Stable Marriage with Ties and Bounded Length Preference Lists[*]

Robert W. Irving[†], David F. Manlove[†,‡,§] and Gregg O'Malley[‡]

Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK.
Email: {rwi,davidm,gregg}@dcs.gla.ac.uk.

Abstract. We consider variants of the classical stable marriage problem in which preference lists may contain ties, and may be of bounded length. Such restrictions arise naturally in practical applications, such as centralised matching schemes that assign graduating medical students to their first hospital posts. In such a setting, weak stability is the most common solution concept, and it is known that weakly stable matchings can have different sizes. This motivates the problem of finding a maximum cardinality weakly stable matching, which is known to be NP-hard in general. We show that this problem is solvable in polynomial time if each man's list is of length at most 2 (even for women's lists that are of unbounded length). However if each man's list is of length at most 3, we show that the problem becomes NP-hard (even if each women's list is of length at most 3) and not approximable within some $\delta > 1$ (even if each woman's list is of length at most 4).

Keywords: Stable marriage problem; ties; incomplete lists; NP-hardness; polynomial-time algorithm

1 Introduction

The Stable Marriage problem (SM) was introduced in the seminal paper of Gale and Shapley [3]. In its classical form, an instance of SM involves n men and n women, each of whom specifies a *preference list*, which is a total order on the members of the opposite sex. A *matching* M is a set of (man,woman) pairs such that each person belongs to exactly one pair. If $(m, w) \in M$, we say that w is m's *partner* in M, and vice versa, and we write $M(m) = w$, $M(w) = m$.

We say that a person x *prefers* y to y' if y precedes y' on x's preference list. A matching M is *stable* if it admits no *blocking pair*, namely a (man,woman) pair (m, w)

such that m prefers w to $M(m)$ and w prefers m to $M(w)$. Gale and Shapley [3] proved that every instance of SM admits at least one stable matching, and described an algorithm – the Gale / Shapley algorithm – that finds such a matching in time that is linear in the input size. In general, there may be many stable matchings (in fact exponentially many in n) for a given instance of SM [13].

Incomplete lists. A variety of extensions of the basic problem have been studied. In the Stable Marriage problem with Incomplete lists (SMI), the numbers of men and women need not be the same, and each person's preference list consists of a subset of the members of the opposite sex in strict order. A (man,woman) pair (m, w) is *acceptable* if each member of the pair appears on the preference list of the other. A *matching M* is now a set of acceptable pairs such that each person belongs to at most one pair. In this context, (m, w) is a blocking pair for a matching M if (a) (m, w) is an acceptable pair, (b) m is either unmatched or prefers w to $M(m)$, and likewise (c) w is either unmatched or prefers m to $M(w)$. Given the definitions of a matching and a blocking pair, we lose no generality by assuming that the preference lists are *consistent* (i.e., given a (man,woman) pair (m, w), m appears on the preference list of w if and only if w appears on the preference list of m). As in the classical case, there is always at least one stable matching for an instance of SMI, and it is straightforward to extend the Gale / Shapley algorithm to give a linear-time algorithm for this case. Again, there may be many different stable matchings, but Gale and Sotomayor [4] showed that every stable matching for a given SMI instance has the same size and matches exactly the same set of people.

Ties. An alternative extension of SM arises if preference lists are allowed to contain *ties*. In the Stable Marriage problem with Ties (SMT) each person's preference list is a partial order over the members of the opposite sex in which indifference is transitive. In other words, each person p's list can be viewed as a sequence of ties, each of length ≥ 1; p prefers each member of a tie to everyone in any subsequent tie, but is indifferent between the members of any single tie. In this context, three definitions of stability have been proposed [6, 11]. Among these three stability criteria, it is *weak stability* that has received the most attention in the literature [15, 19, 7, 8, 9, 10, 16, 17]. A matching M is *weakly stable* if there is no pair (m, w), each of whom prefers the other to his/her partner in M. For a given instance of SMT, a weakly stable matching is bound to exist, and can be found in linear time by breaking all ties in an arbitrary way (i.e. by strictly ranking the members of each tie arbitrarily) and applying the Gale / Shapley algorithm.

Ties and incomplete lists. If we allow both of the above extensions of the classical problem simultaneously, we obtain the Stable Marriage problem with Ties and Incomplete lists (SMTI). In this context a matching M is weakly stable if there is no acceptable pair (m, w), each of whom is either unmatched in M or prefers the other to his/her partner in M. Once again, it is easy to find a weakly stable matching, merely by breaking all the ties in an arbitrary way and applying the Gale / Shapley algorithm. However, the ways in which ties are broken will, in general, affect the size of the resulting matching. It is therefore natural to consider MAX SMTI, the problem of finding a maximum cardinality weakly stable matching (henceforth a maximum weakly stable matching), given an instance of SMTI. MAX SMTI turns out to be NP-hard, even under quite severe restrictions on the number and lengths of ties [19].

Specifically, NP-hardness holds even if ties occur in the men's preference lists only, each tie is of length 2, and each tie comprises the whole of the list in which it appears [19]. (Note that, in the SMTI instance constructed by the reduction in [19], there are men with strictly ordered preference lists of length at least 3.)

The Hospitals/Residents problem. The Hospitals/Residents problem (HR) is a many-to-one generalisation of SMI, so called because of its application in centralised matching schemes for the allocation of graduating medical students, or residents, to hospitals [20]. The best known such scheme is the National Resident Matching Program (NRMP) [22] in the US, but similar schemes exist in Canada [21], in Scotland [12, 23], and in a variety of other countries and contexts. In fact, this extension of SM was also discussed by Gale and Shapley under the name of the College Admissions problem [3]. In an instance of HR, each resident has a preference list containing a subset of the hospitals, and each hospital ranks the residents for which it is acceptable. In addition, each hospital has a *quota* of available posts. In this context, a matching is a set of acceptable (resident,hospital) pairs so that each resident appears in at most one pair and each hospital in a number of pairs that is bounded by its quota. The definition of stability is easily extended to this more general setting (see [6] for details). It is again the case that every problem instance admits at least one stable matching [3], and that all stable matchings have the same size [4]. Clearly SMI is equivalent to the special case of HR in which each hospital has a quota of 1.

The Hospitals / Residents problem with Ties (HRT) allows arbitrary ties in the preference lists. The definition of weak stability can be extended in a natural way to the HRT context [14]. Since HRT is clearly an extension of SMTI, the hardness results for weak stability problems in the latter extend to the former. These results have potentially important implications for large-scale real-world matching schemes. It is unreasonable to expect, say, a large hospital to rank in strict order all of its many applicants, and any artificial rankings, whether submitted by the hospitals themselves, or imposed by the matching scheme administrators, may have significant implications for the number of residents assigned in a stable matching.

Bounded length preference lists. In the context of many large-scale matching schemes, the preference lists of at least one set of agents tend to be short. For example, until recently, students participating in the Scottish medical matching scheme [12, 23] were required to rank just three hospitals in order of preference. This naturally leads to the question of whether the problem of finding a maximum weakly stable matching becomes simpler when preference lists on one or both sides have bounded length.

Let (p, q)-MAX SMTI denote the restriction of MAX SMTI in which each man's list is of length at most p and each woman's list is of length at most q. We use $p = \infty$ or $q = \infty$ to denote the possibility that the men's lists or women's lists respectively are of unbounded length. Halldórsson et al. [8] showed that $(4, 7)$-MAX SMTI is NP-hard and not approximable within some $\delta > 1$ unless P=NP. Halldórsson et al. [9] gave an alternative reduction from Minimum Vertex Cover to MAX SMTI, showing that the latter problem is not approximable within $\frac{21}{19}$ unless P=NP. By starting from the NP-hard restriction of Minimum Vertex Cover to graphs of maximum degree 3 [5], the same reduction shows NP-hardness for $(5, 5)$-MAX SMTI.

In this paper we consider other values of p and q, to identify the 'borderline' between polynomial-time solvability and NP-hardness for (p, q)-MAX SMTI. We show

3

in Section 2 that $(2, \infty)$-MAX SMTI is polynomial-time solvable using a combination of an adapted version of the Gale / Shapley algorithm together with a reduction to the Assignment problem. By contrast, in Section 3 we show that $(3, 3)$-MAX SMTI is NP-hard, even if the ties belong to the preference lists of one sex only. In Section 4 we give an inapproximability result, namely that $(3, 4)$-MAX SMTI is not approximable within some $\delta > 1$ unless P=NP. Finally, in Section 5 we present some concluding remarks.

2 Algorithm for $(2, \infty)$-MAX SMTI

In this section we present a polynomial-time algorithm for MAX SMTI where the preference lists of both men and women may contain ties, the men's lists are of length at most 2 and the women's lists are of unbounded length. Let I be an instance of this problem, and let n_1 and n_2 be the numbers of men and women respectively in I.

Consider the algorithm $(2, \infty)$-MAX-SMTI-`alg` shown in Figure 1. The algorithm consists of three phases, where each phase is highlighted in the figure. We use the term *reduced lists* to refer to participants' lists after any deletions made by the algorithm. Phase 1 of $(2, \infty)$-MAX-SMTI-`alg` is a simple extension of the Gale / Shapley algorithm, and is used to delete certain (man,woman) pairs that can never be part of any weakly stable matching. To "delete the pair (m_i, w_j)", we delete m_i from w_j's list and delete w_j from m_i's list. Phase 1 proceeds as follows. All men are initially unmarked. While some man m_i remains unmarked and m_i has a non-empty reduced list, we set m_i to be marked – it is possible that m_i may again become unmarked at a later stage of the execution. If m_i's reduced list is not a tie of length 2, we let w_j be the woman in first position in m_i's reduced list. Then, for each strict successor m_k of m_i on w_j's list, we delete the pair (m_k, w_j) and set m_k to be unmarked (regardless of whether or not he was already marked).

We remark that the following situation may occur during phase 1. Suppose that some man m_i is indifferent between two women w_j and w_k on his original preference list, and suppose that during some iteration of the while loop he becomes marked. We note that the algorithm does not delete the strict successors of m_i on w_j's list at this stage. Now suppose that, during a subsequent loop iteration, the pair (m_i, w_k) is deleted. Then m_i becomes unmarked, only to be re-marked during a subsequent loop iteration. This re-marking results in the deletions of all pairs (m_r, w_j), where m_r is a strict successor of m_i on w_j's list, as required.

In phase 2 we construct a weighted bipartite graph G and find a minimum cost maximum matching in G using the algorithm in [2]. The graph G is constructed using Algorithm `BuildGraph` shown in Figure 2. That is, each man and woman is represented by a vertex in G, and for each man m_i on woman w_j's reduced list, we add an edge from m_i to w_j with cost $rank(w_j, m_i)$, where $rank(w_j, m_i)$ is the rank of m_i on w_j's reduced list (i.e. 1 plus the number of strict predecessors of m_i on w_j's reduced list). We then find a minimum cost maximum matching M_G in G.

In general, after phase 2, M_G need not be weakly stable in I. In particular, some man m_i who has a reduced list of length 2 that is strictly ordered may be assigned to his second-choice woman w_k in M_G, while his first-choice woman w_j may be unassigned in M_G. Clearly (m_i, w_j) blocks such a matching. To obtain a weakly stable matching M from M_G we execute phase 3. Initially M is set to be equal to M_G.

4

```
/* Phase 1 */
set all men to be unmarked;
while (some man $m_i$ is unmarked and
        $m_i$ has a non-empty reduced list) do
    set $m_i$ to be marked;
    if $m_i$'s reduced list is not a tie of length 2 then
        $w_j$ := woman in first position on $m_i$'s reduced list;
        for each strict successor $m_k$ of $m_i$ on $w_j$'s list do
            set $m_k$ to be unmarked;
            delete the pair $(m_k, w_j)$;

/* Phase 2 */
$G$ := BuildGraph();
$M_G$ := minimum cost maximum matching in $G$;

/* Phase 3 */
$M := M_G$;
while (there exists a man $m_i$ who is assigned
        to his second-choice woman $w_k$ in $M$
        and his first-choice woman $w_j$ is unassigned in $M$) do
    $M := M \setminus \{(m_i, w_k)\}$;
    $M := M \cup \{(m_i, w_j)\}$;
return $M$;
```

Figure 1: Algorithm $(2, \infty)$-MAX-SMTI-alg.

Next, we move each such m_i to his first-choice woman. We note that m_i must be in the *tail* of w_j's reduced list (this is the set of one or more entries tied in last place on w_j's reduced list) since m_i must have been marked during Phase 1, causing all strict successors of m_i on w_j's list to be deleted. Further, we note that there may exist more than one such man in w_j's tail who satisfy the above criterion. Moreover when m_i moves to w_j, w_k becomes unassigned in M. As a result, there may be some other man m_r (who strictly ranks w_k in first place) who now satisfies the loop condition. This process is repeated until no such man exists. Upon termination of phase 3 we will show that the matching M returned is a maximum weakly stable matching.

We begin by showing that the algorithm $(2, \infty)$-MAX-SMTI-alg terminates. It is easy to see that each of phases 1 and 2 is bound to terminate. The following lemma shows that the same is true for phase 3.

Lemma 1. *Phase 3 of* $(2, \infty)$-MAX-SMTI-alg *terminates.*

Proof. We show that the while loop terminates during an execution E of phase 3. For, at a given iteration of the while loop of phase 3, let m_i be some man assigned to his second-choice woman w_k in M and suppose that his first-choice woman w_j is unassigned in M, where m_i's reduced list is of length 2 and is strictly ordered. Then during E, m_i switches from w_k to w_j. Hence each such m_i must strictly improve (in fact m_i can only improve at most once). Therefore since the number of men is finite, phase 3 is bound to terminate. $\square$

```
V := M ∪ W;        // M and W are the sets of men and women in I respectively
E := ∅;
for each man m_i ∈ M do
    for each woman w_j on m_i's reduced list do
        E := E ∪ {(m_i, w_j)};
        cost(m_i, w_j) := rank(w_j, m_i);
G := (V, E);
return G;
```

Figure 2: Algorithm `BuildGraph`.

We next show that phase 1 of $(2, \infty)$-MAX-SMTI-`alg` never deletes a *weakly stable pair*, which is a (man,woman) pair that belongs to some weakly stable matching in I.

Lemma 2. *The algorithm* $(2, \infty)$-MAX-SMTI-`alg` *never deletes a weakly stable pair.*

Proof. Let (m_i, w_j) be a pair deleted during an execution E of $(2, \infty)$-MAX-SMTI-`alg` such that $(m_i, w_j) \in M$, where M is a weakly stable matching in I. Without loss of generality suppose this is the first weakly stable pair deleted during E. Then m_i was deleted from w_j's list during some iteration q of the while loop of phase 1 during E. This deletion was made as a result of w_j being in first position in the reduced list of some man m_r, where m_r's reduced list was not a tie of length 2, and w_j prefers m_r to m_i. Then in M, m_r must obtain a woman w_s such that m_r either prefers w_s to w_j or is indifferent between them, otherwise (m_r, w_j) blocks M. Therefore during E, in both cases (m_r, w_s) must have already been deleted before iteration q, a contradiction. $\square$

Finally we prove that the matching returned by $(2, \infty)$-MAX-SMTI-`alg` is weakly stable in I.

Lemma 3. *The matching returned by algorithm* $(2, \infty)$-MAX-SMTI-`alg` *is weakly stable in I.*

Proof. Suppose for a contradiction that the matching M output by the algorithm $(2, \infty)$-MAX-SMTI-`alg` is not weakly stable. Then there exists a pair (m_i, w_j) that blocks M. We consider the following four cases corresponding to a blocking pair.

Case (i): both m_i and w_j are unassigned in M. Then m_i is unassigned in M_G, and either w_j is unassigned in M_G or becomes unassigned during phase 3. First suppose that w_j is unassigned in M_G. Then the size of the matching M_G could be increased by adding the edge (m_i, w_j) to M_G, contradicting the maximality of M_G. Now suppose that w_j became unassigned as a result of phase 3. Let m_{p_1} denote w_j's partner in M_G. Then during phase 3, m_{p_1} must have become assigned to his first-choice woman w_{q_1}. Suppose w_{q_1} was unassigned in M_G. Then we can find a larger matching by augmenting along the path $(m_i, w_j), (w_j, m_{p_1}), (m_{p_1}, w_{q_1})$, contradicting the maximality of M_G. Therefore w_{q_1} must have been assigned in M_G and became unassigned as a result of phase 3. Hence the man m_{p_2}, to whom w_{q_1} was assigned in M_G, switched to his first-choice woman w_{q_2}. Using an argument similar to that above for w_{q_1}, we can show that w_{q_2} must be assigned in M_G. Therefore some man switched

from w_{q_2} during phase 3 to his first-choice woman. If we continue this process, since each man must strictly improve and the number of men is finite, there exists a finite number of women that can become unassigned as a result of phase 3. Hence at some point there exists a man m_{p_r} who switches to his first-choice woman w_{q_r} and w_{q_r} was already unassigned in M_G. We can then construct an augmenting path in G of the form $(m_i, w_j), (w_j, m_{p_1}), (m_{p_1}, w_{q_1}), (w_{q_1}, m_{p_2}), (m_{p_2}, w_{q_2}), \ldots, (m_{p_r}, w_{q_r})$, which contradicts the maximality of M_G.

Case (ii): m_i is unassigned in M and w_j prefers m_i to her assignee m_k in M. Then m_i is unassigned in M_G. Suppose that w_j is assigned to m_k in M_G. As w_j prefers m_i to m_k, we could obtain a matching with a smaller cost, but with the same size, by removing (m_k, w_j) and adding (m_i, w_j) to M_G, a contradiction. Now suppose that w_j is not assigned to m_k in M_G. Then w_j is either unassigned in M_G or w_j is assigned in M_G to m_r, where $m_r \neq m_k$ and $m_r \neq m_i$. If w_j is unassigned in M_G, we contradict the maximality of M_G. Now suppose w_j is assigned to m_r in M_G. Then since w_j is no longer assigned to m_r in M, m_r must have switched to his first-choice woman w_s during phase 3. Therefore either w_s is unassigned in M_G or w_s became unassigned as a result of some man switching from w_s to his first-choice woman. Again using a similar argument to that in Case (i) we obtain an augmenting path that contradicts the maximality of M_G.

Case (iii): m_i is assigned to w_s in M and m_i prefers w_j to w_s and w_j is unassigned in M. Thus clearly m_i's list is of length 2 and does not contain a tie, and w_j is m_i's first-choice woman. In this situation the loop condition of phase 3 is satisfied. Therefore since the algorithm terminates (Lemma 1) this situation can never arise.

Case (iv): m_i is assigned to w_s in M and m_i prefers w_j to w_s, and w_j is assigned to m_r in M and w_j prefers m_i to m_r. Thus again m_i's list cannot contain a tie, and w_j is his first-choice woman. Therefore either m_i proposed to w_j during phase 1 or w_j was deleted from m_i's list. Hence m_r would have been deleted from w_j's list during phase 1, so it is then impossible that $(m_r, w_j) \in M$. $\square$

Since phase 1 of the algorithm never deletes a weakly stable pair (by Lemma 2), a maximum weakly stable matching must consist of (man,woman) pairs that belong to the reduced lists. We next note that G is constructed from the reduced lists, and since we find a maximum matching in G, the matching output by the algorithm must indeed be a maximum weakly stable matching (by Lemma 3, and since phase 3 does not change the size of the matching output by the algorithm: every man matched in M_G is also matched in M).

The time complexity of the algorithm is dominated by finding the minimum cost maximum matching in $G = (V, E)$. The required matching in G can be constructed in $O(\sqrt{|E|}|V|\log|V|)$ time [2]. Let $n = |V| = n_1 + n_2$. Since $|E| \leq 2n_1 = O(n)$, it follows that $(2, \infty)$-MAX-SMTI-`alg` has time complexity $O(n^{\frac{3}{2}} \log n)$.

We summarise the results of this section in the following theorem.

Theorem 4. *Given an instance I of $(2, \infty)$-MAX SMTI, algorithm $(2, \infty)$-MAX-SMTI-`alg` returns a weakly stable matching of maximum size in $O(n^{\frac{3}{2}} \log n)$ time, where n is the total number of men and women in I.*

$$
\begin{array}{lll}
x_{4i}: & y_{4i} \;\; c(x_{4i}) \;\; y_{4i+1} & (0 \le i \le n-1) \\
x_{4i+1}: & y_{4i+1} \;\; c(x_{4i+1}) \;\; y_{4i+2} & (0 \le i \le n-1) \\
x_{4i+2}: & y_{4i+3} \;\; c(x_{4i+2}) \;\; y_{4i+2} & (0 \le i \le n-1) \\
x_{4i+3}: & y_{4i} \;\; c(x_{4i+3}) \;\; y_{4i+3} & (0 \le i \le n-1) \\
p_j^s: & z_j \;\; c_j^s & (1 \le j \le m \wedge 1 \le s \le 3) \\
q_j: & c_j^1 \;\; c_j^2 \;\; c_j^3 & (1 \le j \le m) \\[1em]
y_{4i}: & (x_{4i} \;\; x_{4i+3}) & (0 \le i \le n-1) \\
y_{4i+1}: & (x_{4i} \;\; x_{4i+1}) & (0 \le i \le n-1) \\
y_{4i+2}: & (x_{4i+1} \;\; x_{4i+2}) & (0 \le i \le n-1) \\
y_{4i+3}: & (x_{4i+2} \;\; x_{4i+3}) & (0 \le i \le n-1) \\
c_j^s: & p_j^s \;\; x(c_j^s) \;\; q_j & (1 \le j \le m \wedge 1 \le s \le 3) \\
z_j: & (p_j^1 \;\; p_j^2 \;\; p_j^3) & (1 \le j \le m)
\end{array}
$$

Figure 3: Preference lists in the constructed instance of $(3,3)$-COM SMTI.

3 NP-hardness of $(3,3)$-MAX SMTI

In this section we show that, in contrast to the case for $(2,\infty)$-MAX SMTI, $(3,3)$-MAX SMTI is NP-hard. The result holds even if the ties belong to the preference lists of one sex only. In fact we will show that $(3,3)$-COM SMTI is NP-complete – this is the problem of deciding, given an instance of SMTI in which all preference lists are of length at most 3, whether a *complete* weakly stable matching (i.e., a weakly stable matching in which everyone is matched) exists. The NP-completeness of $(3,3)$-COM SMTI clearly implies the NP-hardness of $(3,3)$-MAX SMTI.

Our proof of this result uses a reduction from a restricted version of SAT. More specifically, let (2,2)-E3-SAT denote the problem of deciding, given a Boolean formula B in CNF in which each clause contains exactly 3 literals and, for each $v_i \in V$, each of literals v_i and $\bar{v}_i$ appears exactly twice in B, whether B is satisfiable. Berman et al. [1] showed that (2,2)-E3-SAT is NP-complete.

Theorem 5. $(3,3)$-COM SMTI *is NP-complete. The result holds even if the ties belong to the preference lists of one sex only.*

Proof. Let B be an instance of (2,2)-E3-SAT. Let $V = \{v_0, v_1, \ldots, v_{n-1}\}$ and $C = \{c_1, c_2, \ldots, c_m\}$ be the set of variables and clauses respectively in B. Then for each $v_i \in V$, each of literals v_i and $\bar{v}_i$ appears exactly twice in B. (Hence $m = \frac{4n}{3}$.) Also $|c_j| = 3$ for each $c_j \in C$. We form an instance I of $(3,3)$-COM SMTI as follows. The set of men in I is $X \cup P \cup Q$, where $X = \cup_{i=0}^{n-1} X_i$, $X_i = \{x_{4i+r} : 0 \le r \le 3\}$ $(0 \le i \le n-1)$, $P = \cup_{j=1}^{m} P_j$, $P_j = \{p_j^1, p_j^2, p_j^3\}$ $(1 \le j \le m)$ and $Q = \{q_j : c_j \in C\}$. The set of women in I is $Y \cup C' \cup Z$, where $Y = \cup_{i=0}^{n-1} Y_i$, $Y_i = \{y_{4i+r} : 0 \le r \le 3\}$ $(0 \le i \le n-1)$, $C' = \{c_j^s : c_j \in C \wedge 1 \le s \le 3\}$ and $Z = \{z_j : c_j \in C\}$.

The preference lists of the men and women in I are shown in Figure 3. In a given preference list, entries within round brackets are tied. In the preference list of an agent $x_{4i+r} \in X$ $(0 \le i \le n-1$ and $r \in \{0,1\})$, the symbol $c(x_{4i+r})$ denotes the woman $c_j^s \in C'$ such that the $(r+1)$th occurrence of literal v_i appears at position

8

s of c_j. Similarly if $r \in \{2, 3\}$ then the symbol $c(x_{4i+r})$ denotes the woman $c_j^s \in C'$ such that the $(r-1)$th occurrence of literal $\bar{v}_i$ appears at position s of c_j. Also in the preference list of an agent $c_j^s \in C'$, if literal v_i appears at position s of clause $c_j \in C$, the symbol $x(c_j^s)$ denotes the man x_{4i+r-1} where $r = 1, 2$ according as this is the first or second occurrence of literal v_i in B. Otherwise if literal $\bar{v}_i$ appears at position s of clause $c_j \in C$, the symbol $x(c_j^s)$ denotes the man x_{4i+r+1} where $r = 1, 2$ according as this is the first or second occurrence of literal $\bar{v}_i$ in B. Clearly each preference list is of length at most 3, and the ties belong to the women's preference lists only.

For each i ($0 \le i \le n-1$), let $T_i = \{(x_{4i+r}, y_{4i+r}) : 0 \le r \le 3\}$ and $F_i = \{(x_{4i+r}, y_{4i+r+1})\} : 0 \le r \le 3\}$, where addition is taken modulo 4.

We claim that B is satisfiable if and only if I admits a complete weakly stable matching.

For, let f be a satisfying truth assignment of B. Define a complete matching M in I as follows. For each variable $v_i \in V$, if v_i is true under f, add the pairs in T_i to M, otherwise add the pairs in F_i to M. Now let $c_j \in C$. As c_j contains a literal that is true under f, let $s \in \{1, 2, 3\}$ denote the position of c_j in which this literal occurs. Add the pairs (p_j^t, c_j^t) ($1 \le t \ne s \le 3$), (p_j^s, z_j) and (q_j, c_j^s) to M.

As M is a complete matching in I, clearly no woman in $Y \cup Z$ can be involved in a blocking pair of M in I. Nor can a man in P (since he can only potentially prefer a woman in Z) nor a man in Q (since he can only potentially prefer a woman in C, who ranks him last). Now suppose that $(x_{4i+r}, c(x_{4i+r}))$ blocks M, where $0 \le i \le n-1$ and $0 \le r \le 3$. Let $c_j^s = c(x_{4i+r})$, where $1 \le j \le m$ and $1 \le s \le 3$. Then $(q_j, c_j^s) \in M$. If $r \in \{0, 1\}$ then $(x_{4i+r}, y_{4i+r+1}) \in M$, so that v_i is false under f. But literal v_i occurs in c_j, a contradiction, since literal v_i was supposed to be true under f by construction of M. Hence $r \in \{2, 3\}$ and $(x_{4i+r}, y_{4i+r}) \in M$, so that v_i is true under f. But literal $\bar{v}_i$ occurs in c_j, a contradiction, since literal $\bar{v}_i$ was supposed to be true under f by construction of M. Hence M is weakly stable in I.

Conversely suppose that M is a complete weakly stable matching in I. We form a truth assignment f in B as follows. For each i ($0 \le i \le n-1$), $M \cap (X_i \times Y_i)$ is a perfect matching of $X_i \cup Y_i$. If $M \cap (X_i \times Y_i) = T_i$, set v_i to be true under f. Otherwise $M \cap (X_i \times Y_i) = F_i$, in which case we set v_i to be false under f.

Now let c_j be a clause in C ($1 \le j \le m$). There exists some s ($1 \le s \le 3$) such that $(q_j, c_j^s) \in M$. Let $x_{4i+r} = x(c_j^s)$ for some i ($0 \le i \le n-1$) and r ($0 \le r \le 3$). If $r \in \{0, 1\}$ then $(x_{4i+r}, y_{4i+r}) \in M$ by the weak stability of M. Thus variable v_i is true under f, and hence clause c_j is true under f, since literal v_i occurs in c_j. If $r \in \{2, 3\}$ then $(x_{4i+r}, y_{4i+r+1}) \in M$ (where addition is taken modulo 4) by the weak stability of M. Thus variable v_i is false under f, and hence clause c_j is true under f, since literal $\bar{v}_i$ occurs in c_j. Hence f is a satisfying truth assignment of B. $\square$

4 Inapproximability of $(3, 4)$-MAX SMTI

In this section we give an inapproximability result for $(3, 4)$-MAX SMTI. Specifically, we show that $(3, 4)$-MAX SMTI is not approximable within δ, for some $\delta > 1$, unless P=NP. Our proof involves a reduction from a problem involving matchings in graphs. A matching M in a graph G is said to be *maximal* if no proper superset of M is a matching in G. Define MIN-MM to be the problem of finding a minimum cardinality maximal matching, given a graph G. By [8, Theorem 1], MIN-MM is not approximable

within some $\delta_0 > 1$ unless P=NP. The result holds even for subdivision graphs of cubic graphs. (Given a graph G, the *subdivision graph* of G, denoted by $S(G)$, is obtained by subdividing each edge $\{u, w\}$ of G in order to obtain two edges $\{u, v\}$ and $\{v, w\}$ of $S(G)$, where v is a new vertex.)

Theorem 6. $(3, 4)$-MAX SMTI *is NP-hard and not approximable within δ, for some $\delta > 1$, unless P=NP.*

Proof. Let G be an instance of MIN-MM restricted to subdivision graphs of cubic graphs. Then $G = (U, W, E)$ is a bipartite graph where, without loss of generality, each vertex in U has degree 2 and each vertex in W has degree 3. Let $U = \{m_1, \ldots, m_s\}$ and let $W = \{w_1, \ldots, w_t\}$. For each vertex $m_i \in U$, let W_i denote the two vertices adjacent to m_i in G. Similarly for each vertex $w_j \in W$, let U_j denote the three vertices adjacent to w_j in G. We construct an instance I of $(3, 4)$-MAX SMTI as follows: let $U \cup X$ be the set of men and let $W \cup Y$ be the set of women, where $X = \{x_1, \ldots, x_t\}$ and $Y = \{y_1, \ldots, y_s\}$. The preference lists of the men and women in I are as follows:

$$
\begin{array}{llll}
m_i : & (W_i) \; y_i & (1 \le i \le s) & \qquad w_j : \; (U_j) \; x_j \quad (1 \le j \le t) \\
x_i : & w_i & (1 \le i \le t) & \qquad y_j : \; m_j \qquad\quad\; (1 \le j \le s)
\end{array}
$$

In a given preference list, entries within round brackets are tied. Clearly the length of each man's preference list is at most 3, whilst the length of each woman's preference list is at most 4. We claim that $s^+(I) = s + t - \beta_1^-(G)$, where $s^+(I)$ denotes the maximum size of a weakly stable matching in I and $\beta_1^-(G)$ denotes the minimum size of a maximal matching in G.

For suppose that G has a maximal matching M, where $|M| = \beta_1^-(G)$. We construct a matching M' in I as follows. Initially let $M' = M$. There remain $s - |M|$ men in U that are unmatched in M'; denote these men by m_{i_r} $(1 \le i \le s - |M|)$, and add (m_{i_r}, y_{i_r}) to M' for each such m_{i_r}. Finally there remain $t - |M|$ women in W that are unmatched in M'; denote these women by w_{j_r} $(1 \le r \le t - |M|)$, and add (x_{j_r}, w_{j_r}) to M' for each such w_{j_r}. Clearly M' is a matching in I such that $|M'| = |M| + (s - |M|) + (t - |M|) = s + t - \beta_1^-(G)$. It is straightforward to verify that M' is weakly stable in I, and hence $s^+(I) \ge s + t - \beta_1^-(G)$.

Conversely suppose that M' is a weakly stable matching in I, where $|M'| = s^+(I)$. Let $M = M' \cap E$. The weak stability of M' in I implies that M is maximal in G. Moreover, at most $t - |M|$ women in W are matched in M' to men in X, and at most $s - |M|$ men in U are matched in M' to women in Y, and hence $|M'| \le |M| + (t - |M|) + (s - |M|) = s + t - |M|$. Thus $s^+(I) \le s + t - \beta_1^-(G)$. Hence the claim is established.

Theorem 1 of [8] shows that it is NP-hard to distinguish between the cases that $\beta_1^-(G) \le c_0 m$ and $\beta_1^-(G) > \delta_0 c_0 m$, where $c_0 > 0$ is some constant and $m = |E|$. Now if $\beta_1^-(G) \le c_0 m$ then $s^+(I) \ge cs$, whilst if $\beta_1^-(G) > \delta_0 c_0 m$ then $s^+(I) < \delta cs$, where $c = (5 - 6c_0)/3$ and $\delta = (5 - 6\delta_0 c_0)/(5 - 6c_0)$. The result follows by Theorem 1 and Proposition 4 of [8]. □

5   Concluding remarks

In this paper we have presented a polynomial-time algorithm for $(2, \infty)$-MAX SMTI, but have shown that, by contrast, $(3, 3)$-MAX SMTI is NP-hard and $(3, 4)$-MAX SMTI

is not approximable within some $\delta > 1$ unless P=NP.

For the NP-hard variants of (p, q)-MAX SMTI, it remains to investigate the existence of approximation algorithms for these problems that improve on the performance guarantees of those that have already been formulated for the general SMTI case (with no assumptions on the lengths of the preference lists) [19, 9, 16, 17, 18].

Also, the natural extension of (p, q)-MAX SMTI to the many-one HRT case may be formulated: we denote this problem by (p, q)-MAX HRT. It remains to extend the algorithm for $(2, \infty)$-MAX SMTI to the case of $(2, \infty)$-MAX HRT or prove that the latter problem is NP-hard. Clearly Theorem 5 implies that $(3, 3)$-MAX HRT is NP-hard, whilst Theorem 6 implies that $(3, 4)$-MAX HRT is NP-hard and not approximable within some $\delta > 1$ unless P=NP.

Acknowledgements

References

[1] P. Berman, M. Karpinski, and A.D. Scott. Approximation hardness of short symmetric instances of MAX-3SAT. Electronic Colloquium on Computational Complexity Report, number 49, 2003.

[2] H.N. Gabow and R.E. Tarjan. Faster scaling algorithms for network problems. *SIAM Journal on Computing*, 18(5):1013–1036, 1989.

[3] D. Gale and L.S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69:9–15, 1962.

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

[5] M.R. Garey, D.S. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1:237–267, 1976.

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

[7] M. Halldórsson, K. Iwama, S. Miyazaki, and Y. Morita. Inapproximability results on stable marriage problems. In *Proceedings of LATIN 2002: the Latin-American Theoretical INformatics symposium*, volume 2286 of *Lecture Notes in Computer Science*, pages 554–568. Springer, 2002.

[8] 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*, 306(1-3):431–447, 2003.

[9] M.M. Halldórsson, K. Iwama, S. Miyazaki, and H. Yanagisawa. Improved approximation of the stable marriage problem. In *Proceedings of ESA 2003: the 11th Annual European Symposium on Algorithms*, volume 2832 of *Lecture Notes in Computer Science*, pages 266–277. Springer, 2003.

[10] M.M. Halldórsson, K. Iwama, S. Miyazaki, and H. Yanagisawa. Randomized approximation of the stable marriage problem. *Theoretical Computer Science*, 325(3):439–465, 2004.

[11] R.W. Irving. Stable marriage and indifference. *Discrete Applied Mathematics*, 48:261–272, 1994.

[12] R.W. Irving. Matching medical students to pairs of hospitals: a new variation on a well-known theme. In *Proceedings of ESA '98: the Sixth European Symposium on Algorithms*, volume 1461 of *Lecture Notes in Computer Science*, pages 381–392. Springer, 1998.

[13] R.W. Irving and P. Leather. The complexity of counting stable marriages. *SIAM Journal on Computing*, 15(3):655–667, 1986.

[14] R.W. Irving, D.F. Manlove, and S. Scott. The Hospitals/Residents problem with Ties. In *Proceedings of SWAT 2000: the 7th Scandinavian Workshop on Algorithm Theory*, volume 1851 of *Lecture Notes in Computer Science*, pages 259–271. Springer, 2000.

[15] K. Iwama, D. Manlove, S. Miyazaki, and Y. Morita. Stable marriage with incomplete lists and ties. In *Proceedings of ICALP '99: the 26th International Colloquium on Automata, Languages, and Programming*, volume 1644 of *Lecture Notes in Computer Science*, pages 443–452. Springer, 1999.

[16] K. Iwama, S. Miyazaki, and K. Okamoto. A $(2 - c\frac{\log n}{n})$-approximation algorithm for the stable marriage problem. In *Proceedings of SWAT 2004: the 9th Scandinavian Workshop on Algorithm Theory*, volume 3111 of *Lecture Notes in Computer Science*, pages 349–361. Springer, 2004.

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

[18] K. Iwama, S. Miyazaki, and N. Yamauchi. A 1.875–approximation algorithm for the stable marriage problem. In *Proceedings of SODA '07: the 20th ACM-SIAM Symposium on Discrete Algorithms* (to appear). ACM-SIAM, 2007.

[19] D.F. Manlove, R.W. Irving, K. Iwama, S. Miyazaki, and Y. Morita. Hard variants of stable marriage. *Theoretical Computer Science*, 276(1-2):261–279, 2002.

[20] A.E. Roth. The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of Political Economy*, 92(6):991–1016, 1984.

[21] `http://www.carms.ca` (Canadian Resident Matching Service website).

[22] `http://www.nrmp.org/` (National Resident Matching Program website).

[23] `http://www.nes.scot.nhs.uk/sfas` (Scottish Foundation Allocation Scheme website).