



**UNIVERSITY**  
*of*  
**GLASGOW**

Halldorsson, M. and Irving, R.W. and Iwama, K. and Manlove, D. F. and Miyazaki, S. and Morita, Y. and Scott, S. (2003) Approximability results for stable marriage problems with ties. *Theoretical Computer Science* 306(1-3):431-447.

<http://eprints.gla.ac.uk/archive/00000309/>

# Approximability Results for Stable Marriage Problems with Ties

Magnús M. Halldórsson<sup>1</sup>, Robert W. Irving<sup>2</sup>, Kazuo Iwama<sup>\*,3</sup>,  
David F. Manlove<sup>†,2</sup>, Shuichi Miyazaki<sup>4</sup>, Yasufumi Morita<sup>3</sup>,  
and Sandy Scott<sup>2</sup>

<sup>1</sup> Department of Computer Science, University of Iceland, Dunhaga 3, IS-107 Reykjavik, Iceland.  
Email: mmh@hi.is.

<sup>2</sup> Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK.  
Email: {rwi,davidm,sas}@dcs.gla.ac.uk.

<sup>3</sup> School of Informatics, Kyoto University, Kyoto 606-8501, Japan.  
Email: {iwama,ymorita}@kuis.kyoto-u.ac.jp.

<sup>4</sup> Academic Center for Computing and Media Studies, Kyoto University, Kyoto 606-8501, Japan.  
Email: shuichi@kuis.kyoto-u.ac.jp.

## Abstract

We consider instances of the classical stable marriage problem in which persons may include ties in their preference lists. We show that, in such a setting, strong lower bounds hold for the approximability of each of the problems of finding an egalitarian, minimum regret and sex-equal stable matching. We also consider stable marriage instances in which persons may express unacceptable partners in addition to ties. In this setting, we prove that there are constants  $\delta, \delta'$  such that each of the problems of approximating a maximum and minimum cardinality stable matching within factors of  $\delta, \delta'$  (respectively) is NP-hard, under strong restrictions. We also give an approximation algorithm for both problems that has a performance guarantee expressible in terms of the number of lists with ties. This significantly improves on the best-known previous performance guarantee, for the case that the ties are sparse. Our results have applications to large-scale centralised matching schemes.

**Keywords:** Stable marriage problem; Ties; Unacceptable partners; Inapproximability results; Approximation algorithm

## 1 Introduction

An instance  $I$  of the classical Stable Marriage problem (SM) [6, 21, 17] involves  $n$  men and  $n$  women, each of whom ranks all the members of the opposite sex in strict order of preference. A *matching*  $M$  in  $I$  is a bijection between the men and women. We say that a (man,woman)

---

\*Supported in part by Scientific Research Grant, Ministry of Japan, 13480081.

†Supported by award NUF-NAL-02 from the Nuffield Foundation and grant GR/R84597/01 from the Engineering and Physical Sciences Research Council.

pair  $(m, w)$  *blocks*  $M$ , or is a *blocking pair* with respect to  $M$ , if each of  $m$  and  $w$  prefers the other to his/her partner in  $M$ . A matching that admits no blocking pair is said to be *stable*. It is known that every instance of SM admits at least one stable matching [3], and that such a matching can be found in  $O(n^2)$  time using the Gale / Shapley algorithm [3].

The *man-oriented* version of the Gale/Shapley algorithm [3] yields a stable matching called the *man-optimal stable matching*. This is the unique stable matching in which each man has his best possible partner (and each woman her worst) among all stable matchings. Similarly, the *woman-oriented* version leads to the *woman-optimal stable matching* with analogous optimality conditions for the women (and pessimality conditions for the men).

### “Fair” stable matchings

In view of the fact that man-optimal and woman-optimal stable matchings are woman-pessimal and man-pessimal respectively, it is of interest to consider stable matchings that are “fair” to both sexes in a precise sense. Given a matching  $M$  and a person  $q$  in a given SM instance  $I$ , define the *cost* of  $M$  for  $q$ , denoted by  $c_M(q)$ , to be the ranking of  $p_M(q)$  in  $q$ ’s preference list, where  $p_M(q)$  denotes  $q$ ’s partner in  $M$ . In other words,  $c_M(q)$  is one plus the number of persons whom  $q$  prefers to  $p_M(q)$ . Let  $U$  and  $W$  denote the set of men and women in  $I$  respectively, and let  $\mathcal{M}$  denote the set of stable matchings in  $I$ . Define an *egalitarian stable matching* to be a stable matching  $S$  for which  $c(S) = \min_{M \in \mathcal{M}} c(M)$ , where  $c(M) = \sum_{q \in U \cup W} c_M(q)$  for any  $M \in \mathcal{M}$ . Similarly, define a *minimum regret stable matching* to be a stable matching  $S$  for which  $r(S) = \min_{M \in \mathcal{M}} r(M)$ , where  $r(M) = \max_{q \in U \cup W} c_M(q)$  for any  $M \in \mathcal{M}$ . Finally, define a *sex-equal stable matching* to be a stable matching  $S$  for which  $d(S) = \min_{M \in \mathcal{M}} d(M)$ , where

$$d(M) = \left| \sum_{m \in U} c_M(m) - \sum_{w \in W} c_M(w) \right|$$

for any  $M \in \mathcal{M}$ .

Intuitively, an egalitarian stable matching seeks to minimize the total cost of  $M$  taken over all persons in  $I$ , whilst a minimum regret stable matching aims to minimize the maximum cost of  $M$  taken over all persons in  $I$ . Finally in a sex-equal stable matching, the total cost of  $M$  for the men in  $I$  is as close to the total cost of  $M$  for the women in  $I$  as possible.

Denote the problems of finding an egalitarian, minimum regret and sex-equal stable matching by EGALITARIAN SM, MINIMUM REGRET SM and SEX-EQUAL SM respectively, given an instance of SM. It is known that each of EGALITARIAN SM and MINIMUM REGRET SM is polynomial-time solvable [13, 2, 7]. However SEX-EQUAL SM has been shown to be NP-hard [16].

### Ties in the preference lists

A natural generalisation of SM arises when each person need not rank all members of the opposite sex in *strict* order. Some of those might be indifferent among certain members of the opposite sex, so that preference lists may involve *ties*<sup>1</sup>. We use SMT to stand for the variant of SM in which preference lists may include ties. (Henceforth we assume that a tie is of length at least two.) In this context, a matching  $M$  is stable if there is no (man,woman) pair  $(m, w)$ , each of whom *strictly* prefers the other to his/her partner in  $M$ <sup>2</sup>.

---

<sup>1</sup>In this paper, we restrict attention to the case where the indifference takes the form of ties in the preference lists, but the results presented extend to the general case where the preference lists are arbitrary partial orders.

<sup>2</sup>Implicitly here, and henceforth for other stability definitions, such a pair  $(m, w)$  is defined to *block*  $M$ , or to be a *blocking pair* with respect to  $M$ , as for the SM case.

By breaking the ties arbitrarily, an instance  $I$  of SMT becomes an instance  $I'$  of SM, and clearly a stable matching in  $I'$  is also stable in  $I$ . Thus a stable matching in  $I$  can be found using the Gale/Shapley algorithm. (Conversely, given a stable matching  $M$  in  $I$ , it is not difficult to see that there is an instance  $I_M$  of SM in which  $M$  is stable. Hence a matching  $M$  is stable in  $I$  if and only if  $M$  is stable in some instance of SM obtained from  $I$  by breaking the ties.)

The stability criterion considered here is referred to as *weak stability* in [11], where two other notions of stability are formulated for SMT, so-called *strong stability* and *super-stability*. However an instance of SMT need not admit a strongly stable matching or a super-stable matching [11]. By contrast, we have already seen that every instance of SMT admits at least one weakly stable matching. Therefore, perhaps unsurprisingly, of these three definitions, it is weak stability that has received the most attention in the literature [20, 19, 15, 18]. We are concerned exclusively with weak stability in this paper, and henceforth for brevity, the term *stability* will be used to indicate weak stability when ties are present.

The concept of the cost of a matching for a person may easily be extended to the SMT context. Given a matching  $M$  and a person  $q$  in an SMT instance  $I$ ,  $c_M(q)$  is the (possibly joint) ranking of  $p_M(q)$  in  $q$ 's preference list. In other words,  $c_M(q)$  is one plus the number of persons whom  $q$  strictly prefers to  $p_M(q)$ . Given this extension of the definition of  $c_M(q)$ , each of the definitions of an egalitarian, minimum regret and sex-equal stable matching in an instance of SMT follows immediately. Define EGALITARIAN SMT, MINIMUM REGRET SMT and SEX-EQUAL SMT to be the analogous problems to EGALITARIAN SM, MINIMUM REGRET SM and SEX-EQUAL SM respectively, given an instance of SMT.

It is known that each of EGALITARIAN SMT and MINIMUM REGRET SMT is NP-hard, and not approximable within  $n^{1-\varepsilon}$ , for any  $\varepsilon > 0$ , unless P=NP, where  $n$  is the number of persons in a given SMT instance [18]. In this paper we improve these results by demonstrating that a worst possible  $\Omega(n)$  lower bound on the approximability of each of these problems holds. In addition we prove that a similar lower bound holds for SEX-EQUAL SMT.

## Unacceptable partners

An alternative natural extension of SM occurs when persons are permitted to express unacceptable partners. We say that person  $p$  is *acceptable* to person  $q$  if  $p$  appears on the preference list of  $q$ , and *unacceptable* otherwise. If person  $q$  is missing from person  $p$ 's preference list,  $p$  is not prepared to be matched with  $q$ , or to form a blocking pair with  $q$ . We use SMI to stand for this variant of SM where preference lists may be incomplete.

It follows immediately that a matching  $M$  in an instance  $I$  of SMI is now a one-one correspondence between a subset of the men and a subset of the women, such that  $(m, w) \in M$  implies that each of  $m, w$  is acceptable to the other. Also, the revised notion of stability may be defined as follows:  $M$  is stable if there is no (man,woman) pair  $(m, w)$ , each of whom is either unmatched in  $M$  and finds the other acceptable, or prefers the other to his/her partner in  $M$ . (As a consequence of this definition, it follows that from the point of view of finding stable matchings, we may assume, without loss of generality, that  $p$  is acceptable to  $q$  if and only if  $q$  is acceptable to  $p$ .)

A stable matching in  $I$  need not be a complete matching. However, all stable matchings in  $I$  have the same size, and involve exactly the same men and women [4]. Therefore, each of the definitions of an egalitarian, a minimum regret and a sex-equal stable matching in an instance of SMI follows immediately from its SM definition if we discard the unmatched men and women from consideration. In addition, it is a simple matter to extend the Gale/Shapley algorithm to the SMI setting (see [6, Section 1.4.2]).

## Ties and unacceptable partners

The variant of the stable marriage problem which incorporates *both* extensions described above is denoted SMTI. Thus an instance  $I$  of SMTI comprises preference lists, each of which may involve ties and/or unacceptable partners. A combination of the earlier definitions indicates that a matching  $M$  in  $I$  is stable if there is no (man,woman) pair  $(m,w)$ , each of whom is either unmatched in  $M$  and finds the other acceptable, or strictly prefers the other to his/her partner in  $M$ .

As observed above, all stable matchings for a given instance of SMI are of the same size, and all stable matchings for a given instance of SMT are complete (and therefore of the same size). However, for a given instance of SMTI, it is no longer the case that all stable matchings need be of the same size [18]. Furthermore, each of the problems of finding a stable matching of maximum or minimum size, given an SMTI instance, is NP-hard [15, 18]. Therefore one is naturally led to consider the approximability properties of each of these problems. It turns out that each problem admits an approximation algorithm with a performance ratio of 2, since the size of any stable matching is at least half the size of a maximum cardinality stable matching and is at most twice the size of a minimum cardinality stable matching [18]. This has left open the question of whether better approximation algorithms for these problems exist.

In this paper we present both positive and negative results regarding the approximability of each of these problems: we show that the existence of a polynomial-time approximation scheme (PTAS) for either of these problems is unlikely, since there exist constants  $\delta, \delta'$  such that approximating each problem within a factor of  $\delta, \delta'$  (respectively) is NP-hard, under strong restrictions on the instance. However, we also show that, for a given SMTI instance  $I$ , the difference in size between a maximum and a minimum cardinality stable matching is bounded by  $t(I)$ , the number of preference lists that contain ties, and this leads to an approximation algorithm for both problems with a performance guarantee dependent on  $t(I)$ . When  $t(I)$  is relatively small compared to the size of the instance, our result significantly improves on the best-known previous result regarding the approximability of both problems, namely the performance ratio of 2.

## Practical applications

The problems of finding “fair” stable matchings and maximum cardinality stable matchings in a given instance of SMTI have particular significance in practical applications. In a number of countries, large-scale automated matching schemes produce stable matchings of graduating medical students to hospital posts based on the preferences of students over hospitals and vice versa. Examples of such schemes are the National Resident Matching Program (NRMP) [20] in the U.S., the Canadian Resident Matching Service (CaRMS) [1] and the Scottish Pre-registration house officer Allocation scheme (SPA) [12].

The algorithms employed by the NRMP and CaRMS essentially solve a many-one generalisation of SMI called the Hospitals / Residents problem (HR) [6, Section 1.6]. In the context of these two matching schemes, hospitals must rank a possibly large number of applicants in strict order of preference. However, it is unrealistic to expect large and popular hospitals to provide a strict ranking of all of their applicants. The SPA scheme permits hospitals to include ties, a situation which may be modelled by a many-one matching problem known as the Hospitals/Residents problem with Ties (HRT) [14], a generalisation of each of HR and SMTI.

Thus, since the stable matchings in an instance of SMTI may be of different sizes, the same is true for HRT. Yet a prime objective of any matching scheme must be to match as many applicants as possible, and hence this motivates the search for large stable matchings.

In addition, administrators of matching schemes may be interested to find stable matchings that are as fair as possible for both applicants and hospitals alike, and hence this motivates the search for egalitarian, minimum regret and sex-equal stable matchings. Thus our approximability results have implications for matching schemes such as SPA.

## Organisation of the paper

The remainder of this paper is organised as follows. In Section 2 we prove that it is hard to approximate the MIN MAXIMAL MATCHING optimization problem (defined in that section) in a certain class of graphs. This result is required in order to establish, in Section 3, the hardness results for the problems of approximating a maximum or minimum cardinality stable matching in a given instance of SMTI. Then, in Section 4 we present the approximation algorithm for the variants of these problems where, in a given SMTI instance, the number of lists containing ties is bounded. The  $\Omega(n)$  lower bounds for each of the problems of approximating EGALITARIAN SMT, MINIMUM REGRET SMT and SEX-EQUAL SMT are presented in Section 5. Finally, in Section 6 we present some concluding remarks.

## 2 Hardness of approximating MIN MAXIMAL MATCHING

We begin this section with some graph-related definitions. Given a graph  $G = (V, E)$ , a *strongly stable set*  $S$  is a subset of  $V$  such that the distance between every pair of vertices in  $S$  is at least 3. A matching  $M$  in  $G$  is *maximal* if no proper superset of  $M$  is a matching in  $G$ . Let  $\beta_0(G)$ ,  $\beta_{SS}(G)$  and  $\beta_1^-(G)$  denote respectively the sizes of a maximum independent set, a maximum strongly stable set and a minimum maximal matching in  $G$ . Define MIN MAXIMAL MATCHING to be the problem of computing  $\beta_1^-(G)$ , given a graph  $G$ .

MIN MAXIMAL MATCHING is NP-hard, even for subdivision graphs of graphs of maximum degree 3 [10] (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). In this section, we will establish that MIN MAXIMAL MATCHING is hard to approximate in a certain graph class; this result will be required in the next section. In particular, we will prove the following:

**Theorem 2.1.** *It is NP-hard to approximate MIN MAXIMAL MATCHING within  $\delta_0$ , for some  $\delta_0 > 1$ . The result holds even if the instance is restricted to be the subdivision graph of some cubic graph.*

Our proof of Theorem 2.1 involves a chain of reductions starting from MAX-IS. This is the problem of computing  $\beta_0(G)$ , given a graph  $G$ . We denote by MAX-IS( $k$ ) the restriction of MAX-IS in which  $G$  is regular of degree  $k$ .

**Theorem 2.2 ([9]).** *It is NP-hard to approximate MAX-IS(3) within  $\delta_1$ , for some  $\delta_1 < 1$ .*

In fact, there exists a constant  $c_1 > 0$  such that it is NP-hard to distinguish between instances  $G = (V, E)$  of MAX-IS(3) such that  $\beta_0(G) \geq c_1 n$  and  $\beta_0(G) < \delta_1 c_1 n$ , where  $n = |V|$ .

We will use Theorem 2.2 together with the notion of a *gap-preserving reduction* [22, p.308], which may be defined as follows:

**Definition 2.3.** *Let  $\Pi_1$  and  $\Pi_2$  be two optimization problems. Denote by  $OPT_i(x)$  the optimal measure over all feasible solutions for a given instance  $x$  of  $\Pi_i$  ( $i \in \{1, 2\}$ ). Let  $\alpha$  be some constant ( $\alpha \leq 1$  if  $\Pi_1$  is a maximization problem;  $\alpha \geq 1$  otherwise), and let  $g_1$  be a function that maps an instance  $x$  of  $\Pi_1$  to a positive rational number. Then a gap-preserving reduction from  $\Pi_1$  to  $\Pi_2$  is a tuple  $\langle f, \beta, g_2 \rangle$  such that:*

- $f$  maps an instance  $x$  of  $\Pi_1$  to an instance  $f(x)$  of  $\Pi_2$  in polynomial time;
- $\beta$  is a constant ( $\beta \leq 1$  if  $\Pi_2$  is a maximization problem;  $\beta \geq 1$  otherwise);
- $g_2$  maps an instance  $f(x)$  of  $\Pi_2$  to a positive rational number;
- if  $\Pi_1$  and  $\Pi_2$  are maximization problems, then for any instance  $x$  of  $\Pi_1$ :
  - if  $OPT_1(x) \geq g_1(x)$ , then  $OPT_2(f(x)) \geq g_2(f(x))$ ;
  - if  $OPT_1(x) < \alpha g_1(x)$ , then  $OPT_2(f(x)) < \beta g_2(f(x))$ ;
 (if  $\Pi_i$  is a minimization problem, for  $i \in \{1, 2\}$ , then the two inequalities involving  $OPT_i$  in the above conditions should be reversed).

The following proposition is an immediate consequence of Definition 2.3.

**Proposition 2.4.** *Let  $\Pi_1$  and  $\Pi_2$  be two maximization problems, and suppose that there is a gap-preserving reduction from  $\Pi_1$  to  $\Pi_2$ . Assuming the notation of Definition 2.3, suppose further that it is NP-hard to distinguish between instances  $x$  of  $\Pi_1$  such that  $OPT_1(x) \geq g_1(x)$  and  $OPT_1(x) < \alpha g_1(x)$ . Then it is NP-hard to distinguish between instances  $f(x)$  of  $\Pi_2$  such that  $OPT_2(f(x)) \geq g_2(f(x))$  and  $OPT_2(f(x)) < \beta g_2(f(x))$ . (If  $\Pi_i$  is a minimization problem, for  $i \in \{1, 2\}$ , then the two inequalities involving  $OPT_i$  in the above conditions should be reversed). Hence it is NP-hard to approximate  $\Pi_2$  within  $\beta$ .*

Our first gap-preserving reduction involves MAX-SSS. This is the problem of computing  $\beta_{SS}(G)$  for a given graph  $G$ . We denote by MAX-SSS( $k$ ) the restriction of MAX-SSS in which  $G$  is regular of degree  $k$ .

**Theorem 2.5.** *It is NP-hard to approximate MAX-SSS(3) within  $\delta_2$ , for some  $\delta_2 < 1$ .*

*Proof.* Let  $G = (V, E)$  be a cubic graph, given as an instance of MAX-IS(3), where  $n = |V|$  and  $m = |E|$ . We construct a cubic graph  $G' = (V', E')$  as an instance of MAX-SSS(3) as follows. As in the proof of Corollary 3.4 of [10], we initially replace every edge  $\{v, w\}$  of  $G$  by a component comprising the edges  $\{v, u\}, \{u, w\}, \{u, u'\}, \{u', w\}$ . This leaves  $m$  vertices of degree 1 in  $G'$  and  $m$  vertices of degree 2 in  $G'$ .

We may eliminate such vertices as follows. To every vertex  $v$  of degree 1 in  $G'$ , connect the component shown in Figure 1(a). Similarly, for every vertex  $v$  of degree 2 in  $G'$ , connect the component shown in Figure 1(b). It is then clear that the modified graph  $G'$  is cubic.

It is straightforward to verify that  $G$  has an independent set of size  $k$  if and only if  $G'$  has a strongly stable set of size  $3m + k$ , and hence  $\beta_{SS}(G') = \beta_0(G) + 3m$ . Now  $2m = 3n$  as  $G$  is cubic, and it may be verified that  $n' = 22n$ , where  $n' = |V'|$ .

Now let  $c_1$  and  $\delta_1$  be the constants given by Theorem 2.2, such that it is NP-hard to distinguish between the cases  $\beta_0(G) \geq c_1 n$  and  $\beta_0(G) < \delta_1 c_1 n$ . Hence if  $\beta_0(G) \geq c_1 n$ , then  $\beta_{SS}(G') \geq c_2 n'$ , whilst if  $\beta_0(G) < \delta_1 c_1 n$ , then  $\beta_{SS}(G') < \delta_2 c_2 n'$ , where  $c_2 = \frac{2c_1+9}{44}$  and  $\delta_2 = \frac{2\delta_1 c_1 + 9}{2c_1 + 9}$ . The result then follows by Theorem 2.2 and Proposition 2.4.  $\square$

Our second gap-preserving reduction is sufficient to prove Theorem 2.1.

*Proof of Theorem 2.1.* Let  $G = (V, E)$  be a cubic graph, given as an instance of MAX-SSS(3), where  $n = |V|$  and  $m = |E|$ . The constructed instance of MIN MAXIMAL MATCHING is  $S(G)$  (recall that  $S(G)$  is the subdivision graph of  $G$ ). Now by Lemmas 3.1 and 3.2 of [10],  $G$  has a strongly stable set of size  $k$  if and only if  $S(G)$  has a maximal matching of size  $n - k$ . Thus it follows that  $\beta_1^-(S(G)) + \beta_{SS}(G) = n$ . Now  $2m = 3n$  as  $G$  is cubic, and  $m' = 2m$ , where  $m'$  is the number of edges of  $S(G)$ .



Figure 1: Components attached to vertices of degree 1 or 2 in  $G'$ .

Now let  $c_2$  and  $\delta_2$  be the constants given by Theorem 2.5, such that it is NP-hard to distinguish between the cases  $\beta_{SS}(G) \geq c_2n$  and  $\beta_{SS}(G) < \delta_2c_2n$ . Hence if  $\beta_{SS}(G) \geq c_2n$ , then  $\beta_1^-(S(G)) \leq c_0m'$ , whilst if  $\beta_{SS}(G) < \delta_2c_2n$ , then  $\beta_1^-(S(G)) > \delta_0c_0m'$ , where  $c_0 = \frac{1-c_2}{3}$  and  $\delta_0 = \frac{1-\delta_2c_2}{1-c_2}$ . The result then follows by Theorem 2.5 and Proposition 2.4.  $\square$

### 3 Hardness of approximating MAX SMTI and MIN SMTI

Given an instance  $I$  of SMTI, let  $s^+(I)$  (respectively  $s^-(I)$ ) denote the size of a maximum (respectively minimum) cardinality stable matching in  $I$ . Define MAX (respectively MIN) SMTI to be the problem of computing  $s^+(I)$  (respectively  $s^-(I)$ ), given an SMTI instance  $I$ .

Each of MAX SMTI and MIN SMTI is NP-hard [15, 18]. In this section we prove that there exist constants  $\delta, \delta'$  such that each of the problems of approximating MAX SMTI and MIN SMTI within a factor of  $\delta, \delta'$  (respectively) is NP-hard. In each case, the result holds under the restriction that the ties belong to the preference lists of one sex only, and preference lists have constant length. We begin by considering MAX SMTI.

**Theorem 3.1.** *It is NP-hard to approximate MAX SMTI within  $\delta_3$ , for some  $\delta_3 < 1$ . The result holds even if the preference lists in the given instance are of constant length, there is at most one tie per list, and the ties occur on one side only.*

*Proof.* Let  $G = (V, E)$  be the subdivision graph of some cubic graph, given as an instance of MIN MAXIMAL MATCHING. Then  $G$  has a bipartition of  $V$  into the left-hand vertex set  $U$  and the right-hand vertex set  $W$ , where every vertex in  $U$  has degree 3 and every vertex in  $W$  has degree 2.

Let  $U = \{m_1, m_2, \dots, m_s\}$  and  $W = \{w_1, w_2, \dots, w_t\}$ . For each  $i$  ( $1 \leq i \leq s$ ), assume that  $m_i$  is adjacent in  $G$  to the vertices in  $W_i$ , where  $W_i = \{w_{k_{3i}-2}, w_{k_{3i}-1}, w_{k_{3i}}\}$ . Also, assume that  $p_j$  and  $q_j$  are two sequences such that  $p_j < q_j$ ,  $\{m_{p_j}, w_j\} \in E$  and  $\{m_{q_j}, w_j\} \in E$  ( $1 \leq j \leq t$ ).

We form an instance  $I$  of MAX SMTI as follows. Let  $\mathcal{U}$  be the set of men in  $I$ , where  $\mathcal{U} = U \cup X \cup Z$ ,  $X = \{x_1, x_2, \dots, x_t\}$ , and  $Z = \{z_1, z_2, \dots, z_t\}$ . Also, let  $\mathcal{W}$  be the set of women in  $I$ , where  $\mathcal{W} = W \cup W' \cup Y$ ,  $W' = \{w'_1, w'_2, \dots, w'_t\}$ , and  $Y = \{y_1, y_2, \dots, y_s\}$ . For each  $i$  ( $1 \leq i \leq s$ ), let  $W'_i = \{w'_{k_{3i}-2}, w'_{k_{3i}-1}, w'_{k_{3i}}\}$ . Clearly  $|\mathcal{U}| = |\mathcal{W}| = s + 2t$ . Create a preference list for each person in  $I$  as follows:

$$\begin{array}{lll} m_i : (W_i \cup W'_i) \quad y_i & (1 \leq i \leq s) & w_j : z_j \quad m_{p_j} \quad m_{q_j} \quad x_j \quad (1 \leq j \leq t) \\ x_i : w_i & (1 \leq i \leq t) & w'_j : z_j \quad m_{q_j} \quad m_{p_j} \quad (1 \leq j \leq t) \\ z_i : (w_i \quad w'_i) & (1 \leq i \leq t) & y_j : m_j \quad (1 \leq j \leq s) \end{array}$$

Note that, in a given preference list throughout this paper, persons listed within round brackets are tied. Clearly the ties in  $I$  occur in the men's preference lists only and there is at most one tie per list. Also each man's list has length at most 7, whilst each woman's list has length at most 4.

Suppose that  $M$  is a maximal matching in  $G$ , where  $|M| = \beta_1^-(G)$ . We construct a matching  $M'$  in  $I$  as follows. For each  $i$  ( $1 \leq i \leq s$ ), suppose firstly that  $m_i$  is matched in  $M$ , to  $w_j$  say ( $1 \leq j \leq t$ ). If  $i = p_j$ , add the pairs  $(m_i, w_j)$  and  $(z_j, w'_j)$  to  $M'$ . If  $i = q_j$ , add the pairs  $(m_i, w'_j)$  and  $(z_j, w_j)$  to  $M'$ .

On the other hand, if  $m_i$  is unmatched, add the pair  $(m_i, y_i)$  to  $M'$ .

Finally, for any  $j$  ( $1 \leq j \leq t$ ), if  $w_j$  is unmatched, add the pairs  $(x_j, w_j)$  and  $(z_j, w'_j)$  to  $M'$ .

Clearly  $M'$  is a matching in  $I$ , and  $|M'| = 2|M| + (s - |M|) + 2(t - |M|) = s + 2t - |M|$ . It is straightforward to verify that no man in  $X \cup Z$  can belong to a blocking pair of  $M'$ . Now suppose that  $(m_i, w)$  blocks  $M'$  for some  $i$  ( $1 \leq i \leq s$ ) and  $w \in \mathcal{W}$ . Then  $(m_i, y_i) \in M'$ , so that  $w = w_j$  for some  $j$  ( $1 \leq j \leq t$ ) and  $(x_j, w_j) \in M'$ . Thus each of  $m_i$  and  $w_j$  is unmatched in  $M$ , and  $\{m_i, w_j\} \in E$ . Thus  $M \cup \{\{m_i, w_j\}\}$  is a matching in  $G$ , contradicting the maximality of  $M$ . Hence  $M'$  is stable in  $I$ . Also  $s^+(I) \geq s + 2t - |M| = s + 2t - \beta_1^-(G)$ .

Conversely, suppose that  $M'$  is a stable matching in  $I$ , where  $|M'| = s^+(I)$ . For each  $j$  ( $1 \leq j \leq t$ ), either  $(z_j, w_j) \in M'$  or  $(z_j, w'_j) \in M'$ , for otherwise  $(z_j, w_j)$  blocks  $M'$ . Hence

$$M = \left\{ \{m_i, w_j\} : \begin{array}{l} (1 \leq i \leq s) \wedge (1 \leq j \leq t) \wedge \\ ((m_i, w_j) \in M' \vee (m_i, w'_j) \in M') \end{array} \right\}$$

is a matching in  $G$ . Also  $|M'| \leq |M| + (t - |M|) + t + (s - |M|) = s + 2t - |M|$ , for every edge in  $M$  contributes one (man,woman) pair to  $M'$ , and in addition, at most  $(t - |M|)$  men in  $X$  can be matched in  $M'$ , exactly  $t$  men in  $Z$  are matched in  $M'$ , and at most  $(s - |M|)$  women in  $Y$  can be matched in  $M'$  (and everybody who could be matched in  $M'$  has now been counted).

Suppose that  $M$  is not maximal. Then there is some edge  $\{m_i, w_j\}$  in  $G$  such that no edge of  $M$  is incident to either  $m_i$  or  $w_j$ . Thus by definition of  $M$ , either  $m_i$  is unmatched in  $M'$  or  $(m_i, y_i) \in M'$ . Similarly, either (i)  $(x_j, w_j) \in M'$  or  $w_j$  is unmatched, or (ii)  $w'_j$  is unmatched in  $M'$ . In case (i)  $(m_i, w_j)$  blocks  $M'$ , whilst in case (ii)  $(m_i, w'_j)$  blocks  $M'$ , a contradiction. Hence  $M$  is a maximal matching in  $G$ , and  $s^+(I) = |M'| \leq s + 2t - |M| \leq s + 2t - \beta_1^-(G)$ .

Hence  $s^+(I) + \beta_1^-(G) = s + 2t$ . Now  $2t = 3s$ , as  $G$  is the subdivision graph of some cubic graph. Also  $n = s + 2t$  and  $m = 2t$ , where  $n$  is the number of men in  $I$  and  $m$  is the number of edges of  $G$ .

Let  $c_0$  and  $\delta_0$  be the constants given by Theorem 2.1, such that it is NP-hard to distinguish between the cases  $\beta_1^-(G) \leq c_0m$  and  $\beta_1^-(G) > \delta_0c_0m$ . Hence if  $\beta_1^-(G) \leq c_0m$ , then  $s^+(I) \geq c_3n$ , whilst if  $\beta_1^-(G) > \delta_0c_0m$ , then  $s^+(I) < \delta_3c_3n$ , where  $c_3 = \frac{4-3c_0}{4}$  and  $\delta_3 = \frac{4-3\delta_0c_0}{4-3c_0}$ . The result then follows by Theorem 2.1 and Proposition 2.4.  $\square$

We now demonstrate how to modify the proof of Theorem 3.1 in order to establish the hardness of approximating MIN SMTI under the same restrictions.

**Theorem 3.2.** *It is NP-hard to approximate MIN SMTI within  $\delta_4$ , for some  $\delta_4 > 1$ . The result holds even if the preference lists in  $I$  are of constant length, there is at most one tie per list, and the ties occur on one side only.*

*Proof.* The gap-preserving reduction is similar to the one given by the proof of Theorem 3.1, with some small modifications. In the constructed instance  $I$ , the set of men and women no longer includes the persons in  $X \cup Y$ . Any such person is now removed from the preference list of any remaining person in  $I$ . Now each man's preference list is of length at most 6 and each woman's preference list is of length at most 3.

Suppose firstly that  $M$  is a maximal matching in  $G$ , where  $|M| = \beta_1^-(G)$ . The construction of the matching  $M'$  in  $I$  is similar to the previous one; the only difference is as follows.

If  $m_i$  is unmatched in  $M$ , no pair is added to  $M'$ , whilst if  $w_j$  is unmatched in  $M$ , the pair  $(z_j, w_j)$  is added to  $M'$ . It is straightforward to verify that  $M'$  is a stable matching in  $I$  and  $s^-(I) \leq |M'| = t + |M| = t + \beta_1^-(G)$ .

Conversely, suppose that  $M'$  is a stable matching in  $I$ , where  $|M'| = s^-(I)$ . Then using a similar argument to before we may construct a maximal matching  $M$  in  $G$ , where  $s^-(I) = |M'| = t + |M| \geq t + \beta_1^-(G)$ .

Hence  $s^-(I) = t + \beta_1^-(G)$ . Now  $2t = 3s$ , as  $G$  is the subdivision graph of some cubic graph. Also  $n = s + t$  and  $m = 2t$ , where  $n$  is the number of men in  $I$  and  $m$  is the number of edges of  $G$ .

Let  $c_0$  and  $\delta_0$  be the constants given by Theorem 2.1, such that it is NP-hard to distinguish between the cases  $\beta_1^-(G) \leq c_0m$  and  $\beta_1^-(G) > \delta_0c_0m$ . Hence if  $\beta_1^-(G) \leq c_0m$ , then  $s^-(I) \leq c_4n$ , whilst if  $\beta_1^-(G) > \delta_0c_0m$ , then  $s^-(I) < \delta_4c_4n$ , where  $c_4 = \frac{3(1+2c_0)}{5}$  and  $\delta_4 = \frac{1+2\delta_0c_0}{1+2c_0}$ . The result then follows by Theorem 2.1 and Proposition 2.4.  $\square$

It follows immediately from Theorems 3.1 and 3.2 that neither MAX SMTI nor MIN SMTI admits a polynomial-time approximation scheme unless P=NP.

## 4 Approximation algorithm for MAX SMTI and MIN SMTI

As observed earlier, it is shown in [18] that a maximum cardinality stable matching can have size at most twice that of a minimum cardinality stable matching. Hence the obvious polynomial-time algorithm for SMTI – break all ties in an arbitrary way and apply the classical Gale/Shapley algorithm to the resulting instance of SMI – is simultaneously an approximation algorithm for both MAX and MIN SMTI with a performance ratio of 2.

There is no known approximation algorithm for either problem with a stronger performance ratio, even for special cases of the problems in which the ties are restricted to one side, or to the tails of the preference lists. A case of particular interest arises when there is a limit on the number of preference lists that contain ties, and in this section we show that some progress can be made in establishing additional approximation bounds in this setting.

Ideally, in the case of MAX SMTI, one might hope for a bound of the form  $s^+(I)/|M| \leq f(p)$  given an instance  $I$  of SMTI, where  $M$  is a stable matching found by some approximation algorithm (or just any stable matching, found by breaking ties arbitrarily),  $p$  is the proportion of preference lists that contain ties, and  $f(p)$  is a function that decreases to 1 as  $p$  decreases to 0.

However, it is not hard to see that a bound of this form is infeasible. Were such an algorithm to exist, a ‘gap’ argument could be used to show that it could solve instances of MAX SMTI exactly. Given an arbitrary such instance, it could be ‘expanded’ by the addition of new persons, none of whom has a tie in his or her list, and none of whom can be matched in any stable matching. With an appropriate expansion factor, application of the supposed approximation algorithm to this derived instance would solve the original instance exactly.

Instead we derive a bound on the *difference* in size between a maximum (or minimum) cardinality stable matching and an arbitrary stable matching, expressed in terms of the number of preference lists that contain ties. So the usual approximation algorithm – break all ties arbitrarily and apply the Gale/Shapley algorithm – has a performance guarantee, for both MAX SMTI and MIN SMTI, expressible as a difference rather than a ratio. As observed by Garey and Johnson [5, pp.137-138], this form of performance guarantee can reasonably be viewed as being stronger than the more familiar performance ratio form, and there are relatively few NP-hard problems for which approximation algorithms with performance guarantees of this kind are known.

Some additional definitions are necessary before presenting the main results of this section. Let  $M$  and  $M'$  be stable matchings for an instance  $I$  of SMTI. If a person  $p$  strictly prefers his partner in  $M$  to his partner in  $M'$ , or is matched in  $M$  but not in  $M'$ , then we say that  $p$  *strictly prefers  $M$  to  $M'$* . If  $p$  is indifferent between his partners in  $M$  and  $M'$ , or has the same partner in  $M$  as in  $M'$ , or is matched in neither  $M$  nor  $M'$ , then we say that  $p$  is *indifferent between  $M$  and  $M'$* . Define a *tied pair* to be a pair  $(m, w)$  such that  $m$  is in a tie in  $w$ 's list, or  $w$  is in a tie in  $m$ 's list (or both). In what follows,  $tp(M)$  denotes the number of tied pairs in  $M$ , and  $t(I)$  denotes the number of preference lists in  $I$  that contain ties. In general  $tp(M)$  depends on the matching  $M$ , whilst  $t(I)$  is invariant for the given instance  $I$ ; clearly  $tp(M) \leq t(I)$ .<sup>3</sup>

**Lemma 4.1.** *Let  $T$  be a maximum cardinality stable matching for a given instance  $I$  of SMTI. Then if  $M$  is an arbitrary stable matching in  $I$ ,  $|T| \leq |M| + tp(M)$ .*

*Proof.* We construct an undirected graph  $G = G(M, T)$  as follows:  $G$  has a vertex for each person in  $I$ , and two vertices are joined by a blue (respectively red) edge if the corresponding persons are matched in  $T$  but not in  $M$  (respectively in  $M$  but not in  $T$ ). It is clear that the connected components of  $G$  are paths and cycles with edges of alternating colour. Furthermore,  $|T| - |M|$  is at most equal to the number of *blue augmenting paths* in  $G$ , i.e., the number of paths of odd length in which the first and last edges are blue. Further, every such path has at least three edges, since a component that is a path of length one would provide a blocking pair for one of the supposed stable matchings.

We claim that, in every blue augmenting path, at least one of the intermediate vertices represents a person who is indifferent between  $T$  and  $M$ , and is therefore in a tied pair in both  $T$  and  $M$ . This claim, together with the preceding observation, suffices to establish the lemma.

To establish the claim, let  $p_1, q_1, \dots, p_r, q_r$  form a blue augmenting path in  $G$ , for some  $r \geq 2$ . Since  $p_1$  and  $q_r$  are both matched in  $T$  but not in  $M$ , they both strictly prefer  $T$  to  $M$ . Suppose that no person in the path is indifferent between  $T$  and  $M$ . A simple inductive proof starting from  $p_1$  then reveals that  $q_i$  ( $i = 1, 2, \dots, r-1$ ) strictly prefers  $M$  to  $T$ , otherwise  $(p_i, q_i)$  would block  $M$ , and  $p_i$  ( $i = 2, 3, \dots, r$ ) strictly prefers  $T$  to  $M$ , otherwise  $(p_i, q_{i-1})$  would block  $T$ . Thus  $(p_r, q_r)$  blocks  $M$ , a contradiction. Hence at least one of the  $p_i$  ( $2 \leq i \leq r$ ) or  $q_i$  ( $1 \leq i \leq r-1$ ) must be indifferent between  $T$  and  $M$ , as claimed.  $\square$

Since  $tp(M) \leq |M|$ , it follows immediately by Lemma 4.1 that there exists an approximation algorithm for MAX SMTI with performance ratio 2. Using a similar argument to the one employed in the proof of Lemma 4.1, we may deduce that  $|M| \leq |S| + tp(S)$ , where  $S$  is a stable matching of minimum cardinality. Since  $tp(S) \leq |S|$ , it follows immediately that there exists an approximation algorithm for MIN SMTI, also with performance ratio 2. The inequality established by Lemma 4.1 also leads to the following result:

**Theorem 4.2.** *There is an approximation algorithm  $A$  such that, given any instance  $I$  of either MAX SMTI or MIN SMTI,  $A$  finds a stable matching  $M$  in  $I$  satisfying the following inequality:*

$$s^+(I) - t(I) \leq |M| \leq s^-(I) + t(I).$$

*Additionally, we have that  $s^+(I) \leq s^-(I) + t(I)$ .*

---

<sup>3</sup>The results of this section may be extended to the case that preference lists are partially ordered by making the following amendments to two key definitions. In this setting, define a *tied pair* to be a pair  $(m, w)$  such that  $w$  is indifferent between  $m$  and some other man, or  $m$  is indifferent between  $w$  and some other woman (or both). Define  $t(I)$  to be the number of preference lists that are not linearly ordered.

*Proof.* Let  $M$  be defined as in Lemma 4.1. Since  $tp(M) \leq t(I)$ , Lemma 4.1 implies that  $s^+(I) - t(I) \leq |M| \leq s^+(I)$ . Also by Lemma 4.1,  $s^+(I) \leq s^-(I) + t(I)$ , and hence the result follows.  $\square$

We remark that, when the ties in a given instance  $I$  of SMTI are sparse, i.e.  $t(I)$  is small compared to the numbers of men and women in  $I$ , the performance guarantee indicated by Theorem 4.2 is a significant improvement on the best-known previous result, namely the 2-approximation algorithm for each of MAX SMTI and MIN SMTI.

The following instance is an illustration of the worst case for the above theorem. For each  $n \geq 1$ , we define an SMTI instance  $I$  with  $2n$  men, namely  $\{p_1, \dots, p_n, q_1, \dots, q_n\}$ , and  $2n$  women, namely  $\{r_1, \dots, r_n, s_1, \dots, s_n\}$ . For each  $i$  ( $1 \leq i \leq n$ ), define preference lists for  $p_i, q_i, r_i, s_i$  as follows:

$$\begin{array}{ll} p_i : & s_i \quad r_i \\ q_i : & s_i \end{array} \quad \begin{array}{ll} r_i : & p_i \\ s_i : & (p_i \quad q_i) \end{array}$$

There is a stable matching of size  $n$  (namely  $M_1 = \{(p_i, s_i) : 1 \leq i \leq n\}$ ) and one of size  $2n$  (namely  $M_2 = \{(p_i, r_i), (q_i, s_i) : 1 \leq i \leq n\}$ ). Clearly  $s^+(I) = 2n$ , and also  $s^-(I) = n$  since  $|M_2| = 2|M_1|$ . Since the difference between  $s^+(I)$  and  $s^-(I)$  is the number of lists with ties, the bounds given by Theorem 4.2 are tight.

## 5 “Fair” stable matchings in SMT

In this section we give  $\Omega(n)$  lower bounds for the approximability of EGALITARIAN SMT, MINIMUM REGRET SMT and SEX-EQUAL SMT in an instance of SMT with  $n$  men and  $n$  women. We begin by considering EGALITARIAN SMT. Note that, for any matching  $M$  in such an instance of SMT, it follows that  $2n \leq c(M) \leq 2n^2$ . Hence an approximation algorithm with performance guarantee  $n$  is trivial. Our inapproximability result is therefore optimal within a constant factor.

**Theorem 5.1.** *It is NP-hard to approximate EGALITARIAN SMT within  $\delta n$ , for some  $\delta > 0$ , where  $n$  is the number of men in a given SMT instance.*

*Proof.* We give a reduction from an instance  $I$  of MAX SMTI as constructed by the proof of Theorem 3.1. One property of  $I$  is that there exists a constant  $d$  such that the length of each preference list in  $I$  is at most  $d$ . Let  $c_3$  and  $\delta_3$  be the constants given by Theorem 3.1, such that it is NP-hard to distinguish the cases  $s^+(I) \geq c_3 n$  and  $s^+(I) < \delta_3 c_3 n$ , where  $n$  is the number of men in  $I$ .

Let  $X = \{m_1, m_2, \dots, m_n\}$  be the set of men in  $I$  and let  $Y = \{w_1, w_2, \dots, w_n\}$  be the set of women of  $I$ . For each  $i$  ( $1 \leq i \leq n$ ), let  $P_i$  and  $Q_i$  denote the preference lists of  $m_i$  and  $w_i$  in  $I$  respectively. We call the women in  $P_i$  *proper women* for  $m_i$ , and we call the men in  $Q_i$  *proper men* for  $w_i$ .

We transform  $I$  into an instance  $I'$  of EGALITARIAN SMT as follows. Let  $U = X \cup X'$  and  $W = Y \cup Y'$  be the sets of men and women in  $I'$  respectively, where  $X' = \{m'_1, m'_2, \dots, m'_{(1-c_3)n}\}$  and  $Y' = \{w'_1, w'_2, \dots, w'_{(1-c_3)n}\}$ . The preference lists in  $I'$  are constructed as follows:

$$\begin{array}{lll} m_i : & P_i & (Y') \quad [Y \setminus P_i] \\ m'_i : & (W) & (1 \leq i \leq (1 - c_3)n) \\ w_i : & Q_i & (X') \quad [X \setminus Q_i] \\ w'_i : & (U) & (1 \leq i \leq (1 - c_3)n) \end{array}$$

Note that, in a given person's preference list, persons within square brackets are listed in arbitrary strict order where the symbol appears.

Suppose firstly that  $I$  has a stable matching  $M$  such that  $|M| \geq c_3n$ . Then there is a set  $X_u \subseteq X$  of men who are unmatched in  $M$ , where  $|X_u| \leq (1 - c_3)n$ . Similarly there is a set  $Y_u \subseteq Y$  of women who are unmatched in  $M$ , where  $|Y_u| \leq (1 - c_3)n$ . Let  $M_1$  be a matching that assigns each man in  $X_u$  to a woman in  $Y'$ , and let  $M_2$  be a matching that assigns each woman in  $Y_u$  to a man in  $X'$ . Now let  $M_3$  be a perfect matching of the remaining unmatched members of  $X'$  and  $Y'$ . Finally let  $M' = M \cup M_1 \cup M_2 \cup M_3$ . It may be verified that  $M'$  is a stable matching in  $I'$ , and

$$\begin{aligned} c(M') &\leq 2n(d+1) + 2(1 - c_3)n \\ &\leq 2n(d+2). \end{aligned}$$

On the other hand, suppose  $s^+(I) < \delta_3 c_3 n$ . Now let  $M'$  be any stable matching in  $I'$ . Then  $< \delta_3 c_3 n$  men in  $X$  are matched in  $M'$  to one of their proper women. Now at most  $(1 - c_3)n$  of the remaining men in  $X$  can be matched to a woman in  $Y'$ . Hence there are  $> c_3 n(1 - \delta_3)$  men  $u$  in  $X$  such that  $c_{M'}(u) > (1 - c_3)n$ . Similarly there are  $> c_3 n(1 - \delta_3)$  women  $w$  in  $Y$  such that  $c_{M'}(w) > (1 - c_3)n$ . Hence  $c(M') > 2\varepsilon n^2$ , where  $\varepsilon = c_3(1 - c_3)(1 - \delta_3)$ .

Therefore by Theorem 3.1, it is NP-hard to approximate EGALITARIAN SMT within  $\frac{\varepsilon}{d+2}n$ .  $\square$

We now consider MINIMUM REGRET SMT. Note that, for any matching  $M$  in an instance of SMT with  $n$  men and  $n$  women, it follows that  $1 \leq r(M) \leq n$ . Hence an approximation algorithm with performance guarantee  $n$  is trivial. Therefore again, the  $\Omega(n)$  lower bound that we establish is optimal within a constant factor.

**Theorem 5.2.** *It is NP-hard to approximate MINIMUM REGRET SMT within  $\delta n$ , for some  $\delta > 0$ , where  $n$  is the number of men in a given SMT instance.*

*Proof.* We use the same reduction as described in the proof of Theorem 5.1. Let  $I$ ,  $I'$ ,  $n$ ,  $c_3$ ,  $\delta_3$  and  $d$  be as above. If  $s^+(I) \geq c_3n$ , then  $I'$  has a stable matching  $M'$  such that  $r(M') \leq d+1$ . On the other hand, if  $s^+(I) < \delta_3 c_3 n$  then in any stable matching  $M'$  in  $I'$ , at least one man  $u \in X$  satisfies  $c_{M'}(u) > (1 - c_3)n$ . Hence  $r(M') > (1 - c_3)n$ . Therefore by Theorem 3.1, it is NP-hard to approximate MINIMUM REGRET SMT within  $\frac{1-c_3}{d+1}n$ .  $\square$

The final problem that we consider in this section is SEX-EQUAL SMT. We establish an inapproximability result for this problem similar to those of Theorems 5.1 and 5.2.

**Theorem 5.3.** *It is NP-hard to approximate SEX-EQUAL SMT within  $\delta n$ , for some  $\delta > 0$ , where  $n$  is the number of men in a given SMT instance.*

*Proof.* We formulate a reduction similar to the one described in the proof of Theorem 5.1. Let  $I$ ,  $X$ ,  $X'$ ,  $Y$ ,  $Y'$ ,  $P_i$ ,  $Q_i$ ,  $n$ ,  $c_3$ ,  $\delta_3$  and  $d$  be as above. We transform  $I$  into an instance  $I'$  of SEX-EQUAL SMT as follows. Let  $U = X \cup X' \cup S$  and  $W = Y \cup Y' \cup T$  be the sets of men and women in  $I'$  respectively, where  $S = \{s_1, s_2, \dots, s_d\}$  and  $T = \{t_1, t_2, \dots, t_d\}$ . The preference lists in  $I'$  are constructed as follows:

$$\begin{aligned} m_i &: P_i \quad (W \setminus P_i) && (1 \leq i \leq n) \\ m'_i &: (W) && (1 \leq i \leq (1 - c_3)n) \\ s_i &: t_i \quad [W \setminus \{t_i\}] && (1 \leq i \leq d) \\ w_i &: [S] \quad Q_i \quad (X') \quad [X \setminus Q_i] && (1 \leq i \leq n) \\ w'_i &: (U) && (1 \leq i \leq (1 - c_3)n) \\ t_i &: s_i \quad [U \setminus \{s_i\}] && (1 \leq i \leq d) \end{aligned}$$

Clearly in any stable matching  $M'$  in  $I'$ ,  $(s_i, t_i) \in M'$ .

Suppose firstly that  $I$  has a stable matching  $M$  such that  $|M| \geq c_3 n$ . Then we may form  $M'$  as in the proof of Theorem 5.1. Add  $(s_i, t_i)$  to  $M'$  ( $1 \leq i \leq d$ ). It may be verified that  $M'$  is stable in  $I'$ . Also the total cost of  $M'$  for the men is at most  $(d+1)n + (1 - c_3)n + d$ . Similarly the total cost of  $M'$  for the women is at most  $(2d+1)n + (1 - c_3)n + d$ . Hence  $d(M') = |\sum_{u \in U} c_{M'}(u) - \sum_{w \in W} c_{M'}(w)| = |\sum_{u \in X} c_{M'}(u) - \sum_{w \in Y} c_{M'}(w)| \leq \sum_{u \in X} c_{M'}(u) + \sum_{w \in Y} c_{M'}(w) = (3d+2)n$ .

On the other hand, suppose that  $s^+(I) < \delta_3 c_3 n$ . Now let  $M'$  be any stable matching in  $I'$ . As in the previous paragraph, the total cost of  $M'$  for the men is at most  $(d+1)n + (1 - c_3)n + d$ . No woman  $w \in Y$  is matched in  $M'$  to a man in  $S$ , so  $c_{M'}(w) \geq d+1$ . As in the proof of Theorem 5.1, there are  $> c_3 n(1 - \delta_3)$  women  $w$  in  $Y$  such that  $c_{M'}(w) \geq (d+1) + (1 - c_3)n$ . Hence the total cost of  $M'$  for the women is more than

$$(d+1)n + c_3 n(1 - \delta_3)(1 - c_3)n + (1 - c_3)n + d.$$

Thus  $d(M') > \varepsilon n^2$ , where  $\varepsilon$  is as defined in the proof of Theorem 5.1.

Therefore by Theorem 3.1, it is NP-hard to approximate SEX-EQUAL SMT within  $\frac{\varepsilon}{3d+2}n$ .  $\square$

## 6 Concluding remarks

It is interesting to note that the hardness results proved in this paper for approximating both MAX SMTI and MIN SMTI hold for identical restrictions on the positions of ties – there are relatively few examples in the literature of optimization problems having both maximization and minimization versions that are hard to approximate, and fewer still where this property holds for the same restrictions on the instance.

It remains open as to whether there exists an approximation algorithm for either MAX SMTI or MIN SMTI having performance ratio less than 2. However the progress made in this paper indicates that improvements can be obtained when ties are restricted in number. One might hope for further progress when there are additional constraints in place – on the positions and lengths of ties, for example.

## Acknowledgement

We would like to thank the referees for their comments, which have helped to improve the presentation of this paper.

## References

- [1] Canadian Resident Matching Service. How the matching algorithm works. Web document available at <http://www.carms.ca/matching/algorith.htm>.
- [2] T. Feder. A new fixed point approach for stable networks and stable marriages. *Journal of Computer and System Sciences*, 45:233–284, 1992.
- [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.J. Garey and D.S. Johnson. *Computers and Intractability*. Freeman, 1979.

- [6] D. Gusfield and R.W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, 1989.
- [7] D. Gusfield. Three fast algorithms for four problems in stable marriage. *SIAM Journal on Computing*, 16(1):111–128, 1987.
- [8] 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-Verlag, 2002.
- [9] M. Halldórsson and K. Yoshihara. Greedy approximations of independent sets in low degree graphs. In *Proceedings of ISAAC '95: the 6th International Symposium on Algorithms and Computation*, volume 1004 of *Lecture Notes in Computer Science*, pages 152–161. Springer-Verlag, 1995.
- [10] J.D. Horton and K. Kilakos. Minimum edge dominating sets. *SIAM Journal on Discrete Mathematics*, 6:375–387, 1993.
- [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-Verlag, 1998.
- [13] R.W. Irving, P. Leather, and D. Gusfield. An efficient algorithm for the “optimal” stable marriage. *Journal of the Association for Computing Machinery*, 34(3):532–543, 1987.
- [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-Verlag, 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-Verlag, 1999.
- [16] A. Kato. Complexity of the sex-equal stable marriage problem. *Japan Journal of Industrial and Applied Mathematics*, 10:1–19, 1993.
- [17] D.E. Knuth. *Stable Marriage and its Relation to Other Combinatorial Problems*, volume 10 of *CRM Proceedings and Lecture Notes*. American Mathematical Society, 1997. English translation of *Mariages Stables*, Les Presses de L’Université de Montréal, 1976.
- [18] 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.
- [19] E. Ronn. NP-complete stable matching problems. *Journal of Algorithms*, 11:285–304, 1990.
- [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] A.E. Roth and M.A.O. Sotomayor. *Two-sided matching: a study in game-theoretic modeling and analysis*, volume 18 of *Econometric Society Monographs*. Cambridge University Press, 1990.
- [22] V.V. Vazirani. *Approximation Algorithms*. Springer-Verlag, 2001.