Size Versus Stability in the Marriage Problem

Péter Biró, David F. Manlove and Shubham Mittal

Department of Computing Science
University of Glasgow
Glasgow G12 8QQ
UK

Technical Report
TR-2009-297
April 2009

Size Versus Stability in the Marriage Problem*

Péter Biró[1,†], David F. Manlove[1,‡] and Shubham Mittal[2,§]

[1] *Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK. Email:* `{pbiro,davidm}@dcs.gla.ac.uk`.

[2] *Department of Computer Science and Engineering, Block VI, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi 110 016, India. Email:* `cs1040189@cse.iitd.ernet.in`.

Abstract

Given an instance I of the classical Stable Marriage problem with Incomplete preference lists (SMI), a maximum cardinality matching can be larger than a stable matching. In many large-scale applications of SMI, we seek to match as many agents as possible. This motivates the problem of finding a maximum cardinality matching in I that admits the smallest number of blocking pairs (so is "as stable as possible"). We show that this problem is NP-hard and not approximable within $n^{1-\varepsilon}$, for any $\varepsilon > 0$, unless P=NP, where n is the number of men in I. Further, even if all preference lists are of length at most 3, we show that the problem remains NP-hard and not approximable within δ, for some $\delta > 1$. By contrast, we give a polynomial-time algorithm for the case where the preference lists of one sex are of length at most 2. We also extend these results to the cases where (i) preference lists may include ties, and (ii) we seek to minimise the number of agents involved in a blocking pair.

1 Introduction

The Stable Marriage problem (SM) was introduced in the seminal paper of Gale and Shapley [6]. In its classical form, an instance of SM involves n men and n women (whom we collectively refer to as the *agents*), 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 agent 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 $M(m) = w$, $M(w) = m$.

An agent 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 [6] 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 [11].

*A preliminary version of this paper appeared in the Proceedings of WAOA 2008 (the 6th Workshop on Approximation and Online Algorithms).

†Supported by EPSRC grant EP/E011993/1 and by OTKA grant K69027.

‡Supported by EPSRC grants EP/E011993/1 and GR/R84597/01, and by an RSE / Scottish Executive Personal Research Fellowship.

§Supported by a vacation scholarship from the Department of Computing Science, University of Glasgow.

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 agent'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 agent 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 [7] showed that every stable matching for a given SMI instance has the same size and matches exactly the same set of agents.

Motivation. The Hospitals/Residents problem (HR) is a many-to-one generalisation of SMI, so called because of its applications in centralised matching schemes that handle the allocation of graduating medical students, or *residents*, to hospitals [19]. The largest such scheme is the National Resident Matching Program (NRMP) [25] in the US, but similar schemes exist in Canada [24], in Scotland [10, 26], and in a variety of other countries and contexts.

In the 2006-07 run of the Scottish medical matching scheme, called the Scottish Foundation Allocation Scheme (SFAS), there were 781 students and 53 hospitals, with total capacity 789. The matching algorithm (designed and implemented at the Department of Computing Science, University of Glasgow) found a stable matching of size 744, thus leaving 37 students unmatched. Clearly stability is the key property to be satisfied, and it is this that restricts the size of the resultant matching. Nevertheless the administrators asked whether, were the stability criterion to have been relaxed, a larger matching could have been found. We found that a matching of size 781 did exist, but the matching we computed admitted 400 blocking pairs.

"Almost stable" maximum matchings. In practical situations, a blocking pair of a given matching M need not always lead to M being undermined, since the agents involved might be unaware of their potential to improve relative to M. For example, in situations where preference lists are not public knowledge, there may be limited channels of communication that would lead to the awareness of blocking pairs in practice. Nevertheless, it is reasonable to assert that the greater the number of blocking pairs of a given matching M, the greater the likelihood that M would be undermined by a pair of agents in practice. In particular, a maximum cardinality matching (henceforth a maximum matching) for the 2006-07 SFAS data that admits only 10 blocking pairs might be considered to be "more stable" than one with 400 blocking pairs. This motivates the problem of finding a maximum matching that admits the smallest number of blocking pairs (and is therefore, in the sense described above, "as stable as possible"). Eriksson and Häggström [5] also argue that counting the number of blocking pairs of a matching can be an effective way to measure its degree of instability; earlier, this approach had already been taken by Khuller *et al.* [13]. An alternative approach is to count the number of agents who are involved in a blocking pair [22, 5].

Further applications. Further practical applications of "almost stable" maximum matchings arise in similar bipartite settings, where the size of the matching may be considered to be a higher priority than its stability in a particular matching market. Examples include school placement [1] and the allocation of students to projects in a university department [3]. Furthermore, the US Navy has a bipartite matching problem involving the assignment of sailors to billets [17, 23] in which every sailor should be matched to a billet, and meanwhile there are some critical billets that cannot be left vacant.

In non-bipartite contexts, applications arise in kidney exchange settings [21, 27], for example. Here, both the size and the stability of a matching have been considered as being the most important criteria. Centralised programs have been organised in many countries to match incompatible patient-donor pairs, including the US, the Netherlands and the UK. In most programs, the main goal is to maximise the number of transplants (i.e., the first priority is to find a maximum matching) [21]. However other studies [20] consider stability as the first priority. Another example in a non-bipartite setting involves pairing up chess players [14].

Our results. In this paper we present a range of algorithmic results for two problems, namely MAX SIZE MIN BP SMI and MAX SIZE MIN BA SMI. MAX SIZE MIN BP SMI (respectively MAX SIZE MIN BA SMI) is the problem of finding a maximum matching with the smallest number of blocking pairs (respectively *blocking agents*), given an instance of SMI, where an agent is *blocking* if he/she is a member of at least one blocking pair. We firstly show in Section 2 that both problems are NP-hard and not approximable within $n^{1-\varepsilon}$, for any $\varepsilon > 0$, unless P=NP. We then consider special cases of the problems where the preference lists on one or both sides are short (this is motivated in practice by applications such as SFAS, where students are asked to rank six hospitals in order of preference). We show in Section 3 that, even when preference lists on both sides are of length at most 3, each of MAX SIZE MIN BP SMI and MAX SIZE MIN BA SMI is NP-hard and not approximable within δ, for some $\delta > 1$, unless P=NP. On the other hand, for the case where the lists on one side are of length at most 2 (and the lists on the other side are unbounded in length), in Section 4, we give a polynomial-time algorithm for MAX SIZE MIN BP SMI. We show how to modify this algorithm to the case where preference lists may include ties and/or we wish to find a maximum matching with the minimum number of blocking agents, rather than blocking pairs. We remark that ties arise naturally in practice: for example a large hospital with many applicants may be indifferent between those in certain groups. Finally, Section 5 contains concluding remarks.

Related work. Matchings with few blocking pairs have previously been studied from an algorithmic point of view in the context of the Stable Roommates problem (SR), a non-bipartite generalisation of SM, as a means of coping with the fact that, in contrast to the case for SM, an SR instance need not admit a stable matching. Abraham et al. [2] showed that, given an SR instance, the problem of finding a matching with the smallest number of blocking pairs is NP-hard and not approximable within $n^{1/2-\varepsilon}$, for any $\varepsilon > 0$, unless P=NP. In the case that preference lists include ties, the lower bound was strengthened to $n^{1-\varepsilon}$. On the other hand, given a fixed integer K, they showed that the problem of finding a matching with exactly K blocking pairs, or reporting that no such matching exists, is solvable in polynomial time. This paper can be viewed as a counterpart of [2], strengthening its results by moving to the bipartite setting, and answering the remaining previously open questions in a table shown in Section 5.

2 Unbounded length preference lists

Before presenting the main result of this section, we define some notation and terminology relating to matchings and graphs. Given an instance I of SMI, let $\mathcal{M}$ denote the set of matchings in I and let $\mathcal{M}^+$ denote the set of maximum matchings in I. Given a matching $M \in \mathcal{M}$, let $bp_I(M)$ (respectively $ba_I(M)$) denote the set of blocking pairs (respectively blocking agents) with respect to M in I (we omit the subscript when the instance is clear from the context). Let $bp^+(I) = \min\{|bp_I(M)| : M \in \mathcal{M}^+\}$ and let $ba^+(I) = \min\{|ba_I(M)| : M \in \mathcal{M}^+\}$. Define MAX SIZE MIN BP SMI (respectively MAX SIZE MIN BA SMI) to be the problem of finding, given an SMI instance I, a matching $M \in \mathcal{M}^+$ such that $|bp_I(M)| = bp^+(I)$ (respectively $|ba_I(M)| = ba^+(I)$).

Given a graph G, the *subdivision graph* of G, denoted by $S(G)$, is a bipartite graph 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. A matching M in a graph G is said to be *maximal* if no proper superset of M is a matching in G. Let $\beta(G)$ denote the size of a maximum matching in G. Define EXACT-MM to be the problem of deciding, given a graph G and integer K, whether G admits a maximal matching of size exactly K. EXACT-MM is NP-complete, even for subdivision graphs of cubic graphs [16, Lemma 2.2.1]. We now present a gap-introducing reduction from EXACT-MM to MAX SIZE MIN BP SMI.

Theorem 1. MAX SIZE MIN BP SMI *is not approximable within $n^{1-\varepsilon}$, where n is the number of men in a given instance, for any $\varepsilon > 0$, unless P=NP.*

Proof. Let $\varepsilon > 0$ be given. We transform from EXACT-MM restricted to subdivision graphs of cubic graphs, which is NP-complete as noted above. Hence let $G = (V, E)$ (a subdivision graph of some cubic graph G') and K (a positive integer) be an instance of EXACT-MM. Then G is a bipartite graph, and V is a disjoint union of two sets U and W, where each edge $e \in E$ joins a vertex in U to a vertex in W. Let $m = |E|$. We lose no generality by assuming that $K \le \beta(G) \le \min\{|U|, |W|\}$. Suppose that $U = \{u_1, u_2, \ldots, u_{n_1}\}$ and $W = \{w_1, w_2, \ldots, w_{n_2}\}$. Without loss of generality assume that each vertex in U has degree 2 and each vertex in W has degree 3. For each $u_i \in U$, let w_{p_i} and w_{q_i} be the two neighbours of u_i in G, where $p_i < q_i$. Also, for each $w_j \in W$, let u_{r_j}, u_{s_j} and u_{t_j}, be the three neighbours of w_j, where $r_j < s_j < t_j$.

Let $B = \lceil \frac{3}{\varepsilon} \rceil$ and let $C = (n_1 + n_2)^{B+1} - (n_1 + n_2) + 1$. We create an instance I of SMI as follows. The sets of men and women in I are denoted by $\mathcal{U}$ and $\mathcal{W}$ respectively, where $\mathcal{U}$ and $\mathcal{W}$ are as defined in Figure 1. It follows that $|\mathcal{U}| = |\mathcal{W}| = 3n_1 + 4n_2 + 2mC - K$. Let $U^1 = \{u_i^1 : 1 \le i \le n_1\}$ and let $W^1 = \{w_j^1 : 1 \le j \le n_2\}$.

For each $u_i \in U$ and $w_j \in W$ such that $\{u_i, w_j\} \in E$, define $\sigma_{j,i} = 1$ if $w_j = w_{p_i}$ and $\sigma_{j,i} = 2$ if $w_j = w_{q_i}$, and define $\tau_{i,j} = 1$ if $u_i = u_{r_j}$, $\tau_{i,j} = 2$ if $u_i = u_{s_j}$ and $\tau_{i,j} = 3$ if $u_i = u_{t_j}$.

Preference lists for the men and women in I are as shown in Figure 2. In a given agent's preference list, the symbol $[S]$ denotes all members of the set S listed in some arbitrary strict order at the point where the symbol appears, and the symbol $[[S]]$ denotes all members of S listed in increasing subscript order at the point where the symbol appears.

We now give some intuition behind this construction. Suppose that M is a maximal matching of size K in G. For each $\{u_i, w_j\} \in M$, the relevant pair in $U_i \times W_j$ (who rank each other in second place) will be added to a matching M' in I. The $n_1 - K$ men in U (respectively $n_2 - K$ women in W) who are unmatched in M are collectively matched in M' to the women in Y (respectively men in X). The remaining members of U_i (for each $u_i \in U$) and W_j (for each $w_j \in W$) are collectively matched in M' to the members of Z_i and V_j respectively. Each of $U \times Z_i$ and $V_j \times W$ contributes one blocking pair to M'. It is then possible to extend M' to a perfect matching in I without introducing any additional blocking pairs by adding a perfect matching between the members of $G_{i,j} \cup H_{i,j}$ for each $\{u_i, w_j\} \in E$. Hence $|bp(M')| = n_1 + n_2$. Conversely, from a perfect matching M' in I, it is straightforward to extract a matching M in G of size K. If M is not maximal then

$$
\begin{array}{lcll}
\mathcal{U} & = & (\cup_{i=1}^{n_1} U_i) \cup \left(\cup_{\{u_i,w_j\}\in E} G_{i,j}\right) \cup (\cup_{i=1}^{n_2} V_i) \cup X & \\
\mathcal{W} & = & \left(\cup_{j=1}^{n_2} W_j\right) \cup \left(\cup_{\{u_i,w_j\}\in E} H_{i,j}\right) \cup \left(\cup_{j=1}^{n_1} Z_j\right) \cup Y & \\
G_{i,j} & = & G_{i,j}^1 \cup G_{i,j}^2 & (\{u_i,w_j\}\in E) \\
G_{i,j}^d & = & \{g_{i,j}^{c,d} : 1 \le c \le C\} & (\{u_i,w_j\}\in E \wedge 1 \le d \le 2) \\
H_{i,j} & = & H_{i,j}^1 \cup H_{i,j}^2 & (\{u_i,w_j\}\in E) \\
H_{i,j}^d & = & \{h_{i,j}^{c,d} : 1 \le c \le C\} & (\{u_i,w_j\}\in E \wedge 1 \le d \le 2) \\
U_i & = & \{u_i^1, u_i^2, u_i^3\} & (1 \le i \le n_1) \\
V_i & = & \{v_i^1, v_i^2, v_i^3\} & (1 \le i \le n_2) \\
W_j & = & \{w_j^1, w_j^2, w_j^3, w_j^4\} & (1 \le j \le n_2) \\
X & = & \{x_i : 1 \le i \le n_2 - K\} & \\
Y & = & \{y_j : 1 \le j \le n_1 - K\} & \\
Z_j & = & \{z_j^1, z_j^2\} & (1 \le j \le n_1).
\end{array}
$$

Figure 1: Men and women in the constructed instance of MAX SIZE MIN BP SMI.

$$
\begin{array}{lll}
u_i^1 : & z_i^1 \;\; w_{p_i}^{\tau_{i,p_i}} \;\; [H_{i,p_i}^1] \;\; [H_{i,q_i}^1] \;\; [[Y]] & (1 \le i \le n_1) \\
u_i^2 : & z_i^2 \;\; w_{q_i}^{\tau_{i,q_i}} & (1 \le i \le n_1) \\
u_i^3 : & z_i^1 \;\; z_i^2 & (1 \le i \le n_1) \\
g_{i,j}^{c,1} : & h_{i,j}^{c,1} \;\; w_j^1 \;\; h_{i,j}^{c,2} & (\{u_i,w_j\}\in E \wedge 1 \le c \le C) \\
g_{i,j}^{c,2} : & h_{i,j}^{c,2} \;\; h_{i,j}^{c,1} & (\{u_i,w_j\}\in E \wedge 1 \le c \le C) \\
v_i^1 : & w_i^1 \;\; w_i^4 & (1 \le i \le n_2) \\
v_i^2 : & w_i^2 \;\; w_i^4 & (1 \le i \le n_2) \\
v_i^3 : & w_i^3 \;\; w_i^4 & (1 \le i \le n_2) \\
x_i : & [[W^1]] & (1 \le i \le n_2 - K) \\
 & & \\
w_j^1 : & v_j^1 \;\; u_{r_j}^{\sigma_{j,r_j}} \;\; [G_{j,r_j}^1] \;\; [G_{j,s_j}^1] \;\; [G_{j,t_j}^1] \;\; [[X]] & (1 \le j \le n_2) \\
w_j^2 : & v_j^2 \;\; u_{s_j}^{\sigma_{j,s_j}} & (1 \le j \le n_2) \\
w_j^3 : & v_j^3 \;\; u_{t_j}^{\sigma_{j,t_j}} & (1 \le j \le n_2) \\
w_j^4 : & v_j^1 \;\; v_j^2 \;\; v_j^3 & (1 \le j \le n_2) \\
h_{i,j}^{c,1} : & g_{i,j}^{c,2} \;\; u_i^1 \;\; g_{i,j}^{c,1} & (\{u_i,w_j\}\in E \wedge 1 \le c \le C) \\
h_{i,j}^{c,2} : & g_{i,j}^{c,1} \;\; g_{i,j}^{c,2} & (\{u_i,w_j\}\in E \wedge 1 \le c \le C) \\
z_j^1 : & u_j^1 \;\; u_j^3 & (1 \le j \le n_1) \\
z_j^2 : & u_j^2 \;\; u_j^3 & (1 \le j \le n_1) \\
y_j : & [[U^1]] & (1 \le j \le n_1 - K)
\end{array}
$$

Figure 2: Preference lists in the constructed instance of MAX SIZE MIN BP SMI.

there is some $u_i \in U$ and $w_j \in W$, both unmatched in M, such that $\{u_i, w_j\} \in E$. In this case, for each c ($1 \le c \le C$), either $(u_i^1, h_{i,j}^{c,1}) \in bp(M')$ or $(g_{i,j}^{c,1}, w_j^1) \in bp(M')$, and hence $|bp(M')| \ge C$. This introduces the required 'gap' for the inapproximability result.

For the formal argument showing the correctness of the reduction, we firstly show that I admits a perfect matching. As $K \leq \beta(G)$, it follows that G admits a matching M of size K. We form a matching M' in I as follows. Consider each edge $\{u_i, w_j\} \in E$. Suppose firstly that $\{u_i, w_j\} \in M$, where $u_i \in \mathcal{U}$ and $w_j \in \mathcal{W}$. Let $\sigma = 1, 2$ according as w_j is w_{p_i} or w_{q_i} respectively, and let $\tau = 1, 2, 3$ according as u_i is u_{r_j}, u_{s_j} or u_{t_j} respectively. Add (u_i^σ, w_j^τ), $(u_i^{3-\sigma}, z_i^{3-\sigma})$ and (u_i^3, z_i^σ) to M'. Also, for each $w_j^c \in \{w_j^1, w_j^2, w_j^3\} \backslash \{w_j^\tau\}$, add (v_j^c, w_j^c) to M'. Next add (v_j^τ, w_j^4) to M'. Finally, for each c $(1 \leq c \leq C)$, add $(g_{i,j}^{c,1}, h_{i,j}^{c,1})$ and $(g_{i,j}^{c,2}, h_{i,j}^{c,2})$ to M'. Now suppose that $\{u_i, w_j\} \notin M$. If u_i is matched in M, add $(g_{i,j}^{c,1}, h_{i,j}^{c,1})$ and $(g_{i,j}^{c,2}, h_{i,j}^{c,2})$ to M' for each c $(1 \leq c \leq C)$. Otherwise add $(g_{i,j}^{c,1}, h_{i,j}^{c,2})$ and $(g_{i,j}^{c,2}, h_{i,j}^{c,1})$ to M' for each c $(1 \leq c \leq C)$. There are $n_1 - K$ vertices in U that are unmatched in M – denote these vertices by $u_{a_1}, u_{a_2}, \ldots, u_{a_{n_1-K}}$, where $a_1 < a_2 < \ldots < a_{n_1-K}$. Add $(u_{a_i}^1, y_i)$, $(u_{a_i}^2, z_{a_i}^2)$ and $(u_{a_i}^3, z_{a_i}^1)$ to M' $(1 \leq i \leq n_1 - K)$. Similarly there remain $n_2 - K$ vertices in W that are unmatched in M – denote these vertices by $w_{b_1}, w_{b_2}, \ldots, w_{b_{n_2-K}}$, where $b_1 < b_2 < \ldots < b_{n_2-K}$. Add $(x_j, w_{b_j}^1)$, $(v_{b_j}^2, w_{b_j}^2)$, $(v_{b_j}^3, w_{b_j}^3)$ and $(v_{b_j}^1, w_{b_j}^4)$ to M' $(1 \leq j \leq n_2 - K)$. It may be verified that M' is a perfect matching in I.

We now show that if G admits a maximal matching of size K, then $bp(I) \leq n_1 + n_2$. For, suppose that G admits a maximal matching M of size K. As above, we form a perfect matching M' in I. We now show that $|bp(M')| = n_1 + n_2$. For, let $1 \leq i \leq n_1$ be given. Exactly one of (u_i^1, z_i^1), (u_i^2, z_i^2) belongs to $bp(M')$. Now suppose that $(u_i^1, y_b) \in M'$ for some $y_b \in Y$. Then u_i is unmatched in M, so that by the maximality of M in G, each of w_{p_i} and w_{q_i} is matched in M. Hence any $h_{i,j}^{c,d}$ whom u_i^1 prefers to y_b is matched in M' to her first-choice partner, so that $(u_i^1, h_{i,j}^{c,d}) \notin bp(M')$. Similarly $w_{p_i}^{\tau_{i,p_i}}$ is matched in M' to her first-choice partner, so that $(u_i^1, w_{p_i}^{\tau_{i,p_i}}) \notin bp(M')$. Also $(u_i^1, y_{b'}) \notin bp(M')$ for any $y_{b'} \in Y$, for otherwise $b' < b$. But by construction of M', it follows that $(u_{i'}^1, y_{b'}) \in M'$ for some $i' < i$, and hence $y_{b'}$ prefers $u_{i'}^1$ to u_i^1, a contradiction.

Now let $\{u_i, w_j\} \in E$. It is straightforward to verify that no pair of the form $(g_{i,j}^{c,d}, h_{i,j}^{c,d'})$ $(1 \leq c \leq C,\ 1 \leq d, d' \leq 2)$ is in $bp(M')$. Now let $1 \leq i \leq n_2$. Exactly one of (v_i^1, w_i^1), (v_i^2, w_i^2), (v_i^3, w_i^3) belongs to $bp(M')$. Finally let $1 \leq i \leq n_2 - K$ and suppose that $(x_i, w_j^1) \in M'$. For any $j' < j$, either $w_{j'}^1$ is matched in M' to her first or second-choice partner, or $(x_{i'}, w_{j'}^1) \in M'$. In the latter case $i' < i$, so that $(x_i, w_{j'}^1) \notin bp(M')$. Hence $|bp(M')| = n_1 + n_2$ as claimed.

We next show that if G admits no maximal matching of size K then $bp(I) > (n_1 + n_2)^{B+1}$. Suppose that G admits no maximal matching of size K. Let M' be an arbitrary perfect matching in I. We claim that $|bp(M')| > (n_1 + n_2)^{B+1}$. Firstly we note that, for each i $(1 \leq i \leq n_1)$, u_i^3 is matched in M' to some z_i^b $(b = 1, 2)$, and thus $(u_i^b, z_i^b) \in bp(M')$. Similarly for each j $(1 \leq j \leq n_2)$, w_j^4 is matched in M' to some v_j^a $(1 \leq a \leq 3)$, and thus $(v_j^a, w_j^a) \in bp(M')$. Hence $|bp(M')| \geq n_1 + n_2$.

Now let

$$M = \left\{ \{u_i, w_j\} \in E : \begin{array}{l} ((u_i^\sigma, w_j^\tau) \in M' \text{ where } 1 \leq \sigma \leq 2 \wedge 1 \leq \tau \leq 3) \vee \\ ((u_i^1, h_{i,j}^{c,1}) \in M' \text{ where } 1 \leq c \leq C) \end{array} \right\}.$$

We claim that M is a matching in G. For if $(u_i^1, w_j^b) \in M'$ and $(u_i^2, w_{j'}^{b'}) \in M'$ then either z_i^1 or z_i^2 is unmatched in M', a contradiction. Similarly if $(u_i^a, w_j^b) \in M'$ and $(u_{i'}^{a'}, w_j^{b'}) \in M'$ for some $b \neq b'$ then at least one of v_j^1, v_j^2 or v_j^3 is unmatched in M', a contradiction. Finally if $(u_i^1, h_{i,j}^{c,1}) \in M'$ for some $u_i^1 \in U$ and $h_{i,j}^{c,1} \in H$, then $(g_{i,j}^{c,2}, h_{i,j}^{c,2}) \in M'$, which in turn forces $(g_{i,j}^{c,1}, w_j^1) \in M'$. Hence the claim is established.

Also $|M| = K$, for the $n_1 - K$ members of Y are collectively matched in M' to $U_Y^1 \subseteq U^1$. Thus $|U^1 \backslash U_Y^1| = K$. Let $u_i^1 \in U^1 \backslash U_Y^1$. Either $(u_i^1, w_j^b) \in M'$ for some b $(1 \leq b \leq 3)$ and j

$(1 \leq j \leq n_2)$ or $(u_i^1, z_i^1) \in M'$. In the latter case $(u_i^2, w_j^b) \in M'$ for some b $(1 \leq b \leq 3)$ and j $(1 \leq j \leq n_2)$, for otherwise u_i^3 is unmatched in M', a contradiction.

Finally by the hypothesis, M is not maximal in G. Hence there exists some $\{u_i, w_j\} \in E$ such that no edge of M is incident to either u_i or w_j. By construction of M it follows that $(u_i^2, z_i^2) \in M'$, which forces $(u_i^3, z_i^1) \in M'$ and $(u_i^1, y_b) \in M'$ for some $y_b \in Y$. Similarly by construction of M it follows that $(v_j^2, w_j^2) \in M'$ and $(v_j^3, w_j^3) \in M'$, which forces $(v_j^1, w_j^4) \in M'$ and $(x_a, w_j^1) \in M'$ for some $x_a \in X$. Now let c $(1 \leq c \leq C)$ be given. If $\{(g_{i,j}^{c,1}, h_{i,j}^{c,1}), (g_{i,j}^{c,2}, h_{i,j}^{c,2})\} \subseteq M'$ then $(u_i^1, h_{i,j}^{c,1}) \in bp(M')$. Otherwise $\{(g_{i,j}^{c,1}, h_{i,j}^{c,2}), (g_{i,j}^{c,2}, h_{i,j}^{c,1})\} \subseteq M'$, so that $(g_{i,j}^{c,1}, w_j^1) \in bp(M')$. Hence $|bp(M')| \geq n_1 + n_2 + C > (n_1 + n_2)^{B+1}$ as claimed.

Hence the existence of a $(n_1+n_2)^B$-approximation algorithm for MAX SIZE MIN BP SMI implies a polynomial-time algorithm for EXACT-MM in subdivision graphs of cubic graphs, a contradiction unless P=NP. We claim that $(n_1 + n_2)^B \geq n^{1-\varepsilon}$. For, we firstly observe that $n = 3n_1 + 2mC + 4n_2 - K$. Now G is the subdivision graph of a cubic graph G', and n_1 is the number of edges in G', so $2n_1 = 3n_2$. Also $m = 2n_1$. It follows that

$$n = 7n_1 + 4n_2 + 4n_1(n_1 + n_2)^{B+1} - 4n_1(n_1 + n_2) - K. \tag{1}$$

From Equation 1, we may deduce that $n \leq 3(n_1 + n_2)^{B+2}$, and hence

$$(n_1 + n_2)^B \geq 3^{-\frac{B}{B+2}} n^{\frac{B}{B+2}}. \tag{2}$$

By hypothesis $K \leq \min\{n_1, n_2\}$, and without loss of generality we may assume that $n_1 \geq 3$; hence Equation 1 also implies that $n \geq 3^B$, and hence $3^{-\frac{B}{B+2}} \geq n^{-\frac{1}{B+2}}$. But $B + 2 \geq \frac{3}{\varepsilon}$, and hence Inequality 2 implies that $(n_1 + n_2)^B \geq n^{1-\varepsilon}$ as required. □

Let MAX SIZE EXACT BP SMI denote the problem of finding, given an SMI instance I and an integer K', a matching $M \in \mathcal{M}^+$ such that $|bp_I(M)| = K'$.

Corollary 2. MAX SIZE EXACT BP SMI *is NP-complete.*

Proof. We use the same reduction as in the proof of Theorem 1 and set $K' = n_1 + n_2$ and $\varepsilon = \infty$ (i.e. $B = 0$ and $C = 1$). As before G has a maximal matching of size K if and only if I admits a perfect matching M' such that $|bp(M')| \leq K'$. However it is straightforward to verify that any perfect matching M' in I satisfies $|bp(M')| \geq K'$, and hence the result follows. □

Given that SMI is a special case of SR, we may reuse results from [2] to obtain the following theorem.

Theorem 3 ([2]). MAX SIZE EXACT BP SMI *is solvable in polynomial time when K' is fixed.*

We now consider MAX SIZE MIN BA SMI. It turns out that a small modification to the proof of Theorem 1 is sufficient to establish the same inapproximability result for this problem.

Theorem 4. MAX SIZE MIN BA SMI *is not approximable within $n^{1-\varepsilon}$, where n is the number of men in a given instance, for any $\varepsilon > 0$, unless P=NP.*

Proof. We use the same reduction as in the proof of Theorem 1, with the single modification that we now set $C = 2(n_1 + n_2)^{B+1} - 2(n_1 + n_2) + 1$. Using a similar argument to that in the proof of Theorem 1, it follows that if G admits a maximal matching of size K

then I admits a perfect matching M' such that $|ba(M')| = 2(n_1 + n_2)$. Conversely if G does not admit a maximal matching of size K then any perfect matching M' in I satisfies $|ba(M')| \geq 2(n_1 + n_2) + C > 2(n_1 + n_2)^{B+1}$.

Hence the existence of a $(n_1 + n_2)^B$-approximation algorithm for MAX SIZE MIN BA SMI implies a polynomial-time algorithm for EXACT-MM in subdivision graphs of cubic graphs, a contradiction unless P=NP. We claim that $(n_1 + n_2)^B \geq n^{1-\varepsilon}$. As in the proof of Theorem 1, we firstly observe that

$$n \quad = \quad 7n_1 + 4n_2 + 8n_1(n_1 + n_2)^{B+1} - 8n_1(n_1 + n_2) - K. \tag{3}$$

From Equation 3, we may deduce that $n \leq 5(n_1 + n_2)^{B+2}$, and hence

$$(n_1 + n_2)^B \quad \geq \quad 5^{-\frac{B}{B+2}} n^{\frac{B}{B+2}}. \tag{4}$$

By hypothesis $K \leq \min\{n_1, n_2\}$, and without loss of generality we may assume that $n_1 \geq 3$, and hence $n_1 + n_2 \geq 5$ since $2n_1 = 3n_2$. Thus Equation 3 also implies that $n \geq 5^B$, and hence $5^{-\frac{B}{B+2}} \geq n^{-\frac{1}{B+2}}$. But $B + 2 \geq \frac{3}{\varepsilon}$, and hence Inequality 4 implies that $(n_1 + n_2)^B \geq n^{1-\varepsilon}$ as required. □

Let MAX SIZE EXACT BA SMI denote the problem of finding, given an SMI instance I and an integer K', a matching $M \in \mathcal{M}^+$ such that $|ba_I(M)| = K'$.

Corollary 5. MAX SIZE EXACT BA SMI *is NP-complete.*

Proof. We use the same reduction as in the proof of Theorem 1 and set $K' = 2(n_1 + n_2)$ and $\varepsilon = \infty$ (i.e. $B = 0$ and $C = 1$). As before G has a maximal matching of size K if and only if I admits a perfect matching M' such that $|ba(M')| \leq K'$. However it is straightforward to verify that any perfect matching M' in I satisfies $|ba(M')| \geq K'$, and hence the result follows. □

We now turn to the case that K' is fixed in the definition of MAX SIZE EXACT BA SMI. In the following theorem, and in Section 4, we use the following terminology. Let I be an SMI instance in which $\mathcal{U}$ is the set of men and $\mathcal{W}$ is the set of women. The *underlying graph* of I is a bipartite graph $G = (V, E)$, where $V = \mathcal{U} \cup \mathcal{W}$ and E is the set of mutually acceptable pairs.

Theorem 6. MAX SIZE EXACT BA SMI *is solvable in polynomial time when K' is fixed.*

Proof. Let I be an instance of SMI with n men and n women, and let m be the total length of the men's preference lists in I. Let G be the underlying graph of I. We generate the $O(n^{K'})$ subsets of size K' of the agents in I. For each such subset S, we then generate the edge covers of the subgraph of G induced by S; there are at most $O(2^{K'^2})$ such subsets. For each such edge cover B we determine whether I admits a matching M satisfying $bp(M) = B$; this can be accomplished in $O(m)$ time [2]. Such a set of blocking pairs involves precisely the agents in S by construction. Overall this algorithm has $O(mn^{K'})$ complexity. □

3 Preference lists of length at most 3

In this section we consider the case where preference lists in a given instance I of SMI are of bounded length. Given two integers p and q, let MAX SIZE MIN BP (p, q)-SMI (respectively MAX SIZE MIN BA (p, q)-SMI) denote the restriction of MAX SIZE MIN BP SMI (respectively MAX SIZE MIN BA SMI) in which each man's preference list is of length at most p, and each

$$
\begin{array}{rlll}
x_{6i}: & y_{6i} \;\; c(x_{6i}) \;\; y_{6i+1} & & (0 \le i \le n-1) \\
x_{6i+1}: & y_{6i+1} \;\; c(x_{6i+1}) \;\; y_{6i+2} & & (0 \le i \le n-1) \\
x_{6i+2}: & y_{6i+3} \;\; c(x_{6i+2}) \;\; y_{6i+2} & & (0 \le i \le n-1) \\
x_{6i+3}: & y_{6i+4} \;\; c(x_{6i+3}) \;\; y_{6i+3} & & (0 \le i \le n-1) \\
x_{6i+4}: & y_{6i+4} \;\; y_{6i+5} & & (0 \le i \le n-1) \\
x_{6i+5}: & y_{6i} \;\; y_{6i+5} & & (0 \le i \le n-1) \\
p_j^r: & w_j^r \;\; c_j^r & & (1 \le j \le m \wedge 1 \le r \le 3) \\
v_j^r: & w_j^r \;\; z_j & & (1 \le j \le m \wedge 1 \le r \le 3) \\
q_j: & c_j^1 \;\; c_j^2 \;\; c_j^3 & & (1 \le j \le m) \\
y_{6i}: & x_{6i+5} \;\; x_{6i} & & (0 \le i \le n-1) \\
y_{6i+1}: & x_{6i} \;\; x_{6i+1} & & (0 \le i \le n-1) \\
y_{6i+2}: & x_{6i+2} \;\; x_{6i+1} & & (0 \le i \le n-1) \\
y_{6i+3}: & x_{6i+3} \;\; x_{6i+2} & & (0 \le i \le n-1) \\
y_{6i+4}: & x_{6i+4} \;\; x_{6i+3} & & (0 \le i \le n-1) \\
y_{6i+5}: & x_{6i+4} \;\; x_{6i+5} & & (0 \le i \le n-1) \\
c_j^r: & p_j^r \;\; x(c_j^r) \;\; q_j & & (1 \le j \le m \wedge 1 \le r \le 3) \\
w_j^r: & v_j^r \;\; p_j^r & & (1 \le j \le m \wedge 1 \le r \le 3) \\
z_j: & v_j^1 \;\; v_j^2 \;\; v_j^3 & & (1 \le j \le m)
\end{array}
$$

Figure 3: Preference lists in the constructed instance of MAX SIZE MIN BP $(3,3)$-SMI.

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 are of unbounded length, respectively.

We begin by showing that MAX SIZE MIN BP $(3,3)$-SMI is NP-hard and not approximable within some $\delta > 1$ unless P=NP. To prove this, we give a reduction from a restricted version of SAT. Given a Boolean formula B in CNF and a truth assignment f, let $t(f)$ denote the number of clauses of B satisfied simultaneously by f, and let $t(B)$ denote the maximum value of $t(f)$, taken over all truth assignments f of B. Let MAX (2,2)-E3-SAT [4] denote the problem of finding, given a Boolean formula B in CNF in which each clause contains exactly 3 literals and each variable occurs exactly twice as an unnegated literal in B and exactly twice as a negated literal in B, a truth assignment f such that $t(f) = t(B)$.

Theorem 7. *Given any ε $(0 < \varepsilon < \frac{1}{2032})$,* MAX SIZE MIN BP $(3,3)$-SMI *is not approximable within $\frac{3557}{3556+2032\varepsilon}$ unless P=NP.*

Proof. Let ε $(0 < \varepsilon < \frac{1}{2032})$ be given. Let B be an instance of MAX (2,2)-E3-SAT. Let $V = \{v_0, v_1, \dots, v_{n-1}\}$ and $C = \{c_1, c_2, \dots, c_m\}$ be the set of variables and clauses in B respectively. Then for each $v_i \in V$, each of literals v_i and $\bar{v}_i$ appears exactly twice in B. Also $|c_j| = 3$ for each $c_j \in C$.

We form an instance I of MAX SIZE MIN BP SMI as follows. The set of men in I is $X \cup P \cup Q \cup V$ and the set of women in I is $Y \cup C' \cup W \cup Z$, where $X = \cup_{i=0}^{n-1} X_i$, $X_i = \{x_{6i+r} : 0 \le r \le 5\}$ $(0 \le i \le n-1)$, $P = \cup_{j=1}^{m} P_j$, $P_j = \{p_j^1, p_j^2, p_j^3\}$ $(1 \le j \le m)$, $Q = \{q_j : c_j \in C\}$, $V = \cup_{j=1}^{m} V_j$, $V_j = \{v_j^1, v_j^2, v_j^3\}$ $(1 \le j \le m)$, $Y = \cup_{i=0}^{n-1} Y_i$, $Y_i = \{y_{6i+r} : 0 \le r \le 5\}$ $(0 \le i \le n-1)$, $C' = \{c_j^r : c_j \in C \wedge 1 \le r \le 3\}$, $W = \cup_{j=1}^{m} W_j$, $W_j = \{w_j^1, w_j^2, w_j^3\}$ $(1 \le j \le m)$, and $Z = \{z_j : c_j \in C\}$.

The preference lists of the men and women in I are shown in Figure 3. In the preference list of an agent $x_{6i+r} \in X$ $(0 \le i \le n-1$ and $r \in \{0,1\})$, the symbol $c(x_{6i+r})$ denotes the woman $c_j^s \in C'$ such that the $(r+1)$th occurrence of v_i appears at position s of c_j.

Similarly if $r \in \{2, 3\}$ then the symbol $c(x_{6i+r})$ denotes the woman $c_j^s \in C'$ such that the $(r-1)$th occurrence of $\bar{v}_i$ 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_{6i+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_{6i+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.

For each i $(0 \leq i \leq n-1)$, let $T_i = \{(x_{6i+r}, y_{6i+r}) : 0 \leq r \leq 5\}$ and $F_i = \{(x_{6i+r}, y_{6i+r+1})\} : 0 \leq r \leq 5\}$, where addition is taken module $6i$. We firstly note that M is a perfect matching of the men and women in I, where

$$M = \bigcup_{i=0}^{n-1} T_i \cup \{(p_j^1, c_j^1), (v_j^1, w_j^1), (p_j^2, c_j^2), (v_j^2, w_j^2), (q_j, c_j^3), (p_j^3, w_j^3), (v_j^3, z_j) : 1 \leq j \leq m\}.$$

We now give some intuition behind this construction. The agents in $X_i \cup Y_i$ correspond to variable $v_i \in V$, whilst the agents in $P_j \cup V_j \cup W_j \cup \{q_j, c_j^1, c_j^2, c_j^3, z_j\}$ correspond to clause $c_j \in C$. The pairs in T_i are added to a matching M in I if $v_i \in V$ is true under a truth assignment f of B, otherwise the pairs in F_i are added to M. Crucially, if v_i is false under f then each of x_{6i} and x_{6i+1} (corresponding to the first and second occurrences of literal v_i) has his third choice in M. Similarly if v_i is true under f then each of x_{6i+2} and x_{6i+3} (corresponding to the first and second occurrences of literal $\bar{v}_i$) has his third choice in M. Hence if any clause c_j is false under f, then since $(q_j, c_j^s) \in M$ for some $s \in \{1, 2, 3\}$, it follows that $(x(c_j^s), c_j^s) \in bp(M)$. Additionally, regardless of the truth values of V under f, the members of $X_i \times Y_i$ contribute one blocking pair for each $v_i \in V$, as do the members of $V_j \times W_j$ for each $c_j \in C$.

For the formal argument showing the correctness of the reduction, we claim that $t(B) + bp^+(I) = n + 2m$.

For, let f be a truth assignment of B such that $t(f) = t(B)$. We create a perfect 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. In the former case, $bp(M) \cap (X_i \times Y_i) = \{(x_{6i+5}, y_{6i})\}$, whilst in the latter case, $bp(M) \cap (X_i \times Y_i) = \{(x_{6i+4}, y_{6i+4})\}$.

Now let $c_j \in C$. If 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, otherwise set $s = 1$. Add the pairs (p_j^t, c_j^t), (v_j^t, w_j^t) $(1 \leq t \neq s \leq 3)$, (q_j, c_j^s), (p_j^s, w_j^s) and (v_j^s, z_j) to M. Then $bp(M) \cap (V_j \times W_j) = \{(v_j^s, w_j^s)\}$. Now if c_j is not satisfied under f then man $x(c_j^1)$ has his last-choice partner, by construction of M. Hence $(x(c_j^1), c_j^1) \in bp(M)$. Moreover these, together with the $n + m$ blocking pairs in $X \times Y$ and $V \times W$ identified already, are all the blocking pairs of M in I. Hence $|bp(M)| = n + m + (m - t(f))$, i.e. $bp^+(I) + t(B) \leq n + 2m$.

Conversely suppose that I admits a perfect matching M such that $|bp(M)| = bp^+(I)$. We form a truth assignment f in B as follows. For each i $(0 \leq i \leq n-1)$, 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.

We next observe that if $M \cap (X_i \times Y_i) = T_i$, then $bp(M) \cap (X_i \times Y_i) = \{(x_{6i+5}, y_{6i})\}$, whilst if $M \cap (X_i \times Y_i) = F_i$, then $bp(M) \cap (X_i \times Y_i) = \{(x_{6i+4}, y_{6i+4})\}$. For each j $(1 \leq j \leq m)$, $(q_j, c_j^s) \in M$ for some s $(1 \leq s \leq 3)$, so that $\{(p_j^t, c_j^t), (v_j^t, w_j^t)\} \subseteq M$ for each t $(1 \leq t \neq s \leq 3)$, whilst $\{(p_j^s, w_j^s), (v_j^s, z_j)\} \subseteq M$, since M is a perfect matching. Hence $bp(M) \cap (V_j \times W_j) = \{(v_j^s, w_j^s)\}$. Since M is perfect, no agent in $P \cup Q \cup Z$ can be involved in a blocking pair of M.

Now let c_j be a clause in C $(1 \leq j \leq m)$. Exactly one woman $c_j^s \in \{c_j^1, c_j^2, c_j^3\}$ has her last-choice partner in M. If $(x(c_j^s), c_j^s) \in bp(M)$ then the literal occurring at position s of c_j is false. In such a case we claim that the same is true for every literal in c_j. For, suppose not. Then there exists a true literal appearing at position s' of c_j. By construction of f, $x(c_j^{s'})$ has his first-choice partner in M. Let

$$\begin{aligned} M' &= (M\backslash((P_j \cup V_j \cup \{q_j\}) \times (C' \cup W \cup Z))) \\ &\quad \cup\{(q_j, c_j^{s'}), (p_j^{s'}, w_j^{s'}), (v_j^{s'}, z_j)\} \cup \{(p_j^t, c_j^t), (v_j^t, w_j^t) : 1 \leq t \neq s' \leq 3\}. \end{aligned}$$

Then M' is a perfect matching in I and $|bp(M')| < |bp(M)|$, contradicting the choice of M. Hence the claim is established, so that clause c_j is false under f. It follows that $|bp(M)| = n + m + (m - t(f))$, and therefore $t(B) + bp^+(I) \geq n + 2m$. From our earlier inequality it follows that $t(B) + bp^+(I) = n + 2m = \frac{11}{4}m$, since $3m = 4n$.

Berman et al. [4] show that it is NP-hard to distinguish between instances B of MAX (2,2)-E3-SAT for which (i) $t(B) \geq (1 - \varepsilon)m$ and (ii) $t(B) \leq \left(\frac{1015}{1016} + \varepsilon\right) m$. By our construction, it follows that in case (i), $bp^+(I) \leq \left(\frac{3556}{2032} + \varepsilon\right) m$, whilst in case (ii), $bp^+(I) \geq \left(\frac{3558}{2032} - \varepsilon\right) m$. Hence an approximation algorithm for MAX SIZE MIN BP $(3,3)$-SMI with performance guarantee r, for any $r \leq \frac{3557}{3556+2032\varepsilon}$, could be used to decide between cases (i) and (ii) for MAX (2,2)-E3-SAT in polynomial time, which is a contradiction unless P=NP. □

We remark that it is possible to prove Theorem 7 without requiring the agents in the set $V_j \cup W_j \cup \{x_{6i+4}, x_{6i+5}, y_{6i+4}, y_{6i+5} : 0 \leq i \leq n-1\}$, however these agents are included in order to provide a single reduction that is also valid for MAX SIZE MIN BA SMI. We now consider the approximability of this problem.

Theorem 8. *Given any ε $(0 < \varepsilon < \frac{1}{2032})$,* MAX SIZE MIN BA $(3,3)$-SMI *is not approximable within $\frac{3557}{3556+2032\varepsilon}$ unless P=NP.*

Proof. Let ε $(0 < \varepsilon < \frac{1}{2032})$ be given. We use exactly the same reduction as in the proof of Theorem 7 and observe that each blocking pair considered in that proof corresponds to two unique agents. It follows that $2t(B) + ba^+(I) = 2(n + 2m) = \frac{11m}{2}$, since $3m = 4n$.

Berman et al. [4] show that it is NP-hard to distinguish between instances B of MAX (2,2)-E3-SAT for which (i) $t(B) \geq (1 - \varepsilon)m$ and (ii) $t(B) \leq \left(\frac{1015}{1016} + \varepsilon\right) m$. By our construction, it follows that in case (i), $ba^+(I) \leq \left(\frac{3556}{1016} + 2\varepsilon\right) m$, whilst in case (ii), $ba^+(I) \geq \left(\frac{3558}{1016} - 2\varepsilon\right) m$. Hence an approximation algorithm for MAX SIZE MIN BA $(3,3)$-SMI with performance guarantee r, for any $r \leq \frac{3557}{3556+2032\varepsilon}$, could be used to decide between cases (i) and (ii) for MAX (2,2)-E3-SAT in polynomial time, which is a contradiction unless P=NP. □

4 Preference lists on one side of length at most 2

We now consider instances of SMI in which all preference lists on one side are of length at most 2. Let I be an SMI instance in which $\mathcal{U}$ is the set of men and $\mathcal{W}$ is the set of women. Assume without loss of generality that every man has a list of length at most 2. Let G be the underlying graph of I. Let $n = |V(G)|$ and $m = |E(G)|$. Note that $m \leq 2 \cdot |\mathcal{U}| < 2n$.

Define PERFECT MIN BP (p, q)-SMI as follows. An instance of this problem is an SMI instance I in which each man's preference list is of length at most p and each woman's preference list is of length at most q ($p = \infty$ or $q = \infty$ denotes unbounded length preference lists as before). A solution is a perfect matching with the minimum number of blocking pairs in I if I admits a perfect matching, or "no" otherwise.

Lemma 9. PERFECT MIN BP $(2, \infty)$-SMI *is solvable in $O(n)$ time, where n is the number of men in I.*

Proof. If $|\mathcal{U}| \neq |\mathcal{W}|$ then the answer is no. Let $I_1 = I$ the initial instance. If there is a woman in I_1 with an empty lists, then we output "no". If there is a woman w_1 with preference list that contains only one man, say m_1 then w_1 can be matched only to m_1 in a perfect matching. Therefore we add $\{w_1, m_1\}$ to M, we remove w_1 and m_1 from I_1 and obtain an instance I_2. We continue this process, if we find a woman in I_i with empty list then we stop the algorithm with output "no" and if we find a woman w_i with preference list of length 1 then we extend M with the pair (w_i, m_i) and reduce the instance as described above. If for an instance I_k there is no woman with preference list of length at most one then every man and woman must have a preference list of length 2 in I_k, obviously. So the underlying graph of I_k consists of a set of disjoint even cycles. Therefore, to achieve a perfect matching we can choose one from the two possible sets of edges for each of these even cycles. We make these decisions for each even cycle separately after counting the blocking pairs that contain some women in the cycle, for both cases, by considering all the edges of I. This last step can be done also in $O(n)$ time since we have to go through the women's preference lists only once. $\square$

We continue with the related problem MEN COVER MIN BP $(2, \infty)$-SMI. Here, we suppose that the preference lists of the men are of length at most 2, and the problem is to minimize the number of blocking pairs over all matchings that cover the men.

Lemma 10. MEN COVER MIN BP $(2, \infty)$-SMI *is solvable in $O(n^2)$ time, where n is the number of men in I.*

Proof. Suppose that the graph of the instance, $G = (\mathcal{U} \cup \mathcal{W}, E)$ is connected, otherwise, we can solve the problem separately for each component. If the number of men $|\mathcal{U}|$ is greater than the number of women $|\mathcal{W}|$ then we output "no". If $|\mathcal{U}| = |\mathcal{W}|$ then we get an instance of PERFECT MIN BP $(2, \infty)$-SMI. The connectivity of G implies $|\mathcal{W}| \leq |\mathcal{U}| + 1$, so the last possible case is $|\mathcal{W}| = |\mathcal{U}| + 1$. Here, for every $w_j \in \mathcal{W}$ we solve an instance I_j of PERFECT MIN BP $(2, \infty)$-SMI after removing w_j from the graph. Note that if a matching M_j is a minimum solution for I_j then M_j is also a minimum for I between the matchings that does not cover w_j, since in those matchings in I, where w_j is not covered, every man in w_j's list has only one possible partner. Therefore, we can get the optimal solution for I by solving $|\mathcal{W}|$ instances of PERFECT MIN BP $(2, \infty)$-SMI and choosing the minimum of these solutions. $\square$

The problem WOMEN COVER MIN BP $(2, \infty)$-SMI can be defined similarly. Here, we suppose that the preference lists of the men are of length at most 2, and the problem is to minimize the number of blocking pairs over all matchings that cover the women.

Lemma 11. WOMEN COVER MIN BP $(2, \infty)$-SMI *is solvable in $O(n^3)$ time, where n is the number of men in I.*

Proof. Let $G = (\mathcal{U} \cup \mathcal{W}, E)$ be the graph of the instance I and let $bp(M)$ denote the set of blocking pairs for a matching M in I. If there is no such matching that covers $\mathcal{W}$ then we output "no". Otherwise, we deal only with such matchings in this proof that covers $\mathcal{W}$, so we assume this property hereby. Let $bp_{int}(M)$ denote the set of *internal* blocking pairs for M, those blocking pairs that are covered by M. Furthermore, let $bp_{ext}(M)$ denote the *external* blocking pairs, where the men are uncovered by M. Note that $bp(M) = bp_{int}(M) \cup bp_{ext}(M)$.

Our algorithm consists of two cycles. In the first one, we eliminate the external blocking pairs without creating any new internal blocking pair. In the second one, we try to reduce the number of internal blocking edges by switching pairs along augmenting paths and cycles. Finally, we prove that if neither of these steps is possible then the solution is optimal.

Eliminating the external blocking pairs. *Claim 1: Suppose that for a matching M, $bp_{ext}(M) \neq \emptyset$. We can construct a matching M^* such that $bp_{int}(M) \supseteq bp_{int}(M^*) = bp(M^*)$.*

Suppose that $(u_i, w_j) \in bp_{ext}(M)$, and if (u_i, w_k) is also in $bp_{ext}(M)$ then u_i prefers w_j to w_k. Let $M' = (M \setminus \{(M(w_j), w_j)\}) \cup \{(u_i, w_j)\}$. We get $bp_{int}(M') \subseteq bp_{int}(M)$ since only u_i and w_j could be part of a new internal blocking pair. This is because (u_i, w_k) cannot be blocking since either u_i prefers w_j if (u_i, w_k) is blocking for M or (u_i, w_k) is not blocking for M, and w_j received a better partner so she cannot be part of any new blocking pair. Therefore, the set of internal blocking pairs can only reduce. We keep doing this elimination process until obtaining a matching M^* such that $bp_{int}(M^*) = bp(M^*)$. This process must terminate, since the women get better and better partners after each elimination, so no pair can be eliminated twice. The final matching M^* satisfies the required condition.

In order to obtain the $O(n)$ running time for the elimination process we have to ensure that we can find the external blocking pairs efficiently. Here, we use the fact that if an uncovered men u_l is not involved in any blocking pair for a matching M then he cannot be part of a blocking pair for M' either, where M' is obtained after eliminating an external blocking pair as described earlier. So we create a set $\mathcal{N}$ to collect such men that are uncovered by $M^k \cup bp_{ext}(M^k)$, where M^k is the actual matching in the process. Let $\mathcal{N}$ be empty for the initial matching. Whenever we consider an uncovered man $u_i \notin \mathcal{N}$ and we find that u_i does not involved in a blocking pair then we add u_i to $\mathcal{N}$. So in each step we either eliminate an external blocking pair or we extend $\mathcal{N}$ until $\mathcal{N}$ contains every uncovered men. Therefore, the process terminates in $O(n)$ time.

Reducing the number of internal blocking pairs. Let the alternating path P and alternating cycle C be defined as follows. For a matching M, a path $P = \{(u_0, w_1), (w_1, u_1), (u_1, w_2), \dots, (u_{k-1}, w_k), (w_k, u_k)\}$ is an *alternating path* if $(w_i, u_i) \in M$ and $(u_{i-1}, w_i) \notin M$ for every $1 \leq i \leq k$. If $u_0 = u_k$ then we get an *alternating cycle*. Let $M \oplus P$ denote the matching obtained by switching the edges along the alternating path, i.e. by removing the edges (u_i, w_i) from M and adding (u_{i-1}, w_i) to M for every $1 \leq i \leq k$. Furthermore, let $P_{\mathcal{W}}$ and $C_{\mathcal{W}}$ be the women covered by P and C, respectively, and let $P_{\mathcal{U}} = \{u_1, u_2, \dots, u_k\} = M(P_{\mathcal{W}})$ and $P^0_{\mathcal{U}} = \{u_0, u_1, \dots, u_{k-1}\} = (M \oplus P)(P_{\mathcal{W}})$. Finally, let $D(S)$ denote the set of edges incident with the set of vertices S.

Claim 2: Suppose that for a matching M, $bp_{ext}(M) = \emptyset$. If there is an alternating path P such that $|bp_{int}(M \oplus P) \cap D(P_{\mathcal{W}})| < |bp_{int}(M) \cap D(P_{\mathcal{W}})|$ then $|bp_{int}(M \oplus P)| < |bp_{int}(M)|$. Similarly, if there is an alternating cycle C such that $|bp_{int}(M \oplus C) \cap D(C_{\mathcal{W}})| < |bp_{int}(M) \cap D(C_{\mathcal{W}})|$ then $|bp_{int}(M \oplus C)| < |bp_{int}(M)|$.

It is enough to show that if $w_j \notin P_{\mathcal{W}}$ then w_j cannot be involved in any new internal blocking pair for $M \oplus P$. Suppose indirectly that (u_i, w_j) is a new internal blocking pair. If $u_i \notin P^0_{\mathcal{U}}$ then u_i is either uncovered by $M \oplus P$ or has the same partner as in M, so (u_i, w_j) cannot be a new internal blocking pair. If $u_i \in P^0_{\mathcal{U}} \cap P_{\mathcal{U}}$ then $(u_i, w_j) \neq E(G)$ since u_i has only two women in his list and both of them are in $P_{\mathcal{W}}$. Finally, if $u_i = u_0 = P^0_{\mathcal{U}} \setminus P_{\mathcal{U}}$ then (u_0, w_j) cannot be blocking since u_0 was uncovered by M and we supposed that no external blocking pair exists for M, a contradiction.

We show that the construction of all possible alternating paths and cycles together with counting the number of new blocking pairs can be organised in $O(n^2)$ time. Considering the alternating paths, we build up these paths from every uncovered man as follows. Let u be an uncovered men and w^1_1 be the first woman in his lists. We generate the first set of alternating paths starting from u as $P^1_k(u) = \{(u, w^1_1), (w^1_1, u^1_1), (u^1_1, w^1_2), \dots, (u^1_{k-1}, w^1_k), (w^1_k, u^1_k)\}$

for every k while no repetition occur in the sequence, by supposing that $u_i^1 = M(w_i^1)$ and w_i^1 is the other woman in u_{i-1}^1's list beside $M(u_{i-1}^1)$ for every $i \leq k$. We can generate P_k^2, the second set of alternating paths starting from u, similarly. When we count the number of internal blocking pairs that are incident with the women of $P_k^1(u)$ for the new matching $M \oplus P_k^1(u)$ we can use the fact that for a woman w_i^1, the set of internal blocking pairs remains the same in the matchings $M \oplus P_k^1(u)$ for every $k > i$ and can differ only by the pair (w_i^1, m_i^1) for $k = 1$ since m_i^1 is not covered in $M \oplus P_i^1(u)$. Therefore, to count the number of blocking edges for every alternating paths $P_k^1(u)$, we have to go through the preference lists of the women involved in the longest such path only once. Considering the alternating cycles, we try to construct an alternating cycle from every covered vertex u by building up the only alternating path starting from u as described above. If this alternating path returns to $M(u)$ then we find the only alternating cycle in which u may be involved, denoted by $C(u)$. Then we simply count the number of internal blocking pairs incident with the women involved in $C(u)$ for $M \oplus C(u)$.

The optimality. The next claim indicates that if neither of the above improvements is possible then the solution is optimal.

Claim 3: Suppose that $bp_{int}(M) = bp(M)$ and there is a matching M^{opt} such that $|bp(M^{opt})| < |bp(M)|$. Then there must be either an alternating path P such that $|bp_{int}(M \oplus P) \cap D(P_{\mathcal{W}})| < |bp_{int}(M) \cap D(P_{\mathcal{W}})|$ or an alternating cycle C such that $|bp_{int}(M \oplus C) \cap D(C_{\mathcal{W}})| < |bp_{int}(M) \cap D(C_{\mathcal{W}})|$.

By Claim 1 we can suppose that $bp_{int}(M^{opt}) = bp(M^{opt})$. Considering the symmetric difference of M and M^{opt} we get some alternating paths, some alternating cycles and some pairs that remain matched in M^{opt} too. Let $\mathcal{P}_{\mathcal{W}}$ and $\mathcal{C}_{\mathcal{W}}$ denote the set of women that are involved in an alternating path and an alternating cycle, respectively, and let $\mathcal{R}_{\mathcal{W}}$ denote the set of women who get the same partner in M and M^{opt}. Furthermore, let $\mathcal{P}_{\mathcal{U}} = M(\mathcal{P}_{\mathcal{W}})$, $\mathcal{P}_{\mathcal{U}}^0 = M^{opt}(\mathcal{P}_{\mathcal{W}})$, $\mathcal{C}_{\mathcal{U}} = M(\mathcal{C}_{\mathcal{W}})$ and $\mathcal{R}_{\mathcal{U}} = M(\mathcal{R}_{\mathcal{W}})$. Finally, let $\mathcal{DIF} = \mathcal{C}_{\mathcal{U}} \cup (\mathcal{P}_{\mathcal{U}} \cap \mathcal{P}_{\mathcal{U}}^0)$ denote the set of men who are matched with different partners in M and M^{opt}.

First we show that every women w_j in $\mathcal{R}_{\mathcal{W}}$ must be involved in the same internal blocking pairs for M and M^{opt}. Let us consider a pair (u_i, w_j). If $u_i \in \mathcal{R}_{\mathcal{U}}$ then (u_i, w_j) is blocking for M if and only if it is blocking for M^{opt} too, obviously. If $u_i \in \mathcal{DIF}$ then $(u_i, w_j) \notin E(G)$ since u_i has only two women in his list: $M(u_i)$ and $M^{opt}(u_i)$, who are in $\mathcal{P}_{\mathcal{W}} \cup \mathcal{C}_{\mathcal{W}}$. Finally, if $u_i \in \mathcal{P}_{\mathcal{U}}^0 \setminus \mathcal{P}_{\mathcal{U}}$ then u_i is uncovered by M, so (u_i, w_j) cannot be blocking since there is no external blocking pair for M. Similarly, if $u_i \in \mathcal{P}_{\mathcal{U}} \setminus \mathcal{P}_{\mathcal{U}}^0$ then u_i is uncovered by M^{opt}, so (u_i, w_j) cannot be blocking since there is no external blocking pair for M^{opt}.

Therefore, if we sum up the internal blocking pairs according the sets of women involved in the same alternating path or in the same alternating cycle for M and M^{opt}, then we get either an alternating path P or an alternating cycle C such that either $|bp(M^{opt}) \cap D(P_{\mathcal{W}})| < |bp(M) \cap D(P_{\mathcal{W}})|$ or $|bp(M^{opt}) \cap D(C_{\mathcal{W}})| < |bp(M) \cap D(C_{\mathcal{W}})|$.

If for an alternating path P, $|bp(M^{opt}) \cap D(P_{\mathcal{W}})| < |bp(M) \cap D(P_{\mathcal{W}})|$ then we can prove that $\{bp_{int}(M \oplus P) \cap D(P_{\mathcal{W}})\} \subseteq \{bp(M^{opt}) \cap D(P_{\mathcal{W}})\}$ which implies $|bp_{int}(M \oplus P) \cap D(P_{\mathcal{W}})| < |bp_{int}(M) \cap D(P_{\mathcal{W}})|$. To verify this it is enough to show that if for a woman $w_j \in P_{\mathcal{W}}$, (u_i, w_j) is an internal blocking pair for $M \oplus P$ then (u_i, w_j) is an internal blocking pair for M^{opt} too. Note that $M \oplus P(w_j) = M^{opt}(w_j)$, and u_i is from the set of men covered by $M \oplus P$ that is $M \oplus P(\mathcal{W}) = \mathcal{R}_{\mathcal{U}} \cup \mathcal{C}_{\mathcal{U}} \cup (\mathcal{P}_{\mathcal{U}} \setminus P_{\mathcal{U}}) \cup P_{\mathcal{U}}^0 \subseteq \mathcal{R}_{\mathcal{U}} \cup P_{\mathcal{U}}^0 \cup \mathcal{C}_{\mathcal{U}} \cup \mathcal{P}_{\mathcal{U}} = (\mathcal{R}_{\mathcal{U}} \cup P_{\mathcal{U}}^0) \cup (\mathcal{DIF} \setminus P_{\mathcal{U}}^0) \cup (\mathcal{P}_{\mathcal{U}} \setminus \mathcal{P}_{\mathcal{U}}^0)$. If $u_i \in \mathcal{R}_{\mathcal{U}}$ or $u_i \in P_{\mathcal{U}}^0$ then $M \oplus P(u_i) = M^{opt}(u_i)$, so the statement is obvious. If $u_i \in \mathcal{DIF} \setminus P_{\mathcal{U}}^0$ then $(u_i, w_j) \notin E(G)$ since w_j can be neither $M \oplus P(u_i) = M(u_i)$ nor $M^{opt}(u_i)$. Finally, if $u_i \in \mathcal{P}_{\mathcal{U}} \setminus \mathcal{P}_{\mathcal{U}}^0$ then u_i is uncovered by M^{opt}, so again, (u_i, w_j) cannot be blocking for $M \oplus P$ since there is no external blocking pair for M^{opt}.

Similarly, if for an alternating cycle C, $|bp(M^{opt}) \cap D(C_{\mathcal{W}})| < |bp(M) \cap D(C_{\mathcal{W}})|$ then we can prove in the same way that $\{bp_{int}(M \oplus C) \cap D(C_{\mathcal{W}})\} \subseteq \{bp(M^{opt}) \cap D(C_{\mathcal{W}})\}$ which implies $|bp_{int}(M \oplus C) \cap D(C_{\mathcal{W}})| < |bp_{int}(M) \cap D(C_{\mathcal{W}})|$.

Conclusion of the proof. If a matching M is not optimal and there is no external blocking pair then Claim 3 implies that we can find an alternating path or cycle that satisfies the condition described in Claim 2, so by switching the edges along this path or the cycle the number of internal blocking pairs reduces.

The $O(n^3)$ time implementation of the complete algorithm can be obtained as follows. In the first phase of the algorithm we eliminate the external blocking pairs in $O(n)$ time as described in the proof of Claim 1. If their is no more external blocking pair then we try to reduce the number of internal blocking pairs by switching pairs along augmenting paths and cycles as described in Claim 2. This second phase can be done in $O(n^2)$ time, and after each run, either the number of internal blocking edges reduces or we stop, since the solution was optimal. After the second phase we run the first phase again, since new external blocking edges may have been created, and so on. We repeat the first and second phases at most $|\mathcal{W}|$ times, since the number of internal blocking pairs for the initial matching is at most $|\mathcal{W}|$. Therefore, we get $O(n^3)$ for the overall running time of the algorithm. □

Theorem 12. MAX SIZE MIN BP $(2, \infty)$-SMI *is solvable in $O(n^3)$ time, where n is the number of men in I.*

Proof. Let the bipartite graph be $G = (\mathcal{U} \cup \mathcal{W}, E)$, where every man in $\mathcal{U}$ has a preference list of length at most 2. First, we decompose G by using König's theorem. Let $X \subseteq \mathcal{U}$ and $Y \subseteq \mathcal{W}$ be such that $X \cup Y$ is a minimum vertex cover, whose size is equal to the size of a maximum matching of G. Let M be a maximum matching that covers $X \cup Y$. Note that there cannot be an edge (x, y) in M with $(x, y) \in (X \times Y)$.

Let $\mathcal{U}_2$ be a subset of X such that for every $u_i \in \mathcal{U}_2$ there is an alternating path from some $y \in Y$ to u_i, and let $\mathcal{W}_2 = M(\mathcal{U}_2)$. Furthermore, let $\mathcal{U}_3 = X \setminus \mathcal{U}_2$, $\mathcal{U}_1 = \mathcal{U} \setminus X$, $\mathcal{W}_1 = Y$ and $\mathcal{W}_3 = \mathcal{W} \setminus (\mathcal{W}_1 \cup \mathcal{W}_2)$. We claim that $\mathcal{W}_1 \cup \mathcal{W}_2 \cup \mathcal{U}_3$ is also a minimum vertex cover, moreover, the component restricted to the set of vertices $\mathcal{U}_1 \cup \mathcal{U}_2 \cup \mathcal{W}_1 \cup \mathcal{W}_2$ is independent from the component restricted to the set of vertices $\mathcal{U}_3 \cup \mathcal{W}_3$. The fact that $(\mathcal{W}_1 \cup \mathcal{W}_2) \times \mathcal{U}_3$ does not contain any edge is obvious by the definition of $\mathcal{U}_2$. There is no edge between $\mathcal{U}_1$ and $\mathcal{W}_3$ since $X \cup Y$ is a vertex cover. Finally, for every man u_i in $\mathcal{U}_2$, both women in $u_i's$ list must be in $\mathcal{W}_1 \cup \mathcal{W}_2$ by the definition of $\mathcal{U}_2$, so no woman in u_i's list can be from $\mathcal{W}_3$.

Therefore, we can obtain the solution for instance I of MAX SIZE MIN BP $(2, \infty)$-SMI by separately solving a problem of MEN COVER MIN BP $(2, \infty)$-SMI for the subinstance restricted to $\mathcal{U}_3 \cup \mathcal{W}_3$ and a problem of WOMEN COVER MIN BP $(2, \infty)$-SMI for the subinstance restricted to $\mathcal{U}_1 \cup \mathcal{U}_2 \cup \mathcal{W}_1 \cup \mathcal{W}_2$. □

Adapting the algorithm for different models

In the case where the preference lists may include ties, we can use the same algorithm with the following minor modification. When eliminating the external blocking pairs (Claim 1) we supposed that for an unmatched man u_i, if both (u_i, w_j) and (u_i, w_k) are external blocking pairs then u_i prefers w_j to w_k. However in the case of ties, we require that u_i either prefers w_j to w_k or is indifferent between them.

Regarding the problem of minimising the number of agents involved in blocking pairs, we need a little more care. In this case too, we can eliminate every external blocking pair without creating any new internal blocking pair (Claim 1), therefore we can focus on the number of agents involved in internal blocking pairs.

In the second phase of the algorithm, i.e., when we improve the solution by switching edges along augmenting paths and cycles, we need to make the following changes. Instead of counting the internal blocking pairs, we now have to count the agents involved in such pairs. Whenever we counted the internal blocking pairs that were incident with a set of women X in the proof, here instead, we count the women in X who are involved in internal blocking pairs together with those men with whom they form these internal blocking pairs. To be more precise, for a set of women X and a matching M, let $bm_{int}(X, M) := \{u \in \mathcal{U} : \exists w \in X, (u, w) \in bp_{int}(M)\}$ (i.e., the set of men that form internal blocking pairs with some woman from X). If $ba_{int}(M)$ denotes the set of agents involved in internal blocking pairs for a matching M then Claim 2 should be modified as follows.

Claim 2′: Suppose that for a matching M, $bp_{ext}(M) = \emptyset$. If there is an alternating path P such that $|ba_{int}(M \oplus P) \cap (P_{\mathcal{W}} \cup bm_{int}(P_{\mathcal{W}}, (M \oplus P)))| < |ba_{int}(M) \cap (P_{\mathcal{W}} \cup bm_{int}(P_{\mathcal{W}}, M))|$ then $|ba_{int}(M \oplus P)| < |ba_{int}(M)|$. Similarly, if there is an alternating cycle C such that $|ba_{int}(M \oplus C) \cap (C_{\mathcal{W}} \cup bm_{int}(C_{\mathcal{W}}, (M \oplus C)))| < |ba_{int}(M) \cap (C_{\mathcal{W}} \cup bm_{int}(C_{\mathcal{W}}, M))|$ then $|ba_{int}(M \oplus C)| < |ba_{int}(M)|$.

Moreover, in the same way that we could obtain the number of internal blocking pairs by counting them according to some partition of women, we can also obtain the number of agents involved in internal blocking pairs by counting them according to the same partition, since each man can be involved in at most one internal blocking pair. Therefore, the following modified version of Claim 3 can be proved in a similar way.

Claim 3′: Suppose that $bp_{int}(M) = bp(M)$ and there is a matching M^{opt} such that $|ba(M^{opt})| < |ba(M)|$. Then there must be either an alternating path P such that $|ba_{int}(M \oplus P) \cap (P_{\mathcal{W}} \cup bm_{int}(P_{\mathcal{W}}, (M \oplus P)))| < |ba_{int}(M) \cap (P_{\mathcal{W}} \cup bm_{int}(P_{\mathcal{W}}, M))|$ or an alternating cycle C such that $|ba_{int}(M \oplus C) \cap (C_{\mathcal{W}} \cup bm_{int}(C_{\mathcal{W}}, (M \oplus C)))| < |ba_{int}(M) \cap (C_{\mathcal{W}} \cup bm_{int}(C_{\mathcal{W}}, M))|$.

Using a similar modification as described in the first paragraph of this subsection, we can find a maximum matching with the smallest number of blocking agents in the case that preference lists include ties. The running time of these modified algorithms remains $O(n^3)$, since having ties in the lists does not require any significant modification, and counting the agents involved in blocking pairs is not harder than counting the blocking pairs themselves.

Let MAX SIZE MIN BA $(2, \infty)$-SMI be the problem of finding a maximum matching with the smallest number of blocking agents for the case where the lists on one side are of length at most 2, and let MAX SIZE MIN BP $(2, \infty)$-SMTI (respectively MAX SIZE MIN BA $(2, \infty)$-SMTI) be the problem of finding a maximum matching with the smallest number of blocking pairs (respectively blocking agents) for the case where the preference lists on one side are of length at most 2 and ties may occur in the lists on either side. We summarise our results in the following theorem.

Theorem 13. MAX SIZE MIN BP $(2, \infty)$-SMTI, MAX SIZE MIN BA $(2, \infty)$-SMI *and* MAX SIZE MIN BA $(2, \infty)$-SMTI *are solvable in $O(n^3)$ time, where n is the number of men in I.*

5 Concluding remarks

In Table 1 we summarise complexity results for problems involving finding stable matchings and finding matchings with the minimum number of blocking pairs or blocking agents, in the context of instances of SMI and SR. The table is split into columns according to these problems, and further according to whether the preference lists are strictly ordered or include ties.

The problem is to find a matching M	where M is	SMI instances		SR instances	
		strict	with ties	strict	with ties
such that M is stable	arbitrary	P[6]	P [6, 8]	P [9]	N [18, 12]
	maximum	P[6, 7]	N [15]	P[9, 8]	N [18, 12]
such that M has min no. blocking pairs	arbitrary	P (=0) [6]	P (=0) [6, 8]	N [2][1]	N (←)
	maximum	N (*)	N (←)	N (↑)	N (↖)
such that M has min no. blocking agents	arbitrary	P (=0) [6]	P (=0) [6, 8]	N [2][2]	N (←)
	maximum	N (*)	N (←)	N (↑)	N (↖)

Table 1: Complexity results for problems involving finding stable matchings and matchings with the minimum number of blocking pairs / agents.

The rows of the table refer to the case that we seek either a stable matching, or a matching with the minimum number of blocking pairs, or a matching with the minimum number of blocking agents; these rows are further split into the cases that the matching should be of arbitrary or maximum size.

In a given table entry, 'P' denotes that the problem in question is polynomial-time solvable, whilst 'N' denotes NP-hardness. Furthermore, '=0' denotes the fact that an optimal solution admits 0 blocking pairs, whilst '(*)' indicates that the complexity result is established in this paper. The arrows indicate that NP-hardness holds by restriction, given the result in the cell above / to the left / above-left as appropriate.

We conclude with some open problems. The hardness results of Sections 2 and 3 also apply in the cases of HR and its generalisation HRT, where preference lists may include ties. However it remains to extend the algorithms of Section 4 to either of these settings, or to show that the corresponding optimisation problems are NP-hard.

We finally remark that the inapproximability results established by Theorems 7 and 8 leave open the question as to whether there is a c-approximation algorithm for either MAX SIZE MIN BP $(3, 3)$-SMI or MAX SIZE MIN BA $(3, 3)$-SMI, for some constant $c > 1$.

Acknowledgement

We would like to thank Rob Irving and three anonymous referees for helpful comments on earlier versions of this paper.

References

[1] A. Abdulkadiroğlu and T. Sönmez. School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747, 2003.

[2] D.J. Abraham, P. Biró, and D.F. Manlove. "Almost stable" matchings in the Roommates problem. In *Proceedings of WAOA '05: the 3rd Workshop on Approximation and Online Algorithms*, volume 3879 of *Lecture Notes in Computer Science*, pages 1–14. Springer, 2006.

[1] This result holds even if all preference lists are complete.

[2] It may be verified that a straightforward modification of Theorem 1 in [2] establishes the NP-hardness of the problem of finding a matching with the minimum number of blocking agents, given an instance of SR where all preference lists are complete; we omit the details for space reasons.

[3] D.J. Abraham, R.W. Irving, and D.F. Manlove. Two algorithms for the Student-Project allocation problem. *Journal of Discrete Algorithms*, 5(1):79–91, 2007.

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

[5] K. Eriksson and O. Häggström. Instability of matchings in decentralized markets with various preference structures. *International Journal of Game Theory*, 36:409–420, 2008.

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

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

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

[9] R.W. Irving. An efficient algorithm for the "stable roommates" problem. *Journal of Algorithms*, 6:577–595, 1985.

[10] 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.

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

[12] R.W. Irving and D.F. Manlove. The Stable Roommates Problem with Ties. *Journal of Algorithms*, 43:85–105, 2002.

[13] S. Khuller, S.G. Mitchell, and V.V. Vazirani. On-line algorithms for weighted bipartite matching and stable marriages. *Theoretical Computer Science*, 127:255–267, 1994.

[14] E. Kujansuu, T. Lindberg, and E. Mäkinen. The stable roommates problem and chess tournament pairings. *Divulgaciones Matemáticas*, 7(1):19–28, 1999.

[15] 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.

[16] G. O'Malley. *Algorithmic Aspects of Stable Matching Problems*. PhD thesis, University of Glasgow, Department of Computing Science, 2007.

[17] P.A. Robards. Applying two-sided matching processes to the united states navy enlisted assignment process. Master's thesis, Naval Postgraduate School, Monterey, California, 2001.

[18] E. Ronn. NP-complete stable matching problems. *Journal of Algorithms*, 11:285–304, 1990.

[19] 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.

[20] A.E. Roth, T. Sönmez, and M. Utku Ünver. Kidney exchange. *Quarterly Journal of Economics*, 119:457–488, 2004.

[21] A.E. Roth, T. Sönmez, and M. Utku Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125:151–188, 2005.

[22] A.E. Roth and X. Xing. Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists. *Journal of Political Economy*, 105(2):284–329, 1997.

[23] W. Yang, J.A. Giampapa, and K. Sycara. Two-sided matching for the U.S. Navy Detailing Process with market complication. Technical Report CMU-RI-TR-03-49, Robotics Institute, Carnegie-Mellon University, 2003.

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

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

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

[27] `http://www.nepke.org` (New England Program for Kidney Exchange website).