Inapproximability Results on Stable Marriage Problems

Magnús Halldórsson[1], Kazuo Iwama[2⋆],
Shuichi Miyazaki[2] and Yasufumi Morita[2]

[1]Science Institute, University of Iceland
[2]Graduate School of Informatics, Kyoto University
mmh@hi.is, {iwama, shuichi, ymorita}@kuis.kyoto-u.ac.jp

Abstract. The stable marriage problem has received considerable attention both due to its practical applications as well as its mathematical structure. While the original problem has all participants rank all members of the opposite sex in a strict order of preference, two natural variations are to allow for incomplete preference lists and ties in the preferences. Both variations are polynomially solvable by a variation of the classical algorithm of Gale and Shapley. On the other hand, it has recently been shown to be NP-hard to find a maximum cardinality stable matching when both of the variations are allowed.
We show here that it is APX-hard to approximate the maximum cardinality stable matching with incomplete lists and ties. This holds for some very restricted instances both in terms of lengths of preference lists, and lengths and occurrences of ties in the lists. We also obtain an optimal $\Omega(N)$ hardness results for 'minimum egalitarian' and 'minimum regret' variants.

1 Introduction

An instance of the original *stable marriage problem* (*SM*) [5] consists of N men and N women, with each person having a preference list that totally orders all members of the opposite sex. A man and a woman form a *blocking pair* in a matching if both prefer each other to their current partners. A perfect matching is *stable* if it contains no blocking pair. For a matching M containing a pair (m, w), we write that $M(m) = w$ and $M(w) = m$. The stable marriage problem was first studied by Gale and Shapley [2], who showed that every instance contains a stable matching, and gave an $O(N^2)$-time algorithm, so-called the Gale-Shapley algorithm, to find one.

One natural relaxation is to allow for indifference [5, 8], in which each person is allowed to include *ties* in his/her preference. This problem is denoted *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

⋆ Supported in part by Scientific Research Grant, Ministry of Japan, 13480081

pair is called *weakly stable* (or simply "stable"). Variations in which a blocking pair can involve non-strict preference suffer from the fact that a stable matching may not exist, whereas the Gale-Shapley algorithm can be modified to always find a weakly stable matching [5].

Another natural variation is to allow participants to declare one or more unacceptable partners. Thus each person's preference list may be incomplete. Again, the definition of a blocking pair is extended, so that each member of the pair prefers the other over the current partner *or* is currently single and acceptable. The Gale-Shapley algorithm can also be modified to find a stable matching of maximum size in this case [3].

The importance of stability in matchings has been clearly displayed by its success in assigning resident interns to hospitals. For instance, the National Resident Matching Program in the U.S. has used a modified Gale-Shapley algorithm to match more than 95% of its participants for over three decades [5]. Here, residents apply to a subset of hospitals (i.e. incomplete preference lists), with each hospital strictly ranking its applicants. A hospital-resident assignment is actually a many-one matching, but most algorithms and properties carry over from the one-one SM problem that we focus on here.

Strict ranking of all applicants is not reasonable for large hospitals; it is more that they would strictly rank the top candidates, leaving the remainder tied. Irving et. al. [9] report that in a planned Scottish matching scheme SPA, ties are allowed but then resolved using arbitrary tie-breaking. However, different tie-breakings can result in different sizes of stable matchings. Since the objective is to successfully assign as many of the candidates as possible (or to fill as many of the posts as possible, depending on viewpoint), it would be desirable to find an algorithm for maximum cardinality stable matching in the presence of ties and incomplete lists.

This problem *SMTI* (Stable Marriage with Ties and Incomplete lists) was recently considered by Iwama et al. [10] which resolved that it is NP-hard to find a maximum cardinality solution. Previously, such hardness results had been known only for non-bipartite stable matchings, known as the Stable Roommates problem [12]. This NP-hardness result for SMTI was further shown to hold for the restricted case when all ties occur only at the end of a list, occur only in one sex, and are of length only 2 [11]. 2-approximation is easy for this problem since any stable matching is maximal. However, no other approximability results have been shown up to the present.

We study in this paper the approximability of SMTI and several variants. First, we show that Max cardinality SMTI is APX-hard, i.e. hard to approximate within $1+\epsilon$, for some $\epsilon > 0$. The construction applies to a very restricted class of instances, where preference lists are of constant size and are either fully ordered or contain a single tied group. We can further modify it to make ties be of length 2. An important feature of the proof is to establish a "gap location at 1"; namely, that it is NP-hard to distinguish between instances that have a complete stable matching and those where any stable matching leaves a positive constant fraction of participants unmatched.

We then consider two variants: 'minimum egalitarian' SMT and 'minimum regret' SMT. Here preference lists are complete, but the quality of the stable matching depends on how much "less preferred" partner a participant receives, either in an average or a worst case sense. It is known that both problems are NP-hard and cannot be approximated within $N^{1-\epsilon}$ for any small ϵ unless P=NP [10, 11]. We improve these results and show a worst possible $\Omega(N)$ lower bound on the approximability. Note that both problems are solvable in polynomial time if ties are not allowed [4, 5, 7].

Notation. Throughout this paper, instances contain equal number N of men and women. A goodness measure of an approximation algorithm T of an optimization problem is defined as usual: the *performance ratio* of T is the maximum over all instances x of size N of $\max\{T(x)/opt(x), opt(x)/T(x)\}$, where $opt(x)$ ($T(x)$) is the measure of the optimal (algorithm's) solution, respectively. A problem is *hard to approximate within* $f(N)$, if the existence of a polynomial-time algorithm with performance ratio $f(N)$ implies that P=NP.

2 Inapproximability of MAX SMTI

We focus on the maximum cardinality SMTI problem.

Problem: MAX SMTI.
Instance: N men, N women and each person's preference list which may be incomplete and may include ties.
Purpose: Find a stable matching of maximum cardinality.

Theorem 1. *MAX SMTI is hard to approximate within* $1+\epsilon$*, for some constant* $\epsilon > 0$.

Proof. Recall that MAX SAT is the problem of finding a truth assignment to the variables of a given propositional formula in CNF form that satisfies as many clauses as possible. MAX E3SAT(t) is a restriction of MAX SAT, where each clause has exactly three literals, and each variable appears at most t times. It is known that, if P≠NP, there exists a positive constant α and an integer t such that there is no approximation algorithm for MAX E3SAT(t) whose performance ratio is less than $1+\alpha$ [1, 6]. More precisely, the problem has a useful 'gap location' property: Suppose that an instance f of SAT is translated into an instance g of MAX E3SAT(t). If f is satisfiable, then g is also satisfiable. Otherwise, i.e., if f is unsatisfiable, the number of unsatisfied clauses is more than $\delta(=\frac{\alpha}{1+\alpha})$ fraction of clauses of g in any assignment.

We translate an instance f of MAX E3SAT(t) having the above property into an instance $T(f)$ of MAX SMTI. Our reduction has the following property: If f is satisfiable, there is a stable matching for $T(f)$ in which all the people are matched (Lemma 2). If more than δ fraction of clauses of f are unsatisfied in any assignment, then more than $\frac{\delta}{9t}$ fraction of men are single in any stable matching for $T(f)$ (Lemma 4). Hence a polynomial-time $1/(1-\frac{\delta}{9t})$-approximation algorithm implies P=NP.

The reduction is similar to that of [10], with some important simplifications. Let n and l be the numbers of variables and clauses of f, respectively, and let $C_j (1 \leq j \leq l)$ be the jth clause of f. Let t_i be the number of appearances of the variable x_i. (Thus $t = \max\{t_1, t_2, \cdots, t_n\}$.) We construct an instance $T(f)$, namely, the set of men and women, each man's preference list, and each woman's preference list.

The Set of Men and Women $T(f)$ contains $2n + 6l$ men and women. We divide men and women into three groups, respectively, in the following way:

The Set of Men

Group (a): For each clause C_j, we introduce three men a_j, a'_j and a''_j.
Group (b): For each variable x_i, we introduce two men b_i and b'_i.
Group (c): For each literal x_i (or $\overline{x_i}$) in the clause C_j, we introduce a man $c_{i,j}$.

The Set of Women

Group (u): For each variable x_i, we introduce a woman u_i.
Group (v): For each variable x_i, we introduce a woman v_i.
Group (w): For each literal x_i (or $\overline{x_i}$) in the clause C_j, we introduce two women $w^1_{i,j}$ and $w^0_{i,j}$.

Since there are n variables, we have $2n$ group-(b) men, and n group-(u) and group-(v) women. Since there are l clauses and $3l$ literals, there are $3l$ group-(a) and group-(c) men and $6l$ group-(w) women.

Men's Preference Lists We then construct each man's preference list. For better exposition, we use an example of f,

$$f_0 = (x_1 + \overline{x_2} + x_3)(\overline{x_1} + x_2 + x_4)(x_2 + \overline{x_4} + \overline{x_5}).$$

For this instance, men's preference lists will turn out to be as illustrated in Table 1, in which t is equal to 3. Note that Table 1 contains several blanks defined for convenience for constructing the women's preference lists, as detailed later.

For each clause C_j, three men a_j, a'_j and a''_j in Group (a) are introduced. We show how to construct preference lists of men a_1, a'_1 and a''_1 who are associated with $C_1 = (x_1 + \overline{x_2} + x_3)$ of f_0. Since literals x_1, $\overline{x_2}$ and x_3 appear in C_1, six women $w^0_{1,1}$, $w^1_{1,1}$, $w^0_{2,1}$, $w^1_{2,1}$, $w^0_{3,1}$ and $w^1_{3,1}$ have been introduced. a_1 writes $w^1_{1,1}$, $w^0_{2,1}$ and $w^1_{3,1}$ at the first position. (These three women are tied in the list.) Generally speaking, the man a_j writes the woman $w^1_{i,j}$ if x_i appears positively in C_j, and writes $w^0_{i,j}$ if x_i appears negatively in C_j. Both a'_1 and a''_1 write all the above six women at the first position.

Then we construct preference lists of Group-(b) men. We show how to construct preference lists using men b_2 and b'_2 who are associated with the variable x_2 of f_0. The man b_2 writes the woman u_2 at the 2nd position. (The 2nd position is always determined without depending on f.) Then, b_2 writes the woman v_2 at the $t + 4(= 7)$th position. Since x_2 appears in clauses C_1, C_2 and C_3, three women $w^0_{2,1}$, $w^0_{2,2}$ and $w^0_{2,3}$ have been introduced. b_2 writes $w^0_{2,1}$, $w^0_{2,2}$ and $w^0_{2,3}$

at the 3rd, 4th and 5th positions, respectively. Generally speaking, there are t_i women of the form $w^0_{i,j}$ corresponding to the variable x_i. (Recall that t_i is the number of appearances of the variable x_i.) b_i writes these women at 3rd through (t_i+2)th positions. Since $t_i \le t$, these women's positions never reach $(t+4)$th position which is already occupied by v_i.

b'_2's list is similarly constructed. b'_2 writes the woman u_2 at the 1st position and writes the woman v_2 at the $t+3(=6)$th position. There are three women $w^1_{2,1}$, $w^1_{2,2}$ and $w^1_{2,3}$ associated with the variable x_2 since x_2 appears in C_1, C_2 and C_3. b'_2 writes $w^1_{2,1}$, $w^1_{2,2}$ and $w^1_{2,3}$ at the 3rd, 4th and 5th positions, respectively.

Now we move to Group-(c) men. The man $c_{i,j}$ (associated with x_i in C_j) writes women $w^1_{i,j}$ and $w^0_{i,j}$ (associated with x_i in C_j) at the $t+3(=6)$th and $t+4(=7)$th positions, respectively. Now men's lists are completed. Table 1 shows the whole lists of men of $T(f_0)$. As we have mentioned before, men's lists currently contain blanks. Blanks will be removed after we construct women's lists.

Women's Preference Lists We construct women's preference lists automatically from men's preference lists. First, we determine the total order of all men; the position of each man in the order is called his *rank*. The rank of each man of our current example $T(f_0)$ is shown in Table 1, e.g., a_1 is the highest and $c_{5,3}$ is the lowest. Generally speaking, men are lexicographically ordered, where the significance of the indices of $\alpha^\delta_{\beta,\gamma}$ is in the order of α, β, γ and δ, e.g., α is the most significant index and δ is the least significant index. For α, the priority is given to a, b and c in this order. For β and γ, the smaller number precedes the larger number. For δ, fewer primes has more priority.

Women's lists are constructed based on this order. First of all, the preference list of a woman w does not include a man m if w does not appear on m's preference list. Then consider two men m_i and m_j who write w in the list. w strictly prefers m_i to m_j if and only if (1) the rank of m_i is higher than that of m_j, and (2) the position of w in m_i's list is higher than or equal to the position of w in m_j's list. One might think that women's lists can contain partial order in this construction. However, in our translation, each woman's list contains only ties.

It helps much to know that by our construction of women's preference lists, we can determine whether a matching includes a blocking pair only from men's lists. Consider men m_i and m_j matched with w_i and w_j, respectively. Then, (m_i, w_j) is a blocking pair if and only if (i) m_i strictly prefers w_j to w_i, (ii) m_i's rank is higher than m_j's rank, and (iii) the position of w_j in m_i's list is higher than or equal to the position of w_j in m_j's list. Observe that the combination of conditions (ii) and (iii) means that w_j strictly prefers m_i to m_j.

Finally, we remove blanks in men's lists. For each man's list, we simply slide women to the left until no blank remains.

Correctness of the reduction As mentioned before, the correctness follows from two Lemmas 2 and 4.

Lemma 2. *If f is satisfiable, then there is a perfect stable matching for $T(f)$.*

Proof. Suppose that f is satisfied by an assignment A and let $A(x_i) \in \{0, 1\}$ be the value assigned to x_i under A. We construct a stable matching M whose cardinality is $N(= 2n + 6l)$ as follows. Each man in Group (b) is matched according to the assignment A. If $A(x_i) = 0$, then let $M(b_i) = v_i$ and $M(b'_i) = u_i$, and if $A(x_i) = 1$, then let $M(b_i) = u_i$ and $M(b'_i) = v_i$. There is a Group-(c) man $c_{i,j}$ associated with a literal x_i (or $\overline{x_i}$) in the clause C_j. If $A(x_i) = 0$, then let $M(c_{i,j}) = w^1_{i,j}$, and if $A(x_i) = 1$, then let $M(c_{i,j}) = w^0_{i,j}$.

Now we go back to Group-(a) men. Recall that of the two women $w^0_{i,j}$ and $w^1_{i,j}$ associated with the literal x_i (or $\overline{x_i}$) in the clause C_j, one is matched with a man in Group (c) and the other one is still unmatched. Namely, if $A(x_i) = 0$ it is $w^0_{i,j}$ that is unmatched. These unmatched women will be matched with Group-(a) men. Consider a clause C_j with literals z_{i_1}, z_{i_2} and z_{i_3} (i.e., z_{i_k} is x_{i_k} or $\overline{x_{i_k}}$). Since C_j is satisfied by A, at least one of these three literals has the value 1. Without loss of generality, let the literal be z_{i_1}. If $z_{i_1} = x_{i_1}$ then $A(x_{i_1})$ must be 1 and hence, the woman $w^1_{i_1,j}$ is unmatched as mentioned above. By construction of preference lists of Group-(a) men, a_j writes $w^1_{i_1,j}$ in the list because x_{i_1} appears positively in C_j. Otherwise, i.e. if $z_{i_1} = \overline{x_{i_1}}$, then $w^0_{i_1,j}$ is unmatched and a_j writes $w^0_{i_1,j}$ in the list. In either case, a_j can be matched with the woman corresponding to the literal that makes the clause true. There are two other literals z_{i_2} and z_{i_3} in C_j. So there are two unmatched women; one is $w^0_{i_2,j}$ or $w^1_{i_2,j}$, and the other is $w^0_{i_3,j}$ or $w^1_{i_3,j}$, depending on which value x_{i_2} and x_{i_3} receive under A. a'_j and a''_j will be matched with those two women.

Now we have a perfect matching. Since we have shown how to detect blocking pairs, it is easy to check that this matching M is stable. □

Lemma 3. *Let M be an arbitrary stable matching for $T(f)$. If the number of unmatched men in M is k, then there is an assignment for f by which the number of unsatisfied clauses is at most tk.* (The proof is given in Sec. 2.1.)

Lemma 4. *If more than δ fraction of clauses of f are unsatisfied in any assignment, then more than $\frac{\delta}{9t}$ fraction of men are single in any stable matching for $T(f)$.*

Proof. Recall that the number of men in $T(f)$ is $2n+6l$, where n is the number of variables of f and l is the number of clauses of f. Since we can assume that each variable appears at least twice, we have $n \leq 3l/2$. Hence we have $2n + 6l \leq 9l$ men.

Suppose that there is a stable matching for $T(f)$ such that the number of single men is at most $\frac{\delta}{t}l$. Then, by Lemma 3, there must be an assignment for f such that the number of unsatisfied clauses is at most δl, a contradiction. Therefore, more than $\frac{\delta}{t}l$ men are single in any stable matching for $T(f)$. The fraction of men that are single exceeds $(\frac{\delta}{t}l)/9l = \frac{\delta}{9t}$. □

As we have mentioned in the beginning of this proof, there is a positive constant δ such that it is NP-hard to distinguish the following two cases for MAX E3SAT(t): (i) the formula is satisfiable, and (ii) the number of unsatisfied clauses is more than δ fraction in any assignment. By Lemmas 2 and 4, the theorem holds. □

The instances constructed in the proof above have quite restrictive properties. All preference lists are of constant size, or at most $t+2$, where t can be set as small as 5. Also, preference lists are either totally ordered, or totally unordered (i.e. a single tied list). In Sec. 2.2, we give a modification to ensure that ties are all of length 2.

2.1 Proof of Lemma 3

Proof. First of all, it should be noted that Group-(a) men are matched in any stable matching. The reason is as follows: Suppose there is a Group-(a) man m who is single in some stable matching M. Then women written on m's preference list cannot be single since otherwise, that single woman and m form a blocking pair. Thus every woman on m's list must be matched in M. Since we have assumed that m is single, at least one woman w on m's list cannot be matched with Group-(a) men, and hence she is matched with a man in Groups (b) or (c). This can be easily checked by construction. m and w form a blocking pair since Group-(w) women strictly prefers men in Group (a) to men in Groups (b) and (c).

Given an arbitrary stable matching M for $T(f)$, we first determine an *incomplete assignment* A_M which is an assignment to literals of f depending on how Group-(a) men are matched in M. A_M is not an usual truth assignment for variables but an assignment for literals which may contain several contradictions. We denote the literal x_i (resp. $\overline{x_i}$) in the clause C_j by x_i^j (resp. $\overline{x_i^j}$). Suppose x_i appears positively (resp. negatively) in C_j. Then we say that the value of the literal x_i^j (resp. $\overline{x_i^j}$) is *consistent* with the value of the variable x_i if $x_i^j = x_i$ (resp. $\overline{x_i^j} \neq x_i$). We say that two literals associated with x_i are consistent if we can assign the value to x_i so that both literals are consistent with x_i. Note that the consistency of two literals does not depend on the value of variable x_i.

Now we are ready to show how to construct A_M. As we have seen before, every man in Group (a) is matched in M. Suppose that the woman $w_{i,j}^d$ ($d \in \{0,1\}$) is matched with a man in Group (a). This woman exists in Group (w) because the variable x_i appears in the clause C_j. We assign the value to the literal x_i^j (or $\overline{x_i^j}$) so that the value of the literal is consistent with $x_i = d$. Note that, in an incomplete assignment, it can be the case that a literal receives both 0 and 1, or that a literal receives no value.

For example, recall the example in Table 1. Suppose that in a stable matching, say M_0, a_1, a_1' and a_1'' are matched with $w_{1,1}^1$, $w_{1,1}^0$ and $w_{2,1}^1$, respectively. Then, under the incomplete assignment A_{M_0}, x_1^1 receives both 0 and 1, $\overline{x_2^1}$ receives 0 (to be consistent with $x_2 = 1$), and x_3^1 receives no value. Observe that, by this incomplete assignment A_M, each clause C_j contains at least one literal whose value is 1, which corresponds to the woman who is matched with the man a_j.

For each variable x_i, define $CL_i = \{j | x_i \text{ appears in } C_j\}$. Partition CL_i into three subsets according to A_M: $CL_i^2(A_M) = \{j | j \in CL_i$ and x_i in C_j receives both 0 and 1 under $A_M\}$. $CL_i^1(A_M) = \{j | j \in CL_i$ and x_i in C_j receives exactly one value under $A_M\}$. $CL_i^0(A_M) = \{j | j \in CL_i$ and x_i in C_j receives no value under $A_M\}$.

We say that the variable x_i has *Type-I contradiction* if $CL_i^2(A_M) \neq \emptyset$. We say that x_i has *Type-II contradiction* if $CL_i^0(A_M) \neq \emptyset$. We say that x_i has *Type-III contradiction* if there are j_1 and j_2 such that $j_1, j_2 \in CL_i^1(A_M)$ and literals in C_{j_1} and C_{j_2}, associated with x_i, are not consistent. Before proving Lemma 3, we need to prove the following lemmas (Lemmas 5 through 8).

Lemma 5. *(1) If $i_1 \neq i_2$, then (i) $CL_{i_1}^2(A_M) \cap CL_{i_2}^2(A_M) = \emptyset$, and (ii) $CL_{i_1}^0(A_M) \cap CL_{i_2}^0(A_M) = \emptyset$. (2) (i) If $j \in CL_{i_1}^0(A_M)$ for some i_1, then there exists i_2 such that $j \in CL_{i_2}^2(A_M)$. (ii) If $j \in CL_{i_1}^2(A_M)$ for some i_1, then there exists i_2 such that $j \in CL_{i_2}^0(A_M)$.*

Proof. Note that every man in Group (a) is matched in M. Thus, for each clause C_j, the total number of values which three literals in C_j receive is exactly three. Then it is an easy calculation to see that the lemma holds. □

Lemma 6. *Let M be a stable matching for $T(f)$, and A_M be an incomplete assignment constructed from M. Suppose that x_i has Type-II contradiction under A_M and $j \in CL_i^0(A_M)$, namely, x_i in the clause C_j receives no value. Then the man $c_{p,j}$ is single in M. Here p is an integer such that $j \in CL_p^2(A_M)$, whose existence is guaranteed by Lemma 5 (2)-(i).*

Proof. Since $j \in CL_p^2(A_M)$, the literal z_p^j receives two values under A_M. This means that two women $w_{p,j}^1$ and $w_{p,j}^0$ associated with this literal are both matched with some men in Group (a). Then the man $c_{p,j}$ cannot be matched in M since this man writes only these two women in the list. □

Lemma 7. *Let M be a stable matching for $T(f)$, and A_M be an incomplete assignment constructed from M. Suppose that x_i has Type-III contradiction but no Type-II contradiction under A_M. Then at least one man among b_i, b'_i and $c_{i,j}$, where $j \in CL_i^1(A_M)$, is unmatched in M.*

Proof. Suppose that appearance of x_i in C_{j_1} and C_{j_2} causes Type-III contradiction. There are four cases according to the polarity of x_i in C_{j_1} and C_{j_2}. Assume that x_i appears positively in C_{j_1} and negatively in C_{j_2}, namely, each of these literals receive one value such that $x_i^{j_1} = \overline{x_i^{j_2}}$ under A_M. Other three cases are similar to this case and can be omitted. We still have two possibilities: (i) $x_i^{j_1} = \overline{x_i^{j_2}} = 1$, and (ii) $x_i^{j_1} = \overline{x_i^{j_2}} = 0$. We give the proof for (i). The other case is similar.

We assume that all b_i, b'_i and $c_{i,j}$ ($j \in CL_i^1(A_M)$) are matched in M and show a contradiction. Consider a man $c_{i,j}$ for each j such that $j \in CL_i^1(A_M)$. Since $j \in CL_i^1(A_M)$, one of $w_{i,j}^1$ and $w_{i,j}^0$ is matched with a man in Group (a) and the other is matched with the man $c_{i,j}$. Especially, $M(c_{i,j_1}) = w^0_{i,j_1}$, and $M(c_{i,j_2}) = w^1_{i,j_2}$ by our assumption that $x_i^{j_1} = \overline{x_i^{j_2}} = 1$.

Now we turn to two men b_i and b'_i. As discussed above, $w^1_{i,j}$ and $w^0_{i,j}$ ($j \in CL^1_i(A_M)$) are all matched with men in Groups (a) or (c). Also, for each j such that $j \in CL^2_i(A_M)$, $w^1_{i,j}$ and $w^0_{i,j}$ are both matched with men in Group (a). Note that there is no j such that $j \in CL^0_i(A_M)$ since x_i does not contain Type-II contradiction. As a result, all Group-(w) women written on the list of b_i or b'_i are matched with men in Group (a) or (c). Thus if both b_i and b'_i are matched in M, it must be one of the following two cases: (1) $M(b_i) = v_i$ and $M(b'_i) = u_i$, and (2) $M(b_i) = u_i$ and $M(b'_i) = v_i$. In case (1), b_i and w^0_{i,j_1} form a blocking pair. In case (2), b'_i and w^1_{i,j_2} form a blocking pair. In either case, it contradicts the fact that M is stable. □

Lemma 8. *Let M be a stable matching for $T(f)$, and A_M be an incomplete assignment constructed from M. If the number of variables that have Type-II and/or Type-III contradiction under A_M is k, then there are at least k single men in M.*

Proof. Suppose that variables x_{i_1} and x_{i_2} ($i_1 \neq i_2$) have Type-II or Type-III contradiction. By Lemmas 6 and 7, at least one man associated with each variable is single. Let them be m_{i_1} and m_{i_2}, respectively. All we have to show is that $m_{i_1} \neq m_{i_2}$. We consider the following four cases:

Case 1: Both x_{i_1} and x_{i_2} have Type-II contradiction. By Lemma 6, m_{i_1} is $c_{p,r}$ and m_{i_2} is $c_{q,s}$ for some p, q, r and s. Recall the proof of Lemma 6. This means that x_{i_1} in the clause C_r receives no value and so, x_p in C_r receives both 0 and 1. Also x_{i_2} in C_s receives no value and so, x_q in C_s receives both 0 and 1. Namely, $r \in CL^0_{i_1}(A_M)$ and $s \in CL^0_{i_2}(A_M)$. By Lemma 5 (1)-(ii), $r \neq s$ and hence $m_{i_1} \neq m_{i_2}$.

Case 2: Only x_{i_1} has Type-II contradiction. In this case, m_{i_1} is $c_{p,r}$ where $r \in CL^2_p(A_M)$, and m_{i_2} is one of b_{i_2}, b'_{i_2} and $c_{i_2,j}$, where $j \in CL^1_{i_2}(A_M)$. If m_{i_2} is b_{i_2} or b'_{i_2}, then clearly $m_{i_1} \neq m_{i_2}$. Suppose m_{i_2} is $c_{i_2,j}$ for some j such that $j \in CL^1_{i_2}(A_M)$. If $m_{i_1} = m_{i_2}$, p and r must be equal to i_2 and j, respectively. This is impossible because it results that $r \in CL^1_p(A_M)$ and $r \in CL^2_p(A_M)$.

Case 3: Only x_{i_2} has Type-II contradiction. Same to Case 2.

Case 4: Neither x_{i_1} nor x_{i_2} have Type-II contradiction. By Lemma 7, m_{i_1} must be one of b_{i_1}, b'_{i_1} and c_{i_1,j_1} for some j_1, and m_{i_2} must be one of b_{i_2}, b'_{i_2} and c_{i_2,j_2} for some j_2. Clearly $m_{i_1} \neq m_{i_2}$ because $i_1 \neq i_2$. □

Now we are ready to prove Lemma 3. Suppose there are k unmatched men in M. Then, by Lemma 8, the number of variables that have Type-II or III contradiction is at most k. We construct a truth assignment A'_M of f from A_M in the following way: If x_i contains Type-II or Type-III contradiction, determine $A'_M(x_i)$ arbitrarily. If x_i does not contain any type of contradiction, all literals associated with x_i are consistent under A_M. We determine $A'_M(x_i)$ so that all the literals become consistent with x_i. Otherwise, if x_i contains only Type-I contradiction, then $CL_i = CL^1_i(A_M) \cup CL^2_i(A_M)$, namely, each literal associated with x_i receives both 1 and 0, or exactly one value because x_i does not contain Type-II contradiction. Furthermore, all literals which receive one value are consistent since x_i does not contain Type-III contradiction. We determine $A'_M(x_i)$ so that those literals become consistent with x_i.

Recall that, under the incomplete assignment A_M, every clause has at least one literal to which the value 1 is assigned. We count an upper bound on the number of clauses that become 0 by changing A_M into A'_M. Let L be the set of all clauses that contain a variable having Type-II or Type-III contradiction and let $\overline{L}$ be the set of all remaining clauses. Since there are at most k variables that have Type-II or III contradiction, and since each variable appears at most t times, it turns out that $|L| \leq tk$. We claim that clauses in $\overline{L}$ are all satisfied by A'_M. Let C_j be a clause in $\overline{L}$ and z_i^j (which is x_i^j or $\overline{x_i^j}$) be a literal in C_j. If x_i does not contain any contradiction under A_M, then the value of z_i^j is equivalent under A_M and A'_M. Suppose x_i contains only Type-I contradiction, i.e., $j \in CL_i^1(A_M)$ or $j \in CL_i^2(A_M)$. If $j \in CL_i^1(A_M)$, then again the value of z_i^j is equivalent under A_M and A'_M by definition of $A_{M'}$. We then claim that $j \notin CL_i^2(A_M)$: If $j \in CL_i^2(A_M)$ then there must be p such that $j \in CL_p^0(A_M)$ by Lemma 5 (2)-(ii). So C_j contains a variable containing Type-II contradiction and hence C_j must be in L.

Now, for any $C_j \in \overline{L}$, the value of every literal in C_j under A'_M is equivalent to the value of it under A_M. Again, recall that all the clauses have at least one literal having the value 1 under A_M. So every clause in $\overline{L}$ is satisfied by A'_M. □

2.2 Hardness for Restricted Instances

In this section, we show how to modify the construction in the proof of Theorem 1 to obtain the same hardness result for restricted instances where the length of ties is at most two.

Theorem 9. *MAX SMTI is hard to approximate within $1 + \epsilon$, for some $\epsilon > 0$, even if each person writes at most one tie of length two.*

Proof. In the problem MAX ONE-IN-THREE E3SAT(t), we are given a CNF formula such that each clause contains exactly three literals and each variable appears at most t times. A clause is satisfied if and only if exactly one literal in the clause is true. The purpose of this problem is to find an assignment which satisfies a maximum number of clauses. By a simple polynomial-time reduction [13] from MAX E3SAT(t), we can show that there exists a constant $\delta > 0$ such that it is NP-hard to distinguish the following two cases: (i) There is an assignment that satisfies all the clauses of f. (ii) In any assignment, at least δ fraction of the whole clauses are unsatisfied.

We will slightly modify the reduction in the proof of Theorem 1 in the following way. In the reduction in the proof of Theorem 1, men a'_j and a''_j write six women corresponding to literals in the jth clause. In the new reduction, these two men write only three women, that is, three women among six, who do not appear in a_j's list. We can similarly show that the resulting MAX SMTI instance has a gap property. Observe that, in men's side, ties appear only in Group-(a) men's lists, each of length three.

Let I be an instance of SMTI constructed as above. We modify I and construct a new instance I' with preserving the gap property. Consider a Group-(a) man m who writes three women w_1, w_2 and w_3 in the list. We replace this man m with two men m_1 and m_2 and a woman y whose preference lists are as follows:

$$\begin{array}{llll} m_1\colon & y \ \ (w_1 \ \ w_2) & \qquad & y\colon \ \ (m_1 \ \ m_2) \\ m_2\colon & y \ \ (w_2 \ \ w_3) & & \end{array}$$

Here, persons within a parenthesis are tied. In w_1's list, m is replaced by m_1, and in w_3's list, m is replaced by m_2. In w_2's list, m is replaced by m_1 and m_2 in this order of perference. We call these three persons m_1, m_2 and y a *block* of m. The size of I' (i.e. the number of men in I') is bounded by a constant times of the size of I. The correctness follows from following two lemmas:

Lemma 10. *If there is a perfect stable matching for I, then there is a perfect stable matching for I'.*

Proof. Let M be a perfect stable matching for I. We construct a perfect stable matching M' for I'. Consider a Group-(a) man m of I who is replaced by m_1, m_2 and y as above. Recall that all Group-(a) men are matched in any stable matching. Hence m is matched in M. If m is matched with w_1 (w_2) in M, then m_1 is matched with w_1 (w_2) and m_2 is matched with y in M'. If m is matched with w_3 in M, then m_2 is matched with w_3 and m_1 is matched with y in M'. Men in Group (b) or (c) are matched in the same way as M. It is not hard to see that M' is stable in I'. □

Lemma 11. *If more than k men are unmatched in any stable matching for I, then more than k men are unmatched in any stable matching for I'.*

Proof. Suppose that there is a stable matching M' for I' in which k men are unmatched. We construct a stable matching M for I in which k men are unmatched.

Suppose that a man m of I is replaced by m_1, m_2 and y as above. Then, it is not hard to see that y is matched with m_1 or m_2 in any stable matching for I'. Hence exactly one man (m_1 or m_2) is unmatched with a woman in m's block. Although details are omitted, we can show that this man has a partner in any stable matching for I', namely, he is matched with a woman outside m's block.

Now we construct a matching M from M'. Each man except for Group-(a) men is matched with the same woman as M', or unmatched if he is unmatched in M'. Consider a Group-(a) man m. As discussed above, in M', there is one man, say m_i, who is matched with outside m's block. In M, m is matched with the woman with whom m_i is matched in M'. We can easily verify that M is stable. The number of unmatched men is same in M and M'. □

We can further show that the hardness result holds for instances in which ties appear only in one sex. Because of the space restriction, we give only a rough sketch of its proof.

Theorem 12. *MAX SMTI is hard to approximate within* $1+\epsilon$*, for some* $\epsilon > 0$*, even if ties appear in only men's lists, each man writes at most one tie of length at most three.*

Proof. As in the proof of Theorem 9, we modify an SMTI instance, say I, translated from MAX ONE-IN-THREE E3SAT(t). Recall that in I, ties of length three appear in Group-(a) men's lists and ties of length two appear in Group-(u) and Group-(v) women's lists. For each i ($1 \le i \le n$), consider two men b_i, b'_i and two women u_i, v_i:

$$\begin{array}{llll} b_i\colon & u_i \ \cdots \ v_i & \qquad u_i\colon & (b_i \ \ b'_i) \\ b'_i\colon & u_i \ \cdots \ v_i & \qquad v_i\colon & (b_i \ \ b'_i) \end{array}$$

In our reduction, these four people will be replaced by eight people whose preference lists are shown in the following figure:

$$\begin{array}{llll} s_u : & (u_{i1} \ \ u_{i2}) & \qquad u_{i1} : & s_u \ \ b_i \\ s_v : & (v_{i1} \ \ v_{i2}) & \qquad u_{i2} : & s_u \ \ b'_i \\ b_i : & u_{i1} \ \cdots \ v_{i1} \ \ v_{i2} & \qquad v_{i1} : & s_v \ \ b_i \ \ b'_i \\ b'_i : & u_{i2} \ \cdots \ v_{i2} \ \ v_{i1} & \qquad v_{i2} : & s_v \ \ b'_i \ \ b_i \end{array}$$

The correctness follows from a similar argument as in the proof of Theorem 9. □

3 Inapproximability of MIN Egalitarian SMT and MIN Regret SMT

The *regret* of a person p in a matching M is defined to be the rank (within p's preference list) of p's partner in M. Namely, $regret_M(p)$ is the number of persons that p *strictly prefers* to his/her current partner $M(p)$ plus one. Given that there can be many possible solutions to a stable marriage instance, it is natural to seek a solution that maximizes the overall satisfaction with the assignment, or alternatively minimizes the dissatisfaction.

Minimum egalitarian stable matching The cost of a stable matching M is defined to be $\sum_p regret_M(p)$, where the sum is taken over all persons. For the classical stable marriage instances without ties, a polynomial-time algorithm is known for finding an optimal stable matching [5, 7]. However, when ties are allowed, the problem becomes intractable even with complete preference lists. Denote this problem, with ties but complete lists, as *MIN egalitarian SMT*. We give here a lower bound $\Omega(N)$ on the approximation ratio. Note that the cost of a matching is between $2N$ and $2N^2$. Hence an approximation ratio N is trivial. Our inapproximability result is optimal within a constant factor.

Theorem 13. *MIN egalitarian SMT is hard to approximate within ϵN, for some $\epsilon > 0$.*

Proof. Let I be an instance of SMTI constructed in the proof of Theorem 1. Let $X = \{m_1, m_2, \cdots m_N\}$ be the set of men and $Y = \{w_1, w_2, \cdots w_N\}$ be the set of women of I. Let P_i be the preference list of m_i and Q_i that of w_i. For each $1 \leq i \leq N$, we call women in P_i *proper women* for m_i and men in Q_i *proper men* for w_i.

We translate I into an instance I' of MIN egalitarian SMT. I' consists of X and Y, along with new men $X' = \{m'_1, m'_2, \cdots, m'_N\}$ and women $Y' = \{w'_1, w'_2, \cdots, w'_N\}$. Preference lists of I' are constructed as follows:

m'_i :	w'_i [other $2N-1$ women arbitrarily]	$(1 \leq i \leq N)$
m_i :	P_i [women in Y' arbitrarily] [other women in Y arbitrarily]	$(1 \leq i \leq N)$
w'_i :	m'_i [other $2N-1$ men arbitrarily]	$(1 \leq i \leq N)$
w_i :	Q_i [men in X' arbitrarily] [other men in X arbitrarily]	$(1 \leq i \leq N)$

Note that each m'_i is matched with w'_i in any stable matching for I' since they write each other strictly first in their lists. Thus, there is a one-to-one correspondence between stable matchings for I and I'. We know that either I has a stable matching of size N, or the size of any stable matching for I is at most $(1-\delta)N$ for a constant δ. If I has a stable matching of size N, then there is a stable matching, say M', for I', where each man in X is matched with a proper woman and each woman in Y is matched with a proper man. Since all preference lists in I are of constant length, say, at most d, the regret of each person with respect to M' is constant and the total cost of M' is at most $2N + 2dN$.

On the other hand, suppose that the size of any stable matching for I is at most $(1-\delta)N$. Then, any stable matching for I' has at least δN men and women that cannot be matched with proper persons. Since they cannot be matched with persons in $X' \cup Y'$, their regret must be larger than N, and hence the sum of their regrets is at least $2 \times \delta N^2$. Hence, a $\frac{\delta}{d+1}N$-approximation algorithm would solve an NP-complete problem. ◻

Minimum regret stable matching Another measure of general satisfaction with the assignment would be to measure the worst case *regret*, i.e. $\max_p regret_M(p)$. This problem is solvable in polynomial time for complete lists without ties [4]. Here we show an optimal inapproximability of problem when ties are allowed. Refer to this problem as *MIN regret SMT*.

Theorem 14. *MIN regret SMT is hard to approximate within ϵN, for some $\epsilon > 0$.*

Proof. We use the same reduction as described in the proof of Theorem 13. Let I and I' be as above. Hence I has N men and N women, and I' has $2N$ men and $2N$ women. If I has a perfect stable matching, then I' has a stable matching in which all persons are matched with proper persons. In this matching, every person's cost is constant and hence the optimal cost is constant, say, d. If I does not have a perfect stable matching, then there is at least one person who is not matched with a proper person and his/her cost is at least N. Therefore, a polynomial-time $\frac{N}{d}$-approximation algorithm implies P=NP. $\square$

References

1. S. Arora and C. Lund, "Hardness of Approximations," Chapter in the book Approximation Algorithms for NP-hard problems, D. Hochbaum editor, PWS Publishing, 1996.
2. D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *Amer. Math. Monthly*, Vol.69, pp.9-15, 1962.
3. D. Gale and M. Sotomayor, "Some remarks on the stable matching problem," *Discrete Applied Mathematics*, Vol.11, pp.223-232, 1985.
4. D. Gusfield, "Three fast algorithms for four problems in stable marriage," SIAM Journal on Computing, Vol. 16, pp. 111–128, 1987.
5. D. Gusfield and R. W. Irving, "The Stable Marriage Problem: Structure and Algorithms," MIT Press, Boston, MA, 1989.
6. J. Håstad, "Some optimal inapproximability results," *Proc. STOC 97*, pp. 1–11, 1997.
7. R. W. Irving, P. Leather and D. Gusfield, "An efficient algorithm for the "optimal" stable marriage," Journal of the A.C.M., Vol. 34, pp. 532–543, 1987.
8. R. W. Irving, "Stable marriage and indifference," *Discrete Applied Mathematics*, Vol.48, pp.261-272, 1994.
9. R. W. Irving, D. F. Manlove and S. Scott, "The hospital/residents problem with ties," *Proc. SWAT 2000*, LNCS 1851, pp. 259–271, 2000.
10. K. Iwama, D. Manlove, S. Miyazaki, and Y. Morita, "Stable marriage with incomplete lists and ties," *Proc. ICALP'99*, LNCS 1644, pp. 443-452, 1999.
11. D. Manlove, R. W. Irving, K. Iwama, S. Miyazaki, Y. Morita, "Hard variants of stable marriage," Technical Report TR-1999-43, Computing Science Department of Glasgow University, September 1999 (to appear in Theoretical Computer Science).
12. E. Ronn,"NP-complete stable matching problems," *J. Algorithms*, Vol.11, pp.285-304, 1990.
13. T. J. Schaefer, "The complexity of satisfiability problems," *Proc. STOC 78*, pp. 216–226, 1978.

a_1	$w^1_{1,1}$	$w^0_{2,1}$	$w^1_{3,1}$				
a'_1	$w^0_{1,1}$	$w^0_{2,1}$	$w^0_{3,1}$	$w^1_{1,1}$	$w^1_{2,1}$	$w^1_{3,1}$	
a''_1	$w^0_{1,1}$	$w^0_{2,1}$	$w^0_{3,1}$	$w^1_{1,1}$	$w^1_{2,1}$	$w^1_{3,1}$	
a_2	$w^0_{1,2}$	$w^1_{2,2}$	$w^1_{4,2}$				
a'_2	$w^0_{1,2}$	$w^0_{2,2}$	$w^0_{4,2}$	$w^1_{1,2}$	$w^1_{2,2}$	$w^1_{4,2}$	
a''_2	$w^0_{1,2}$	$w^0_{2,2}$	$w^0_{4,2}$	$w^1_{1,2}$	$w^1_{2,2}$	$w^1_{4,2}$	
a_3	$w^1_{2,3}$	$w^0_{4,3}$	$w^0_{5,3}$				
a'_3	$w^0_{2,3}$	$w^0_{4,3}$	$w^0_{5,3}$	$w^1_{2,3}$	$w^1_{4,3}$	$w^1_{5,3}$	
a''_3	$w^0_{2,3}$	$w^0_{4,3}$	$w^0_{5,3}$	$w^1_{2,3}$	$w^1_{4,3}$	$w^1_{5,3}$	
b_1		u_1	$w^0_{1,1}$	$w^0_{1,2}$			v_1
b'_1	u_1		$w^1_{1,1}$	$w^1_{1,2}$		v_1	
b_2		u_2	$w^0_{2,1}$	$w^0_{2,2}$	$w^0_{2,3}$		v_2
b'_2	u_2		$w^1_{2,1}$	$w^1_{2,2}$	$w^1_{2,3}$	v_2	
b_3		u_3	$w^0_{3,1}$				v_3
b'_3	u_3		$w^1_{3,1}$			v_3	
b_4		u_4	$w^0_{4,2}$	$w^0_{4,3}$			v_4
b'_4	u_4		$w^1_{4,2}$	$w^1_{4,3}$		v_4	
b_5		u_5	$w^0_{5,3}$				v_5
b'_5	u_5		$w^1_{5,3}$			v_5	
$c_{1,1}$						$w^0_{1,1}$	$w^1_{1,1}$
$c_{1,2}$						$w^0_{1,2}$	$w^1_{1,2}$
$c_{2,1}$						$w^0_{2,1}$	$w^1_{2,1}$
$c_{2,2}$						$w^0_{2,2}$	$w^1_{2,2}$
$c_{2,3}$						$w^0_{2,3}$	$w^1_{2,3}$
$c_{3,1}$						$w^0_{3,1}$	$w^1_{3,1}$
$c_{4,2}$						$w^0_{4,2}$	$w^1_{4,2}$
$c_{4,3}$						$w^0_{4,3}$	$w^1_{4,3}$
$c_{5,3}$						$w^0_{5,3}$	$w^1_{5,3}$

Table 1. Preference lists of men of $T(f_0)$

$$f_0 = (x_1 + \overline{x_2} + x_3)(\overline{x_1} + x_2 + x_4)(x_2 + \overline{x_4} + \overline{x_5})$$