sum of the men's score is as close as possible to that of the women's. This problem is known to be strongly NP-hard. We give a polynomial time algorithm for finding *a near optimal* solution in the sex-equal stable marriage problem. Furthermore, we consider the problem of optimizing additional criterion: among stable matchings that are near optimal in terms of the sex-equality, find a minimum egalitarian stable matching. We show that this problem is NP-hard, and give a polynomial time algorithm whose approximation ratio is less than two.

In Chapter 4 and Chapter 5, we consider general settings of the original stable marriage problem. While the original stable marriage problem requires all participants to rank all members of the opposite sex in a strict order, two natural variations

are to allow for incomplete preference lists and ties in the preferences. Either variation is polynomially solvable, but it was shown to be NP-hard to find a maximum cardinality stable matching when both of the variations are allowed. It is easy to see that in the generalized variant, the size of any two stable matchings differ by at most a factor of two, and so, an approximation algorithm with a factor two is trivial.

In Chapter 4, we give the first nontrivial result for approximation of factor less than two. Our algorithm achieves an approximation ratio of $2/(1+L^{-2})$ for instances in which only men have ties of length at most L. When both men and women are allowed to have ties, but the lengths are limited to two, we show a ratio of $13/7$ (< 1.858). We also improve the lower bound on the approximation ratio to $33/29$ (> 1.1379).

In Chapter 5, we give a randomized approximation algorithm and show that its expected approximation ratio is at most $10/7$ (< 1.4286) for a restricted but still NP-hard case, where ties occur in only men's lists, each man writes at most one tie, and the length of ties is two. We also show that our analysis is nearly tight by giving a lower bound $32/23$ (> 1.3913) for it. Furthermore, we show that these restrictions except for the last one can be removed without increasing the approximation ratio too much.

Contents

List of Figures

Acknowledgments

The author would like to express his attitude of gratitude toward his supervisor, Prof. Kazuo Iwama for offering him the opportunity to work on the topic of this thesis and his invaluable guidance throughout my studies in Kyoto University. His enthusiasm and passion for research and teaching was contagious and a source of inspiration.

The author would like to express his sincere appreciation to Prof. Shuichi Miyazaki for his guidance and support during the completion of this work. This thesis would not be completed without his valuable advice at any time of the day or night and without a lot of fruitful discussions with him.

The author is deeply grateful to thank Prof. Magnús M. Halldórsson at Iceland University for their valuable technical assistance in various segments of this work. The author would like to thank Prof. Masao Fukushima and Prof. Hiroshi Nagamochi for their proofreadings.

The author would like to thank all the members of Iwama laboratory for giving him helpful advice and heartfelt encouragements for his research work. The author would like to thank all of his friends for having a meaningful and enjoyable time together.

Finally, the author would like to acknowledge IBM Tokyo Research Laboratory for its financial support and wishes to thank his colleagues at the laboratory for encouraging him to study at the university to receive PhD.

Chapter 1

Introduction

1.1 Background

An instance I of the original *stable marriage problem* (*SM*) [GS62, GI89] consists of N men, N women, and each person's preference list. A preference list is a totally ordered list including all members of the opposite sex depending on his/her preference. For a matching M between men and women, a pair of a man m and a woman w is called a *blocking pair* if both prefer each other to their current partners. A matching with no blocking pair is called *stable*. The problem is to find a stable matching for a given instance. This problem was first studied by Gale and Shapley [GS62], who showed that every instance contains at least one stable matching, and gave an $O(N^2)$-time algorithm to find one, which is known as the Gale-Shapley algorithm.

The stable marriage problem has great practical significance. One of the best known applications is to assign medical students to hospitals based on the preferences of students over hospitals and vice versa, examples of which are NRMP in the US [GI89], CaRMS in Canada, SPA in Scotland [Irv98], and JRMP in Japan [JRMP]. Another application, reported in [TST99], is the assignment of students to secondary schools in Singapore.

In general, there are many different stable matchings for a single instance, and the Gale-Shapley algorithm finds only one of them (*man-optimal* or *woman-optimal*) with an extreme property: In the man-optimal stable matching, each man is matched with his best possible partner, while each woman gets her worst possible partner, among all stable matchings. Therefore, it is natural to try to obtain a matching which is not only stable but also "good" in some criterion.

There are three major optimization criteria for the quality of stable matchings: the *minimum regret* stable marriage problem, the *minimum egalitarian* stable marriage problem, and the *sex-equal* stable marriage problem. The minimum regret stable marriage problem is to find a stable matching which minimizes the largest regret among all participants. The minimum egalitarian stable marriage problem is to find a stable matching such that the sum of each person's score is smallest. The sex-equal stable marriage problem is to find a stable matching with the property that the sum of the men's score is as close as possible to that of the women's. Note that the number of stable matchings for one instance grows exponentially in general (see [IL86], e.g.). Nevertheless, for the first two problems, Gusfield [Gus87], and Irving, Leather and Gusfield [ILG87], respectively, proposed polynomial time algorithms by exploiting a lattice structure which is of polynomial size but contains information of all stable matchings. In contrast, it is hard to obtain a sex-equal stable matching. The question of its complexity was posed by Gusfield and Irving [GI89], and was later proved to be strongly NP-hard by Kato [Kat93]. Thus, the next step should be its approximability for which we have no knowledge so far. We discuss this problem in Chapter 3.

Next, we consider general settings of the original stable marriage problem, because strict rankings of all members may not be reasonable in practice. For instance, if there are many hosiptals in a students-hosiptals assignment, the students may have unacceptable hospitals or may have difficulties in ranking all the hospitals strictly.

One natural relaxation is to allow for indifference [GI89, Irv94], in which each person is allowed to include *ties* in his/her preference. This problem is denoted by *SMT* (Stable Marriage with Ties). When ties are allowed, the definition of stability needs to be extended. A man and a woman form a blocking pair if each *strictly* prefers the other to his/her current partner. A matching without such a blocking pair is called *weakly stable* (or simply "stable" in this thesis). Variations in which a blocking pair can involve non-strict preferences (referred as *super-stability* and *strong stability*) suffer from the fact that a stable matching may not exist, but there are polynomial-time algorithms to find such matchings when they exist [Irv94]. In contrast, a weakly stable matching always exists and the Gale-Shapley algorithm can be modified to find one [GI89].

Another natural variation is to allow participants to declare one or more *unacceptable partners*. If m appears in w's list, we say that m is *acceptable* to w, and vice versa. We refer to this problem as SMI (Stable Marriage with Incomplete lists). Since each person's preference list may be incomplete, matchings may not be necessarily perfect. Again, the definition of a blocking pair is extended: m and w form a blocking pair for M if the following three conditions are met. (i) m and w are not matched in M, but each is acceptable to the other, (ii) either m is single in M, or m prefers w to his partner in M, and (iii) either w is single in M, or w prefers m to her partner in M. A stable matching may not be perfect but it is known that all stable matchings for an SMI instance are of the same size [GS85], and again the Gale-Shapley algorithm can be modified to find a stable matching [GI89].

However, the situation changes if we allow both relaxations simultaneously, which is denoted by *SMTI* (Stable Marriage with Ties and Incomplete lists). In this case, one instance can have stable matchings of different size and the problem of finding a maximum stable matching (denoted by *MAX SMTI*) is NP-hard [IMMM99, MII+02, HIMM02]. This NP-hardness also holds for several restricted cases such as the case that all ties occur only in one sex, are of length two and every person's list contains at most one tie [MII+02]. The hardness result has been further extended to APX-hardness [HII+03]. Since a stable matching is a maximal matching, the sizes of any two stable matchings for an instance differ by a factor at most two. Hence, any stable matching is a 2-approximation; yet, there was no known approximation algorithm with approximation ratio strictly better than 2. In Chapter 4 and Chapter 5, we give deterministic and randomized approximation algorithms for restricted cases of MAX SMTI. The approximation algorithms for MAX SMTI has been studied recently [IMO06, IMY05], and the current best result is 1.875-approximation [IMY07] without any restriction.

1.2 Overview of this Thesis

In Chapter 3, we give a polynomial time algorithm for finding *a near optimal* solution in the sex-equal stable marriage problem. The algorithm exploits an underlying structure "rotation poset" (a partially-ordered set of rotations) of the stable marriage problem, which is originally defined in [IL86]. The main idea is to start from the man-optimal stable matching and find a set of rotations that turns it into an improved matching. To guarantee a polynomial time computation, we show that the number of rotations where we need to try all of its subsets is sufficiently small. Next, we consider the problem of optimizing additional criterion: among stable matchings that are near optimal in terms of the sex-equality, find a minimum egalitarian stable matching. We show that this problem is NP-hard by the reduction from the clique problem. Then, we give a polynomial time algorithm whose approximation ratio is less than two for this problem. This algorithm is an improved version of the former algorithm: the improved algorithm uses the minimum egalitarian stable matching as an initital solution, while the former algorithm uses the man-optimal stable matching as an initial solution.

In Chapter 4, we give nontrivial upper and lower bounds on approximating MAX SMTI. On the negative side, we show that the problem is hard to approximate within a factor of 33/29 (> 1.1379). We also show that the problem is hard to approximate within a factor of 21/19 (> 1.1052), even for a strongly restricted case. These results are obtained via an approximability relation with Minimum Vertex Cover. If the strong conjecture of $(2 - \epsilon)$-hardness for Minimum Vertex Cover holds, then our lower bound will be improved to 4/3 (> 1.3333) and 5/4 ($= 1.25$), respectively. On the positive side, we give an algorithm called ShiftBrk, which is based on the following simple idea. Suppose, for simplicity, that all ties are of the same length L (but the following results hold for the case where all ties are of length *at most* L). Then ShiftBrk first breaks all the ties into an arbitrary order and obtains a stable marriage instance *without ties*. Then we "shift" cyclically the order of all the originally tied women in each man's list simultaneously, creating L different instances. For each of them, we in turn apply the shift operation against the ties of women's lists, obtaining L^2 instances in total. We finally compute L^2 stable matchings for these L^2

instances in polynomial time, all of which are stable in the original instance [GI89], and select a largest solution. We prove the following: (i) ShiftBrk achieves an approximation ratio of $2/(1 + L^{-2})$ (1.6 and 1.8 when $L = 2$ and 3, respectively) if the given instance includes ties in only men's (or women's) lists. (Note that in this case, ShiftBrk constructs only L instances.) We also give a tight example for this analysis. (ii) When both men and women are allowed to include ties, but only of length two, in their preference lists, it achieves an approximation ratio of 13/7 (< 1.858).

In Chapter 5, we give a randomized approximation algorithm for MAX SMTI, which is based on the following simple idea: given an SMTI instance, obtain an SMI instance by breaking each tie in the SMTI instance with equal probability, compute a stable matching for the SMI instance, and output it. We show that its expected approximation ratio is at most 10/7 (< 1.4286) for a restricted but still NP-hard case, where ties occur in only men's lists, each man writes at most one tie, and the length of ties is two. We also show that our analysis is nearly tight by giving a lower bound 32/23 (> 1.3913) for our algorithm. Furthermore, we show that these restrictions except for the last one can be removed without increasing the approximation ratio too much.

Chapter 2

The Stable Marriage Problem

2.1 The Original Definition

An instance of the stable marriage problem consists of N men, N women, and each person's preference list. (Throughout this thesis, instances contain equal number N of men and women.) A preference list is a totally ordered list including all members of the opposite sex depending on his/her preference. The following instance I is an example for $N = 4$.

$$\begin{array}{llcll} m_1: & w_1\ w_2\ w_3\ w_4 & \quad & w_1: & m_2\ m_1\ m_3\ m_4 \\ m_2: & w_2\ w_3\ w_4\ w_1 & & w_2: & m_1\ m_2\ m_3\ m_4 \\ m_3: & w_4\ w_3\ w_2\ w_1 & & w_3: & m_3\ m_4\ m_2\ m_1 \\ m_4: & w_3\ w_2\ w_1\ w_4 & & w_4: & m_1\ m_2\ m_3\ m_4 \end{array}$$

Here, m_i's represent men and w_i's represent women (for $1 \leq i \leq 4$). Instance I shows that man m_1 prefers w_1 to w_2, w_2 to w_3, and w_3 to w_4. The other preference lists are written in a similar manner.

A *matching* is a set of disjoint pairs of a man and a woman (m, w). If a pair (m, w) is in a matching M, we say m is *matched* with w (w is matched with m) in M and we write $M(m) = w$ ($M(w) = m$). Given a matching M, a man m and a woman w are said to form a *blocking pair* for M (or simply "(m, w) *blocks* M") if m prefers w to $M(m)$ and w prefers m to $M(w)$. For example, a matching $M_1 = \{(m_1, w_1), (m_2, w_2), (m_3, w_3), (m_4, w_4)\}$ for I contains a blocking pair (m_3, w_4).

A matching is called *stable* if it contains no blocking pair. For example, $M_2 = \{(m_1, w_1), (m_2, w_2), (m_3, w_4), (m_4, w_3)\}$ is a stable matching for I. (In general, an instance has more than one stable matchings. For example, $M_3 = \{(m_1, w_2), (m_2, w_1), (m_3, w_4), (m_4, w_3)\}$ is another stable matching for I.)

2.2 Ties and Incomplete Lists

For SMTI (Stable Marriage with Ties and Incomplete lists), we extend some notations. A *matching* is a set of disjoint pairs of a man and a woman (m, w) such that m is acceptable to w and vice versa. If a man m (or a woman w) is not matched in M, we say m (w) is *single*. Given a matching M, a man m and a woman w are said to form a *blocking pair* for M (or simply "(m, w) *blocks* M") if all the following conditions are met:

1. m and w are not matched together in M but are acceptable to each other,
2. m is either single in M or prefers w to $M(m)$, and
3. w is either single in M or prefers m to $M(w)$.

A matching is called *weakly stable* (or simply *stable*) if it contains no blocking pair.

The following instance $\hat{I}$ is an example of SMTI instance.

m_1:	$w_2\ (w_3\ w_4)$	w_1:	m_4
m_2:	$w_2\ w_3$	w_2:	$m_2\ (m_1\ m_3)\ m_4$
m_3:	$w_2\ w_4$	w_3:	$m_4\ m_2\ m_1$
m_4:	$(w_1\ w_2\ w_4)\ w_3$	w_4:	$m_1\ m_4\ m_3$

Persons in parentheses show that they are tied, i.e., instance $\hat{I}$ shows that m_1 does not determine relative preferences between w_3 and w_4. Since m_1 does not include w_1 in his preference list, w_1 is not acceptable for m_1. (Throughout this thesis, we may assume without loss of generality that acceptability is mutual, i.e. that the occurrence of w in m's preference list implies the occurrence of m in w's list, and vice versa.) A matching $M_4 = \{(m_1, w_3), (m_2, w_2), (m_4, w_4)\}$ is one of the stable matchings for $\hat{I}$.

When we seek a stable matching for an instance of SMTI, we sometimes make an SMI instance by using a tie-breaking, because it is easy to see that a stable matching for the SMI instance obtained by any tie-breaking is also stable for the original SMTI instance (see [GI89] for example). The following instance I' is a result of breaking ties in $\hat{I}$.

$$
\begin{array}{llll}
m_1: & w_2\ [w_3\ w_4] & \qquad w_1: & m_4 \\
m_2: & w_2\ w_3 & \qquad w_2: & m_2\ [m_3\ m_1]\ m_4 \\
m_3: & w_2\ w_4 & \qquad w_3: & m_4\ m_2\ m_1 \\
m_4: & [w_1\ w_4\ w_2]\ w_3 & \qquad w_4: & m_1\ m_4\ m_3
\end{array}
$$

The brakets shows that there are ties in those places in $\hat{I}$. For example, while m_1 does not determine relative preference betwenn w_3 and w_4 in $\hat{I}$, m_1 strictly prefers w_3 to w_4 in I'.

Let M_1 and M_2 be two stable matchings for an SMTI instance $\hat{I}$. We sometimes use the bipartite graph G_{M_1,M_2} defined as follows: Each vertex of G_{M_1,M_2} is associated with a person in $\hat{I}$. We include an edge between vertices m and w if and only if m and w are matched in M_1 or M_2 (if they are matched in both, we include two edges between them; hence G_{M_1,M_2} is a multigraph). Observe that the degree of each vertex is then at most two, and each connected component of G_{M_1,M_2} is a simple path, a cycle or an isolated vertex.

2.3 Notations

If m strictly prefers w_i to w_j in an instance I, we write "$w_i \succ w_j$ in m's list in I," or simply, "$w_i \succ w_j$ in m's I-list." If the instance I and/or the man m are clear from the context, we often omit them.

Let $\hat{I}$ be an SMTI instance and let p be a person in $\hat{I}$ whose preference list contains a tie which includes persons $q_1, q_2, \cdots, q_k$. In this case, we write "$(\cdots q_1 \cdots q_2 \cdots q_k \cdots)$ in p's $\hat{I}$-list." Suppose that the tie $(\cdots q_1 \cdots q_2 \cdots q_k \cdots)$ in p's $\hat{I}$-list is broken into $\cdots q_1 \succ \cdots \succ q_2 \succ \cdots \succ q_k \cdots$ in I. Then we write "$[\cdots q_1 \cdots q_2 \cdots q_k \cdots]$ in p's I-list." That is, "$[\cdots q_1 \cdots q_2 \cdots]$ in p's I-list" means that q_1 precedes q_2 in p's list in I but they are tied in p's list in $\hat{I}$.

A goodness measure of an approximation algorithm is defined as usual: Let P be a maximization/minimization problem, and A be an approximation algorithm for P. A is said to be an $r(N)$-approximation algorithm if $\max\{opt(x)/A(x), A(x)/opt(x)\} \leq r(N)$ for all instances x of size N, where $opt(x)$ and $A(x)$ are the size of the optimal and A's solution for x, respectively. The problem P is said to be *approximable* within $r(N)$ if there is a polynomial-time $r(N)$-approximation algorithm for P. P is *NP-hard to approximate within* $r(N)$, if the existence of an $r(N)$-approximation algorithm for P implies P=NP.

Chapter 3

Approximation Algorithms for the Sex-Equal Stable Marriage Problem

3.1 Introduction

In this chapter, we consider the quality of a solution for the original stable marriage problem. There are three major optimization criteria for the quality of stable matchings. Let $p_m(w)$ ($p_w(m)$, respectively) denote the position of woman w in man m's preference list (the position of man m in woman w's preference list, respectively). For a stable matching M, define a *regret cost* $r(M)$ to be

$$r(M) = \max_{(m,w)\in M} \max\{p_m(w), p_w(m)\}.$$

Also, define an *egalitarian cost* $c(M)$ to be

$$c(M) = \sum_{(m,w)\in M} p_m(w) + \sum_{(m,w)\in M} p_w(m),$$

and a *sex-equalness cost* $d(M)$ to be

$$d(M) = \sum_{(m,w)\in M} p_m(w) - \sum_{(m,w)\in M} p_w(m).$$

The minimum regret stable marriage problem (*the minimum egalitarian stable marriage problem* and *the sex-equal stable marriage problem*, respectively) is to find a stable matching M with minimum $r(M)$ ($c(M)$ and $|d(M)|$, respectively) [GI89].

While there are polynomial time algorithm for the first two problems, the sex-equal stable marriage problem is NP-hard. Therefore, we consider finding *near optimal* solutions for the sex-equal stable marriage problem. Let M_0 and M_z be the man-optimal and the woman-optimal stable matchings, respectively. Note that $d(M_0) \leq d(M) \leq d(M_z)$ for any stable matching M (see Fig. 3.1). Our goal is to obtain a stable matching M such that $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$ for a given constant ϵ, where $\Delta = \min\{|d(M_0)|, |d(M_z)|\}$. Namely, we define the following problem called *Near SexEqual* (*NSE* for short). Given a stable marriage instance I and a positive constant ϵ, it asks to find a stable matching M such that $|d(M)| \leq \epsilon\Delta$ if such M exists, or answer "No" otherwise. We give a polynomial time algorithm for NSE, which runs in time $O(N^{3+\frac{1}{\epsilon}})$ where N is the number of men and women.

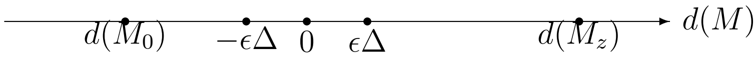


Figure 3.1: The sex-equalness costs of stable matchings

NSE asks to find an *arbitrary* stable matching whose sex-equalness cost lies within some range. However, we may want to find a good one if there are several solutions in the range. In fact, there is an instance I that has two stable matchings M and M' such that $d(M) = d(M') = 0$ but $c(M) \ll c(M')$. This motivates us to consider the following corresponding optimization problem MinESE (*Minimum Egalitarian Sex-Equal stable marriage problem*): Given a stable marriage instance I and a positive constant ϵ, find a stable matching M which minimizes $c(M)$ under the condition that $|d(M)| \leq \epsilon\Delta$, (or answer "No" if none exists). We show that MinESE is NP-hard, and give a polynomial time $(2 - (\epsilon - \delta)/(2 + 3\epsilon))$-approximation algorithm for an arbitrary δ such that $0 < \delta < \epsilon$, whose running time is $O(N^{4+\frac{1+\epsilon}{\delta}})$.

Our results also hold for the *weighted* versions of the above problems, in which $p_m(w)$ ($p_w(m)$, respectively) represents not simply a rank of w in m's preference list, but an arbitrary score of m for w (of w for m), where $p_m(w) > 0$ ($p_w(m) > 0$) and $p_m(w) < p_m(w')$ if and only if m prefers w to w' ($p_w(m) < p_w(m')$ if and only if w

prefers m to m') for all m and w.

Related Results. As mentioned above, the minimum regret stable marriage problem and the minimum egalitarian stable marriage problem can be solved in polynomial time [Gus87, ILG87, GI89], but the sex-equal stable marriage problem is strongly NP-hard [Kat93]. If we allow ties in preference lists, all these problems become hard, even to approximate, if we seek for optimal *weakly* stable matching: For each problem, there exists a positive constant δ such that there is no polynomial-time approximation algorithm with approximation ratio δn unless P=NP [HII+03].

3.2 Rotation Poset

In this section, we explain a *rotation poset (partially-ordered set)*, originally defined in [IL86], which is an underlying structure of stable matchings. Here, we give only a brief sketch necessary for understanding the algorithms given later. Readers can refer to [GI89] for further details.

We fix an instance I. Let M be a stable matching for I. For each such M, we can associate a *reduced list*, which is obtained from the original preference lists by removing entries by some rule. One property of the reduced list associated with M is that, in M, each man is matched with the first woman in the reduced list, and each woman is matched with the last man. A *rotation exposed in M* is an ordered list of pairs $\rho = (m_0, w_0), (m_1, w_1), \ldots, (m_{r-1}, w_{r-1})$ such that, for every i $(0 \leq i \leq r-1)$, m_i and w_i are matched in M, and w_{i+1} is at the second position in m_i's reduced list, where $i+1$ is taken modulo r. There exists at least one rotation for any stable matching except for the woman-optimal stable matching M_z.

For a stable matching M and a rotation $\rho = (m_0, w_0), (m_1, w_1), \ldots, (m_{r-1}, w_{r-1})$ exposed in M, *eliminating ρ from M* means to replace m_i's partner from w_i to w_{i+1} for each i $(0 \leq i \leq r-1)$, (and to update a reduced list accordingly). Note that by eliminating a rotation, men become worse off while women become better off. The resulting matching is denoted by M/ρ. It is well known that M/ρ is also stable for I. If a rotation is exposed in M/ρ, then we can similarly obtain another stable matching by eliminating it.

Now, let $\mathcal{M}$ be the set of all stable matchings for I, and Π be the set of rotations ρ such that ρ is exposed in some stable matching in $\mathcal{M}$. Then, it is known that $|\Pi| \leq N^2$. The *rotation poset* $(\Pi, \prec)$, which is uniquely determined for instance I, is the set Π with a partial order $\prec$ defined for elements in Π. For two rotations ρ_1 and ρ_2 in Π, $\rho_1 \prec \rho_2$ intuitively means that ρ_1 must be eliminated before ρ_2, or ρ_2 is never exposed until ρ_1 is eliminated. It is known that the rotation poset can be constructed in $O(N^2)$ time.

A *closed subset* R of the rotation poset $(\Pi, \prec)$ is a subset of Π such that if $\rho \in R$ and $\rho' \prec \rho$ then $\rho' \in R$. There is a one-to-one correspondence between $\mathcal{M}$ and the set of closed subsets of $(\Pi, \prec)$: Let R be a closed subset. Starting from the man-optimal stable matching M_0, if we eliminate all rotations in R successively in a proper order defined by $\prec$, then we can obtain a stable matching. Conversely, any stable matching can be obtained by this procedure for some closed subset. We denote the stable matching corresponding to a closed subset R by M_R. For simplicity, we sometimes write $c(R)$ and $d(R)$ instead of $c(M_R)$ and $d(M_R)$, respectively. Especially, the empty subset corresponds to the man-optimal stable matching M_0, and the set Π itself corresponds to the woman-optimal stable matching M_z. From M_0, if we eliminate all rotations according to the order $\prec$, then we eventually reach M_z.

For a rotation $\rho = (m_0, w_0), (m_1, w_1), \ldots, (m_{r-1}, w_{r-1})$, we define $w_c(\rho)$ and $w_d(\rho)$, which represent the cost change of egalitarian and sex-equalness, respectively, by eliminating ρ:

$$w_c(\rho) = \sum_{i=0}^{r-1}(p_{m_i}(w_{i+1}) - p_{m_i}(w_i)) + \sum_{i=0}^{r-1}(p_{w_i}(m_{i-1}) - p_{w_i}(m_i)),$$

$$w_d(\rho) = \sum_{i=0}^{r-1}(p_{m_i}(w_{i+1}) - p_{m_i}(w_i)) - \sum_{i=0}^{r-1}(p_{w_i}(m_{i-1}) - p_{w_i}(m_i)).$$

Here, note that $w_d(\rho) > 0$ for all ρ since by eliminating a rotation, some men become worse off, and a some women become better off, and other people remain matched with the same partners. Now, let ρ be a rotation exposed in a stable matching M. Then, it is obvious from the definition that $c(M/\rho) = c(M) + w_c(\rho)$ and $d(M/\rho) =$

$d(M) + w_d(\rho)$. Also, it is easy to see that for any closed subset R,

$$c(M_R) = c(M_0) + \sum_{\rho \in R} w_c(\rho) \quad \text{and} \quad d(M_R) = d(M_0) + \sum_{\rho \in R} w_d(\rho).$$

Hence, the minimum egalitarian stable marriage problem (the sex-equal stable marriage problem, respectively) is equivalent to the problem of finding a closed subset R such that $c(M_0) + \sum_{\rho \in R} w_c(\rho)$ ($|d(M_0) + \sum_{\rho \in R} w_d(\rho)|$, respectively) is minimum. For example, the algorithm for finding a minimum egalitarian stable matching in [ILG87] efficiently finds such R by exploiting network flow.

3.3 The Sex-Equal Stable Marriage Problem

Recall that M_0 is the man-optimal stable matching and M_z is the woman-optimal stable matching. Note that any stable matching M satisfies $d(M_0) \leq d(M) \leq d(M_z)$. Thus, this problem is trivial if $d(M_0) \geq 0$ or $d(M_z) \leq 0$, namely, if $d(M_0) \geq 0$, M_0 is optimal, while if $d(M_z) \leq 0$, M_z is optimal. Therefore, we consider the case where $d(M_0) < 0 < d(M_z)$. Recall that $\Delta = \min\{|d(M_0)|, |d(M_z)|\}$. In the following, we assume without loss of generality that $|d(M_0)| \leq |d(M_z)|$ since otherwise, we can exchange the role of men and women. Hence, $\Delta = \min\{|d(M_0)|, |d(M_z)|\} = |d(M_0)|$.

We first briefly give the underlying idea of our algorithm presented in this section. Recall that, for a given instance I and ϵ, we are to find a stable matching M such that $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$ if any. As an easy case, assume that all rotations ρ of I satisfy $w_d(\rho) \leq 2\epsilon\Delta$. Now, we construct a rotation poset $(\Pi, \prec)$ of I, and starting from M_0, we eliminate rotations in an order of any linear extension of $\prec$. Recall that by eliminating a rotation, the sex-equalness cost increases, but by at most $2\epsilon\Delta$ by assumption. Note that $d(M_0) < 0 < d(M_z)$, and recall that if we eliminate all rotations from M_0, we eventually reach M_z. Then, in this sequence, we certainly meet a desirable stable matching at some point.

However, this procedure fails if there is a rotation with large sex-equalness cost: If we eliminate such a rotation, then we may "jump" from M to M' such that $d(M) < -\epsilon\Delta$ and $d(M') > \epsilon\Delta$ even if there is a feasible solution. To resolve this problem, we will try all combinations of selecting such "large" rotations, and treat "small" rotations in the above manner. To evaluate the time complexity, we show that the number of large rotations is limited.

Before giving a description of our algorithm, we give a couple of notations. Let R be any (not necessarily closed) subset of a poset $(\Pi, \prec)$. Then $R_{\min} = R \cup \{\rho \mid \text{there exists a } \rho' \text{ such that } \rho' \in R \text{ and } \rho \prec \rho'\}$. That is, $R_{\min}$ is the minimal closed subset of Π satisfying $R_{\min} \supseteq R$. Similarly, $R_{\max} = R \cup \{\rho \mid \text{there exists a } \rho' \text{ such that } \rho' \in R \text{ and } \rho' \prec \rho\}$.

Algorithm 1

1. Construct the rotation poset $(\Pi, \prec)$.

2. Let R^L be the set of rotations ρ such that $w_d(\rho) > 2\epsilon\Delta$, and R^S be $\Pi \setminus R^L$.

3. For each set R in 2^{R^L} such that $|R| \leq \frac{1+\epsilon}{2\epsilon}$, do,

  (a) If $R_{\min} \cap (R^L \setminus R)_{\max} \neq \emptyset$, then go to **3** and choose the next R.

  (b) If $-\epsilon\Delta \leq d(R_{\min}) \leq \epsilon\Delta$, then output $M_{R_{\min}}$.

  (c) Fix an arbitrary order $\rho_1, \rho_2, \cdots, \rho_k \in R^S \setminus (R_{\min} \cup (R^L \setminus R)_{\max})$ which is consistent with $\prec$.

  (d) For $i = 1$ to k, if $-\epsilon\Delta \leq d(R_{\min} \cup \{\rho_1, \rho_2, \cdots, \rho_i\}) \leq \epsilon\Delta$, then output $M_{R_{\min} \cup \{\rho_1, \rho_2, \cdots, \rho_i\}}$ and halt.

4. Output "No," and halt.

Theorem 3.1 *There is an algorithm for NSE whose running time is $O(N^{3+\frac{1}{\epsilon}})$.*

Proof. **Correctness Proof.** Clearly, if there is no M such that $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$, then the algorithm answers "No." On the other hand, suppose that there is M_X such that $-\epsilon\Delta \leq d(M_X) \leq \epsilon\Delta$, where X is the set of rotations corresponding to M_X. Let $X^L = X \cap R^L$ and $X^S = X \cap R^S$. Then, $d(X^L) \leq d(M_X) \leq \epsilon\Delta$. Because $w_d(\rho) > 2\epsilon\Delta$ for any rotation $\rho \in X^L$, $|X^L| < \frac{d(X^L)-d(M_0)}{2\epsilon\Delta} \leq \frac{|d(M_0)|+\epsilon\Delta}{2\epsilon\Delta} = \frac{1+\epsilon}{2\epsilon}$. So, Algorithm 1 selects X^L at Step 3 as R, and we consider this particular execution of Step 3.

First, note that $d((X^L)_{\min}) \leq \epsilon\Delta$ since otherwise, $d(M_X) \geq d((X^L)_{\min}) > \epsilon\Delta$, a contradiction. If $-\epsilon\Delta \leq d((X^L)_{\min}) \leq \epsilon\Delta$, then Algorithm 1 outputs $M_{(X^L)_{\min}}$ at Step 3(b). Finally, suppose that $d((X^L)_{\min}) < -\epsilon\Delta$. Note that $d((X^L)_{\min} \cup \{\rho_1, \rho_2, \cdots, \rho_k\}) \geq d(M_X) \geq -\epsilon\Delta$ and that any rotation ρ_i $(1 \leq i \leq k)$ satisfies $w_d(\rho_i) \leq 2\epsilon\Delta$. Hence there must be j $(1 \leq j \leq k)$ such that $-\epsilon\Delta \leq d((X^L)_{\min} \cup \{\rho_1, \rho_2, \cdots, \rho_j\}) \leq \epsilon\Delta$.

Time Complexity. Steps 1 and 2 can be performed in $O(N^2)$. Inside the loop of Step 3 can be performed in $O(N^2)$ since the number of rotations is at most $O(N^2)$. Clearly, Step 4 can be performed in constant time.

We consider the number of repetitions of Step 3, i.e., the number of R satisfying the condition at Step 3. Let this number be t. Recall that the number of rotations is at most N^2 as mentioned in Sec. 3.2. So, $|R^L| \leq N^2$. Since $|R| \leq \frac{1+\epsilon}{2\epsilon}$,

$$t = \sum_{k=1}^{\lfloor \frac{1+\epsilon}{2\epsilon} \rfloor} \binom{N^2}{k} \leq \sum_{k=1}^{\lfloor \frac{1+\epsilon}{2\epsilon} \rfloor} \frac{(N^2)^{\lfloor \frac{1+\epsilon}{2\epsilon} \rfloor}}{k!} = O(N^{\frac{1+\epsilon}{\epsilon}}).$$

Hence the time complexity of Algorithm 1 is $O(N^2) \cdot O(N^{\frac{1+\epsilon}{\epsilon}}) = O(N^{3+\frac{1}{\epsilon}})$. □

Remark. Recall that Δ is defined as $\Delta = \min\{|d(M_0)|, |d(M_z)|\}$. However, we can modify Algorithm 1 so that it works even if Δ is defined as $\Delta = \min\{\min\{|d(M_0)|, |d(M_z)|\}, \frac{1}{\log N} \max\{|d(M_0)|, |d(M_z)|\}\}$. The idea is as follows. Note that M_0 and M_z can be found in polynomial time. First, we compare $\min\{|d(M_0)|, |d(M_z)|\}$ and $\frac{1}{\log N} \max\{|d(M_0)|, |d(M_z)|\}$. If the former is smaller, then we apply Algorithm 1. If the latter is smaller, then we modify Algorithm 1 so that it executes Step 3 for all subsets of 2^{R^L}, and apply it. Note that, from the discussion in Sec. 3.2,

$$d(M_z) = d(M_0) + \sum_{\rho \in \Pi} w_d(\rho).$$

Hence $|R^L| < \frac{d(M_z) - d(M_0)}{2\epsilon\Delta} \leq \frac{2 \max\{|d(M_0)|, |d(M_z)|\}}{2\epsilon\Delta} = \frac{\log N}{\epsilon}$. Thus the number of repetitions of Step 3 is at most $2^{|R^L|} = N^\epsilon$, which is polynomial. Note that the modified algorithm is stronger when $|d(M_0)|$ and $|d(M_z)|$ are close, more precisely, when they differ at most $\log N$ factor.

Remark. There are several goodness measures of an approximation algorithm A for a minimization problem. The usual measure is the *approximation ratio* of A, which is defined as $\max\{A(x)/opt(x)\}$ over all instances x, where $opt(x)$ and $A(x)$ are the costs of the optimal and the algorithm's solutions, respectively. However, this measure cannot be used for the sex-equal stable marriage problem, because $opt(x)$ can be zero. For such a case, there is another measure: the *relative accuracy* [CW04, Nes98], which is defined as $\max\{(\max(x) - opt(x))/(\max(x) - A(x))\}$ over all instances x, where $opt(x)$, $\max(x)$, and $A(x)$ are the cost of the optimal solution, the worst solution, and the algorithm's solution, respectively. By using the modified version of Algorithm 1 in the above remark, we can construct an algorithm which achieves the relative accuracy $1 + \epsilon/\log N$ for an arbitrary constant $\epsilon > 0$.

Let M_{opt} be an optimal solution for the sex-equal stable marriage problem. Recall that we are considering the case where $d(M_0) < 0 < d(M_z)$. If $|d(M_{opt})| > D/2$, where $D = \max\{|d(M_0)|, |d(M_z)|\}$, M_{opt} can be obtained in polynomial time in the following way: Let M_a be the stable matching such that no other stable matching M satisfies $d(M_a) < d(M) < -D/2$ and let M_b be the stable matching such that no other stable matching M satisfies $D/2 < d(M) < d(M_b)$. Then, M_{opt} is either M_a or M_b. Since $d(M_b) - d(M_a) > D$, there exists a rotation ρ_H such that $w_d(\rho_H) > D$ (otherwise there must be a stable matching M_c such that $d(M_a) < d(M_c) < d(M_b)$). Also, this ρ_H is unique because

$$\sum_{\rho \in \Pi} w_d(\rho) = d(M_z) - d(M_0) \leq 2D.$$

It is easy to see that the maximum closed subset which does not contain ρ_H corresponds to M_a and that the minimum closed subset which contains ρ_H corresponds to M_b. Thus, M_a and M_b can be obtained in polynomial time. Finally, assume that $|d(M_{opt})| \leq D/2$. For each i such that $i = 1, 2, \ldots, \left\lceil \frac{\log N}{\epsilon} \right\rceil$, we find a stable matching with sex-equalness cost between $-\frac{\epsilon}{2\log N}Di$ and $\frac{\epsilon}{2\log N}Di$ if any using modified version of Algorithm 1, and output the best one. Then, it is easy to see that an output matching M satisfies $|d(M)| - |d(M_{opt})| \leq \frac{\epsilon}{2\log N}D$. Now, the relative accuracy is

$$\frac{\max(x) - opt(x)}{\max(x) - T(x)} = 1 + \frac{T(x) - opt(x)}{\max(x) - T(x)} \leq 1 + \frac{(\epsilon/2\log N)D}{D/2} = 1 + \frac{\epsilon}{\log N}.$$

3.4 The Minimum Egalitarian Sex-Equal Stable Marriage Problem

In NSE, we are asked to find a stable matching whose sex-equalness cost is in some range close to 0. However, if there are several stable matchings satisfying the condition, there might be good ones and bad ones. In fact, there is an instance I that has two stable matchings M and M' whose sex-equalness costs are the same (0), but egalitarian costs are significantly different.

3.4.1 Construction of Instance I

I is constructed in the following steps. First consider the following instance I_1 consisting of $2n$ men and $2n$ women:

m_1:	$w_1\ w_{n+1}\ w_{n+2}\ \cdots\ w_{2n}\ w_2$	w_1:	$m_n\ m_1$
m_2:	$w_2\ w_{n+1}\ w_{n+2}\ \cdots\ w_{2n}\ w_3$	w_2:	$m_1\ m_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m_{n-1}:	$w_{n-1}\ w_{n+1}\ w_{n+2}\ \cdots\ w_{2n}\ w_n$	w_{n-1}:	$m_{n-2}\ m_{n-1}$
m_n:	$w_n\ w_{n+1}\ w_{n+2}\ \cdots\ w_{2n}\ w_1$	w_n:	$m_{n-1}\ m_n$
m_{n+1}:	w_{n+1}	w_{n+1}:	$m_{n+1}\ m_1\ \cdots\ m_n$
m_{n+2}:	w_{n+2}	w_{n+2}:	$m_{n+2}\ m_1\ \cdots\ m_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m_{2n-1}:	w_{2n-1}	w_{2n-1}:	$m_{2n-1}\ m_1\ \cdots\ m_n$
m_{2n}:	w_{2n}	w_{2n}:	$m_{2n}\ m_1\ \cdots\ m_n$

If a preference list is not complete, then add any missing persons at the tail of the list in an arbitrary order. Instance I_1 has two stable matchings: $M_1 = \{(m_1, w_1), (m_2, w_2), \ldots, (m_{n-1}, w_{n-1}), (m_n, w_n), (m_{n+1}, w_{n+1}), (m_{n+2}, w_{n+2}), \ldots, (m_{2n-1}, w_{2n-1}), (m_{2n}, w_{2n})\}$ and $M_2 = \{(m_1, w_2), (m_2, w_3), \ldots, (m_{n-1}, w_n), (m_n, w_1), (m_{n+1}, w_{n+1}), (m_{n+2}, w_{n+2}),$

$\ldots, (m_{2n-1}, w_{2n-1}), (m_{2n}, w_{2n})\}$. Note that $c(M_1) = 5n$, $d(M_1) = -n$, $c(M_2) = n^2 + 5n$, and $d(M_2) = n^2 + n$.

Let I_2 be the instance obtained from I_1 by exchanging men and women (see the following).

m_1:	$w_n\ w_1$	w_1:	$m_1\ m_{n+1}\ m_{n+2} \cdots m_{2n}\ m_2$
m_2:	$w_1\ w_2$	w_2:	$m_2\ m_{n+1}\ m_{n+2} \cdots m_{2n}\ m_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m_{n-1}:	$w_{n-2}\ w_{n-1}$	w_{n-1}:	$m_{n-1}\ m_{n+1}\ m_{n+2} \cdots m_{2n}\ m_n$
m_n:	$w_{n-1}\ w_n$	w_n:	$m_n\ m_{n+1}\ m_{n+2} \cdots m_{2n}\ m_1$
m_{n+1}:	$w_{n+1}\ w_1 \cdots w_n$	w_{n+1}:	m_{n+1}
m_{n+2}:	$w_{n+2}\ w_1 \cdots w_n$	w_{n+2}:	m_{n+2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m_{2n-1}:	$w_{2n-1}\ w_1 \cdots w_n$	w_{2n-1}:	m_{2n-1}
m_{2n}:	$w_{2n}\ w_1 \cdots w_n$	w_{2n}:	m_{2n}

Let I be the instance obtained by putting I_1 and I_2 together and padding missing persons at the tail of preference lists to make them complete. For more detail, the set of men of I is the union of the sets of men in I_1 and I_2, and the set of women of I is similar. The preference list of a man m in I, who came from I_1, is constructed by adding women in I_2 in an arbitrary order to the tail of the list of m in I_1. Preference lists of other people are constructed similarly. Then, I has four stable matchings M_3, M_4, M_5, and M_6, where $c(M_3) = 10n$, $d(M_3) = 0$, $c(M_4) = n^2 + 10n$, $d(M_4) = -n^2 - 2n$, $c(M_5) = n^2 + 10n$, $d(M_5) = n^2 + 2n$, $c(M_6) = 2n^2 + 10n$, and $d(M_6) = 0$. We see that $d(M_3) = d(M_6) = 0$, while $c(M_3)$ is small but $c(M_6)$ is large.

This motivates us to consider the following problem, *MinESE* (*the Minimum Egalitarian Sex-Equal stable marriage problem*): Given an instance I and a constant ϵ such that $0 < \epsilon < 1$, find a stable matching M with minimum $c(M)$, under the condition that $|d(M)| \leq \epsilon\Delta$, (or answer "No" if no such solution exists). First, in Sec. 3.4.2, we show that MinESE is NP-hard. Then, in Sec. 3.4.3, we give an approximation algorithm for MinESE.

3.4.2 NP-hardness of MinESE

It turned out that there is a polynomial-time algorithm for obtaining a stable matching M such that (a) $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$ *or* (b) $c(M)$ is minimum. Interestingly, it is hard to obtain M satisfying (a) *and* (b).

Theorem 3.2 *MinESE is NP-hard.*

Proof. We will prove the theorem by a reduction from a variant of the clique problem. First, we give the definition of the clique problem.

The clique problem with a specified vertex

Input: A graph $G(V, E)$, a vertex $v_* \in V$, and a positive integer k.

Output: "Yes," if $G(V, E)$ has a clique of size k containing v_*.

"No," otherwise.

This problem is NP-complete since we can test if $G'(V \cup \{v_*\}, E)$ has a $k+1$-clique using an algorithm for this problem, where G' is made by adding a vertex v_* and the edges between v_* and all the vertices in V to a graph $G(V, E)$ for a clique problem. Also, it is NP-complete even if we restrict the problem so that $G(V, E)$ is d-regular with $d > 2k^2$ (by adding dummy vertices and edges). In the following, we assume that the given graph satisfies this property. First, we construct a poset from the clique problem. Given a graph $G = (V, E)$, $v_* \in V$, and an integer k, we construct a poset $(\Pi, \prec)$ in a similar manner as the construction in [JN83]. Let Π be $V \cup E \cup \{\rho_+\}$, where ρ_+ is an additional element. Define the precedence relation $\prec$ as follows: $\rho_+ \prec v$ for all $v \in V$, and $v \prec e$ if and only if $v \in V$ is incident to $e \in E$ in $G(V, E)$. Then, ρ_+ has outdegree $|V|$, each element $v \in V$ has outdegree d and indegree 1, and each element $e \in E$ has indegree 2. Note that, to construct a non-empty closed subset, we need to choose ρ_+, and if we select k elements from V, we can take at most $k(k-1)/2$ elements from E.

Next, we construct a MinESE instance I' from the poset $(V \cup E \cup \{\rho_+\}, \prec)$ obtained by the above construction. We first construct I from $(V \cup E \cup \{\rho_+\}, \prec)$

using the same construction as [Kat93]. The obtained stable marriage instance I has the following properties:

(a) The rotation poset of I is exactly $(V \cup E \cup \{\rho_+\}, \prec)$.

(b) The rotation ρ_+ involves $|V|+1$ men and $|V|+1$ women.

(c) Let ρ_v be a rotation corresponding to $v \in V$. Then, ρ_v involves $d+1$ men and $d+1$ women.

(d) Let ρ_e be a rotation corresponding to $e \in E$. Then, ρ_e involves three men and three women.

(e) For each rotation $\rho = \{(m_1, w_1), \ldots, (m_r, w_r)\}$ of I and for each i, w_{i+1} is the next to w_i in m_i's list and m_{i-1} is the next to m_i in w_i's list, where $i+1$ and $i-1$ are taken modulo r.

It should be noted that $w_c(\rho) = 0$ for all rotations ρ in I by property (e). Then, we modify I and construct I' so that the following holds in I':

(1) $w_c(\rho_+) = 0$ and $w_d(\rho_+) = ((1+\epsilon)D - 2B)/(1-\epsilon)$, where $B = 2dk + 5k^2 - 3k - 1$ and $D = (2d+k+1)|V| + 8|E| - 1$.

(2) For any rotation ρ_v corresponding to $v \in V \setminus \{v_*\}$, $w_c(\rho_v) = k-1$ and $w_d(\rho_v) = 2d+k+1$.

(3) For the rotation ρ_{v_*} corresponding to $v_* \in V$, $w_c(\rho_{v_*}) = k-2$ and $w_d(\rho_{v_*}) = 2d+k$.

(4) For any rotation ρ_e corresponding to $e \in E$, $w_c(\rho_e) = -2$ and $w_d(\rho_e) = 8$.

(5) The man-optimal stable matching M_0 has the sex-equalness cost $-(D-B)/(1-\epsilon)$, namely, $d(M_0) = -(D-B)/(1-\epsilon)$. Here, ϵ is any constant such that $0 < \epsilon < 1$, which is given as an input of MinESE (note that B and D are defined in (1)).

The first four conditions (1), (2), (3), and (4) can be satisfied by padding "dummy" persons in preference lists as in [Kat93]: We pad $((1+\epsilon)D-2B)/2(1-\epsilon)-(|V|+1)$ (dummy) women between w_1 and w_2 in m_1's list in I' and $((1+\epsilon)D-2B)/2(1-\epsilon)-(|V|+1)$ (dummy) men between m_1 and m_2 in w_1's list in I' for rotation $\rho_+ = (m_1, w_1), \ldots, (m_{|V|+1}, w_{|V|+1})$. For each rotation $\rho_v = (m_1, w_1), \ldots, (m_{d+1}, w_{d+1})$ corresponding to $v \in V \setminus \{v_*\}$, we pad $k-1$ (dummy) women between w_1 and w_2 in m_1's list in I'. Then $w_c(\rho_v) = k-1$ and $w_d(\rho_v) = 2d+k+1$ as required. For the rotation $\rho_{v_*} = (m_1, w_1), \ldots, (m_{d+1}, w_{d+1})$ corresponding to $v_* \in V$, we pad $k-2$ (dummy) women between w_1 and w_2 in m_1's list, resulting that $w_c(\rho_{v_*}) = k-2$ and $w_d(\rho_{v_*}) = 2d+k$ in I'. Similarly, for each rotation $\rho_e = (m_1, w_1), (m_2, w_2), (m_3, w_3)$ corresponding to $e \in E$, we pad 2 (dummy) men between m_1 and m_2 in w_1's list, resulting that $w_c(\rho_e) = -2$ and $w_d(\rho_e) = 8$ in I'. The last condition (5) can be satisfied by adding "unbalanced" pairs to I appropriately (similarly as [Kat93]). The produced instance of MinESE is (I', ϵ). Note that these construction can be done in polynomial time. It is easy to observe that in the instance I', $d(M_z) = d(M_0) + w_d(\rho_+) + D = (D-B)/(1-\epsilon)$. Hence, $|d(M_0)| = |d(M_z)|$ and $\epsilon\Delta = \epsilon \min\{|d(M_0)|, |d(M_z)|\} = \epsilon|d(M_0)| = \epsilon(D-B)/(1-\epsilon) = d(M_0) + w_d(\rho_+) + B$.

We will show that G has a k-clique which contains v_* if and only if there is a stable matching M in I' such that $-\epsilon\Delta \le d(M) \le \epsilon\Delta$ and $c(M) < c(M_0)$. If this is true, we can show that MinESE is NP-hard as follows: Given an instance G of the clique problem with a specified vertex, we construct a MinESE instance I' and ϵ by the above reduction. Then, we find an optimal solution M and the man-optimal stable matching M_0. Finally, we compare $c(M_0)$ and $c(M)$: If $c(M) < c(M_0)$, the answer to the clique problem is "yes," otherwise, "no."

We first show "only if" part. Suppose that G has a k-clique C which contains v_*. In the rotation poset of I', let R be the set of rotations corresponding to k vertices and $k(k-1)/2$ edges of C. Then, $R' = R \cup \{\rho_+\}$ is a closed subset and $c(R') = 0 + (k-1)(k-1) + (k-2) - 2k(k-1)/2 = -1$. Hence $c(M) = c(M_0) + c(R') < c(M_0)$. $d(R') = w_d(\rho_+) + (k-1)(2d+k+1) + (2d+k) + 8k(k-1)/2 = w_d(\rho_+) + B$, so $d(M) = d(M_0) + d(R') = \epsilon\Delta$.

We then show "if" part. Suppose that G does not have a k-clique containing

v_*. Let M be any stable matching of I' such that $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$, and R be the set of rotations of I' corresponding to M. First, note that R contains ρ_+, since otherwise, $R = \emptyset$ and so $d(M) = d(M_0) = -\Delta$, which contradicts the assumption that $d(M) \geq -\epsilon\Delta$. Let v_M and e_M be the numbers of the rotations in R which correspond to elements in V and E, respectively. Then, $d(R) \geq w_d(\rho_+) + (2d+k+1)(v_M-1) + (2d+k) + 8e_M$. If $v_M > k$, $d(R) > w_d(\rho_+) + B$ since $d > 2k^2$. Hence, $d(M) = d(M_0) + d(R) > \epsilon\Delta$, again contradicting the assumption. Therefore, v_M must satisfy $0 \leq v_M \leq k$. Recall that $e_M \leq v_M(v_M-1)/2$. First, suppose that $v_M = 0$. Then, $e_M = 0$ and $c(M) = c(M_0)$. Next, suppose that $1 \leq v_M \leq k-1$. Then, $c(M) \geq c(M_0) + (k-1)(v_M-1) + (k-2) - 2e_M \geq c(M_0) + v_M(k-v_M) - 1 \geq c(M_0)$. Finally, suppose that $v_M = k$. If R does not contain ρ_{v_*}, then $c(M) \geq c(M_0) + (k-1)v_M - 2e_M \geq c(M_0)$. If R contains ρ_{v_*}, then $e_M < v_M(v_M-1)/2$, because otherwise, R corresponds to a k-clique containing v_*, a contradiction. Then, we can similarly show that $c(M) \geq c(M_0)$. □

Remark. Note that the reduction in the NP-hardness proof produces an instance (I, ϵ) of MinESE such that $|d(M_0)| = |d(M_z)|$ in I, and ϵ is any constant such that $0 < \epsilon < 1$. Observe that if $|d(M_0)| = |d(M_z)|$ and $\epsilon = 1$, then MinESE is equivalent to the minimum egalitarian stable marriage problem, which can be solved in polynomial time.

Remark. We can modify the reduction so that it preserves the gap of the Dense Subgraph Problem (DSP) (see [FPK01, AHI02], e.g.). Although some PTASs are known for DSP in some settings of parameters, existence of PTAS is not known for general case. Feige [Fei02] and Khot [Kho04] provided evidence that DSP may be hard to approximate within some constant factor. Therefore, we conjecture that MinESE either does not have a PTAS.

3.4.3 Approximation Algorithms for MinESE

Here, we give a $(2 - (\epsilon - \delta)/(2 + 3\epsilon))$-approximation algorithm for MinESE for an arbitrary δ such that $0 < \delta < \epsilon$. Similarly as Sec. 3.3, we assume that $|d(M_0)| \leq |d(M_z)|$. In this section, we prove two simple but important lemmas that link the egalitarian cost and the sex-equalness cost, whose proofs are given later. (i) For any stable matching M, $|d(M)| < c(M)$ (Lemma 3.1). (ii) For any stable matching M and a rotation ρ exposed in M, by eliminating ρ from M, the cost change in the egalitarian cost is at most the cost change in the sex-equalness cost (Lemma 3.2).

To illustrate an idea of the algorithm, we first consider a restricted case and show that our algorithm achieves 2-approximation. For a fixed $\delta > 0$, suppose that all rotations satisfy $w_d(\rho) \leq \delta\Delta$. Given I and ϵ, we first find a minimum egalitarian stable matching M_{eg}, which can be done in polynomial time. If $-\epsilon\Delta \leq d(M_{eg}) \leq \epsilon\Delta$, then we are done since M_{eg} is an optimal solution for MinESE. If $d(M_{eg}) < -\epsilon\Delta$, then we eliminate rotations one by one as Algorithm 1 until the sex-equalness cost first becomes $-\epsilon\Delta$ or larger. If $d(M_{eg}) > \epsilon\Delta$, then we "add" rotations one by one until the sex-equalness cost first becomes $\epsilon\Delta$ or smaller. Here, "adding a rotation" means the reverse operation of eliminating a rotation. If we do not reach a feasible solution by this procedure, then we can conclude that there is no feasible solution, by a similar argument as in Sec. 3.3. If we find a stable matching M such that $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$, then we can show that this is a 2-approximation, namely, $c(M) \leq 2c(M_{eg})$ using (i) and (ii) above (note that the optimal cost is at least $c(M_{eg})$): Suppose, for example, that $d(M_{eg}) < -\epsilon\Delta$ (see Fig. 3.2). Then, by (ii), $c(M) - c(M_{eg}) \leq d(M) - d(M_{eg})$, and by (i), $|d(M_{eg})| < c(M_{eg})$. Also, since the costs of rotations are at most $\delta\Delta$, and since M is the first feasible solution found by this procedure, $d(M) \leq -(\epsilon - \delta)\Delta < 0$. Putting these together, we have that $c(M)/c(M_{eg}) < 2$.

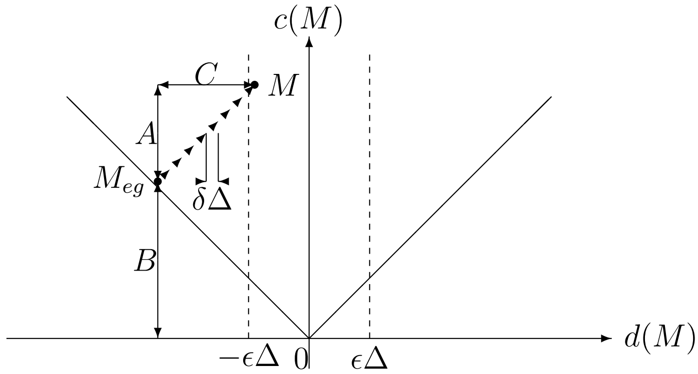


Figure 3.2: $C \leq B$ by (i) and $A \leq C$ by (ii). Hence $A + B \leq C + B \leq 2B$.

However, we may have rotations of large costs. Then we take a similar approach as in Sec. 3.3: Let R^L be the set of such large rotations. Then, for any partition R_1 and R_2 of R^L ($R_1 \cup R_2 = R^L$ and $R_1 \cap R_2 = \emptyset$), we want to obtain a minimum egalitarian stable matching whose corresponding closed subset A contains all rotations in R_1 but none in R_2. For this purpose, we need to solve the following problem: Given an instance I and disjoint subsets of rotations R_1 and R_2 of R^L, find a minimum egalitarian stable matching M_A under the condition that the corresponding closed subset A satisfies $A \supseteq R_1$ and $A \cap R_2 = \emptyset$. For this problem, we can use the same algorithm for the minimum egalitarian stable marriage problem in [GI89]. We denote this procedure by minEgalitarian(R_1, R_2). First, we review the following proposition described in [GI89]:

Proposition 3.3 *[GI89] Given a poset* $(\Pi, \prec)$*, there is an* $O(N^4)$*-time algorithm which finds a minimum-weight closed subset of* $(\Pi, \prec)$ *with respect to the egalitarian cost.*

Our procedure minEgalitarian(R_1, R_2) is as follows: Without loss of generality, assume that there are no elements such that $r_2 \prec r_1$ ($r_1 \in R_1$ and $r_2 \in R_2$) since there exists no solution in such a case. Construct the poset $(\Pi', \prec)$ by removing all the rotations in $(R_1)_{\min}$ and $(R_2)_{\max}$ from $(\Pi, \prec)$ (recall the definitions of $R_{\min}$ and $R_{\max}$ given before Algorithm 1), and let R' be the subset obtained by using Proposition 3.3 to $(\Pi', \prec)$. Then, it is easy to see that $(R_1)_{\min} \cup R'$ is an optimal solution for minEgalitarian(R_1, R_2). Now, we are ready to give the algorithm for MinESE.

Algorithm 2

1. Construct the rotation poset $(\Pi, \prec)$.

2. Let $M_{best} = NULL$.

3. Let R^L be the set of rotations ρ such that $w_d(\rho) > \delta\Delta$, and R^S be $\Pi \setminus R^L$.

4. For each set R in 2^{R^L} such that $|R| \leq \frac{1+\epsilon}{\delta}$, do,

(a) Let $A = $ minEgalitarian$(R, R^L \setminus R)$. If $d(A) < -\epsilon\Delta$, go to **(b)**. If $d(A) > \epsilon\Delta$, go to **(c)**. If $-\epsilon\Delta \leq d(A) \leq \epsilon\Delta$, go to **(d)**.

(b) Fix an arbitrary order $\rho_1, \rho_2, \cdots, \rho_k \in R^S \setminus (A \cup (R^L \setminus R)_{\max})$ which

---

is consistent with $\prec$.

For $i = 1$ to k, if $-\epsilon\Delta \leq d(A \cup \{\rho_1, \rho_2, \cdots, \rho_i\}) \leq \epsilon\Delta$, then let $A = A \cup \{\rho_1, \rho_2, \cdots, \rho_i\}$ and go to **(d)**.

(c) Fix an arbitrary order $\rho_1, \rho_2, \cdots, \rho_k \in (A \cap R^S) \setminus R_{\min}$ which is consistent with $\prec$.

For $i = k$ to 1, if $-\epsilon\Delta \leq d(A \setminus \{\rho_i, \rho_{i+1}, \cdots, \rho_k\}) \leq \epsilon\Delta$, then let $A = A \setminus \{\rho_i, \rho_{i+1}, \cdots, \rho_k\}$ and go to **(d)**.

(d) If $c(A) < c(M_{best})$, then let $M_{best} = M_A$.

5. If $M_{best} \neq NULL$, then output M_{best}, otherwise output "No," and halt.

Theorem 3.4 *There is a* $(2-(\epsilon-\delta)/(2+3\epsilon))$*-approximation algorithm for MinESE whose running time is* $O(N^{4+\frac{1+\epsilon}{\delta}})$ *for an arbitrary* δ *such that* $0 < \delta < \epsilon$.

Proof. **Correctness Proof.** Clearly, if there is no M such that $-\epsilon\Delta \leq d(M) \leq \epsilon\Delta$, then the algorithm answers "No." On the other hand, suppose that there is a feasible solution, and let M_{opt} be an optimal solution. We first show that Algorithm 2 finds a feasible solution. Let OPT be the rotation set corresponding to M_{opt}, and $OPT^L = OPT \cap R^L$. Then, $d(OPT^L) \leq d(M_{opt}) \leq \epsilon\Delta$. Because $w_d(\rho) > \delta\Delta$ for any rotation $\rho \in OPT^L$, $|OPT^L| < \frac{d(OPT^L)-d(M_0)}{\delta\Delta} \leq \frac{|d(M_0)|+\epsilon\Delta}{\delta\Delta} = \frac{1+\epsilon}{\delta}$. So, Algorithm 2 selects OPT^L at Step 4 as R, and we consider this particular execution of Step 4. We show that in this execution, Algorithm 2 finds a feasible solution. Let $A_{opt} = \text{minEgalitarian}(OPT^L, R^L \setminus OPT^L)$. There are three cases:

(i) $\boldsymbol{-\epsilon\Delta \leq d(A_{opt}) \leq \epsilon\Delta}$**.** $M_{A_{opt}}$ is selected as M_{best} at Step 4(d).

(ii) $\boldsymbol{d(A_{opt}) < -\epsilon\Delta}$**.** Note that $d(A_{opt} \cup \{\rho_1, \rho_2, \cdots, \rho_k\}) \geq d(M_{opt}) \geq -\epsilon\Delta$ and that any rotation ρ_i $(1 \leq i \leq k)$ satisfies $w_d(\rho_i) \leq \delta\Delta$. Hence there must be j $(1 \leq j \leq k)$ such that $-\epsilon\Delta \leq d(A_{opt} \cup \{\rho_1, \rho_2, \cdots, \rho_j\}) \leq -(\epsilon-\delta)\Delta$. (See Fig. 3.3.)

(iii) $\boldsymbol{d(A_{opt}) > \epsilon\Delta}$**.** Note that $d(A_{opt} \setminus \{\rho_1, \rho_2, \cdots, \rho_k\}) \leq d(M_{opt}) \leq \epsilon\Delta$ and that any rotation ρ_i $(1 \leq i \leq k)$ satisfies $w_d(\rho_i) \leq \delta\Delta$. Hence there must be j $(1 \leq j \leq k)$ such that $(\epsilon-\delta)\Delta \leq d(A_{opt} \setminus \{\rho_j, \rho_{j+1}, \cdots, \rho_k\}) \leq \epsilon\Delta$.

Next, we analyze the approximation ratio. Let M^* be the matching found in this particular execution of Step 4. We show that $c(M^*) \leq (2-(\epsilon-\delta)/(2+3\epsilon))c(M_{opt})$, which gives a proof for the approximation ratio. We first prove the following two lemmas:

Lemma 3.1 *For any stable matching M, $|d(M)| < c(M)$.*

Proof. If $d(M) \geq 0$, then $c(M) - |d(M)| = 2\sum_{(m,w)\in M} p_w(m) > 0$. Otherwise, $c(M) - |d(M)| = 2\sum_{(m,w)\in M} p_m(w) > 0$. □

Lemma 3.2 *Let $R = \{\rho_1, \ldots, \rho_{r-1}\}$ be a set of rotations and let $M_1, \cdots, M_r$ be stable matchings such that $M_{i+1} = M_i/\rho_i$ for $1 \leq i < r$. Then, $|c(M_r) - c(M_1)| \leq d(M_r) - d(M_1)$.*

Proof. Suppose that $m = M_i(w) = M_{i+1}(w')$ and $w = M_i(m) = M_{i+1}(m')$ for a fixed i. By the properties of the rotation [GI89], m prefers w to w' and w prefers m' to m. Let $d(m) = p_m(w') - p_m(w)$ and $d(w) = p_w(m) - p_w(m')$. Then $d(m) \geq 0$ and $d(w) \geq 0$, and it follows that

$$|c(M_{i+1}) - c(M_i)| = \left|\sum_m d(m) - \sum_w d(w)\right| \leq \left|\sum_m d(m) + \sum_w d(w)\right| = d(M_{i+1}) - d(M_i).$$

By summing up the above inequality for all i, we have

$$|c(M_r) - c(M_1)| \leq \sum_{i=1}^{r-1} |c(M_{i+1}) - c(M_i)| \leq \sum_{i=1}^{r-1} (d(M_{i+1}) - d(M_i)) = d(M_r) - d(M_1).$$

□

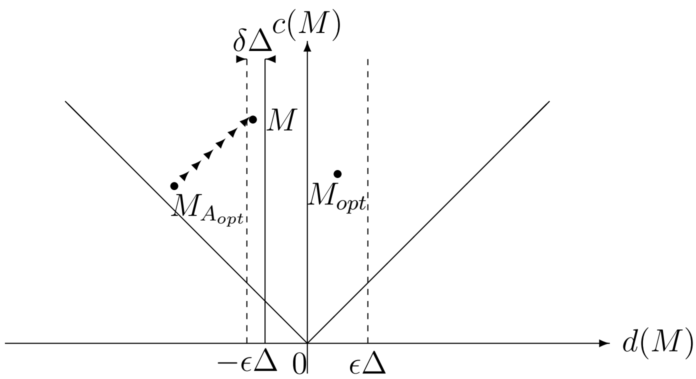


Figure 3.3: Finding a feasible solution

Note that $A_{opt} = \text{minEgalitarian}(OPT^L, R^L \setminus OPT^L)$. So, $c(A_{opt}) \le c(M_{opt})$ since OPT, the rotation set correspoding to M_{opt}, is one of the candidates for A_{opt}. We will consider the following four cases (note that $d(A_{opt}) \ge -\Delta$ for any stable matching M):

Case (i): $-\epsilon\Delta \le d(A_{opt}) \le \epsilon\Delta$. In this case, $M^* = M_{A_{opt}}$, which is an optimal solution since $c(A_{opt}) = c(M_{opt})$.

Case (ii): $\epsilon\Delta < d(A_{opt}) \le (2+3\epsilon)\Delta$. In this case, Step 4(b) of Algorithm 2 is executed. We have $|c(A_{opt}) - c(M^*)| \le d(A_{opt}) - d(M^*)$ by Lemma 3.2. Since $d(M^*) \ge (\epsilon - \delta)\Delta$ and (ii) hold, $|c(A_{opt}) - c(M^*)| \le (1 - (\epsilon - \delta)/(2+3\epsilon))d(A_{opt})$. Since $|d(A_{opt})| < c(A_{opt})$ by Lemma 3.1 and $c(A_{opt}) \le c(M_{opt})$, $c(M^*) < (2 - (\epsilon - \delta)/(2+3\epsilon))c(M_{opt})$.

Case (iii): $(2+3\epsilon)\Delta < d(A_{opt})$. Since both M_{opt} and M^* can be obtained by repeatedly eliminating rotations from M_0, $|c(M_{opt}) - c(M_0)| \le d(M_{opt}) - d(M_0)$ and $|c(M^*) - c(M_0)| \le d(M^*) - d(M_0)$ by Lemma 3.2. Since both $d(M_{opt})$ and $d(M^*)$ are at most $\epsilon\Delta$, $c(M^*) - c(M_{opt}) \le 2(1+\epsilon)\Delta$ (note that $|d(M_0)| = \Delta$). It follows that $c(M^*) - c(M_{opt}) \le 2(1+\epsilon)d(A_{opt})/(2+3\epsilon) = (2 - \epsilon/(2+3\epsilon))d(A_{opt})$. Since we have $|d(A_{opt})| < c(A_{opt})$ by Lemma 3.1 and $c(A_{opt}) \le c(M_{opt})$, $c(M^*) \le (2 - \epsilon/(2+3\epsilon))c(M_{opt})$.

Case (iv): $-\Delta \le d(A_{opt}) < -\epsilon\Delta$. The same as Case (ii).

Time Complexity. Steps 1, 2, 3, and 5 can be executed in $O(N^2)$ time. Step 4(a) is performed in the same time complexity as finding a minimum egalitarian stable matching, namely, $O(N^4)$. We can see that Steps 4(b) through 4(d) can be performed in time $O(N^2)$ by a similar analysis of Algorithm 1. The number of repetitions of Step 4 can be analyzed in the same way as the proof of Theorem 3.1, which is $O(N^{\frac{1+\epsilon}{\delta}})$. Hence the time complexity of Algorithm 2 is $O(N^{4+\frac{1+\epsilon}{\delta}})$. □

3.5 Concluding Remarks

In this chapter, we gave a polynomial time algorithm for finding near optimal sex-equal stable matching. Furthermore, we proved NP-hardness and developed a polynomial time approximation algorithm whose approximation ratio is less than 2 for

MinESE. Our future work is to improve the approximation ratio of MinESE.

Chapter 4

Improved Approximation Results for the Stable Marriage Problem

4.1 Introduction

In this chapter, we give nontrivial upper and lower bounds on approximating MAX SMTI.

On the negative side, we show that the problem is hard to approximate within a factor of 33/29 (> 1.1379). We also show that the problem is hard to approximate within a factor of 21/19 (> 1.1052), even for a strongly restricted case. These results are obtained via an approximability relation with Minimum Vertex Cover. If the strong conjecture of $(2 - \epsilon)$-hardness for Minimum Vertex Cover holds, then our lower bound will be improved to 4/3 (> 1.3333) and 5/4 ($= 1.25$), respectively.

On the positive side, we give an approximation algorithm ShiftBrk, where an input instance contains ties of lengths at most L. We prove the following: (i) ShiftBrk achieves an approximation ratio of $2/(1 + L^{-2})$ (1.6 and 1.8 when $L = 2$ and 3, respectively) if the given instance includes ties in only men's (or women's) lists. (Note that in this case, ShiftBrk constructs only L instances.) We also give a tight example for this analysis. (ii) When both men and women are allowed to include ties, but only of length two, in their preference lists, it achieves an approximation ratio of 13/7 (< 1.858). We conjecture that ShiftBrk also achieves a factor of less than two for general instances with $L \geq 3$.

4.2 Inapproximability Results

In this section, we obtain a lower bound on the approximation ratio of MAX SMTI using a reduction from the Minimum Vertex Cover problem (MVC for short). Let $G = (V, E)$ be a graph. A vertex cover C for G is a set of vertices in G such that every edge in E has at least one endpoint in C. MVC is to find, for a given graph G, a vertex cover with the minimum number of vertices, which is denoted by $VC(G)$. Dinur and Safra [DS05] gave an improved lower bound of $10\sqrt{5} - 21$ on the approximation ratio of MVC using the following proposition with $p = \frac{3-\sqrt{5}}{2} - \delta$ for arbitrarily small δ. We shall, however, see that the value $p = 1/3$ is optimal for our purposes.

Proposition 4.1 *[DS05] For any $\epsilon > 0$ and $p < \frac{3-\sqrt{5}}{2}$, the following holds: If there is a polynomial-time algorithm that, given a graph $G = (V, E)$, distinguishes between the following two cases, then P=NP.*

(1) $|VC(G)| \leq (1 - p + \epsilon)|V|$.
(2) $|VC(G)| > (1 - \max\{p^2, 4p^3 - 3p^4\} - \epsilon)|V|$.

For a MAX SMTI instance $\hat{I}$, let $OPT(\hat{I})$ be a maximum cardinality stable matching and $|OPT(\hat{I})|$ be its size.

Theorem 4.2 *For any $\epsilon > 0$ and $p < \frac{3-\sqrt{5}}{2}$, the following holds: If there is a polynomial-time algorithm that, given a MAX SMTI instance $\hat{I}$ of size N, distinguishes between the following two cases, then P=NP.*

(1) $|OPT(\hat{I})| \geq \frac{2+p-\epsilon}{3} N$.
(2) $|OPT(\hat{I})| < \frac{2+\max\{p^2, 4p^3-3p^4\}+\epsilon}{3} N$.

Proof. Given a graph $G = (V, E)$, we will construct, in polynomial time, an SMTI instance $\hat{I}(G)$ with N men and N women. Our reduction satisfies the following two conditions: (i) $N = 3|V|$, and (ii) $|OPT(\hat{I}(G))| = 3|V| - |VC(G)|$. Then, it is not hard to see that Proposition 4.1 implies Theorem 4.2.

Now we show the reduction. For each vertex v_i of G, we construct three men v_i^A, v_i^B and v_i^C, and three women v_i^a, v_i^b and v_i^c. Hence, there are $3|V|$ men and $3|V|$

women in total. Suppose that the vertex v_i is adjacent to d vertices $v_{i_1}, v_{i_2}, \ldots, v_{i_d}$. Then, preference lists of six people corresponding to v_i are as follows.

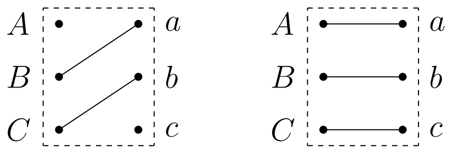
$$
\begin{array}{llll}
v_i^A\colon & v_i^a & v_i^a\colon & v_i^B\ v_{i_1}^C\ \cdots\ v_{i_d}^C\ v_i^A \\[2ex]
v_i^B\colon & (v_i^a\ v_i^b) & v_i^b\colon & v_i^B\ v_i^C \\[2ex]
v_i^C\colon & v_i^b\ v_{i_1}^a\ \cdots\ v_{i_d}^a\ v_i^c & v_i^c\colon & v_i^C
\end{array}
$$

The orders of persons in preference lists of v_i^C and v_i^a are determined as follows: $v_p^a \succ v_q^a$ in v_i^C's list if and only if $v_p^C \succ v_q^C$ in v_i^a's list. Clearly, this reduction can be performed in polynomial time. It is not hard to see that condition (i) holds.

We show that condition (ii) holds. Given a minimum vertex cover $VC(G)$ for G, we construct a stable matching M for $\hat{I}(G)$ as follows: For each vertex v_i, if $v_i \in VC(G)$, let $M(v_i^B) = v_i^a$, $M(v_i^C) = v_i^b$, and leave v_i^A and v_i^c single. If $v_i \notin VC(G)$, let $M(v_i^A) = v_i^a$, $M(v_i^B) = v_i^b$, and $M(v_i^C) = v_i^c$. Fig. 4.1 shows a part of M corresponding to v_i.

It is straightforward to verify that M is stable in $\hat{I}(G)$. It is easy to see that there is no blocking pair consisting of a man and a woman associated with the same vertex. Suppose there is a blocking pair associated with different vertices v_i and v_j. Then it must be (v_i^C, v_j^a) or (v_j^C, v_i^a), and without loss of generality, we assume that it is (v_i^C, v_j^a). Then, v_i^C and v_j^a are acceptable to each other, and so, by the construction of preference lists, v_i and v_j must be adjacent in G. As a result, either or both are contained in $VC(G)$. By the construction of the matching, this implies that either v_i^C or v_j^a is matched with a person at the top of his/her preference list, which is a contradiction. Hence, there is no blocking pair for M. Observe that $|M| = 2|VC(G)| + 3(|V| - |VC(G)|) = 3|V| - |VC(G)|$. Hence $|OPT(\hat{I}(G))| \geq |M| = 3|V| - |VC(G)|$.

Conversely, let M be a maximum stable matching for $\hat{I}(G)$. (We use M instead of $OPT(\hat{I}(G))$ for simplicity.) Consider a vertex $v_i \in V$ and the corresponding six persons. Note that v_i^B is matched in M, as otherwise (v_i^B, v_i^b) would block M. We consider two cases according to his partner.

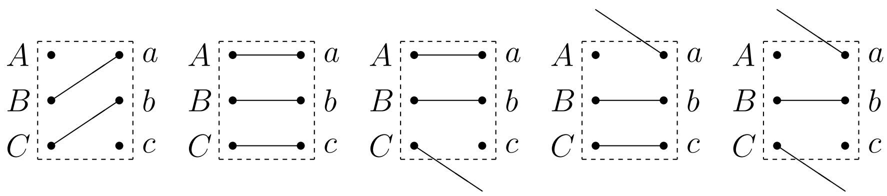


$v_i \in VC(G)$ $\qquad$ $v_i \in V \setminus VC(G)$

Figure 4.1: A part of matching M

Case (1) $M(v_i^B) = v_i^a$ Then, v_i^b is matched in M, as otherwise (v_i^C, v_i^b) blocks M. Since v_i^B is already matched with v_i^a, $M(v_i^b) = v_i^C$. Then, both v_i^A and v_i^c must be single in M. In this case, we say that "v_i causes a pattern 1." A diagrammatic representation of a pattern 1 is given in Fig. 4.2.

Case (2) $M(v_i^B) = v_i^b$ Then, v_i^a is matched in M, as otherwise (v_i^A, v_i^a) blocks M. Since v_i^B is already matched with v_i^b, there remain two cases: (a) $M(v_i^a) = v_i^A$ and (b) $M(v_i^a) = v_{i_j}^C$ for some j. Similarly, for v_i^C, there are two cases: (c) $M(v_i^C) = v_i^c$ and (d) $M(v_i^C) = v_{i_j}^a$ for some j. Hence, we have four cases in total. These cases are referred to as patterns 2 through 5 (see Fig. 4.2). For example, a combination of cases (b) and (c) corresponds to pattern 4.



pattern 1 $\quad$ pattern 2 $\quad$ pattern 3 $\quad$ pattern 4 $\quad$ pattern 5

Figure 4.2: Five patterns caused by v_i

Lemma 4.1 *Each vertex causes a pattern 1, 2 or 5.*

Proof. Suppose that a vertex v causes a pattern 3. Then, there is a sequence of vertices $v_{i_1}(= v), v_{i_2}, \ldots, v_{i_\ell} (\ell \geq 2)$ such that $M(v_{i_1}^A) = v_{i_1}^a$, $M(v_{i_j}^C) = v_{i_{j+1}}^a$ $(1 \leq j \leq \ell - 1)$ and $M(v_{i_\ell}^C) = v_{i_\ell}^c$, so that v_{i_1} causes a pattern 3, v_{i_2} through $v_{i_{\ell-1}}$ cause a pattern 5, and v_{i_ℓ} causes a pattern 4. (If we assume that v causes a pattern 4, then by the same argument, we can show the existence of a sequence of vertices with the same property as above, namely, $v_{i_\ell}(=v), v_{i_{\ell-1}}, \ldots, v_{i_1}(\ell \geq 2)$ such that v_{i_ℓ} causes a pattern 4, $v_{i_{\ell-1}}$ through v_{i_2} cause a pattern 5, and v_{i_1} causes a pattern 3.)

First, consider the case of $\ell \geq 3$. We show that, for each $2 \leq j \leq \ell-1$, $v^a_{i_{j+1}} \succ v^a_{i_{j-1}}$ in $v^C_{i_j}$'s list. We will prove this fact by induction.

Since $v^a_{i_1}$ is matched with $v^A_{i_1}$, the man at the tail of her list, $M(v^C_{i_2})(=v^a_{i_3}) \succ v^a_{i_1}$ in $v^C_{i_2}$'s list; otherwise, $(v^C_{i_2}, v^a_{i_1})$ blocks M. Hence, the statement is true for $j=2$. Suppose that the statement is true for $j=k$, namely, $v^a_{i_{k+1}} \succ v^a_{i_{k-1}}$ in $v^C_{i_k}$'s list. By the construction of preference lists, $v^C_{i_{k+1}} \succ v^C_{i_{k-1}}$ in $v^a_{i_k}$'s list. Then, if $v^a_{i_k} \succ v^a_{i_{k+2}}$ in $v^C_{i_{k+1}}$'s list, $(v^C_{i_{k+1}}, v^a_{i_k})$ blocks M. Hence, the statement is true for $j=k+1$.

Now, it turns out that $v^a_{i_\ell} \succ v^a_{i_{\ell-2}}$ in $v^C_{i_{\ell-1}}$'s list, which implies that $v^C_{i_\ell} \succ v^C_{i_{\ell-2}}$ in $v^a_{i_{\ell-1}}$'s list. Then, $(v^C_{i_\ell}, v^a_{i_{\ell-1}})$ blocks M since $M(v^C_{i_\ell}) = v^c_{i_\ell}$, a contradiction.

It is straightforward to verify that, when $\ell=2$, $(v^C_{i_2}, v^a_{i_1})$ blocks M, a contradiction. □

By Lemma 4.1, each vertex v_i causes a pattern 1, 2 or 5. Construct the subset C of vertices in the following way: If v_i causes a pattern 1 or pattern 5, then let $v_i \in C$, otherwise, let $v_i \notin C$.

We show that C is actually a vertex cover for G. Suppose not. Then, there are two vertices v_i and v_j in $V \backslash C$ such that $(v_i, v_j) \in E$ and both of them cause pattern 2, i.e., $M(v^C_i) = v^c_i$ and $M(v^A_j) = v^a_j$. Then (v^C_i, v^a_j) blocks M, contradicting the stability of M. Hence, C is a vertex cover for G. It is easy to see that $|M|(=|OPT(\hat{I}(G))|) = 2|C|+3(|V|-|C|) = 3|V|-|C|$. Thus $|VC(G)| \leq |C| = 3|V|-|OPT(\hat{I}(G))|$. Hence, condition (ii) holds. □

Theorem 4.2 implies the following inapproximability result:

Corollary 4.3 *MAX SMTI is NP-hard to approximate within a factor* $\frac{21}{19}-\delta$ *for any constant* $\delta > 0$.

Proof. By letting $p=\frac{1}{3}$ in Theorem 4.2, we know that the existence of a polynomial-time algorithm that distinguishes between the following two cases implies P=NP for an arbitrary small positive constant ϵ:

(1) $|OPT(\hat{I})| \geq \frac{21-\epsilon}{27}N$.

(2) $|OPT(\hat{I})| < \frac{19+\epsilon}{27}N$.

Now, suppose that there is a polynomial-time approximation algorithm T for MAX SMTI whose approximation ratio is at most $\frac{21}{19} - \delta$ for some δ. Then, consider the above statement with fixed constant ϵ such that $\epsilon < \frac{361\delta}{40-19\delta}$.

If an instance of the case (1) is given to T, it outputs a solution whose size is at least $\frac{21-\epsilon}{27}N\frac{1}{\frac{21}{19}-\delta}$. If an instance of the case (2) is given to T, it outputs a solution whose size is less than $\frac{19+\epsilon}{27}N$. It is easy to observe that $\frac{21-\epsilon}{27}N\frac{1}{\frac{21}{19}-\delta} > \frac{19+\epsilon}{27}N$ by the definition of ϵ. Hence, using T, we can distinguish between the cases (1) and (2), which implies P=NP. This completes the proof. $\square$

Observe that Theorem 4.2 and Corollary 4.3 hold for the restricted case where ties occur only in the preference lists of one sex and are of length only two. Furthermore, each preference list is either totally ordered or consists of a single tied pair.

Next, we show a further improved result of inapproximability by a similar proof for Theorem 4.2, using the fact that Proposition 4.1 holds even if we restrict the graph $G(V, E)$ has a perfect matching [CC07]. This result holds in the weaker restriction: ties are of length only two. Note that this result does not hold when ties occur in the preference lists of both sex.

Theorem 4.4 *For any $\epsilon > 0$ and $p < \frac{3-\sqrt{5}}{2}$, the following holds: If there is a polynomial-time algorithm that, given a MAX SMTI instance $\hat{I}$ of size N, distinguishes between the following two cases, then P=NP.*

(1) $|OPT(\hat{I})| \geq \frac{1.5+p-\epsilon}{2.5}N$.
(2) $|OPT(\hat{I})| < \frac{1.5+\max\{p^2,4p^3-3p^4\}+\epsilon}{2.5}N$.

Proof. Given a graph $G = (V, E)$ which has a perfect matching M_p, we will construct, in polynomial time, an SMTI instance $\hat{I}(G)$ with N men and N women. Our reduction satisfies the following two conditions: (i) $N = \frac{5}{2}|V|$, and (ii) $|OPT(\hat{I}(G))| = \frac{5}{2}|V| - |VC(G)|$. Then, it is not hard to see that Proposition 4.1 implies Theorem 4.4.

Now we show the reduction. For each pair $e_{ij} = (v_i, v_j)$ in M_p, we construct five men v_i^A, v_j^A, v_i^B, v_j^B, and e_{ij}^C, and five women v_i^a, v_j^a, v_i^b, v_j^b, and e_{ij}^c. Hence, there are $5|M_p|$ men and $5|M_p|$ women in total. Suppose that the vertex v_i is adjacent to

$di+1$ vertices $v_{i_1}, v_{i_2}, \ldots, v_{i_{di}}$ and v_j and the vertex v_j is adjacent to $dj+1$ vertices $v_{j_1}, v_{j_2}, \ldots, v_{j_{dj}}$ and v_i. Then, preference lists of ten people corresponding to e_{ij} are as follows.

$$
\begin{array}{llll}
v_j^A: & v_j^b & \qquad v_i^a: & v_i^B \\[2ex]
v_i^B: & (e_{ij}^c\ v_j^b)\ v_{i_1}^b\ \cdots\ v_{i_{di}}^b\ v_i^a & \qquad v_j^b: & e_{ij}^C\ v_i^B\ v_{j_1}^B\ \cdots\ v_{j_{dj}}^B\ v_j^A \\[2ex]
e_{ij}^C: & e_{ij}^c\ (v_i^b\ v_j^b) & \qquad e_{ij}^c: & (v_j^B\ v_i^B)\ e_{ij}^C \\[2ex]
v_j^B: & (e_{ij}^c\ v_i^b)\ v_{j_1}^b\ \cdots\ v_{j_{dj}}^b\ v_j^a & \qquad v_i^b: & e_{ij}^C\ v_j^B\ v_{i_1}^B\ \cdots\ v_{i_{di}}^B\ v_i^A \\[2ex]
v_i^A: & v_i^b & \qquad v_j^a: & v_j^B
\end{array}
$$

The orders of persons in preference lists of v_i^B and v_i^b are determined as follows: $v_p^b \succ v_q^b$ in v_i^B's list if and only if $v_p^B \succ v_q^B$ in v_i^b's list. Similarly, $v_p^b \succ v_q^b$ in v_j^B's list if and only if $v_p^B \succ v_q^B$ in v_j^b's list. Clearly, this reduction can be performed in polynomial time. Since $|V| = 2|M_p|$, it is not hard to see that condition (i) holds.

We show that condition (ii) holds. Given a minimum vertex cover $VC(G)$ for G, we construct a stable matching M for $\hat{I}(G)$ as follows: For each edge $e_{ij} = (v_i, v_j) \in M_p$, at least one of v_i and v_j is in $VC(G)$. If both v_i and v_j are in $VC(G)$, let $M(v_i^B) = v_j^b$, $M(v_j^B) = v_i^b$, $M(e_{ij}^C) = e_{ij}^c$, and leave v_i^A, v_j^A, v_i^a, and v_j^a single. If $v_i \in VC(G)$ and $v_j \notin VC(G)$, let $M(v_i^B) = e_{ij}^c$, $M(v_j^A) = v_j^b$, $M(v_j^B) = v_j^a$, $M(e_{ij}^C) = v_i^b$, and leave v_i^A and v_i^a single. Similarly, if $v_j \in VC(G)$ and $v_i \notin VC(G)$, let $M(v_j^B) = e_{ij}^c$, $M(v_i^A) = v_i^b$, $M(v_i^B) = v_i^a$, $M(e_{ij}^C) = v_j^b$, and leave v_j^A and v_j^a single. Fig. 4.3 shows a part of M corresponding to e_{ij}.

It is straightforward to verify that M is stable in $\hat{I}(G)$. It is easy to see that there is no blocking pair consisting of a man and a woman associated with the same pair in M_p. Suppose there is a blocking pair associated with different vertices v_k and v_l such that $(v_k, v_l) \notin M_p$. Then it must be (v_k^B, v_l^b). Then, v_k^B and v_l^b are acceptable to each other, and so, by the construction of preference lists, v_k and v_l must be adjacent in G. As a result, either or both are contained in $VC(G)$. By the construction of the matching, this implies that either v_k^b or v_l^b is matched with a person at the top-two of her preference list, which is a contradiction. Hence, there is no blocking pair for M. Observe that $|M| = 3(|VC(G)| - |V|/2) + 4(|V| - |VC(G)|) = 2.5|V| - |VC(G)|$. Hence $|OPT(\hat{I}(G))| \geq |M| = 2.5|V| - |VC(G)|$.

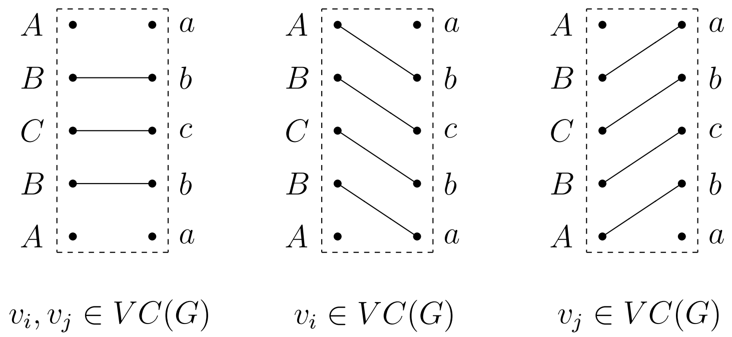


Figure 4.3: A part of matching M

Conversely, let M be a maximum stable matching for $\hat{I}(G)$. (We use M instead of $OPT(\hat{I}(G))$ for simplicity.) Consider a pair $e_{ij} = (v_i, v_j) \in M_p$ and the corresponding ten persons. Note that e^c_{ij} is matched in M, as otherwise (e^C_{ij}, e^c_{ij}) would block M. We consider three cases according to his partner.

Case (1) $M(e^c_{ij}) = e^C_{ij}$ Then, v^B_i is matched with v^b_j in M, as otherwise (v^B_i, v^b_j) blocks M. Similarly, v^B_j is matched with v^b_i in M. Then, all of v^A_i, v^A_j, v^a_i, and v^a_j must be single in M. In this case, we say that "e_{ij} causes a pattern 1" or equivalently "v_i (v_j) causes a pattern 1." A diagrammatic representation of a pattern 1 is given in Fig. 4.4.

Case (2) $M(e^c_{ij}) = v^B_j$ Then, e^C_{ij} is matched in M, as otherwise (e^C_{ij}, v^b_j) blocks M. Since e^c_{ij} is already matched with v^B_j, there remain two cases: (x) $M(e^C_{ij}) = v^b_i$ and (y) $M(e^C_{ij}) = v^b_j$. In case (x), v^B_i is matched with v^b_j in M, as otherwise (v^B_i, v^b_j) blocks M. Then, all of v^A_i, v^A_j, v^a_i, and v^a_j must be single in M. In this case, we say that "e_{ij} causes a pattern 2" or "v_i (v_j) causes a pattern 2." In case (y), v^B_i is matched in M, as otherwise (v^B_i, v^a_i) blocks M. Since e^c_{ij} and v^b_j are already matched with v^B_j and e^C_{ij} respectively, there remain two cases: (a) $M(v^B_i) = v^a_i$ and (b) $M(v^B_i) = v^b_{i_k}$ for some k. Similarly, for v^b_i, there are two cases: (c) $M(v^b_i) = v^A_i$ and (d) $M(v^b_i) = v^B_{j_l}$ for some l. Hence, we have four cases in total. These cases are referred to as patterns 3 through 6 (see Fig. 4.4). For example, a combination of cases (b) and (c) corresponds to pattern 4.

Case (3) $M(e^c_{ij}) = v^B_i$ By exchanging the role of v_i and v_j, this case is same with Case (2).

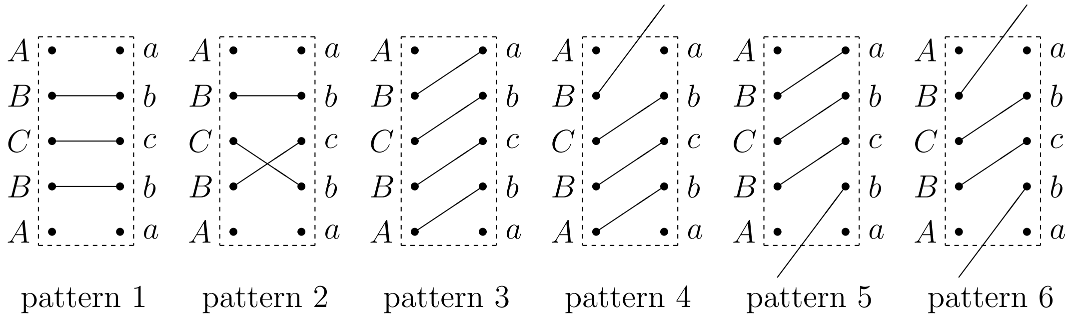


Figure 4.4: Five patterns caused by v_i

Lemma 4.2 *Each vertex causes a pattern 1, 2, 3 or 6.*

Proof. Suppose that a pair e_{ij} causes a pattern 4. Then, there is a sequence of vertices $v_{i_1}(= v), v_{i_2}, \ldots, v_{i_\ell}(\ell \geq 2)$ such that $M(v^A_{i_1}) = v^b_{i_1}$, $M(v^B_{i_j}) = v^b_{i_{j+1}}$ $(1 \leq j \leq \ell - 1)$ and $M(v^B_{i_\ell}) = v^a_{i_\ell}$, so that v_{i_1} causes a pattern 4, v_{i_2} through $v_{i_{\ell-1}}$ cause a pattern 6, and v_{i_ℓ} causes a pattern 5. (If we assume that e_{ij} causes a pattern 5, then by the same argument, we can show the existence of a sequence of vertices with the same property as above, namely, $v_{i_\ell}(= v), v_{i_{\ell-1}}, \ldots, v_{i_1}(\ell \geq 2)$ such that v_{i_ℓ} causes a pattern 5, $v_{i_{\ell-1}}$ through v_{i_2} cause a pattern 6, and v_{i_1} causes a pattern 4.)

First, consider the case of $\ell \geq 3$. We show that, for each $2 \leq j \leq \ell - 1$, $v^b_{i_{j+1}} \succ v^b_{i_{j-1}}$ in $v^B_{i_j}$'s list. We will prove this fact by induction.

Since $v^b_{i_1}$ is matched with $v^A_{i_1}$, the man at the tail of her list, $M(v^B_{i_2})(= v^b_{i_3}) \succ v^b_{i_1}$ in $v^B_{i_2}$'s list; otherwise, $(v^B_{i_2}, v^b_{i_1})$ blocks M. Hence, the statement is true for $j = 2$. Suppose that the statement is true for $j = k$, namely, $v^b_{i_{k+1}} \succ v^b_{i_{k-1}}$ in $v^B_{i_k}$'s list. By the construction of preference lists, $v^B_{i_{k+1}} \succ v^B_{i_{k-1}}$ in $v^b_{i_k}$'s list. Then, if $v^b_{i_k} \succ v^b_{i_{k+2}}$ in $v^B_{i_{k+1}}$'s list, $(v^B_{i_{k+1}}, v^b_{i_k})$ blocks M. Hence, the statement is true for $j = k + 1$.

Now, it turns out that $v^b_{i_\ell} \succ v^b_{i_{\ell-2}}$ in $v^B_{i_{\ell-1}}$'s list, which implies that $v^B_{i_\ell} \succ v^B_{i_{\ell-2}}$ in $v^b_{i_{\ell-1}}$'s list. Then, $(v^B_{i_\ell}, v^b_{i_{\ell-1}})$ blocks M since $M(v^B_{i_\ell}) = v^b_{i_\ell}$, a contradiction.

It is straightforward to verify that, when $\ell = 2$, $(v^B_{i_2}, v^b_{i_1})$ blocks M, a contradiction. $\square$

By Lemma 4.2, each pair e_{ij} in M_p causes a pattern 1, 2, 3, or 6. Construct the subset C of vertices in the following way: If e_{ij} causes a pattern 1, pattern 2, or pattern 6, then let $v_i, v_j \in C$, otherwise, let $v_j \in C$ and $v_i \notin C$.

We show that C is actually a vertex cover for G. Suppose not. Then, there are two vertices v_p and v_q in $V \setminus C$ such that $(v_p, v_q) \in E$ and both of them cause pattern 3, i.e., $M(v_p^B) = v_p^a$ and $M(v_q^B) = v_q^a$. Then (v_p^B, v_q^b) blocks M, contradicting the stability of M. Hence, C is a vertex cover for G. It is easy to see that $|M|(= |OPT(\hat{I}(G))|) = 3(|C| - |V|/2) + 4(|V| - |C|) = 2.5|V| - |C|$. Thus $|VC(G)| \leq |C| = 2.5|V| - |OPT(\hat{I}(G))|$. Hence, condition (ii) holds. □

Theorem 4.4 implies the following inapproximability result:

Corollary 4.5 *MAX SMTI is NP-hard to approximate within any factor less than or equal to* $\frac{33}{29}$.

Proof. Similar with the proof of Corollary 4.3. □

Remark. A long-standing conjecture states that MVC is hard to approximate within a factor of $2 - \epsilon$ [KR03]. We obtain lower bounds 5/4 (= 1.25) and 4/3 (> 1.3333) for MAX SMTI, modulo this conjecture as follows.

First, we consider the strongly restricted case as in Theorem 4.2. Let $G = (V, E)$ be a graph such that $|VC(G)| \geq \frac{|V|}{2}$. (Due to Nemhauser and Trotter [NT75], approximability of MVC for general graphs is equivalent to approximability of MVC for graphs $G = (V, E)$ with this restriction.) We have already proved that $|OPT(\hat{I}(G))| = 3|V| - |VC(G)|$. Also, it is easy to see that given any stable matching M for $\hat{I}(G)$, we can obtain a vertex cover C for G with $|C| \leq 3|V| - |M|$. Combining these facts with $|VC(G)| \geq \frac{|V|}{2}$ and $|OPT(\hat{I}(G))|/|M| \leq r$, we have that $|C| \leq (6 - \frac{5}{r})|VC(G)|$. Hence, we can construct a polynomial-time $(6 - \frac{5}{r})$-approximation algorithm for MVC using a polynomial-time r-approximation algorithm for MAX SMTI for the strong restricted case.

Similarly, we can construct a polynomial-time $(5 - \frac{4}{r})$-approximation algorithm for MVC using a polynomial-time r-approximation algorithm for MAX SMTI for the weaker restricted case as in Theorem 4.4.

4.3 Approximation Algorithm ShiftBrk

In this section, we give an approximation algorithm ShiftBrk for MAX SMTI. We define first some notation regarding ties.

Suppose that, in an SMTI instance $\hat{I}$, a man m has a tie T of length ℓ consisting of women $w_1, w_2, \ldots, w_\ell$. Also, suppose that this tie T is broken into $[w_1\ w_2 \cdots w_\ell]$ in m's I-list. We say "shift tie T in I" to obtain a new SMI instance I' in which only the tie T is changed to $[w_2 \cdots w_\ell\ w_1]$ and other preference lists are the same as in I. If I' is the result of shifting all broken ties in men's lists in I, then we write "$I' = \mathit{Shift}_m(I)$." Similarly, if I' is the result of shifting all broken ties in women's lists in I, then we write "$I' = \mathit{Shift}_w(I)$." Let L be the maximum length of ties in $\hat{I}$. The full description of ShiftBrk is given in Fig. 4.5.

Algorithm ShiftBrk($\hat{I}$)

$I_{1,1}$:= an SMI instance obtained by breaking all ties of $\hat{I}$ in an arbitrary order;
for $i := 2$ **to** L
    $I_{i,1} := \mathit{Shift}_m(I_{i-1,1})$;
for $i := 1$ **to** L
    for $j := 2$ **to** L
        $I_{i,j} := \mathit{Shift}_w(I_{i,j-1})$;
for $i := 1$ **to** L
    for $j := 1$ **to** L
        $M_{i,j}$:= stable matching for $I_{i,j}$;
Output a largest matching among all $M_{i,j}$'s;

Figure 4.5: Algorithm ShiftBrk

Since stable matchings for SMI instances can be obtained in polynomial time using the Gale-Shapley algorithm, ShiftBrk runs in time polynomial in N. It is easy to see that all $M_{i,j}$ are stable for $\hat{I}$ (see [GI89] for example). Hence, ShiftBrk outputs a feasible solution.

4.3.1 Bidominating Pairs

Before analyzing the approximation ratio, we will define an important notion, a bidominating pair[1], which plays an important role in our analysis. Let $\hat{I}$ be an SMTI instance and M_{opt} be a largest stable matching for $\hat{I}$. Let I be an SMI instance obtained by breaking all ties of $\hat{I}$ and M be a stable matching for I. A pair (m, w) is said to be a *bidominating pair for* M if they are matched together in M, both are matched to other people in M_{opt}, and both prefer each other (in I) to their partners in M_{opt}. That is, (a) $M(m) = w$, (b) m is matched in M_{opt} and $w \succ M_{opt}(m)$ in m's I-list, and (c) w is matched in M_{opt} and $m \succ M_{opt}(w)$ in w's I-list.

Lemma 4.3 *Let* (m, w) *be a bidominating pair for* M*. Then, one or both of the following holds:*
(i) $[\cdots w \cdots M_{opt}(m) \cdots]$ *in* m*'s* I*-list; (ii)* $[\cdots m \cdots M_{opt}(w) \cdots]$ *in* w*'s* I*-list.*

Proof. If the strict preferences hold also in $\hat{I}$, i.e. $w \succ M_{opt}(m)$ in m's $\hat{I}$-list and $m \succ M_{opt}(w)$ in w's $\hat{I}$-list, then (m, w) blocks M_{opt} in $\hat{I}$. Thus, at least one of these preferences in I must have been caused by the breaking of ties in $\hat{I}$. □

Fig. 4.6 shows a simple example of a bidominating pair. (A dotted line means that the endpoints are matched in M_{opt} and a solid line means the same in M. In m_3's list, w_2 and w_3 are tied in $\hat{I}$ and this tie is broken into $[w_2\ w_3]$ in I.)

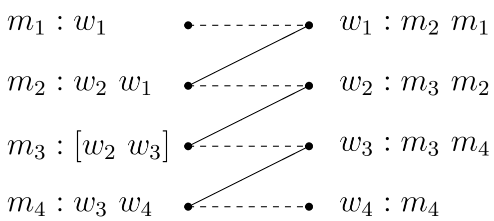


Figure 4.6: A bidominating pair (m_3, w_2) for M

Lemma 4.4 $|M| \geq |M_{opt}| - k$*, where* k *is the number of bidominating pairs for* M*.*

[1] In an earlier version of this article, this was denoted *annoying pair*.

Proof. Consider a connected component of the bipartite graph $G_{M,M_{opt}}$ (see Section 2.2 for the definition), which is either a cycle or a path. A cycle contains equally many edges from M and M_{opt}, while a path can contain one more edge from one of the matchings. We show that each path in $G_{M,M_{opt}}$ contains at least one bidominating pair for M. This will imply the lemma.

Consider a path $m_1, w_1, m_2, w_2, \ldots, m_\ell, w_\ell$, where $w_s = M_{opt}(m_s)$ $(1 \leq s \leq \ell)$ and $m_{s+1} = M(w_s)$ $(1 \leq s \leq \ell - 1)$. (This path begins with a man and ends with a woman. Other cases can be proved in a similar manner.) Suppose that this path does not contain a bidominating pair for M. Since m_1 is single in M, $m_2 \succ m_1$ in w_1's I-list (otherwise, (m_1, w_1) blocks M). Then, consider the man m_2. Since we assume that (m_2, w_1) is not a bidominating pair, $w_2 \succ w_1$ in m_2's I-list. We can continue the same argument to show that $m_3 \succ m_2$ in w_2's I-list and $w_3 \succ w_2$ in m_3's I-list, and so on. Finally, we have that $w_\ell \succ w_{\ell-1}$ in m_ℓ's I-list. Since w_ℓ is single in M, (m_ℓ, w_ℓ) blocks M, a contradiction. Hence, every path must contain at least one bidominating pair. □

4.4 ShiftBrk for Instances Where Only Men Have Ties

In this section, we consider SMTI instances such that only men have ties each with length at most L. Note that we do not restrict the *number* of ties in the list; one man can have more than one tie, as long as each tie is of length at most L. We show that ShiftBrk achieves an approximation ratio of $2/(1 + L^{-2})$.

Let $\hat{I}$ be an SMTI instance. We fix a largest stable matching M_{opt} for $\hat{I}$ and denote $n = |M_{opt}|$. All preferences in this section are with respect to $\hat{I}$ unless otherwise stated. Since women do not have ties, ShiftBrk produces L instances $I_{1,1}, I_{2,1}, \ldots, I_{L,1}$. Let us denote them for simplicity as $I_1, I_2, \ldots, I_L$. Let $M_1, M_2, \ldots, M_L$ be corresponding stable matchings obtained in line 9 of ShiftBrk.

As shown in Lemma 4.4, if the number of bidominating pairs for a stable matching M is small, then M is relatively large. We will show that among L stable matchings, there is at least one stable matching that contains a small number of bidominating pairs. To do so, we observe the following property: Suppose that a man m is matched in M_{opt} and all of $M_1, M_2, \ldots, M_L$. If m does not contain $M_{opt}(m)$ in a tie, then m cannot be a part of a bidominating pair in any of L matchings. If m contains $M_{opt}(m)$ in a tie, then m cannot be a part of a bidominating pair in M_i, where in I_i, m's tie is broken so that $M_{opt}(m)$ comes the first place in the tie. In the proof, we estimate the lower bound on the number of such men in terms of the optimal size n.

Let V_{opt} and W_{opt} be the set of men and women, respectively, that are matched in M_{opt}. Let V_a be the subset of V_{opt} such that each man $m \in V_a$ has a partner in all of $M_1, \ldots, M_L$. Let $W_b = \{w | M_{opt}(w) \in V_{opt} \setminus V_a\}$. Note that, by definition, $W_b \subseteq W_{opt}$ and $|V_a| + |W_b| = n$. For each woman w, let $best(w)$ be the man that w prefers the most among $M_1(w), \ldots, M_L(w)$; if she is single in each of $M_1, \ldots, M_L$, then $best(w)$ is not defined.

Lemma 4.5 *Let w be in W_b. Then $best(w)$ exists and is in V_a, and is preferred by w over $M_{opt}(w)$. That is, $best(w) \in V_a$ and $best(w) \succ M_{opt}(w)$ in w's $\hat{I}$-list.*

Proof. By the definition of W_b, $M_{opt}(w) \in V_{opt} \setminus V_a$. By the definition of V_a, there is a matching M_i in which $M_{opt}(w)$ is single. Since M_i is a stable matching for $\hat{I}$, w has a partner in M_i and further, that partner $M_i(w)$ is preferred over $M_{opt}(w)$ (as otherwise, $(M_{opt}(w), w)$ blocks M_i). Since w has a partner in M_i, $best(w)$ is defined and differs from $M_{opt}(w)$. By the definition of $best(w)$, w prefers $best(w)$ over $M_{opt}(w)$. That implies that $best(w)$ is matched in M_{opt}, i.e. $best(w) \in V_{opt}$, as otherwise $(best(w), w)$ blocks M_{opt}. Finally, $best(w)$ must be matched in each $M_1, \ldots, M_L$, i.e. $best(w) \in V_a$, as otherwise $(best(w), w)$ blocks M_i in which $best(w)$ is single. □

Lemma 4.6 *Consider a man m and two women w_1, w_2, where $m = best(w_1) = best(w_2)$. Then w_1 and w_2 are tied in m's $\hat{I}$-list.*

Proof. Since $m = best(w_1) = best(w_2)$, there are matchings M_i and M_j such that $m = M_i(w_1) = M_j(w_2)$. First, suppose that $w_1 \succ w_2$ in m's list. Since $m = M_j(w_2)$, w_1 is not matched with m in M_j. By the definition of $best(w)$, w_1 is either single or matched with a man below m in her list, in the matching M_j. In either case, (m, w_1)

blocks M_j in $\hat{I}$, a contradiction. By exchanging the role of w_1 and w_2, we can show that it is not the case that $w_2 \succ w_1$ in m's list. Hence, w_1 and w_2 must be tied in m's $\hat{I}$-list. □

By Lemma 4.6, each man can be $best(w)$ for at most L women w because the length of ties is at most L. Let us partition V_a into V_t and $\overline{V_t}$, where V_t is the set of men m such that m is $best(w)$ for exactly L women $w \in W_b$ and $\overline{V_t} = V_a \setminus V_t$.

Lemma 4.7 *There is a matching M_k for which the number of bidominating pairs is at most $|M_k| - (|V_t| + \frac{|\overline{V_t}|}{L})$.*

Proof. Consider a man $m \in V_t$. By definition, there are L women $w_1, \dots, w_L$ such that $m = best(w_1) = \cdots = best(w_L)$, $M_i(w_i) = m$ for $1 \leq i \leq L$, and all these women are in W_b. By Lemma 4.6, all these women are tied in m's $\hat{I}$-list. By Lemma 4.5, each woman w_i prefers $best(w_i)(= m)$ to $M_{opt}(w_i)$ so that $m \neq M_{opt}(w_i)$ for any i. This means that none of these women can be $M_{opt}(m)$. For m to form a bidominating pair for M_i, $w_i(= M_i(m))$ and $M_{opt}(m)$ must be tied in m's list, due to Lemma 4.3 (i) (note that the case (ii) of Lemma 4.3 does not happen because women do not have ties). Hence, m cannot form a bidominating pair for any of M_1 through M_L.

Next, consider a man $m \in \overline{V_t}$. If $M_{opt}(m)$ is not in a tie in m's list, m cannot form a bidominating pair for any of M_1 through M_L, by the same argument as above. If m includes $M_{opt}(m)$ in a tie, there exists an instance I_i such that $M_{opt}(m)$ lies in first place in the broken tie of m's I_i-list. This means that m does not constitute a bidominating pair for M_i by Lemma 4.3.

Hence, there is a matching M_k for which at least $|V_t| + \frac{|\overline{V_t}|}{L}$ men, among those matched in M_k, do not form a bidominating pair. Hence, the number of bidominating pairs is at most $|M_k| - (|V_t| + \frac{|\overline{V_t}|}{L})$. □

Lemma 4.8 $|V_t| + \frac{|\overline{V_t}|}{L} \geq \frac{n}{L^2}$.

Proof. By the definition of V_t, a man in V_t is $best(w)$ for L different women in W_b, while a man in $\overline{V_t}$ is $best(w)$ for at most $L - 1$ women in W_b. Recall that by

Lemma 4.5, for each woman w in W_b, there is a man in V_a that is *best*(w). Thus, W_b contains at most $|V_t|L + |\overline{V_t}|(L-1)$ women. Since $|V_a| + |W_b| = n$, we have that

$$n \leq |V_a| + |V_t|L + |\overline{V_t}|(L-1) = L|V_a| + |V_t|.$$

Now,

$$\begin{aligned}
|V_t| + \frac{|\overline{V_t}|}{L} &= |V_t| + \frac{|V_a| - |V_t|}{L} \\
&= \frac{1}{L}|V_a| + \frac{L-1}{L}|V_t| \\
&\geq \frac{1}{L}\frac{n - |V_t|}{L} + \frac{L-1}{L}|V_t| \\
&= \frac{n}{L^2} + \frac{L^2 - L - 1}{L^2}|V_t| \\
&\geq \frac{n}{L^2}.
\end{aligned}$$

The last inequality is due to the fact that $L^2 - L - 1 > 0$ since $L \geq 2$. □

Theorem 4.6 *The approximation ratio of* ShiftBrk *is at most* $2/(1 + L^{-2})$ *for instances where only men have ties, and these ties are of length at most* L.

Proof. By Lemmas 4.7 and 4.8, there is a matching M_k for which the number of bidominating pairs is at most $|M_k| - n/L^2$. By Lemma 4.4, $|M_k| \geq n - \left(|M_k| - \frac{n}{L^2}\right)$, which implies that $|M_k| \geq \frac{L^2+1}{2L^2}n = \frac{1+L^{-2}}{2}n$. □

Remark The same result holds for men's preference lists being arbitrary partial order. Suppose that each man m's list is a partial order with width at most L, i.e., such that the maximum number of mutually incomparable women for m is at most L. Then, we can partition its partial order into L chains by Dilworth's theorem [Dil50]. In each "shift," we give the priority to one of L chains and the resulting totally ordered preference list is constructed so that it satisfies the following property: Each member (woman) of the chain with the priority lies top among all women mutually incomparable with her for m in the original partial order. It is not hard to see that the theorem holds for this case.

Lower Bounds for ShiftBrk We show that the upper bound of Theorem 4.6 is tight. We give an example for $L = 4$ to explain how it works. The following preference lists illustrate part of a worst-case example. There are $2L$ men and $2L$ women:

$$
\begin{array}{llll}
m_1^1: & (\ w_1^1\ \ w_1^2\ \ w_1^3\ \ w_1^4\ ) & \qquad w_1^1: & m_1^1 \\
m_2^1: & (\ w_2^1\ \ w_2^2\ \ w_2^3\ \ w_2^4\ ) & \qquad w_2^1: & m_2^1 \\
m_1^2: & (\ w_2^2\ \ w_1^2\ \ w_2^3\ \ w_2^4\ ) & \qquad w_1^2: & m_1^1\ \ m_1^2 \\
m_2^2: & w_2^2 & \qquad w_2^2: & m_2^1\ \ m_1^2\ \ m_2^2\ \ m_1^3\ \ m_1^4 \\
m_1^3: & (\ w_2^2\ \ w_2^3\ \ w_1^3\ \ w_2^4\ ) & \qquad w_1^3: & m_1^1\ \ m_1^3 \\
m_2^3: & w_2^3 & \qquad w_2^3: & m_2^1\ \ m_1^3\ \ m_2^3\ \ m_1^4\ \ m_1^2 \\
m_1^4: & (\ w_2^2\ \ w_2^3\ \ w_2^4\ \ w_1^4\ ) & \qquad w_1^4: & m_1^1\ \ m_1^4 \\
m_2^4: & w_2^4 & \qquad w_2^4: & m_2^1\ \ m_1^4\ \ m_2^4\ \ m_1^2\ \ m_1^3
\end{array}
$$

The largest stable matching for this instance is of size $2L$ (m_j^i is matched with w_j^i for all i, j). Consider to apply ShiftBrk to the above instance. If the initial SMI instance I_1 is constructed by breaking ties in the same order as written above, then ShiftBrk produces L stable matchings $M_1, \ldots, M_L$, where $|M_1| = L + 1$ and $|M_2| = |M_3| = \cdots = |M_L| = L$.

To make the complete example, let $I^1, \ldots, I^L$ be L copies of the above instance and let I^{all} be the instance constructed by putting $I^1, \ldots, I^L$ together. Then, in the worst case initial tie-breaking, ShiftBrk produces L matchings each of which has size $(L+1)\cdot 1 + L\cdot(L-1) = L^2+1$, while a largest stable matching for I^{all} is of size $2L^2$. Hence, the approximation ratio of ShiftBrk for I^{all} is $2L^2/(L^2+1) = 2/(1+L^{-2})$, which proves the tightness of the analysis.

Finally, we briefly sketch how to make a sub-instance for general L. Prepare $2L$ men m_j^i and $2L$ women w_j^i ($1 \leq i \leq L$ and $1 \leq j \leq 2$). For $j = 1, 2$, m_j^1's list consists of one tie including women $w_j^1, w_j^2, \ldots, w_j^L$ in this order (this "order" is important only for considering worst case tie-breaking). For $2 \leq i \leq L$, m_2^i's list includes only one woman w_2^i. For $2 \leq i \leq L$, m_1^i's list consists of one tie of length L,

($w_2^2\ w_2^3\ \cdots\ w_2^{i-1}\ w_2^i\ w_1^i\ w_2^{i+1}\ \cdots\ w_2^L$). Woman w_j^1 ($j = 1, 2$) includes the man m_j^1. For $2 \le i \le L$, woman w_1^i includes m_1^1 and m_1^i in this order. For $2 \le i \le L$, woman w_2^i's list is as follows: $m_2^1\ m_1^i\ m_2^i\ m_1^{i+1}\ m_1^{i+2}\ \cdots\ m_1^L\ m_1^2\ m_1^3\ \cdots\ m_1^{i-1}$.

4.5 ShiftBrk for Instances Where Both Men And Women Have Ties

In this section we show that when $L = 2$, the performance ratio of ShiftBrk is better than two even if we allow women's lists to include ties. ShiftBrk creates four SMI instances $I_{1,1}, I_{1,2}, I_{2,1}$ and $I_{2,2}$ from the SMTI input instance $\hat{I}$, since $L = 2$. Note that men have the same lists in $I_{i,1}$ and $I_{i,2}$ ($i \in \{1, 2\}$), while women have the same lists in $I_{1,j}$ and $I_{2,j}$ ($j \in \{1, 2\}$). Let $M_{i,j}$ ($i \in \{1, 2\}$, $j \in \{1, 2\}$) be a stable matching for $I_{i,j}$ obtained in line 9 of ShiftBrk. We fix an optimal solution M_{opt}, a largest stable matching for $\hat{I}$, and let $n = |M_{opt}|$ as before.

Let V and W be the sets of men and women in $\hat{I}$, respectively. For $i, j \in \{1, 2\}$, let $V_{i,j}$ ($W_{i,j}$, respectively) be the set of men (women, respectively) that are matched in $M_{i,j}$. Observe that $|V_{i,j}| = |W_{i,j}| = |M_{i,j}|$. Define $A_{i,j}$ to be the set of pairs $(m, w) \in M_{i,j}$ that are bidominating for $M_{i,j}$, and $B_{i,j}$ the set of pairs $(m, w) \in M_{i,j}$ that are matched but not bidominating in $M_{i,j}$, that is, $B_{i,j} = M_{i,j} \setminus A_{i,j}$. Observe that $|A_{i,j}| + |B_{i,j}| = |M_{i,j}|$.

The following lemma shows that the number of matched men for a given broken instance, but are not matched when the instance is shifted, is at most the number of non-bidominating pairs.

Lemma 4.9 *For $j = 1, 2$, $|V_{1,j} \setminus V_{2,j}| \le |B_{1,j}|$.*

Proof. Let m_1 be in $V_{1,j} \setminus V_{2,j}$. Consider the bipartite graph $G_{M_{1,j},M_{2,j}}$ (see Section 2.2 for the definition). Since m_1 is matched in $M_{1,j}$ but single in $M_{2,j}$, there is a path in $G_{M_{1,j},M_{2,j}}$, starting from m_1. Assume that the path ends with a man (the other case can be discussed similarly and hence will be omitted). Let the path be $m_1, w_1, \ldots, m_\ell$, where $w_s = M_{1,j}(m_s)$ and $m_{s+1} = M_{2,j}(w_s)$ for $1 \le s \le \ell - 1$. We show that the path contains a pair $(m_i, w_i) \in B_{1,j}$.

Suppose that the path $m_1, w_1, \ldots, m_\ell$ does not contain a pair in $B_{1,j}$. $m_2 \succ m_1$ in w_1's $I_{2,j}$-list; otherwise, (m_1, w_1) blocks $M_{2,j}$. Then, since $I_{1,j} = \mathit{Shift}_m(I_{2,j})$, (a) $m_2 \succ m_1$ in w_1's $I_{1,j}$-list. By the assumption that $(m_1, w_1) \notin B_{1,j}$, (m_1, w_1) is a bidominating pair for $M_{1,j}$. By the definition of bidominating pairs, $m_1 \succ M_{opt}(w_1)$ in w_1's $I_{1,j}$-list. Using the above (a), we have that (b) $m_2 \succ m_1 \succ M_{opt}(w_1)$ in w_1's $I_{1,j}$-list.

Then we show the claim: For $1 \leq i \leq \ell - 2$, if $m_{i+1} \succ m_i \succ M_{opt}(w_i)$ in w_i's $I_{1,j}$-list, the following (x) and (y) hold.

(x) $w_{i+1} \succ w_i$ in m_{i+1}'s $I_{2,j}$-list.

(y) $m_{i+2} \succ m_{i+1} \succ M_{opt}(w_{i+1})$ in w_{i+1}'s $I_{1,j}$-list.

If the above claim holds, we can apply (x) and (y) repeatedly to (b), obtaining $m_\ell \succ m_{\ell-1} \succ M_{opt}(w_{\ell-1})$ in $w_{\ell-1}$'s $I_{1,j}$-list. Then, $(m_\ell, w_{\ell-1})$ blocks $M_{1,j}$, a contradiction.

Hence, in the following, we prove the above claim. Since $m_{i+1} \succ m_i$ in w_i's $I_{1,j}$-list by the assumption, (c) $w_{i+1} \succ w_i$ in m_{i+1}'s $I_{1,j}$-list (otherwise, (m_{i+1}, w_i) blocks $M_{1,j}$). Also, since we assume that $(m_{i+1}, w_{i+1}) \notin B_{1,j}$, (d) $w_{i+1} \succ M_{opt}(m_{i+1})$ in m_{i+1}'s $I_{1,j}$-list. Considering (c) and (d), we have following two cases: (A) $w_{i+1} \succ M_{opt}(m_{i+1}) \succ w_i$ in m_{i+1}'s $I_{1,j}$-list. (B) $w_{i+1} \succ w_i \succ M_{opt}(m_{i+1})$ in m_{i+1}'s $I_{1,j}$-list. (Note that $M_{opt}(m_{i+1}) \neq w_i$ because we assume that $m_{i+1} \succ m_i \succ M_{opt}(w_i)$ in w_i's $I_{1,j}$-list; hence $M_{opt}(w_i) \neq m_{i+1}$.)

Case (A): Since the length of ties is two, w_{i+1} and w_i are not tied in m_{i+1}'s list. Hence, the relative order of w_{i+1} and w_i is the same in $I_{1,j}$ and $I_{2,j}$. Thus (x) holds.

Case (B): Suppose that w_i and $M_{opt}(m_{i+1})$ are not tied in m_{i+1}'s $\hat{I}$-list. Then, since we assume that $w_i \succ M_{opt}(m_{i+1})$ in m_{i+1}'s $I_{1,j}$-list, this relation also holds for $\hat{I}$. Furthermore, we assume that $m_{i+1} \succ m_i \succ M_{opt}(w_i)$ in w_i's $I_{1,j}$-list. Hence, $m_{i+1} \succ M_{opt}(w_i)$ in w_i's $\hat{I}$-list. Then, (m_{i+1}, w_i) blocks M_{opt} in $\hat{I}$. Thus w_i and $M_{opt}(m_{i+1})$ are tied in m_{i+1}'s $\hat{I}$-list. Since $w_{i+1} \succ [w_i \ M_{opt}(m_{i+1})]$ in m_{i+1}'s $I_{1,j}$-list, $w_{i+1} \succ [M_{opt}(m_{i+1}) \ w_i]$ in m_{i+1}'s $I_{2,j}$-list. Thus (x) holds.

Since (x) holds, $m_{i+2} \succ m_{i+1}$ in w_{i+1}'s $I_{2,j}$-list (otherwise, (m_{i+1}, w_{i+1}) blocks $M_{2,j}$). This means that $m_{i+2} \succ m_{i+1}$ in w_{i+1}'s $I_{1,j}$-list because every woman's list is the same in $I_{1,j}$ and $I_{2,j}$. Since $(m_{i+1}, w_{i+1}) \notin B_{1,j}$, (m_{i+1}, w_{i+1}) is a bidominating pair for $M_{1,j}$, so that $m_{i+1} \succ M_{opt}(w_{i+1})$ in w_{i+1}'s $I_{1,j}$-list. Hence, $m_{i+2} \succ m_{i+1} \succ M_{opt}(w_{i+1})$ in w_{i+1}'s $I_{1,j}$-list. Thus (y) holds. □

By exchanging the role of men and women, we have the following lemma:

Lemma 4.10 $|W_{1,1} \setminus W_{1,2}| \leq |B_{1,1}|$.

Consider a matching $M_{i,j}$ ($i = 1, 2$ and $j = 1, 2$). We will define two subsets of men (women, respectively) participating in bidominating pairs for $M_{i,j}$, namely, $P^m_{i,j}$ and $Q^m_{i,j}$ ($P^w_{i,j}$ and $Q^w_{i,j}$, respectively). By Lemma 4.3, at least one of the following holds for a bidominating pair (m, w) for $M_{i,j}$: (i) $[w\ M_{opt}(m)]$ in m's $I_{i,j}$-list. (ii) $[m\ M_{opt}(w)]$ in w's $I_{i,j}$-list. If (i) holds, let $m \in P^m_{i,j}$ and $w \in P^w_{i,j}$. If (ii) holds, then let $m \in Q^m_{i,j}$ and $w \in Q^w_{i,j}$. Note that, $|P^m_{i,j}| = |P^w_{i,j}|$ and $|Q^m_{i,j}| = |Q^w_{i,j}|$. Also, notice that $P^m_{i,j}$ and $Q^m_{i,j}$ ($P^w_{i,j}$ and $Q^w_{i,j}$, respectively) are not necessarily disjoint.

In order to explain the following lemma, let us say that a woman w is *anxious* in a given matching if she is a part of a bidominating pair (m, w) and $[m\ M_{opt}(w)]$ in w's list. The next lemma says that the set of anxious women when the women's preference lists are broken in one way is disjoint from the set of anxious women when the preference lists are broken in the other way.

Lemma 4.11 $(Q^w_{1,1} \cup Q^w_{2,1}) \cap (Q^w_{1,2} \cup Q^w_{2,2}) = \emptyset$.

Proof. Suppose that $w \in (Q^w_{1,1} \cup Q^w_{2,1}) \cap (Q^w_{1,2} \cup Q^w_{2,2})$. Then, there are $i \in \{1, 2\}$ and $j \in \{1, 2\}$ such that $w \in Q^w_{i,1} \cap Q^w_{j,2}$. Since $w \in Q^w_{i,1}$, $[m\ M_{opt}(w)]$ in w's $I_{i,1}$-list for some man m. Since $w \in Q^w_{j,2}$, $[m'\ M_{opt}(w)]$ in w's $I_{j,2}$-list for some man m'. However, this is impossible because women's ties are broken in different ways in $I_{i,1}$ and $I_{j,2}$, a contradiction. □

The following lemma is the key to our argument that one of the matchings we find must be relatively large. By Lemma 4.4, it suffices to focus on the case when the number of bidominating pairs, $|A_{i,j}|$, is large, close to $\frac{n}{2}$. Then, $|B_{i,j}|$ is small, close to zero. In that case, the lemma shows that a certain subset of the matched women is large, close to $\frac{n}{2}$.

Lemma 4.12 *For* $j = 1, 2$, $|W_{1,j} \cap (Q^w_{1,j} \cup Q^w_{2,j})| \geq |A_{1,j}| - |V_{1,j} \setminus V_{2,j}| - |B_{2,j}|$.

Proof. Let $j \in \{1, 2\}$. Define X_j to be the set of men that belong to bidominating pairs both for $M_{1,j}$ and for $M_{2,j}$:

$$X_j = \{m \mid (m, M_{1,j}(m)) \in A_{1,j} \cap (m, M_{2,j}(m)) \in A_{2,j}\}.$$

We first show that $|W_{1,j} \cap (Q^w_{1,j} \cup Q^w_{2,j})| \geq |X_j|$. For this purpose, we consider a man in $m \in X_j$, and show that there is a woman in $W_{1,j} \cap (Q^w_{1,j} \cup Q^w_{2,j})$ corresponding to m. We then show that a woman does not correspond to different men. When considering a man m in X_j, we consider two cases: $m \in X_j \cap P^m_{1,j}$ and $m \in X_j \setminus P^m_{1,j}$.

First, consider a man m_p in $X_j \cap P^m_{1,j}$. Since $m_p \in P^m_{1,j}$, $[M_{1,j}(m_p)\ M_{opt}(m_p)]$ in m_p's $I_{1,j}$-list. Since $I_{2,j} = \mathit{Shift}_m(I_{1,j})$, $[M_{opt}(m_p)\ M_{1,j}(m_p)]$ in m_p's $I_{2,j}$-list. Since $(m_p, M_{2,j}(m_p))$ is a bidominating pair for $M_{2,j}$ by the definition of X_j, $M_{2,j}(m_p) \succ M_{opt}(m_p)$ in m_p's $I_{2,j}$-list. It then follows that

$$M_{2,j}(m_p) \succ [M_{opt}(m_p)\ M_{1,j}(m_p)] \text{ in } m_p\text{'s } I_{2,j}\text{-list.} \quad \cdots \quad (*)$$

Thus, by Lemma 4.3, $[m_p\ M_{opt}(M_{2,j}(m_p))]$ in $M_{2,j}(m_p)$'s $I_{2,j}$-list. Hence, $M_{2,j}(m_p) \in Q^w_{2,j}$. Since $I_{1,j} = \mathit{Shift}_m(I_{2,j})$ and by $(*)$ above, $M_{2,j}(m_p) \succ [M_{1,j}(m_p)\ M_{opt}(m_p)]$ in m_p's $I_{1,j}$-list. It is easy to see that $M_{2,j}(m_p) \in W_{1,j}$ because otherwise, $(m_p, M_{2,j}(m_p))$ blocks $M_{1,j}$. Hence, $M_{2,j}(m_p) \in W_{1,j} \cap Q^w_{2,j}$.

Next, consider a man m_r in $X_j \setminus P^m_{1,j}$. By the definition of $W_{1,j}$, $M_{1,j}(m_r) \in W_{1,j}$, and since $m_r \notin P^m_{1,j}$, $M_{1,j}(m_r) \in Q^w_{1,j}$ by Lemma 4.3. Hence, $M_{1,j}(m_r) \in W_{1,j} \cap Q^w_{1,j}$.

Finally, we show that the above projection is an injection. To see this, it suffices to show that there is no woman w such that $w = M_{2,j}(m_p) = M_{1,j}(m_r)$ for different men $m_p \in X_j \cap P^m_{1,j}$ and $m_r \in X_j \setminus P^m_{1,j}$. Suppose such a woman w exists. Then, $w \in Q^w_{2,j} \cap Q^w_{1,j}$ by the above observation. Since $w \in Q^w_{2,j}$, $[m_p\ M_{opt}(w)]$ in w's $I_{2,j}$-list. Since $w \in Q^w_{1,j}$, $[m_r\ M_{opt}(w)]$ in w's $I_{1,j}$-list. Since $I_{1,j} = \mathit{Shift}_m(I_{2,j})$, women's lists are the same in $I_{1,j}$ and $I_{2,j}$, which implies that $m_p = m_r$. This is a contradiction.

Hence, we have that $|W_{1,j} \cap (Q^w_{1,j} \cup Q^w_{2,j})| \geq |X_j|$. Finally, we will show that $|X_j| \geq |A_{1,j}| - |V_{1,j} \setminus V_{2,j}| - |B_{2,j}|$. Let $A^m_{1,j}$ and $A^m_{2,j}$ be the sets of men participating $A_{1,j}$ and $A_{2,j}$, respectively, so that $|A^m_{1,j}| = |A_{1,j}|$ and $|A^m_{2,j}| = |A_{2,j}|$. Then, by the definition of X_j, $X_j = A^m_{1,j} \cap A^m_{2,j}$. It then follows that

$$\begin{aligned}
|X_j| &= |A^m_{1,j}| + |A^m_{2,j}| - |A^m_{1,j} \cup A^m_{2,j}| \\
&= |A_{1,j}| + |A_{2,j}| - |A^m_{1,j} \cup A^m_{2,j}| \\
&= |A_{1,j}| + |M_{2,j}| - |B_{2,j}| - |A^m_{1,j} \cup A^m_{2,j}| \\
&= |A_{1,j}| + |V_{2,j}| - |B_{2,j}| - |A^m_{1,j} \cup A^m_{2,j}| \\
&= |A_{1,j}| + |V_{1,j} \cup V_{2,j}| - |V_{1,j} \setminus V_{2,j}| - |B_{2,j}| - |A^m_{1,j} \cup A^m_{2,j}| \\
&\geq |A_{1,j}| - |V_{1,j} \setminus V_{2,j}| - |B_{2,j}|.
\end{aligned}$$

The last inequality follows from the fact that $|V_{1,j} \cup V_{2,j}| \geq |A^m_{1,j} \cup A^m_{2,j}|$, which can be easily verified by definition. □

From the above lemmas, the following theorem holds.

Theorem 4.7 *The approximation ratio of* ShiftBrk *is at most* 13/7 *for instances where the length of ties is two.*

Proof. We prove that $\max\{|M_{1,1}|, |M_{1,2}|, |M_{2,1}|, |M_{2,2}|\} \geq \frac{7}{13}n$. For this purpose, we prove the following statement: If all of $|M_{1,1}|$, $|M_{2,1}|$ and $|M_{2,2}|$ are smaller than $\frac{7}{13}n$, then $|M_{1,2}| \geq \frac{7}{13}n$.

By Lemma 4.11, $(Q^w_{1,1} \cup Q^w_{2,1}) \cap (Q^w_{1,2} \cup Q^w_{2,2}) = \emptyset$. Hence, $|W_{1,1} \cup W_{1,2}| \geq |W_{1,1} \cap (Q^w_{1,1} \cup Q^w_{2,1})| + |W_{1,2} \cap (Q^w_{1,2} \cup Q^w_{2,2})|$. So, $|W_{1,2}| = |W_{1,1} \cup W_{1,2}| - |W_{1,1} \setminus W_{1,2}| \geq |W_{1,1} \cap (Q^w_{1,1} \cup Q^w_{2,1})| + |W_{1,2} \cap (Q^w_{1,2} \cup Q^w_{2,2})| - |W_{1,1} \setminus W_{1,2}|$. By Lemma 4.12, $|W_{1,j} \cap (Q^w_{1,j} \cup Q^w_{2,j})| \geq |A_{1,j}| - |V_{1,j} \setminus V_{2,j}| - |B_{2,j}|$ for $j = 1, 2$. By Lemma 4.10, $|W_{1,1} \setminus W_{1,2}| \leq |B_{1,1}|$. By Lemma 4.9, $|V_{1,j} \setminus V_{2,j}| \leq |B_{1,j}|$. Hence,

$$\begin{aligned}
|W_{1,2}| &\geq |W_{1,1} \cap (Q^w_{1,1} \cup Q^w_{2,1})| + |W_{1,2} \cap (Q^w_{1,2} \cup Q^w_{2,2})| - |W_{1,1} \setminus W_{1,2}| \\
&\geq |A_{1,1}| - |V_{1,1} \setminus V_{2,1}| - |B_{2,1}| + |A_{1,2}| - |V_{1,2} \setminus V_{2,2}| - |B_{2,2}| - |B_{1,1}| \\
&\geq |A_{1,1}| - |B_{1,1}| - |B_{2,1}| + |A_{1,2}| - |B_{1,2}| - |B_{2,2}| - |B_{1,1}|.
\end{aligned}$$

Now, since $|M_{1,1}| < \frac{7}{13}n$, $|A_{1,1}| > \frac{6}{13}n$ by Lemma 4.4, and $|B_{1,1}| = |M_{1,1}| - |A_{1,1}| <$

$\frac{1}{13}n$. For the same reason, $|B_{2,1}| < \frac{1}{13}n$ and $|B_{2,2}| < \frac{1}{13}n$. Hence,

$$|W_{1,2}| \geq |A_{1,2}| - |B_{1,2}| + \frac{2}{13}n.$$

Recall that $|W_{1,2}| = |M_{1,2}| = |A_{1,2}| + |B_{1,2}|$. Putting $|W_{1,2}| = |M_{1,2}|$ and $|B_{1,2}| = |M_{1,2}| - |A_{1,2}|$ to the above inequality, it follows that $|A_{1,2}| \leq |M_{1,2}| - \frac{1}{13}n$. By Lemma 4.4, we have that $|M_{1,2}| \geq n - |A_{1,2}|$. Hence, $|M_{1,2}| \geq \frac{7}{13}n$. □

4.6 Concluding Remarks

In this chapter, we presented improved approximability and inapproximability results for the restricted instances of MAX SMTI. An obvious open problem is to close the gap between the upper and lower bounds on the approximation ratio in the general case. One direction would be to try to extend the result of Sec. 4.5 to arbitrary tie-lengths.

Chapter 5

Randomized Approximation of the Stable Marriage Problem

5.1 Introduction

In Chapter 4, we showed that it is NP-hard to approximate MAX SMTI with the ratio 21/19 even if

(R1) all ties occur only in one sex,

(R2) each person writes at most one tie, and

(R3) ties are of length only two.

We also showed an approximation algorithm for MAX SMTI, whose performance ratio is better than two, for certain restricted cases. In particular, when restrictions (R1) through (R3) above hold, its performance ratio is 1.6.

In this chapter, we give a simple randomized approximation algorithm whose performance ratio is better than the above algorithm. We show that for instances for which the above restrictions (R1) through (R3) hold, its expected approximation ratio is at most 10/7 (< 1.4286).

This chapter is organized as follows. In Sec. 5.2, we present our randomized approximation algorithm RandBrk. In Sec. 5.3, we give an analysis for an easier upper bound of 5/3 on the approximation ratio of RandBrk. This helps to understand a complicated analysis for an improved bound of 10/7 given in Sec. 5.4.

This bound almost matches the lower bound of 32/23 (> 1.3913) shown in Sec. 5.2. Finally, in Sec. 5.5, we analyze the effects of lifting some of restrictions (R1) through (R3).

5.2 Algorithm RandBrk and Basic Facts

Recall that our SMTI instances satisfy three conditions (R1) through (R3). Algorithm RandBrk, which receives such an instance $\hat{I}$ and produces a stable matching for $\hat{I}$, consists of the following two steps:

Step 1. For each man m who writes a tie in $\hat{I}$, break the tie with equal probability, namely, if women w_1 and w_2 are tied in m's list, then w_1 precedes w_2 with probability 1/2, and vice versa. Let I be the resulting SMI instance.

Step 2. Find a stable matching M for I by the Gale-Shapley algorithm and output it.

Since the Gale-Shapley algorithm runs in deterministic polynomial time, RandBrk is a (randomized) polynomial time algorithm. We already know several basic facts about its correctness and performance. Let S denote the set of men who write a tie in $\hat{I}$ and $SMI(\hat{I})$ denote the set of $2^{|S|}$ different SMI instances obtained by breaking ties in $\hat{I}$ (recall that the length of ties is two).

Lemma 5.1 *[GI89] For any $I \in SMI(\hat{I})$, any stable matching for I is also stable for $\hat{I}$. (Namely* RandBrk *outputs a feasible solution.)*

Lemma 5.2 *[GS85, GI89] Let M_1 and M_2 be arbitrary stable matchings for the same SMI instance. Then (i) $|M_1| = |M_2|$ (where $|M|$ denotes the size of the matching M) and (ii) the set of men (women, resp.) matched in M_1 is exactly the same as the set of men (women, resp.) matched in M_2.*

Thus the performance of RandBrk depends only on Step 1. By this lemma, we can define $cost(I)$ for an SMI instance I as the (unique) size of stable matchings for I. Also, let $OPT(\hat{I})$ denote the size of a largest stable matching for SMTI instance $\hat{I}$.

Lemma 5.3 *[MII+02] (i) There exists $I_1 \in SMI(\hat{I})$ such that cost$(I_1) = OPT(\hat{I})$ and (ii) for any $I_2 \in SMI(\hat{I})$, cost$(I_2) \geq OPT(\hat{I})/2$.*

The reason for (ii) is easy: Suppose that m and w are matched in a largest stable matching for $\hat{I}$. Then at least one of them has a partner in any stable matching for I_2. Otherwise they are clearly a blocking pair for that matching.

It follows that any stable matching is a two-approximation of optimal size, and hence the approximation ratio of RandBrk does not exceed two. Its true value, denoted by $Cost_{RB}(\hat{I})$, is obtained by calculating the expected value for $cost(I)$, namely, by calculating

$$Cost_{RB}(\hat{I}) = \frac{1}{2^{|S|}} \sum_{I \in SMI(\hat{I})} cost(I).$$

Before analyzing $Cost_{RB}(\hat{I})$, we show a lower bound 32/23 (> 1.3913) on the approximation ratio of RandBrk. Hence, our analysis 10/7 is almost tight. Consider the following SMTI instance $\hat{I}$:

$$\begin{array}{llcll}
m_1: & (w_1\ w_4) & \qquad & w_1: & m_1 \\
m_2: & (w_2\ w_3) & & w_2: & m_2 \\
m_3: & (w_3\ w_4) & & w_3: & m_2\ m_3 \\
m_4: & w_4 & & w_4: & m_1\ m_3\ m_4
\end{array}$$

Two women in the parenthesis are tied in the list. The largest stable matching for this instance is of size 4 (m_i is matched with w_i for $1 \leq i \leq 4$). There are eight SMI instances in $SMI(\hat{I})$. The size of stable matchings for each of those eight instances are 4, 3, 3, 3, 3, 3, 2 and 2. Hence, the expected size is $(4+3+3+3+3+3+2+2)/8 = 23/8 = 23/32 \cdot OPT(\hat{I})$.

5.3 Analysis for Upper Bound 5/3

First, we give some notations and conventions. Let us fix an arbitrary SMTI instance $\hat{I}$. As defined above, S always denotes the set of men whose preference list includes a tie. We frequently say that "flip the tie of a man m of I" (although I is an SMI instance), which means that we obtain a new SMI instance I' by changing $[w_i \ w_j]$ in m's I-list into $[w_j \ w_i]$. For SMI instances I_1 and I_2, if I_1 is obtained from I_2 by flipping the tie of m, then we write $I_1 = fp(I_2, m)$ (equivalently, $I_2 = fp(I_1, m)$).

To evaluate $Cost_{RB}(\hat{I})$, we introduce the following *deterministic* algorithm called TreeGen. TreeGen accepts an SMI instance $I \in SMI(\hat{I})$ and a subset A of S, and produces a binary tree T. Each vertex v of T is associated with some instance I' in $SMI(\hat{I})$ and a subset A' of S. It should be noted that the introduction of TreeGen is only for the purpose of analysis; we are not interested in actually running it or other features, such as its time complexity (which is clearly exponential).

Procedure TreeGen(I, A)**.** (Given an SMI instance $I \in SMI(\hat{I})$ and a subset $A \subseteq S$ of men, construct a binary tree T.)

(1) Create a vertex v whose label is (I, A).

(2) If $A = \emptyset$, return v.

(3) Else, select a man m, denoted by $flip(v)$, in A, and let TreeGen$(I, A - \{m\})$ and TreeGen$(fp(I, m), A - \{m\})$ be the left child and the right child of v, respectively. (How to select $flip(v)$ will be specified later.)

We are interested in the behavior of TreeGen for the special input (I_{opt}, S), where I_{opt} is an SMI instance in $SMI(\hat{I})$ such that $cost(I_{opt}) = OPT(\hat{I})$ (its existence is due to Lemma 5.3). Then the tree T_{opt} generated by TreeGen from I_{opt} looks as follows. The root is associated with I_{opt}, which can produce an optimal stable matching M_{opt} of the original $\hat{I}$. Let $v = (I, A)$ be a vertex in T_{opt}. Then, if we select a man m in A and if we go to the left child, the associated instance does not change (and of course the associated matching size does not change). However, if we go to the right child, then its associated instance receives a single flip of m and its matching size may decrease. In both cases, m is removed from A as a "touched" man. Now the next lemma is important, which guarantees that the amount of loss in the size of matching when we go to the right child is at most one.

Lemma 5.4 *[MII+02] Let I_1 and I_2 be in $SMI(\hat{I})$ and m^* be a man in S such that $I_1 = fp(I_2, m^*)$ (equivalently, $I_2 = fp(I_1, m^*)$). Also let M_1 and M_2 be stable matchings for I_1 and I_2, respectively. Then $||M_2| - |M_1|| \leq 1$.*

Our analysis uses TreeGen in the following way. Note that the generated tree T_{opt} has exactly $2^{|S|}$ leaves, one for each instance in $SMI(\hat{I})$. Hence, $Cost_{RB}(\hat{I})$ is equal to the average value of $cost(I)$ for all $I \in SMI(\hat{I})$ since RandBrk produces each instance in $SMI(\hat{I})$ with equal probability. Let $v = (I, A)$ be a vertex of T_{opt}. Then $size(v)$ is defined to be the size of a stable matching associated with v, namely, $size(v) = cost(I)$. Now we define $ave(v)$ as follows. (i) If v is a leaf (i.e., $A = \emptyset$), then $ave(v) = size(v)$. (ii) Otherwise, $ave(v) = \frac{1}{2}(ave(l(v)) + ave(r(v)))$, where $l(v)$ (resp. $r(v)$) is the left child (resp. right child) of v. (We use these notations, $l(v)$ and $r(v)$, throughout this chapter.) The following observation is now immediate:

Observation 5.1 *$Cost_{RB}(\hat{I}) = ave(v_0)$, where v_0 is the root of T_{opt}.*

Thus all we have to do is to evaluate $ave(v_0)$. Remember that T_{opt} has the property that if we move to the left child, then the size of the stable matching is preserved and if we go to the right child, then the size may decrease by one. Then one might be curious about what kind of result can be obtained for the value of $ave(v_0)$ if we assume this worst case, i.e., if we always lose one when moving to the right. Unfortunately, the result of this analysis is very poor, or we can only guarantee a half of the size of a maximum stable matching, which means that the approximation ratio is as bad as two.

Our basic idea to avoid this worst-case scenario is as follows:

(i) If $|S|$ (= the number of ties) is small compared to $\sigma = cost(\hat{I}) = cost(I_{opt})$, say $|S| = \sigma/2$, then one can show that the above simple analysis guarantees a (good) approximation ratio of 4/3.

(ii) If $|S|$ is relatively large, then we can select a "good" man m as $flip(v)$ in Step (3) of TreeGen in the following sense: If we flip the tie of m, then either we do not lose the size of matching (Lemma 5.6), or if we do lose the size of matching then we can always select m' in the next round such that flipping his tie does not make the size decrease (Lemma 5.5).

For a vertex v in T_{opt}, let $height(v)$ be the height of v in T_{opt}, namely, if the label of v is (I, A), then $height(v) = |A|$. The proof of following lemmas will be given in Sec. 5.3.1.

Lemma 5.5 *Let $v = (I, A)$ be an arbitrary vertex in T_{opt} such that $height(v) > size(v)/2$. Suppose that for any man $m \in A$, selecting m as $flip(v)$ implies that $size(r(v)) = size(v) - 1$. Then there exist two men m_α and m_β in A such that $size(l(v)) = size(v)$, $size(r(v)) = size(v) - 1$, and $size(l(r(v))) = size(r(r(v))) = size(v) - 1$, by choosing $flip(v) = m_\alpha$ and $flip(r(v)) = m_\beta$. (Case 2-(ii) of Fig. 5.1 illustrates how $size(v)$ changes by flipping the ties of m_α and m_β.)*

Lemma 5.6 *Let $v = (I, A)$ be an arbitrary vertex in T_{opt} such that $height(v) > size(v)$. Then there exists a man $m \in A$ such that selecting m as $flip(v)$ makes $size(r(v)) = size(v)$.*

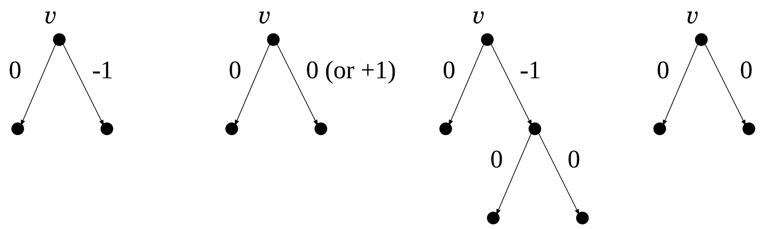


Figure 5.1: Each case of the rule

Now we select $flip(v)$ in TreeGen by the following rule (see Fig. 5.1):

Case 1. $height(v) \leq size(v)/2$. In this case, set $flip(v)$ to be an arbitrary man in A. (In this case, we assume the worst case, i.e., the size decreases in every step.)

Case 2. $size(v)/2 < height(v) \leq size(v)$.

Case 2-(i): If there exists a man $m \in A$ such that letting $flip(v) = m$ makes $size(r(v)) \geq size(v)$, then set $flip(v) = m$.

Case 2-(ii): Otherwise, set $flip(v) = m_\alpha$ and $flip(r(v)) = m_\beta$ whose existence is guaranteed by Lemma 5.5.

Case 3. $height(v) > size(v)$. Let $flip(v) = m$ that makes $size(r(v)) = size(v)$ whose existence is guaranteed by Lemma 5.6.

By the above rule, we can obtain the following lemma whose proof is given in Sec. 5.3.2.

Lemma 5.7 *For any vertex v in T_{opt}, $ave(v) \geq \frac{3}{5} size(v)$.*

By applying Lemma 5.7 to the root vertex v_0 of T_{opt}, we have that $ave(v_0) \geq \frac{3}{5} size(v_0)$. Since $size(v_0)$ is the optimal cost and $ave(v_0)$ ($= Cost_{RB}(\hat{I})$ by Observation 5.1) is the expected cost of RandBrk's output, we have the following theorem.

Theorem 5.2 *The approximation ratio of* RandBrk *is at most $\frac{5}{3}$.*

5.3.1 Proofs of Lemmas

Proof of Lemma 5.5

By the assumption of the lemma, $size(r(v)) = size(v) - 1$ no matter how we choose $flip(v)$. Clearly, $size(l(v)) = size(v)$ and $size(l(r(v))) = size(v) - 1$. We only need to show that we can choose m_α and m_β that makes $size(r(r(v))) = size(v) - 1$.

We prove some properties of G_{M_1,M_2} (see Section 2.2 for the definition) when M_1 and M_2 are closely related, whose proof is essentially given in [MII+02].

Lemma 5.8 *Let I_1 and I_2 be in $SMI(\hat{I})$, such that $I_2 = fp(I_1, m^*)$. Let M_1 and M_2 be stable matchings for I_1 and I_2, respectively. Furthermore, suppose that $[w_+ \ w_-]$ in m^*'s I_1-list. If G_{M_1,M_2} contains a path, the path contains m^*, and $M_1(m^*) = w_+$ and $M_2(m^*) = w_-$.*

Proof. Assume that there is a path which does not contain m^*, and suppose that the path starts from a man and ends with a woman. (For other cases, we can do similar arguments.) Now let the path be $m_1, w_1, m_2, w_2, \ldots, m_k, w_k$. Assume that $(m_1, w_1), (m_2, w_2), \ldots, (m_k, w_k)$ are pairs matched in M_1 and $(m_2, w_1), (m_3, w_2), \ldots,$ (m_k, w_{k-1}) are pairs matched in M_2. (See Fig. 5.2 (1).) It should be noted that, since this path does not contain m^*, preference lists of all these persons are same in I_1 and I_2.

Since m_1 is matched with w_1 in M_1, m_1's list contains w_1. Then, $m_2 \succ m_1$ in w_1's list, since otherwise, m_1 (who is single in M_2) and w_1 form a blocking pair for M_2. For the same reason, $w_2 \succ w_1$ in m_2's list. Continuing this argument along with the path, we have that $w_k \succ w_{k-1}$ in m_k's list. Also, w_k writes m_k in her list. Then it follows that m_k and w_k form a blocking pair for M_2 in I_2, which contradicts the stability of M_2.

Hence, the path is unique, say P, and P contains m^*. If m^* is an endpoint of P, we can conclude that M_1 or M_2 is unstable by doing the same argument as above, a contradiction. So, P can be written as $m_1, w_1, m_2, w_2, \ldots, w_{i-1}, m_i(= m^*), w_i, \ldots, m_{k-1}, w_{k-1}, m_k, w_k$. (Again, we can do same arguments when two endpoints are both men or both women.) Assume that $(m_1, w_1), (m_2, w_2), \ldots, (m_k, w_k)$ are pairs matched in M_1 and $(m_2, w_1), (m_3, w_2), \ldots, (m_k, w_{k-1})$ are pairs matched in M_2. (See Fig. 5.2 (2).) We will show that $w_i = w_+$ and $w_{i-1} = w_-$.

Let P_1 and P_2 be paths such that $P_1 = m_1, w_1, m_2, w_2, \ldots, w_{i-1}$ and $P_2 = w_k, m_k, w_{k-1}, m_{k-1}, \ldots, w_i$. By doing the same argument as above from m_1 to m^*, we can see that each man in P_1 gets a better partner in M_1 than in M_2, and each woman in P_1 gets a better partner in M_2 than in M_1. Similarly, each man in P_2 gets a better partner in M_2 and each woman in P_2 gets a better partner in M_1. Especially, $m^* \succ m_{i-1}$ in w_{i-1}'s list and $m^* \succ m_{i+1}$ in w_i's list. Then, it is not hard to see that $w_i \succ w_{i-1}$ in m^*'s I_1-list, as otherwise, (m^*, w_{i-1}) is a blocking pair for M_1. For the same reason, $w_{i-1} \succ w_i$ in m^*'s I_2-list. This is possible only when $w_i = w_+$ and $w_{i-1} = w_-$. Now the proof is completed. □

The following lemma is immediate from the above proof.

Lemma 5.9 *Let I_1, I_2, M_1, M_2, m^*, w_+ and w_- be same as Lemma 5.8. Suppose that G_{M_1,M_2} contains a path $P = p_1, \ldots, w_-, m^*, w_+, \ldots, p_2$, and let P_1 be the path from p_1 to w_- and P_2 be the path from p_2 to w_+. Then all men in P_1 and all women in P_2 have a better partner in M_1 than in M_2. All men in P_2 and all women in P_1 have a better partner in M_2 than in M_1.*

Lemma 5.10 *Let I_1, I_2, M_1, M_2, m^*, w_+ and w_- be same in Lemma 5.8. If $|M_1| \neq |M_2|$ then $M_1(m^*) = w_+$.*

Proof. Since $|M_1| \neq |M_2|$, G_{M_1,M_2} contains a path. Then, the proof follows from Lemma 5.8. □

The following lemma is immediate from Lemma 5.10.

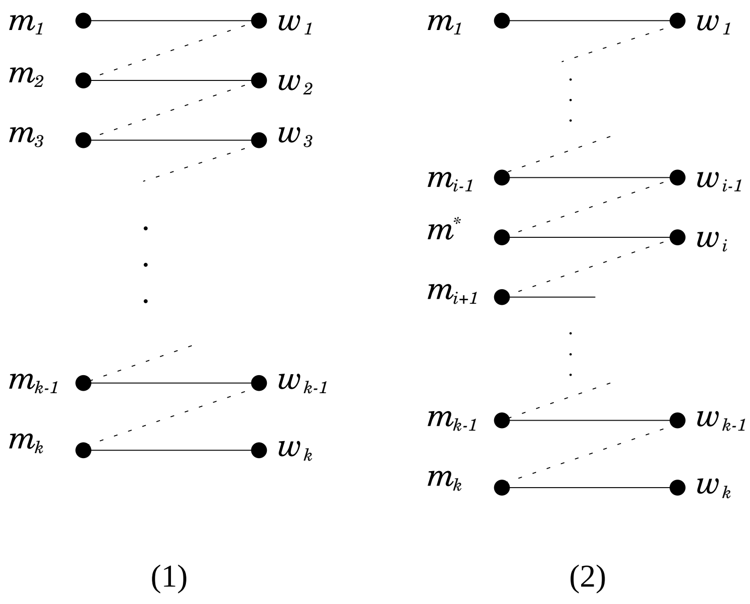


Figure 5.2: Paths in the proof of Lemma 5.8 (Solid lines represent edges in M_1, and dotted lines represent edges in M_2)

Lemma 5.11 *Let I_1, I_2, M_1, M_2, m^*, w_+ and w_- be same in Lemma 5.8. Then M_1 is stable in I_2 and hence $|M_1| = |M_2|$ if (i) m^* is single in M_1 or (ii) $M_1(m^*) \neq w_+$.*

Now we are ready to prove Lemma 5.5. Consider a vertex $v = (I, A)$ in T_{opt} satisfying the assumption of Lemma 5.5. Let M be an arbitrary stable matching for I. For a man m_i in A, denote by $[w_{i_a}\ w_{i_b}]$ the tie in m_i's preference list in I. We claim that (1) $M(m_i) = w_{i_a}$ and (2) w_{i_b} is matched in M.

Proof of Claim (1). By the assumption of Lemma 5.5, setting $flip(v) = m_i$ makes $size(r(v)) = size(v) - 1$, namely, $size(v) \neq size(r(v))$. Then Lemma 5.10 implies that $M(m_i) = w_{i_a}$. □

Proof of Claim (2). Suppose that w_{i_b} is single in M. Set $flip(v) = m_i$ and let $I_r(=fp(I, m_i))$ be an SMI instance associated with $r(v)$, the right child of v. Let M_r be a stable matching for I_r. Since we have assumed that $|M_r| = |M| - 1$, $M_r(w_{i_b}) = m_i$ by Lemma 5.10. Thus w_{i_b} is matched in M_r.

Since w_{i_b} is matched in M_r but single in M, there is a path in G_{M,M_r}, but this is the only path in G_{M,M_r}, by Lemma 5.8. This path starts from w_{i_b}, which means that the number of M_r-edges is greater than or equal to the number of M-edges in this path. Since those numbers are equal in all the other cycles, $|M_r| \geq |M|$, which contradicts the assumption that $|M_r| = |M| - 1$. $\square$

Thus we have shown that for any man m_i in A who has a tie $[w_{i_a}\ w_{i_b}]$, $M(m_i) = w_{i_a}$ and w_{i_b} is matched in M. Now let us take another man m_j in A who has a tie $[w_{j_a}\ w_{j_b}]$. We say that m_i and m_j are disjoint if $\{w_{i_a}, w_{i_b}\} \cap \{w_{j_a}, w_{j_b}\} = \emptyset$. Suppose that all pairs of m_i and m_j are disjoint. Then since none of those w_{i_a}, w_{i_b}, w_{j_a} and w_{j_b} is single as proved above, the matching size $|M|$ (=$size(v)$) is at least $2 \cdot |A|$, which is equal to $2 \cdot height(v)$. This implies $size(v) \geq 2 \cdot height(v)$, which contradicts the assumption of this lemma. Hence there must be a pair of m_i and m_j that are not disjoint. Without loss of generality, we consider the following two cases: (1) $w_{i_b} = w_{j_a}$ or (2) $w_{i_b} = w_{j_b}$. (Note that w_{i_a} and w_{j_a} are matched in M with m_i and m_j, respectively, namely, $w_{i_a} \neq w_{j_a}$.)

Case (1): In this case, set $flip(v) = m_i$ and $flip(r(v)) = m_j$. Let I_r and I_{rr} be SMI instances associated with $r(v)$ and $r(r(v))$, respectively, and let M_r and M_{rr} be stable matchings for I_r and I_{rr}, respectively. By Lemma 5.10, $M_r(m_i) = w_{i_b}(= w_{j_a})$. This means that $M_r(m_j) \neq w_{j_a}$. Hence by Lemma 5.11, $|M_{rr}| = |M_r|(= |M| - 1)$ as desired.

Case (2): For clarity, let w_b denote $w_{i_b}(= w_{j_b})$. Without loss of generality, suppose that $m_i \succ m_j$ in w_b's list. Then we set $flip(v) = m_i$ and $flip(r(v)) = m_j$. Let I_r, I_{rr}, M_r and M_{rr} be same as Case (1). Note that, by Lemma 5.10, $M_r(m_i) = w_b$. Then it turns out that M_r is stable in I_{rr}. (Reason: Assume that M_r is stable in I_r but not stable in I_{rr}. An easy observation shows that the blocking pair must be (m_j, w_b). However, $M_r(m_i) = w_b$ as mentioned above. So it is impossible for this pair to block M_r in I_{rr} because $m_i \succ m_j$ in w_b's list.) Hence $|M_{rr}| = |M_r|$ by Lemma 5.2 (i). $\square$

Proof of Lemma 5.6

Consider a vertex $v = (I, A)$ with $height(v) > size(v)$. Then, there must be a man in A who is single in a stable matching for I. By Lemma 5.11 (i), we can select this man as $flip(v)$, resulting in $size(r(v)) = size(v)$. $\square$

5.3.2 Performance analysis

We introduce the following function:

$$f(s,h) = \begin{cases} s & \text{for } h = 0, \\ (f(s,h-1) + f(s-1,h-1))/2 & \text{for } 0 < h \le \frac{s}{2}, \\ (f(s,h-1) + f(s-1,h-2))/2 & \text{for } \frac{s}{2} < h \le s, \\ f(s,s) & \text{for } h > s. \end{cases}$$

Observe that the function is properly defined for all integers $s \ge 2$, $h \ge 0$.

Using Lemmas 5.4, 5.5, and 5.6, it is straightforward to argue by induction that f bounds the size of the solution found by RandBrk from below.

Lemma 5.12 *For any vertex v in T_{opt}, $ave(v) \ge f(size(v), height(v))$.*

Now, define

$$g(s,h) = \begin{cases} s - \frac{h}{2} & \text{for } 0 \le h \le \frac{s}{2}, \\ \frac{9}{10}s - \frac{3}{10}h & \text{for } \frac{s}{2} < h \le s, \\ \frac{3}{5}s & \text{for } h > s. \end{cases}$$

Lemma 5.13 *$f(s,h) \ge g(s,h)$, for all $s, h \ge 0$.*

Proof. Proof is by induction on h. When $h = 0$, $f(s,h) = s = g(s,h)$. When $h \le \frac{s}{2}$, we have by the inductive hypothesis that

$$\begin{aligned} f(s,h) &= \frac{1}{2}(f(s,h-1) + f(s-1,h-1)) \\ &\ge \frac{1}{2}(g(s,h-1) + g(s-1,h-1)) \\ &= \frac{1}{2}(s - \frac{1}{2}(h-1) + (s-1) - \frac{1}{2}(h-1)) \\ &= s - \frac{1}{2}h \\ &= g(s,h). \end{aligned}$$

When $s/2 < h \le s$, we have that

$$f(s,h) \ \ge \ \frac{1}{2}(g(s,h-1) + g(s-1,h-2))$$

$$\begin{aligned}
&\geq \frac{1}{2}\left(\frac{9}{10}s - \frac{3}{10}(h-1) + \frac{9}{10}(s-1) - \frac{3}{10}(h-2)\right) \\
&= g(s,h)
\end{aligned}$$

by noting that

$$g(s,h-1) = \begin{cases} \frac{9}{10}s - \frac{3}{10}(h-1) & \text{for } \frac{s}{2}+1 < h \leq s, \\ s - \frac{h-1}{2} \geq \frac{9}{10}s - \frac{3}{10}(h-1) & \text{for } \frac{s}{2} < h \leq \frac{s}{2}+1, \end{cases}$$

and

$$g(s-1,h-2) = \begin{cases} \frac{9}{10}(s-1) - \frac{3}{10}(h-2) & \text{for } \frac{s}{2}+2 < h \leq s, \\ (s-1) - \frac{h-2}{2} \geq \frac{9}{10}(s-1) - \frac{3}{10}(h-2) & \text{for } \frac{s}{2} < h \leq \frac{s}{2}+2. \end{cases}$$

Finally, when $s < h$, we have that $f(s,h) = f(s,s) \geq g(s,s) = \frac{3}{5}s = g(s,h)$. □

Observe that $g(s,h) \geq 3s/5$ for all h. Thus, $g(s,h) \geq 3s/5$. Lemma 5.7 now follows from Lemmas 5.12 and 5.13.

5.4 Analysis for Upper Bound 10/7

In the analysis of the previous section, we assumed that after we reach a vertex v with $height(v) = size(v)/2$, the matching size decreases when traversing to the right child. However, such a worst-case scenario occurs only when the associated stable matching looks as in Fig. 5.3.

For the corresponding vertex $v = (I, A)$, $height(v) = 4$ and $size(v) = 8$. Solid lines represent the current matching. $m_{2i+1} \in A$ $(0 \leq i \leq 3)$, and $[w_{2i+1}\ w_{2i+2}]$ in m_{2i+1}'s list. If we flip m_{2i+1}'s tie, then m_{2i+1} will be matched with w_{2i+2} and the matching size decreases. It is not hard to see that in the subtree rooted at v, any right edge produces a size decrease by one. Dotted lines represent the matching associated with the rightmost leaf of this subtree. In the previous section, we assumed that for any vertex v with $height(v) = size(v)/2$, an associated stable matching has this structure. However, as one can suspect, this case is not likely to happen starting from the same root vertex of T_{opt} since this structure is very special. We shall resolve

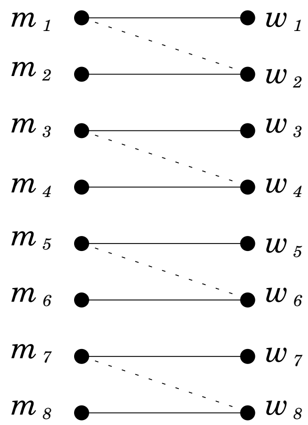


Figure 5.3: A worst case example

this by considering not only the size of the matching at each vertex but also its structure.

5.4.1 Outline of Analysis

In the previous section, we used only two parameters for the label of each vertex of T_{opt}. In this section, we use more sophisticated label

$$L(v) = (I, A, M, Y),$$

where I and A are as before the associated SMI instance and the set of not-yet-selected men, respectively, M is a stable matching for I, and Y is a subset of the women in $\hat{I}$. The basic idea of using this label for a better analysis is as follows.

Consider a vertex v with $L(v) = (I, A, M, Y)$. Recall that when creating T_{opt}, TreeGen selects a man m from A, and creates left and right children so that the matching size $|M|$ does not decrease too much, even in traversing to the right. To this end, we have derived some properties on stable matchings according to the size of M and A, namely, $size(v)$ and $height(v)$.

This time, we introduce an invariant $\delta(I, A, M, Y)$ that ensures that if we select a man in $M(Y) \cap A$, then the current matching M is also stable in the instance $fp(I, m)$ (Lemma 5.16). Here $M(Y)$ is defined to be $\{M(w) | w \in Y\}$. Therefore, TreeGen selects a man m from $M(Y) \cap A$ if the set is not empty, that does not decrease the size of matching in the right child.

Otherwise, i.e., if $M(Y) \cap A$ is empty, TreeGen selects a man m from $A - M(Y)$. Let $L(r(v)) = (I_r, A_r, M_r, Y_r)$ be the label of the right child of v. If the size of matching M_r does decrease, then we show that we can at least add something to Y, namely, $M_r(Y_r) \cap A_r \neq \emptyset$. We also show that $\delta(I_r, A_r, M_r, Y_r)$ is maintained, which allows us to select m from $M_r(Y_r) \cap A_r$ in the right child.

After completing this labeling, we show a lower bound of $ave(v)$. (Proof is given in Sec. 5.4.4.)

Lemma 5.14 *For any vertex v in T_{opt}, $ave(v) \geq \frac{7}{10} size(v)$.*

Applying this lemma to the root of T_{opt}, the main result of this chapter is immediate.

Theorem 5.3 *The approximation ratio of* RandBrk *is at most $\frac{10}{7}$.*

What we have to do in the rest of the chapter is to show that we can design TreeGen so that it follows the above scenario and the invariant δ is maintained at every vertex of the resulting tree T_{opt}.

5.4.2 Details of TreeGen

When TreeGen processes a vertex v with $L(v) = (I, A, M, Y)$, it selects a man m in A and creates a left child $l(v)$ and a right child $r(v)$ with $L(l(v)) = (I, A - \{m\}, M_l, Y_l)$ and $L(r(v)) = (fp(I, m), A - \{m\}, M_r, Y_r)$. To describe the behavior of TreeGen, we need to specify m, M_l, Y_l, M_r and Y_r. In what follows, we first introduce a directed graph $G(I, A, M)$ to define invariant $\delta(I, A, M, Y)$. Then the detailed description of TreeGen follows.

Graph $G(I, A, M)$. Let $I \in SMI(\hat{I})$, $A \subseteq S$ and M be a stable matching for I. (Recall that S is the set of all men who write ties in the original SMTI instance $\hat{I}$.) Then the directed graph $G(I, A, M)$ is defined as follows: The set of vertices of $G(I, A, M)$ is the set of women in I who have a partner in M. There is a directed edge from vertex w_i to vertex w_j (denoted by "$w_i \to w_j$") if and only if the following conditions are met:

(i) $M(w_i) \in A$,

(ii) $[w_i\ w_j]$ in $M(w_i)$'s I-list, and

(iii) $M(w_i) \succ M(w_j)$ in w_j's I-list.

The following illustration shows what preference lists satisfying (ii) and (iii) look like, where $m_i = M(w_i)$ and $m_j = M(w_j)$.

$$
\begin{array}{llcll}
m_i\text{:} & \cdots [\ w_i \quad w_j\ ] \cdots & \qquad & w_i\text{:} & \cdots\ m_i \quad \cdots\cdots\cdots \\
m_j\text{:} & \cdots\ w_j \quad \cdots\cdots & \qquad & w_j\text{:} & \cdots\ m_i \cdots\ m_j \quad \cdots
\end{array}
$$

One can see that if $w_i \to w_j$, then flipping w_i and w_j in m_i's list creates a blocking pair (m_i, w_j) for M. Also observe that the outdegree of each vertex is at most one but the indegree may be more than one. For a vertex w of $G(I, A, M)$, let $P(w)$ denote the set $\{w' | w' \to w\}$. Suppose that $P(w) = \{w_{i_1}, w_{i_2}, \ldots, w_{i_k}\}$ and $M(w_{i_1}) \succ M(w_{i_2}) \succ \cdots \succ M(w_{i_k})$ in w's I-list. Then we write $w_{i_j} = P_j(w)$ $(1 \leq j \leq k)$. The following lemma represents a key property of the graph $G(I, A, M)$.

Lemma 5.15 *Consider a graph $G(I, A, M)$ and its vertex w_i. Suppose that $[w_i\ w_j]$ in $M(w_i)$'s I-list, w_j is matched in M, and $M(w_i) \in A$. If either w_i or w_j is isolated in $G(I, A, M)$ (both indegree and outdegree zero), then M is stable in $fp(I, M(w_i))$.*

Proof. If $M(w_i) \succ M(w_j)$ in w_j's list, then $w_i \to w_j$ by definition, contradicting the assumption. Hence, $M(w_j) \succ M(w_i)$ in w_j's list. If a blocking pair for M is created by flipping w_i and w_j in the list of $M(w_i)$, the blocking pair must be $(M(w_i), w_j)$. But this is impossible since $M(w_j) \succ M(w_i)$ in w_j's list as shown above. □

Invariant $\delta(I, A, M, Y)$. Using $G(I, A, M)$, we define an invariant $\delta(I, A, M, Y)$ consisting of three statements $(\delta 1)$ through $(\delta 3)$:

$(\delta 1)$ Any woman in Y is matched in M.

$(\delta 2)$ For any woman w in Y, w is an isolated vertex of $G(I, A, M)$.

($\delta 3$) For any woman $w \in Y$, if $[w_i\ w_j]$ in $M(w)$'s I-list and if $M(w) \in A$, then $w \neq w_i$ or w_j is matched in M. (The following preference lists show condition ($\delta 3$) for $w \in Y$.)

$$m\colon\ \cdots [\ w_i\ w_j\ ] \cdots \qquad\qquad w\colon\ \cdots\ m\ \cdots$$

The following lemma illustrates the advantage of $\delta(I, A, M, Y)$.

Lemma 5.16 *Let v be a vertex of T_{opt}, and suppose that $L(v) = (I, A, M, Y)$, $|M(Y) \cap A| \geq 1$ and $\delta(I, A, M, Y)$ holds. Then selecting any man m from $M(Y) \cap A$ makes the same M stable in $fp(I, m)$.*

Proof. Since $m \in M(Y) \cap A$, m is matched in M, say, with w, where $w \in Y$. Suppose that $[w_i\ w_j]$ in m's I-list. Since $\delta(I, A, M, Y)$ holds, $w \neq w_i$ or w_j is matched in M. If $w \neq w_i$, M is stable in $fp(I, m)$ by Lemma 5.11 (ii). Otherwise, suppose that $w = w_i$ and w_j is matched in M. Observe that w is isolated in $G(I, A, M)$ since $w \in Y$ by assumption (see ($\delta 2$) above). Hence, we can apply Lemma 5.15 to show that M is stable in $fp(I, m)$. □

Algorithm SELECT. Now we are ready to give algorithm SELECT which determines a man to be flipped when expanding a vertex of TREEGEN. Before giving details, we will show a useful lemma which is referred to several times when explaining the behavior of SELECT. The lemma is a sophisticated version of Lemma 5.4.

Lemma 5.17 *Let I_1 and I_2 be in $SMI(\hat{I})$ and m^* be a man in $\hat{I}$ such that $I_2 = fp(I_1, m^*)$. Let M_1 and M_2 be stable matchings for I_1 and I_2, respectively. Then following (i) through (v) hold.*

(i) m^* is matched in M_1 if and only if m^* is matched in M_2.

(ii) $|M_2| - |M_1| = 1$ or 0 or -1.

(iii) If $|M_2| - |M_1| = 1$, then there exists a unique woman that is matched in M_2 but single in M_1, that is, all women matched in M_1 are matched in M_2.

(iv) If $|M_2| = |M_1|$, then either (a) the set of women matched in M_2 and M_1 are the same or (b) there is a unique woman that is matched in M_1 but single in M_2 and another unique woman that is matched in M_2 but single in M_1.

(v) If $|M_2| - |M_1| = -1$, then there is a unique woman that is matched in M_1 but single in M_2.

Proof. (i) follows from the fact that m^* cannot be an endpoint of a path in G_{M_1,M_2} (see the proof of Lemma 5.8). (ii) through (v) are immediate from the fact that G_{M_1,M_2} contains at most one path (Lemma 5.8). □

Now we are ready to explain algorithm SELECT. There are several cases and each case consists of a condition part denoted by *Cond*, an operation part *Op*, and a *Comment* part. If the condition given in the *Cond* part is met, then TREEGEN executes the operation given in the *Op* part. Otherwise, TREEGEN goes to the next case, where we can assume that all the previous conditions are unsatisfied. Initially, we give the root v_0 the label $L(v_0) = (I_{opt}, S, M_{opt}, \emptyset)$ (recall that I_{opt} is an SMI instance in $SMI(\hat{I})$ whose stable matching is of the same size to the size of a largest stable matching for $\hat{I}$, and M_{opt} is an arbitrary stable matching for I_{opt}). Now, suppose that we are expanding a vertex v with $L(v) = (I, A, M, Y)$.

Case 1. *Cond:* There is a man $m \in A$ such that M is stable in $fp(I, m)$. *Op:* Select this man m and let $L(l(v)) = (I, A - \{m\}, M, Y)$ and $L(r(v)) = (fp(I, m), A - \{m\}, M, Y)$. *Comment:* If $M(Y) \cap A$ is not empty, then any m in $M(Y) \cap A$ satisfies *Cond*. Even if $M(Y) \cap A = \emptyset$, there might be an m which satisfies the *Cond*.

Case 2. *Cond:* The graph $G(I, A, M)$ contains a directed cycle, say, $w_0, w_1, w_2, \ldots, w_{k-1}, w_0$. *Op:* Construct a matching M' from M as follows: Since vertices of $G(I, A, M)$ is women who are matched in M, M has matched pairs $(m_0, w_0), (m_1, w_1), (m_2, w_2), \ldots$, and (m_{k-1}, w_{k-1}). Remove these pairs and add new pairs $(m_0, w_1), (m_1, w_2), (m_2, w_3), \ldots, (m_{k-1}, w_0)$. Select an arbitrary p $(0 \le p \le k-1)$, and let $L(l(v)) = (I, A - \{m_p\}, M, Y)$ and $L(r(v)) = (fp(I, m_p), A - \{m_p\}, M', Y)$. *Comment:* The above matching M' is stable in $fp(I, m_p)$ as will be proved in Lemma 5.18.

Case 3. *Cond:* There is a man $m \in A$ such that $P(M(m)) = \emptyset$ and $|M'| = |M|$ where M' is a stable matching for $fp(I, m)$. *Op:* In this case, $M' \neq M$ because otherwise Case 1 can be applied. We are in the situation of Lemma 5.17 (iv). (a) If sets of matched women in M and M' are the same, we let $Y_r = Y$. (b) Otherwise, there is a woman w_i who is matched in M but single in M'. (b-1) If $w_i \notin Y$, then we let $Y_r = Y$. (b-2) If $w_i \in Y$, then we let $Y_r = Y - \{w_i\}$. Set $L(l(v)) = (I, A - \{m\}, M, Y \cup \{M(m)\})$ and $L(r(v)) = (fp(I, m), A - \{m\}, M', Y_r)$.

Case 4. *Cond:* There is a man $m \in A$ satisfying $|M'| - |M| = 1$, where M' is a stable matching for $fp(I, m)$ (i.e., we *gain* one). *Op:* Let $L(l(v)) = (I, A - \{m\}, M, Y)$ and $L(r(v)) = (fp(I, m), A - \{m\}, M', Y,)$.

Case 5. This case is a bit complicated. One can see that, for any man $m \in A$ such that $P(M(m)) = \emptyset$, $|M'| - |M| = -1$, where M' is a stable matching for $fp(I, m)$. Also $G(I, A, M)$ has no directed cycle because otherwise we can apply Case 2. We can assume that $G(I, A, M)$ has at least one edge, as otherwise we can select any man in A and apply Case 1.

So we can select a woman (vertex) w whose outdegree is zero and indegree is non-zero. We apply P_1 repeatedly to this w, i.e., traverse to $P_1(w)$, then to $P_1(P_1(w))$, and so on. Finally, we get to w_a such that $P(w_a) = \emptyset$ (recall that $G(I, A, M)$ does not include cycles). Let k be the length of the path from w_a to w and select w such that this value becomes maximum (if there are more than one such w, take one arbitrarily), and let $m_a = M(w_a)$. Now we consider two cases according to the value k where the second case has three subcases (see Fig. 5.4).

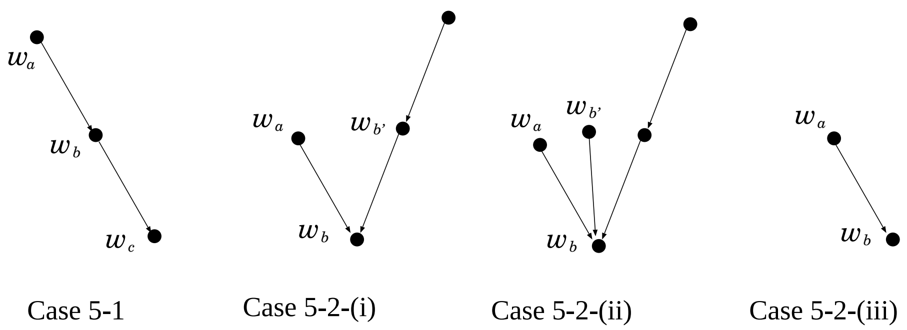


Figure 5.4: Conditions for Case 5

Case 5-1. *Cond:* $k \geq 2$. *Op:* Let w_b and w_c be women such that $w_a \to w_b$ and $w_b \to w_c$ (note that $w_c = w$ if $k = 2$). Let M' be a stable matching for $fp(I, m_a)$. Note that $P(w_a) = \emptyset$ and hence that $|M'| = |M| - 1$; otherwise, we can apply one of Cases 1 through 4 to v. Let w_i be the unique woman who is matched in M but single in M' (see Lemma 5.17 (v)).

(a) If $w_i \notin Y$ then let $Y_r = Y \cup \{w_b, w_c\}$. (b) Otherwise, i.e., if $w_i \in Y$, let $Y_r = Y \cup \{w_a, w_b, w_c\} - \{w_i\}$. We set $L(l(v)) = (I, A - \{m_a\}, M, Y \cup \{w_a\})$ and $L(r(v)) = (fp(I, m_a), A - \{m_a\}, M', Y_r)$.

Case 5-2. *Cond:* $k = 1$. *Comment:* Again, let $w_a \to w_b$. This time, $w_b = w$. Note that $|P(w_b)| \geq 1$ since $w_a \in P(w_b)$. Also, if $|P(w_b)| \geq 2$, there exists $P_2(w_b)$. We have following three subcases (see Fig. 5.4):

Case 5-2-(i). *Cond:* $|P(w_b)| \geq 2$ and $P(P_2(w_b)) \neq \emptyset$. *Op:* Let M' be a stable matching for $fp(I, m_a)$. Let w_i be the woman who is matched in M but single in M' (Lemma 5.17 (v)). (a) If $w_i \notin Y$, let $Y_r = Y \cup \{w_b\}$. (b) If $w_i \in Y$, then let $Y_r = Y \cup \{w_a, w_b\} - \{w_i\}$.

We set $L(l(v)) = (I, A - \{m_a\}, M, Y \cup \{w_a\})$ and $L(r(v)) = (fp(I, m_a), A - \{m_a\}, M', Y_r)$. (We can show that in this case, Case 5-1 can be applied to this new child $l(v)$ when determining labels of $l(l(v))$ and $r(l(v))$. This is important for Lemma 5.26 to hold. We prove this property in the proof of Lemma 5.26.)

Case 5-2-(ii). *Cond:* $|P(w_b)| \geq 2$ and $P(P_2(w_b)) = \emptyset$. *Op:* In this case, we determine not only $L(l(v))$ and $L(r(v))$, but also $L(l(l(v)))$, $L(r(l(v)))$, $L(l(r(v)))$ and $L(r(r((v)))$. Let $w_{b'} = P_2(w_b)$ and $m_{b'} = M(w_{b'})$.

First, set $L(l(v)) = (I, A - \{m_a\}, M, Y)$ and $L(l(l(v))) = (I, A - \{m_a, m_{b'}\}, M, Y \cup \{w_a, w_{b'}\})$. To determine $L(r(l(v)))$, let $I_1 = fp(I, m_{b'})$, and M_1 be a stable matching for I_1. Then $|M_1| = |M| - 1$ (this holds by the condition for Case 5 to be applied) and by Lemma 5.17 (v), there is one woman w_{i_1} who is matched in M but single in M_1. If $w_{i_1} \notin Y$, let $Y' = Y \cup \{w_a, w_b\}$. Otherwise, let $Y' = Y \cup \{w_a, w_b, w_{b'}\} - \{w_{i_1}\}$. Set $L(r(l(v))) = (I_1, A - \{m_a, m_{b'}\}, M_1, Y')$.

Finally, we determine $L(r(v))$, $L(l(r(v)))$ and $L(r(r(v))))$. Let $I_2 = fp(I, m_a)$ and $I_3 = fp(I_2, m_{b'})$. Let M_2 be a stable matching for I_2. Similarly as above, by Lemma 5.17 (v), $|M_2| = |M| - 1$ and there is one woman w_{i_2} who is matched in M but single in M_2. If $w_{i_2} \notin Y$ then let $Y' = Y \cup \{w_b, w_{b'}\}$. Otherwise, let $Y' = Y \cup \{w_a, w_b, w_{b'}\} - \{w_{i_2}\}$. Set $L(r(v)) = (I_2, A - \{m_a\}, M_2, Y')$, $L(l(r(v))) = (I_2, A - \{m_a, m_{b'}\}, M_2, Y')$ and $L(r(r(v))) = (I_3, A - \{m_a, m_{b'}\}, M_2, Y')$. *Comment:* The stability of M_2 in I_3 will be proved in Lemma 5.19 in the next section.

Case 5-2-(iii). *Cond:* $|P(w_b)| = 1$. *Op:* Let M' be a stable matching for $fp(I, m_a)$. Let w_i be the unique woman who is matched in M but single in M'. (a) If $w_i \notin Y$, then let $Y_r = Y \cup \{w_b\}$. (b) If $w_i \in Y$, then let $Y_r = Y \cup \{w_a, w_b\} - \{w_i\}$.

Set $L(l(v)) = (I, A - \{m_a\}, M, Y \cup \{w_a, w_b\})$ and $L(r(v)) = (fp(I, m_a), A - \{m_a\}, M', Y_r)$.

5.4.3 Validity of Labeling

In this section, we prove that the labeling scheme of algorithm Select is valid, namely, for each vertex v such that $L(v) = (I, A, M, Y)$, M is stable in I and the invariant $\delta(I, A, M, Y)$ holds.

Matching Stability

Note that nontrivial cases are $r(v)$ of Case 2 and $r(r(v))$ of Case 5-2-(ii).

Lemma 5.18 *The matching M' constructed in Case 2 is stable in $fp(I, M(w_p))$.*

Proof. First, we show that M' is stable in I. By the definition of $G(I, A, M)$, $[w_i\ w_{i+1}]$ in m_i's I-list, and $m_i \succ m_{i+1}$ in w_{i+1}'s list, for $0 \le i \le k-1$. (m_k and w_k represent m_0 and w_0, respectively.) Since M is stable in I, if there were a blocking pair for M', it must be (m_i, w_i) for some i. But this is impossible because w_i prefers the partner in M' $(= m_{i-1})$ to the partner in M $(= m_i)$ as shown above. Hence M' is stable in I. (The stability of M' can be seen from the fact that M' is the result of *eliminating a rotation* from M. See [GI89] for example.)

Notice that $[w_p\ w_{p+1}]$ in m_p's I-list, and $M'(m_p) = w_{p+1}$, namely, $M'(m_p) \neq w_p$. Hence by Lemma 5.11 (ii), M' is also stable in $fp(I, m_p)$. □

Lemma 5.19 *The matching M_2 constructed in Case 5-2-(ii) is stable in I_3.*

Proof. Note that $w_a \to w_b$ and $w_{b'} \to w_b$ in $G(I, A, M)$, which means that $[w_a\ w_b]$ in m_a's list and $[w_{b'}\ w_b]$ in $m_{b'}$'s I-list. Recall that I_2 is the result of flipping the list of m_a in I. Hence, $[w_b\ w_a]$ in m_a's I_2-list. Since $|M_2| = |M| - 1$ by assumption, $M_2(w_b) = m_a$ by Lemma 5.10.

Now, I_3 is obtained by flipping the tie of $m_{b'}$ of I_2, from $[w_{b'}\ w_b]$ to $[w_b\ w_{b'}]$. If M_2, which is stable in I_2, were not stable in I_3, the blocking pair must be $(m_{b'}, w_b)$.

Recall that $P_1(w_b) = w_a$ and $P_2(w_b) = w_{b'}$, which means that $m_a \succ m_{b'}$ in w_b's list. Then, it is impossible for $(m_{b'}, w_b)$ to be a blocking pair for M_2 since $M_2(w_b) = m_a$ as shown above. □

Invariant $\delta(I, A, M, Y)$

Here, we show that invariant δ is maintained for all vertices in T_{opt}. This can be done by investigating each case of algorithm SELECT. As one can see later, the hardest part of the proof is to show that (δ2) and (δ3) hold. Before proving the main part, we give some useful lemmas that help to simplify the proof.

For the following lemmas (Lemma 5.20 through Lemma 5.23) we assume the following:

(A) v and v' are vertices in T_{opt} such that v' is a child of v.

(B) $L(v) = (I, A, M, Y)$, $L(v') = (I', A', M', Y')$ and $\delta(I, A, M, Y)$ holds.

(C) $L(v')$ is determined by applying one of Cases 2 through 5 in Sec. 5.4.2 to v.

Notice that we cannot apply Case 1 to v because of (C) above. This fact is often used in the proof of Lemmas 5.20 through 5.24 hence should be kept in mind.

Lemma 5.20 *For any man m and woman w, if the following (a) or (b) is true, then (c) or (d) is true.*

(a) m is single in M.

(b) $M(m) \not\to w$ (which means that there is no edge from $M(m)$ to w) in $G(I, A, M)$.

(c) m is single in M'.

(d) $M'(m) \not\to w$ in $G(I', A', M')$.

Proof. First, consider the case that $m \notin A'$. In this case, we can conclude (c) or (d) directly (without assuming (a) or (b)) as follows: If m is single in M', then (c) follows. So suppose that m is matched in M'. Since we assume that $m \notin A'$, there is no edge from $M'(m)$ in $G(I', A', M')$, and (d) follows. (Recall the definition of graph $G(I', A', M')$. If there is an edge from a woman w', then her partner $M'(w')$ must be in A'.)

Next, suppose that $m \in A'$. Then $m \in A$ since $A' \subset A$. First, assume that condition (a) is true. Then, by Lemma 5.11 (i), M is stable in $fp(I, m)$, contradicting the assumption that Case 1 cannot be applied to v. Next, assume condition (b), namely $M(m) \not\rightarrow w$ in $G(I, A, M)$. Suppose that $[w_+ \ w_-]$ in m's I-list. Then, by the definition of the graph G, at least one of the following four conditions are true: (b1) $m \notin A$, (b2) $M(m) \neq w_+$, (b3) $w \neq w_-$, (b4) $M(w) \succ m$ in w's list. Case (b1) is impossible since $A' \subset A$ and because we assume that $m \in A'$. If (b2) holds, we have the same contradiction as for the case assuming condition (a). Now suppose that (b3) holds. Since $m \in A'$, m is not selected as a flipped man, that is, m's list is same in I and I'. Hence it is not the case that $[w_* \ w]$ in m's I'-list, where w_* is any woman. If m is single in M', then (c) holds. Otherwise, if m is matched in M', there is no edge from $M'(m)$ to w in $G(I', A', M')$ and (d) holds. (Note that, for an edge $M'(m) \rightarrow w$ to exist, there must be $[w_* \ w]$ in m's I'-list for some woman w_*, by definition.) Finally, suppose that (b2) and (b3) do not hold but (b4) holds. Then, again M is stable in $fp(I, m)$, a contradiction. □

Lemma 5.21 *Suppose that a woman w has different partners in M and M', i.e., $M(w) \neq M'(w)$. Then, at v', ($\delta 3$) holds for w. Namely, assuming that $[w_i \ w_j]$ in $M'(w)$'s I'-list, if $M'(w) \in A'$ then $w \neq w_i$ or w_j is matched in M'.*

Proof. Suppose $M'(w) \in A'$. Since Case 1 cannot be applied to v by assumption, it results from Lemma 5.11 that $M(M'(w)) = w_i$. Since $M(w) \neq M'(w)$ by assumption, we have that $w \neq w_i$ and hence the lemma follows. □

Lemma 5.22 *If $w \in Y$ and $w \in Y'$, then following (i) and (ii) hold.*

(i) At v', ($\delta 2$) holds for w. Namely, w is isolated in $G(I', A', M')$.

(ii) At v', ($\delta 3$) holds for w. Namely, assuming that $[w_i\ w_j]$ in $M'(w)$'s I'-list, if $M'(w) \in A'$ then $w \neq w_i$ or w_j is matched in M'.

Proof. (i) Suppose that w has an edge in $G(I', A', M')$, then there is a man $m \in A'$ such that his list is $[w_+\ w_-]$ in I', where $w_+ = w$ or $w_- = w$. In the following, we show that $w_+ \notin Y$ and $w_- \notin Y$, which is a contradiction since we assume that $w \in Y$.

First, suppose that $w_+ \in Y$. As Case 1 cannot be applied to v, $M(m) = w_+$ by Lemma 5.11. Also, since $A' \subset A$, $m \in A$. Then, it results that $m \in A \cap M(Y)$. We can see that M is stable in $fp(I, m)$ by Lemma 5.16, a contradiction. Hence $w_+ \notin Y$.

Next, assume that $w_- \in Y$. Then, w_- has a partner in M. Since $\delta(I, A, M, Y)$ holds, w_- is isolated in $G(I, A, M)$. Also $m \in A$. Then, we can apply Lemma 5.15 and show that M is stable in $fp(I, m)$, leading to the same contradiction as above. Consequently, $w_- \notin Y$.

(ii) Since $w \in Y$ and Case 1 cannot be applied to v, $M(w) \notin A$ by Lemma 5.16. So, if $M'(w) = M(w)$, then $M'(w) \notin A'$ because $A' \subset A$, and hence we are done. If $M'(w) \neq M(w)$, we are done by Lemma 5.21. $\square$

Lemma 5.23 *Suppose that $w \rightarrow w'$ in $G(I, A, M)$, and w' is matched in M'. Then, at v', ($\delta 3$) holds for w. Namely, assuming that $[w_i\ w_j]$ in $M'(w)$'s I'-list, if $M'(w) \in A'$ then $w \neq w_i$ or w_j is matched in M'.*

Proof. If $M'(w) \neq M(w)$, then we are done by Lemma 5.21. So suppose that $M'(w) = M(w)$. Since $w \rightarrow w'$ in $G(I, A, M)$, it must be $[w\ w']$ in $M(w)$'s I-list. Suppose $M'(w) \in A'$. Then, we have $[w\ w']$ in $M'(w)$'s I'-list because $M'(w)$'s list is same in I and I', namely, $w_j = w'$. Since we assume that $w'(= w_j)$ is matched in M', ($\delta 3$) holds. $\square$

The following lemma assumes same conditions (A) and (B) as the four preceding lemmas, but for condition (C), "applying one of Cases 2 through 5" should be replaced by "applying Case 5."

Lemma 5.24 *Let v and v' be vertices in T_{opt} such that v' is the right child of v. Let $L(v) = (I, A, M, Y)$ and $L(v') = (I', A', M', Y')$. Suppose that $\delta(I, A, M, Y)$*

holds and that $L(v')$ *is determined by applying Case 5 in Sec. 5.4.2 to* v *(Hence* $I' = fp(I, M(w_a)))$*. Then following (i) and (ii) hold.*

(i) Suppose that there is a woman w_b *in* $G(I, A, M)$ *such that* $w_a = P_1(w_b)$*. Then* w_b *is isolated in* $G(I', A', M')$ *and* $M(w_b) \neq M'(w_b)$*.*

(ii) Furthermore, if there is a woman w_c *in* $G(I, A, M)$ *such that* $w_b = P_1(w_c)$*, namely, in Case 5-1, then* w_c *is isolated in* $G(I', A', M')$ *and* $M(w_c) \neq M'(w_c)$*.*

Proof. Let $m_a = M(w_a)$, $m_b = M(w_b)$ and $m_c = M(w_c)$. Since $w_a \to w_b \to w_c$ in $G(I, A, M)$, these six persons' preference lists in I look as follows:

$$\begin{array}{llll} m_a: & \cdots [\ w_a \ \ w_b\ ] \cdots & w_a: & \cdots m_a \cdots\cdots\cdots \\ m_b: & \cdots [\ w_b \ \ w_c\ ] \cdots & w_b: & \cdots m_a \cdots m_b \cdots \\ m_c: & \cdots\cdots w_c \cdots\cdots & w_c: & \cdots m_b \cdots m_c \cdots \end{array}$$

(i) Since $|M'| = |M| - 1$, $m_a = M'(w_b)$ by Lemma 5.10. Because $m_b \neq m_a$, $M(w_b) \neq M'(w_b)$. Next, we will prove that w_b is isolated in $G(I', A', M')$. Since $m_a \in A - A'$, $m_a \notin A'$. Noting that $M'(w_b) = m_a$, we see that there is no edge in $G(I', A', M')$ going from w_b. To show that there is no edge entering w_b, we consider three types of men, and show that, for each type, their partners in M' have no edge to w_b in $G(I', A', M')$. The first type is a man who is single in M, and the second type is a man m such that $M(m) \not\to w_b$ in $G(I, A, M)$. For such m, we can apply Lemma 5.20 to show that $M'(m) \not\to w_b$ in $G(I', A', M')$. (More precisely, either $M'(m)$ does not exist or $M'(m) \not\to w_b$.) Next, we consider m such that $M(m) \to w_b$ in $G(I, A, M)$. If $m = m_a$, we have shown that $M'(m_a) = w_b$, and hence, it is clear that $M'(m) \not\to w_b$ in $G(I', A', M')$. If $m \neq m_a$, $m = M(P_k(w_b))$ for $k \geq 2$. Let this man be m_k for simplicity. Note that w_b is matched with m_a in M'. Hence, for $M'(m_k)$ to have an edge to w_b in $G(I', A', M')$, m_k must precede m_a in w_b's list. But this is impossible because $m_a = P_1(w_b)$ in $G(I, A, M)$, which means that $m_a \succ m_k$ in w_b's list.

(ii) We consider two cases: $M'(m_b) = w_c$ and $M'(m_b) \neq w_c$.

First suppose that $M'(m_b) = w_c$. Then clearly $M'(w_c) \neq M(w_c)$ because $M(w_c) = m_c$ and $m_b \neq m_c$. Next we show that w_c is isolated in $G(I', A', M')$. First, let us see that there is no edge from w_c. Since $w_b \to w_c$ in $G(I, A, M)$, $[w_b\ w_c]$ in m_b's I-list.

Also, since $m_b \notin A - A'$, m_b's list is same in I and I'. Furthermore, our assumption here is that $m_b = M'(w_c)$. Hence $[w_b\ w_c]$ in $M'(w_c)$'s I'-list. Notice that for w_c to have an edge from it in $G(I', A', M')$, it must be the case that $[w_c\ w_*]$ for some woman w_* in $M'(w_c)$'s I'-list. Hence we can see that there is no edge from w_c. The fact that there is no edge entering w_c can be proved similarly as in (i), replacing w_a and w_b above with w_b and w_c.

Next, suppose that $M'(m_b) \neq w_c$. We first prove that $M(w_c) \neq M'(w_c)$. Since $M(w_c) = m_c$, it suffices to show that $M'(w_c) \neq m_c$. Because $w_b \to w_c$ in $G(I, A, M)$, $m_b \succ m_c$ in w_c's list. Now, consider the graph $G_{M,M'}$ whose definition is given in Sec. 5.3.1. Since $|M'| = |M| - 1$, the connected component of $G_{M,M'}$ that contains m_a is a path. Then by Lemma 5.9, m_b prefers $M(m_b)$ to $M'(m_b)$, namely, $w_b \succ M'(m_b)$ in m_b's list (both in I and I'). Then, if $M'(w_c) \neq m_c$, (m_b, w_c) blocks M', a contradiction.

Then, we show that w_c is isolated in $G(I', A', M')$. First we show that there is no edge from w_c. Suppose that $[w_+\ w_-]$ in $M'(w_c)$'s I'-list. If there were an edge from w_c, then w_c is w_+. We show that this is not the case. We can see that, by Lemma 5.11, $w_+ = M(M'(w_c))$. But we have already proven above that $M(w_c) \neq M'(w_c)$, namely, $w_c \neq M(M'(w_c))$. Hence $w_+ \neq w_c$.

Next, we show that there is no edge entering w_c in $G(I', A', M')$. As before, we first consider a man m such that m is single in M or $M(m) \not\to w_c$ in $G(I, A, M)$. For this man m, we can apply Lemma 5.20 to show that $M'(m) \not\to w_c$ in $G(I', A', M')$. Next we consider m such that $M(m) \to w_c$ in $G(I, A, M)$. Suppose that $m = m_b$, namely, $M(m) = w_b$. Since $w_b \to w_c$ in $G(I, A, M)$, $[w_b\ w_c]$ in m_b's I-list. Since $m_b \notin A - A'$, m_b's list is same in I and I', namely, $[w_b\ w_c]$ also in I'. Now if there were an edge $M'(m_b) \to w_c$ in $G(I', A', M')$, $M'(m_b)$ must be w_b. But this is impossible because $M'(w_b) = m_a$ as we have shown in (i).

Finally suppose that $m \neq m_b$. Since $w_b = P_1(w_c)$, $m_b \succ m$ in w_c's list. Again, by considering $G_{M,M'}$ and Lemma 5.9, we can show that $M'(w_c) \succ m_b$ in w_c's list. Thus $M'(w_c) \succ m$ in w_c's list. Now, by definition, there is no edge from $M'(m)$ to w_c in $G(I', A', M')$. $\square$

Now we are ready to give the main lemma of this section.

Lemma 5.25 *Suppose that T_{opt} is constructed by* TreeGen *using algorithm* Select *described in Sec. 5.4.2. Then, for any vertex v of T_{opt} with $L(v) = (I, A, M, Y)$, $\delta(I, A, M, Y)$ holds. Namely ($\delta 1$) Any woman in Y is matched in M. ($\delta 2$) For any woman w in Y, w is an isolated vertex of $G(I, A, M)$. ($\delta 3$) For any woman $w \in Y$, let $[w_i\ w_j]$ in $M(w)$'s I-list. If $M(w) \in A$ then $w \neq w_i$ or w_j is matched in M.*

Proof. It is not hard to verify that ($\delta 1$) holds for every vertex. Hence we prove that ($\delta 2$) and ($\delta 3$) hold.

First of all, we show that the invariant holds for the root vertex v_0. This is obvious because $L(v_0) = (I_{opt}, S, M_{opt}, \emptyset)$, i.e., the fourth argument is $\emptyset$.

Let $L(v) = (I, A, M, Y)$, $L(l(v)) = (I_l, A_l, M_l, Y_l)$ and $L(r(v)) = (I_r, A_r, M_r, Y_r)$. Suppose that $\delta(I, A, M, Y)$ holds. We show that statements ($\delta 2$) and ($\delta 3$) also holds for $l(v)$ and $r(v)$. (In Case 5-2-(ii), we will give proofs for the four grandchildren of v.)

Case 1. Suppose that $L(l(v))$ and $L(r(v))$ are determined using Case 1 of Sec. 5.4.2 (we sometimes say that "Case 1 is applied to v" for this meaning). Then $M_l = M_r = M$ and $Y_l = Y_r = Y$.

We first show that ($\delta 2$) holds for $l(v)$. First of all, notice that the set of vertices is same in $G(I, A, M)$ and $G(I_l, A_l, M_l)$ because $M_l = M$. Since $Y_l = Y$, all we need to show is that if a vertex is isolated in $G(I, A, M)$, it is also isolated in $G(I_l, A_l, M_l)$. So, we show that if there is an edge from w_i to w_j in $G(I_l, A_l, M_l)$, then there is an edge from w_i to w_j in $G(I, A, M)$. Suppose that $w_i \to w_j$ in $G(I_l, A_l, M_l)$. Then, by definition for $w_i \to w_j$, we have that (a) $M_l(w_i) \in A_l$, (b) $[w_i\ w_j]$ in $M_l(w_i)$'s I_l-list and (c) $M_l(w_i) \succ M_l(w_j)$ in w_j's I_l-list. By combining the above (a) with the facts that $M_l = M$ and $A_l \subset A$, we have (a') $M(w_i) \in A$. Because $I_l = I$, (b') $[w_i\ w_j]$ in $M(w_i)$'s I-list, and (c') $M(w_i) \succ M(w_j)$ in w_j's I-list. A combination of conditions (a') through (c') is exactly the definition for $w_i \to w_j$ in $G(I, A, M)$. Hence ($\delta 2$) holds. (We can do the same argument for $r(v)$.)

Next, we show that ($\delta 3$) holds for $l(v)$. Let $w \in Y_l$ such that $M_l(w) \in A_l$. Since $Y = Y_l$, $w \in Y$. Because $\delta(I, A, M, Y)$ holds, if $M(w) \in A$, then $w \neq w_i$ or w_j is matched in M, where $[w_i\ w_j]$ in $M(w)$'s I-list. Since $M_l = M$ and $M_l(w) \in A_l$, it must be the case that $[w_i\ w_j]$ in $M_l(w)$'s I_l-list. $M_l(w) \in A_l$ means that $M(w) \in A$, and hence $w \neq w_i$ or w_j is matched in M. Noting that $M_l = M$ and $Y_l = Y$, we can conclude that $w \neq w_i$ or w_j is matched in M_l, and hence ($\delta 3$) follows. (Again, we can do the same argument for $r(v)$.)

From now on, we prove for Cases 2 through 5. Recall that these cases are applied when Case 1 cannot be applied to v. Suppose that we want to prove that ($\delta 2$) and ($\delta 3$) hold for Y_l. Then we have to prove that some properties are satisfied for *all* women in Y_l. However, by Lemma 5.22, we know that if $w \in Y_l$ is also in Y, then those properties are satisfied for w. Hence what we need to do is to verify properties for women in the difference $Y_l - Y$. This argument also holds for Y_r.

Cases 2 and 4. Note that, in these cases, $Y_l = Y_r = Y$. Hence the differences are empty and we are done by Lemma 5.22.

Case 3. We first give a proof for $l(v)$. Since $Y_l - Y = \{M(m)\}$, where $m \in A - A_l$, we will check properties for this woman. First of all, note that $M_l = M$ and hence $M_l(m) = M(m)$. By condition for Case 3, $P(M(m)) = \emptyset$ in $G(I, A, M)$. Hence $P(M_l(m)) = \emptyset$ in $G(I_l, A_l, M_l)$ by Lemma 5.20, namely, there is no edge entering $M_l(m)$. Also, since $m \notin A_l$, there is no edge going from $M_l(m)$ in $G(I_l, A_l, M_l)$. Consequently, $M_l(m)$ is isolated in $G(I_l, A_l, M_l)$. Hence ($\delta 2$) holds for $M_l(m)$. Since $m \notin A_l$, ($\delta 3$) clearly holds for $M_l(m)$.

For $r(v)$, note that $Y_r \subseteq Y$, namely, $Y_r - Y = \emptyset$. As mentioned before, we are done (by Lemma 5.22).

Case 5-1. For $l(v)$, we can prove in exactly the same way as in Case 3 above, by replacing m, $M(m)$ and $M_l(m)$ by $M(w_a)$, w_a and w_a, respectively.

Next, we consider $r(v)$. We consider women in $Y_r - Y$. In Case 5-1 (a), $Y_r - Y = \{w_b, w_c\}$. Since $I_r = fp(I, m_a)$, w_b and w_c are isolated in $G(I_r, A_r, M_r)$ by Lemma 5.24. Thus ($\delta 2$) holds for w_b and w_c. Also, we have that $M(w_b) \neq M_r(w_b)$, $M(w_c) \neq M_r(w_c)$ by Lemma 5.24. Hence we can apply Lemma 5.21 to w_b and w_c to show that ($\delta 3$) holds for w_b and w_c.

Then, consider Case 5-1 (b). In this case, $Y_r - Y = \{w_a, w_b, w_c\}$. For w_b and w_c, we can do the same argument as above. Hence we consider only w_a. Since $P(w_a) = \emptyset$ in $G(I, A, M)$, we have that $P(w_a) = \emptyset$ in $G(I_r, A_r, M_r)$ by Lemma 5.20.

Furthermore, since $M(w_a) \notin A_r$, w_a's outdegree is 0 in $G(I_r, A_r, M_r)$. Hence w_a is isolated, and ($\delta 2$) holds. Since $w_a \to w_b$ and $w_b \in Y_r$ (namely w_b is matched in M_r), we can use Lemma 5.23 to show that ($\delta 3$) holds for w_a.

Case 5-2-(i). For $l(v)$, the proof is almost the same as in Case 5-1. For $r(v)$, $Y_r - Y = \{w_b\}$ in Case (a) and $Y_r - Y = \{w_a, w_b\}$ in Case (b). One can prove in the similar way as Case 5-1.

Case 5-2-(ii). Let Y_{ll}, Y_{rl}, Y_{lr} and Y_{rr} be the fourth arguments of $L(l(l(v)),$ $L(r(l(v))$, $L(l(r(v))$ and $L(r(r(v))$, respectively.

Note that $Y_l = Y$, $Y_{ll} - Y_l = \{w_a, w_{b'}\}$, $Y_{rl} - Y_l = \{w_a, w_b\}$ or $\{w_a, w_b, w_{b'}\}$, $Y_r - Y = \{w_a, w_{b'}\}$ or $\{w_a, w_b, w_{b'}\}$, $Y_{rl} = Y_{rr} = Y_r$. By considering Lemma 5.22, all we have to show is that ($\delta 2$) and ($\delta 3$) hold for w_a, w_b and $w_{b'}$ at corresponding vertices. This might be a long proof but the method is exactly same as we have done until now. Hence it is omitted.

Case 5-2-(iii). This can be done in exactly the same way as in Case 5-2-(i). □

5.4.4 Performance Analysis

Consider any vertex v of T_{opt} with the label $L(v) = (I, A, M, Y)$, and recall the definition of $size(v)$ and $ave(v)$ (which is given in Sec. 5.3): For any vertex v, $size(v) = |M|$. If v is a leaf of T_{opt}, then $ave(v) = size(v)$, and if v is a non-leaf vertex, then $ave(v) = \frac{1}{2}(ave(l(v)) + ave(r(v)))$.

For non-negative integers m and y, define

$$g(m, y) = \frac{7}{10}m + \frac{3}{10}y.$$

In the following, we show that $ave(v) \geq g(|M|, |Y|)$. Then, Lemma 5.14 is immediate since $ave(v) \geq g(|M|, |Y|) \geq \frac{7}{10}(|M|) = \frac{7}{10}size(v)$.

Lemma 5.26 *For any vertex v such that $L(v) = (I, A, M, Y)$, $ave(v) \geq g(|M|, |Y|)$.*

Proof. We will prove the lemma by induction. First, suppose that v is a leaf of T_{opt}. Then $ave(v) = |M| \geq \frac{7}{10}|M| + \frac{3}{10}|Y| = g(|M|, |Y|)$. (Note that by condition ($\delta 1$), any woman Y is matched in M. Therefore, $|M| \geq |Y|$.)

Next, consider a non-leaf vertex v and assume that the claim is true for all descendants of v. We will show that the claim is true for v. Let the labels of v, $l(v)$ and $r(v)$ be $L(v) = (I, A, M, Y)$, $L(l(v)) = (I_l, A_l, M_l, Y_l)$ and $L(r(v)) = (I_r, A_r, M_r, Y_r)$, respectively.

Define integers m_l, y_l, m_r and y_r as follows:

$$m_l = |M_l| - |M|,\ y_l = |Y_l| - |Y|,\ m_r = |M_r| - |M| \text{ and } y_r = |Y_r| - |Y|.$$

Also, define an integer s,

$$s = 7(m_l + m_r) + 3(y_l + y_r).$$

Then,

$$\begin{aligned} ave(v) &= \frac{1}{2}ave(l(v)) + \frac{1}{2}ave(r(v)) \\ &\geq \frac{1}{2}g(|M_l|, |Y_l|) + \frac{1}{2}g(|M_r|, |Y_r|) \\ &= \frac{1}{2}g(|M| + m_l, |Y| + y_l) + \frac{1}{2}g(|M| + m_r, |Y| + y_r) \\ &= g(|M|, |Y|) + \frac{s}{20}. \end{aligned}$$

So, it suffices to show that $s \geq 0$ for each case of SELECT. Suppose that labels of $l(v)$ and $r(v)$ are determined using Case 1. Since $|M_l| = |M_r| = |M|$ and $|Y_l| = |Y_r| = |Y|$, clearly $s = 0$. By the same method, we can calculate the value of s in each case (see Table 5.1).

For cases 5-2-(i) and 5-2-(ii), we need some special consideration:

Case 5-2-(i). First of all, we show that in this case Case 5-1 is always applied to $l(v)$. To this end, we show that *Cond* of Cases 1 through 4 do not hold for $l(v)$ but *Cond* of Case 5-1 holds.

Recall that $I_l = I$, $M_l = M$ and $A_l \subset A$, which means that associated instances

	m_l	y_l	m_r	y_r	s
Case 1	0	0	0	0	0
Case 2	0	0	0	0	0
Case 3 (a), (b-1)	0	1	0	0	3
Case 3 (b-2)	0	1	0	-1	0
Case 4	0	0	1	0	7
Case 5-1	0	1	-1	2	2
Case 5-2-(iii)	0	2	-1	1	2

Table 5.1: Values of m_l, y_l, m_r, y_r and s in each case

are the same in v and $l(v)$, and the possibility of selecting a flipping man strictly decreases. If Cases 1, 3 or 4 were able to be applied to $l(v)$ by selecting a man, say m, we could have applied to the same operation to v by selecting the same man m, a contradiction. Also, one can see that the only difference of associated graphs, G and G_l, is that w_a disappears in G_l. So, if G does not contain cycles, then neither does G_l, namely, we cannot apply Case 2 to $l(v)$. In G_l, w_a disappears, which means that $P_1(w_b) = w_{b'}$ in G_l and hence $P_1(P_1(w_b))$ exists. This is *Cond* of Case 5-1.

Now,

$$\begin{aligned} ave(v) &= \frac{1}{4}ave(l(l(v))) + \frac{1}{4}ave(r(l(v))) + \frac{1}{2}ave(r(v)) \\ &\geq \frac{1}{4}g(|M|, |Y|+2) + \frac{1}{4}g(|M|-1, |Y|+3) + \frac{1}{2}g(|M|-1, |Y|+1) \\ &= g(|M|, |Y|). \end{aligned}$$

Case 5-2-(ii). In this case, we determine all four grandchildren.

$$\begin{aligned} ave(v) &= \frac{1}{4}ave(l(l(v))) + \frac{1}{4}ave(r(l(v))) + \frac{1}{4}ave(l(r(v))) + \frac{1}{4}ave(r(r(v))) \\ &\geq \frac{1}{4}g(|M|, |Y|+2) + \frac{1}{4}g(|M|-1, |Y|+2) \qquad (5.1) \\ &\quad + \frac{1}{4}g(|M|-1, |Y|+2) + \frac{1}{4}g(|M|-1, |Y|+2) \\ &= g(|M|, |Y|) + \frac{3}{40}. \end{aligned}$$

□

5.5 More General Instances

Recall that we imposed the following restrictions to SMTI instances: (R1) Ties appear only in men's lists, (R2) each man's list includes at most one tie and (R3) the length of ties is two. In this section, we analyze the performance of RandBrk when some of these restrictions are removed.

5.5.1 Ties in Both Sides

If ties appear in both men and women's lists, the approximation factor of RandBrk becomes worse. However, if the number of people who write a tie is limited, then it is still better than two. In this section, we consider the case that the number of such people is at most the size of a largest stable matching. Given an SMTI instance $\hat{I}$, let S_m and S_w be the set of men and women, respectively, who write a tie in $\hat{I}$. We modify the analysis in Sec. 5.3. Construct a tree T_{opt} by TreeGen($I_{opt}, S_m \cup S_w$), where I_{opt} is the same as before, i.e., an SMI instance corresponding to a largest stable matching for $\hat{I}$. We can prove the following lemma which is similar to Lemma 5.5.

Lemma 5.27 *Let $v = (I, A)$ be an arbitrary vertex in T_{opt} such that $height(v) > \frac{2}{3}size(v)$. Suppose that for any person $p \in A$, flipping p's list implies that $size(r(v)) = size(v) - 1$. Then there exists a pair of persons p_α and p_β in A such that $size(l(v)) = size(v)$, $size(r(v)) = size(v) - 1$, and $size(l(r(v))) = size(r(r(v))) = size(v) - 1$, by choosing $flip(v) = p_\alpha$ and $flip(r(v)) = p_\beta$.*

Proof. Since we assume that flipping p's list implies that $size(r(v)) = size(v) - 1$, people in A are all matched (see Lemma 5.11). Partition A into A_m and A_w, the sets of men and women in A, respectively. Without loss of generality, assume that $|A_m| \geq |A_w|$. Let W be the multiset of women who appear in A_w or in the ties of men in A_m. Then $|W| = 2|A_m| + |A_w| \geq \frac{3}{2}height(v) > size(v)$. As we have discussed in the claims (1) and (2) in Sec. 5.3.1, all these women are matched. Hence at least one woman, say w, appears at least twice in W. If w appears in two men's ties, we can do the same argument in Cases (1) and (2) of Sec. 5.3.1. So, assume that w appears in some man $m (\in A_m)$'s tie and in A_w.

If $[w\ w']$ for some w' in m's I-list, then it must be the case that $[m\ m']$ for some m' in w's I-list, and that m and w are matched in a stable matching for I (Lemma 5.10). Then, choose m and w as p_α and p_β, respectively. Let I_r and I_{rr} be SMI instances associated with $r(v)$ and $r(r(v))$, and M_r and M_{rr} be stable matchings for these instances, respectively. Since $[w'\ w]$ in m's I_r-list, m is matched with w' in M_r by Lemma 5.10, namely, w is not matched with m. Hence, by Lemma 5.11 (ii), $|M_{rr}| = |M_r|$.

For the other case, namely the case that $[w'\ w]$ in m's I-list, we can do a similar argument by selecting m and w as p_α and p_β. □

We can obtain a 7/4 upper bound by modifying $g(s,h)$ in Sec. 5.3.2 as follows:

$$g(s,h) = \begin{cases} s - \frac{h}{2} & \text{for } 0 \le h \le \frac{2}{3}s, \\ \frac{6}{7}s - \frac{2}{7}h & \text{for } \frac{2}{3}s < h \le s. \end{cases}$$

5.5.2 Multiple Ties for Each Man

The upper bound does not change for this generalization. In the previous analysis, each vertex in T_{opt} is labeled with (I, A, M, Y), where A is the set of men whose preference list has not been touched yet. Here we generalize A to be the set of *ties* which have not been touched yet. T_{opt} is constructed by TREEGEN$(I_{opt}, S, M_{opt}, \emptyset)$, where S is the set of all ties in $\hat{I}$. Then, at some vertex v in T_{opt}, if a man m includes two or more untouched ties, then we can apply Case 1 by selecting his tie which does not include $M(m)$ (see Lemma 5.11). Hence we can apply Case 1 as long as there are two or more ties belonging to the same man. So, it is not hard to see that the same upper bound holds for this case.

5.5.3 Longer Ties

In this section, we show that the performance of RANDBRK becomes poor if we allow arbitrary length of ties. More precisely, we show that there exists an SMTI instance which contains ties of length ℓ, for which the approximation ratio of RANDBRK is at least $2 - O(\frac{\log \ell}{\ell})$. Here is the instance:

$$\begin{array}{llll}
m_1\colon & (w_1\ w'_1\ \cdots\ w'_{\ell-1}) & w_1\colon & m_1 \\
\vdots & \vdots & \vdots & \vdots \\
m_{\ell-1}\colon & (w_{\ell-1}\ w'_1\ \cdots\ w'_{\ell-1}) & w_{\ell-1}\colon & m_{\ell-1} \\
m'_1\colon & w'_1 & w'_1\colon & m_1\ \cdots\ m_{\ell-1}\ m'_1 \\
\vdots & \vdots & \vdots & \vdots \\
m'_{\ell-1}\colon & w'_{\ell-1} & w'_{\ell-1}\colon & m_1\ \cdots\ m_{\ell-1}\ m'_{\ell-1}
\end{array}$$

First of all, one can see that there is a stable matching of size $2\ell-2$ for this example where m_i is matched with w_i and m'_i is matched with w'_i, for $1 \le i \le \ell-1$. We then analyze the expected size of a stable matching obtained by RandBrk. Suppose that ties of men m_1 through $m_{\ell-1}$ are randomly broken in Step 1 of RandBrk. To obtain a stable matching for this SMI instance, we apply the following algorithm: First, fix the order of men to be processed as $m_1, m_2, \ldots, m_{\ell-1}, m'_1, m'_2, \ldots, m'_{\ell-1}$. Then, for each man m who has not been processed yet, match m with the woman at the highest position in m's list among those who are currently single. If there is no such woman, m remains single. It is not hard to prove that the stable matching obtained by this algorithm is exactly the same as the one obtained by the Gale-Shapley algorithm (the intuitive reason is that every woman's preference list is consistent with the order of the men as defined above).

Now let us evaluate the expected number of women w_i ($1 \le i \le \ell-1$) matched in the resulting matching. Note that m_i is the only man that writes w_i in the list. Consider man m_1. With probability $\frac{1}{\ell}$, w_1 is at the top of m_1's list. Hence m_1 is matched with w_1 with probability $\frac{1}{\ell}$, and m_1 is matched with one of $w'_1, \ldots, w'_{\ell-1}$ with probability $1-\frac{1}{\ell}$. Next, observe that m_2 will be matched with w_2 if and only if, (i) w_2 lies at the top of m_2's list, or (ii) for some w'_t, w'_t is at the top of m_2's list, w'_t is matched with m_1, and w_2 is at the second position of m_2's list. Hence the probability that m_2 is matched with w_2 is at most $\frac{1}{\ell-1}$. Similarly, the probability that m_k is matched with w_k is at most $\frac{1}{\ell-k+1}$. Hence, after $m_{\ell-1}$ is processed, the expected number of women w_i ($1 \le i \le \ell-1$) currently matched is at most

$$\sum_{k=1}^{\ell-1} \frac{1}{\ell-k+1} \le \int_1^{\ell} \frac{1}{x} dx = \ln \ell.$$

In the rest of the execution of the algorithm, w_i will never be matched since no m'_j writes her. Hence the expected size of matching is at most $\ln \ell + (\ell - 1)$. Since the maximum stable matching is of size $2\ell - 2$, the approximation ratio of RandBrk is at least

$$\frac{2\ell - 2}{\ell + \ln \ell} = 2 - O\left(\frac{\log \ell}{\ell}\right).$$

Chapter 6

Concluding Remarks

In this thesis, we studied approximability and inapproximability of the stable marriage problems. In Chapter 3, we considered the sex-equal stable marriage problem and its variant. In Chapter 4 and Chapter 5, we considered MAX SMTI.

In Chapter 3, we developed approximation algorithms for the sex-equal stable marriage problem (NSE) and MinESE. Let M_0 and M_z be the man-optimal and the woman-optimal stable matchings, respectively. NSE is the problem to obtain a stable matching M such that $|d(M)| \leq \epsilon\Delta$ for a given constant, where $\Delta = \min\{|d(M_0)|, |d(M_z)|\}$. We gave a polynomial time algorithm for NSE, which runs in time $O(n^{3+\frac{1}{\epsilon}})$. Furthermore, we considered another problem MinESE. MinESE is to find a stable matching M which minimizes $c(M)$ under the condition that $|d(M)| \leq \epsilon\Delta$ for a given constant ϵ. We showed that MinESE is NP-hard, and gave a polynomial time $(2-(\epsilon-\delta)/(2+3\epsilon))$-approximation algorithm for an arbitrary δ such that $0 < \delta < \epsilon$, whose running time is $O(n^{4+\frac{1+\epsilon}{\delta}})$.

In Chapter 4, we gave the first nontrivial result for approximation of factor less than two for MAX SMTI. Our algorithm ShiftBrk achieves an approximation ratio of $2/(1+L^{-2})$ for instances in which only men have ties of length at most L. When both men and women are allowed to have ties, but the lengths are limited to two, then we show a ratio of 13/7 (< 1.858). We also improved the lower bound on the approximation ratio to 33/29 (> 1.1379).

In Chapter 5, we gave a randomized approximation algorithm RandBrk for MAX SMTI and showed that its expected approximation ratio is at most 10/7 (< 1.4286)

for a restricted but still NP-hard case, where ties occur in only men's lists, each man writes at most one tie, and the length of ties is two. We also showed that our analysis is nearly tight by giving a lower bound 32/23 (> 1.3913) for RandBrk. Furthermore, we showed that these restrictions except for the last one can be removed without increasing the approximation ratio too much.

Our future work is to improve the gap between inapproximability results and approximability results for the stable marriage problems. With respect to MinESE, we conjecture that the lower bound of MinESE is close to two for a small $\epsilon > 0$, but we did not succeed even to show that MinESE does not admit PTAS. With respect to MAX SMTI, there remains a large gap between the current best upper bound 1.875 and the current best lower bound 33/29 (> 1.1379). Even for the restricted case, our best deterministic algorithm achieves 1.6-approximation, so further study will be necessary.

Bibliography

[AHI02] Y. Asahiro, R. Hassin, and K. Iwama. Complexity of finding dense subgraphs. *Discrete Applied Mathematics*, Vol. 121, Issue 1–3, pp. 15–26, 2002.

[CC07] M. Chlebík and J. Chlebíková. Minimum 2SAT-DELETION: Inapproximability results and relations to Minimum Vertex Cover. *Discrete Applied Mathematics*, Vol. 155, Issue 2, pp. 172–179, 2007.

[CW04] M. Charikar and A. Wirth. Maximizing quadratic programs: extending Grothendieck's inequality. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pp. 54–60, 2004.

[Dil50] R. P. Dilworth. A decomposition theorem for partially ordered sets. *Annals of Mathematics*, Vol. 51, pp. 161–166, 1950.

[DS05] I. Dinur and S. Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, Vol. 162, No. 1, pp. 439–485, 2005.

[Fei02] U. Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pp. 534–543, 2002.

[FPK01] U. Feige, D. Peleg, and G. Kortsarz. The dense k-subgraph problem. *Algorithmica*, Vol. 29, No. 3, pp. 410–421, 2001.

[Gus87] D. Gusfield. Three fast algorithms for four problems in stable marriage. *SIAM Journal on Computing*, Vol. 16, Issue 1, pp. 111–128, 1987.

[GI89] D. Gusfield and R. W. Irving. *The Stable Marriage Problem: Structure and Algorithms.* MIT Press, Boston, MA, 1989.

[GS62] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, Vol. 69, No. 1, pp. 9–15, 1962.

[GS85] D. Gale and M. Sotomayor. Some remarks on the stable matching problem. *Discrete Applied Mathematics*, Vol. 11, pp. 223–232, 1985.

[HII+03] M. M. Halldórsson, R. W. Irving, K. Iwama, D. F. Manlove, S. Miyazaki, Y. Morita, and S. Scott. Approximability results for stable marriage problems with ties. *Theoretical Computer Science*, Vol. 306, pp. 431–447, 2003.

[HIMM02] M. M. Halldórsson, K. Iwama, S. Miyazaki, and Y. Morita. Inapproximability Results on Stable Marriage Problems. In *Proceedings of the 5th Latin American Symposium on Theoretical Informatics*, Lecture Notes in Computer Science Vol. 2286, pp. 554–568, 2002.

[Irv94] R. W. Irving. Stable marriage and indifference. *Discrete Applied Mathematics*, Vol. 48, Issue 3, pp. 261–272, 1994.

[Irv98] R. W. Irving. Matching medical students to pairs of hospitals: a new variation on a well-known theme. In *Proceedings of the 6th European Symposium on Algorithms*, Lecture Notes in Computer Science Vol. 1461, pp. 381–392, 1998.

[IL86] R. W. Irving and P. Leather. The complexity of counting stable marriages. *SIAM Journal on Computing*, Vol. 15, No. 3, pp. 655–667, 1986.

[ILG87] R. W. Irving, P. Leather, and D. Gusfield. An efficient algorithm for the "optimal" stable marriage. *Journal of the ACM*, Vol. 34, Issue 3, pp. 532–543, 1987.

[IMMM99] K. Iwama, D. F. Manlove, S. Miyazaki, and Y. Morita. Stable marriage with incomplete lists and ties. In *Proceedings of the 26th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science Vol. 1644, pp. 443–452, 1999.

[IMO06] K. Iwama, S. Miyazaki, and K. Okamoto. A $(2-c\log N/N)$-Approximation Algorithm for the Stable Marriage Problem. *IEICE TRANSACTIONS on Information and Systems*, Vol. E89-D, No. 8, pp. 2380–2387, 2006.

[IMY05] K. Iwama, S. Miyazaki, and N. Yamauchi. A $(2 - c\frac{1}{\sqrt{N}})$-approximation algorithm for the stable marriage problem. In *Proceedings of the 16th International Symposium on Algorithms and Computation*, Lecture Notes in Computer Science Vol. 3827, pp. 902–914, 2005.

[IMY07] K. Iwama, S. Miyazaki, and N. Yamauchi. A 1.875–approximation algorithm for the stable marriage problem. In *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms*, pp. 288–297, 2007.

[JN83] D. S. Johnson and K. A. Niemi. On knapsacks, partitions, and a new dynamic programming technique for trees. *Mathematics of Operations Research*, Vol. 8, No. 1, pp. 1–14, 1983.

[JRMP] Japanese Resident Matching Program (JRMP), `http://www.jrmp.jp/`

[Kat93] A. Kato. Complexity of the sex-equal stable marriage problem. *Japan Journal of Industrial and Applied Mathematics (JJIAM)*, Vol. 10, pp. 1–19, 1993.

[Kho04] S. Khot. Ruling out PTAS for graph min-bisection, densest subgraph and bipartite clique. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pp. 136–145, 2004.

[KR03] S. Khot and O. Regev. Vertex cover might be hard to approximate to within $2 - \epsilon$. In *Proceedings of the 18th IEEE Annual Conference on Computational Complexity*, pp. 379–386, 2003.

[MII+02] D. F. Manlove, R. W. Irving, K. Iwama, S. Miyazaki, and Y. Morita. Hard variants of stable marriage. *Theoretical Computer Science*, Vol. 276, Issue 1–2, pp. 261–279, 2002.

[Nes98] Y. Nesterov. Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software*, Vol. 9, pp. 141–160, 1998.

[NT75] G. L. Nemhauser and L. E. Trotter. Vertex packing: structural properties and algorithms. *Mathematical Programming*, Vol. 8, pp. 232–248, 1975.

[TST99] C. P. Teo, J. V. Sethuraman, and W. P. Tan. Gale-Shapley stable marriage problem revisited: strategic issues and applications. In *Proceedings of the 7th Conference on Integer Programming and Combinatorial Optimization*, Lecture Notes in Computer Science Vol. 1610, pp. 429–438, 1999.

Publication List

- Magnús M. Halldórsson, Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa. Improved Approximation Results for the Stable Marriage Problem. *ACM Transactions on Algorithms*, Vol. 3, Issue 3, Article No. 30, August, 2007.

- Magnús M. Halldórsson, Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa. Randomized Approximation of the Stable Marriage Problem. *Theoretical Computer Science*, Vol. 325, No. 3, pp. 439–465, October, 2004.

- Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa. Approximation Algorithms for the Sex-Equal Stable Marriage Problem. In *Proceedings of the 10th Workshop on Algorithms and Data Structures* (WADS 2007), Lecture Notes in Computer Science Vol. 4619, pp. 201–213, Halifax, Canada, August, 2007.

- Magnús M. Halldórsson, Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa. Improved Approximation of the Stable Marriage Problem. In *Proceedings of the 11th annual European Symposium on Algorithms* (ESA 2003), Lecture Notes in Computer Science Vol. 2832, pp. 266–277, Budapest, Hungary, September, 2003.

- Magnús M. Halldórsson, Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa. Randomized Approximation of the Stable Marriage Problem. In *Proceedings of the 9th annual International Computing and Combinatorics Conference* (COCOON 2003), Lecture Notes in Computer Science Vol. 2697, pp. 339–350, Big Sky, USA, July, 2003.