Biró, P. and Manlove, D.F. and Mittal, S. (2009) Size versus stability in the marriage problem. Lecture Notes in Computer Science, 5426 . pp. 1528. ISSN 0302-9743
http://eprints.gla.ac.uk/5930/

Deposited on: 01 June 2009

# Size Versus Stability in the Marriage Problem 

Péter Birón ${ }^{1 \star}$, David F. Manlove ${ }^{1 * \star}$ and Shubham Mittal ${ }^{2 \star * *}$<br>${ }^{1}$ Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK. Email: $\{$ pbiro, davidm\}@dcs.gla.ac.uk.<br>${ }^{2}$ Department of Computer Science and Engineering, Block VI, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi 110 016, India.<br>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 $\mathrm{P}=\mathrm{NP}$, 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 $\delta$, 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.


## 1 Introduction

The Stable Marriage problem (SM) was introduced in the seminal paper of Gale and Shapley [7]. In its classical form, an instance of SM involves $n$ men and $n$ women, each of whom specifies a preference list, which is a total order on the members of the opposite sex. A matching $M$ is a set of (man,woman) pairs such that each person belongs to exactly one pair. If $(m, w) \in M$, we say that $w$ is $m$ 's partner in $M$, and vice versa, and we write $M(m)=w, M(w)=m$.

A person $x$ prefers $y$ to $y^{\prime}$ if $y$ precedes $y^{\prime}$ on $x^{\prime}$ '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 [7] 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 [12].

[^0]Incomplete lists. A variety of extensions of the basic problem have been studied. In the Stable Marriage problem with Incomplete lists (SMI), the numbers of men and women need not be the same, and each person's preference list consists of a subset of the members of the opposite sex in strict order. A (man,woman) pair $(m, w)$ is acceptable if each member of the pair appears on the preference list of the other. A matching $M$ is now a set of acceptable pairs such that each person belongs to at most one pair. In this context, $(m, w)$ is a blocking pair for a matching $M$ if (a) $(m, w)$ is an acceptable pair, (b) $m$ is either unmatched or prefers $w$ to $M(m)$, and likewise (c) $w$ is either unmatched or prefers $m$ to $M(w)$. Given the definitions of a matching and a blocking pair, we lose no generality by assuming that the preference lists are consistent (i.e., given a (man,woman) pair $(m, w), m$ appears on the preference list of $w$ if and only if $w$ appears on the preference list of $m$ ). As in the classical case, there is always at least one stable matching for an instance of SMI, and it is straightforward to extend the Gale / Shapley algorithm to give a linear-time algorithm for this case. Again, there may be many different stable matchings, but Gale and Sotomayor [8] showed that every stable matching for a given SMI instance has the same size and matches exactly the same set of people.
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 [20]. The largest such scheme is the National Resident Matching Program (NRMP) [24] in the US, but similar schemes exist in Canada [25], in Scotland [11,26], 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 200607 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 [6] 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. [14].

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 [18, 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 [22, 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) [22]. However other studies [21] consider stability as the first priority. Another example in a non-bipartite setting involves pairing up chess players [15].

Our results. In this paper we present a range of algorithmic results for max SIZE MIN BP SMI, the problem of finding a maximum matching with the smallest number of blocking pairs, given an instance of Smi. We firstly show in Section 2 that this problem is NP-hard and not approximable within $n^{1-\varepsilon}$, for any $\varepsilon>0$, unless $\mathrm{P}=\mathrm{NP}$. We then consider special cases of the problem 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, MAX CARD MIN BP SMI is NP-hard and not approximable within $\delta$, for some $\delta>1$, unless $\mathrm{P}=\mathrm{NP}$. 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 Card min bp Smi. 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 $\mathrm{P}=\mathrm{NP}$. 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 $b p_{I}(M)$ denote the set of blocking pairs with respect to $M$ in $I$ (we omit the subscript when the instance is clear from the context). Let $b p^{+}(I)=\min \left\{\left|b p_{I}(M)\right|: M \in \mathcal{M}^{+}\right\}$. Define max sIze MIN BP SMI to be the problem of finding, given an SMI instance $I$, a matching $M \in \mathcal{M}^{+}$such that $\left|b p_{I}(M)\right|=b p^{+}(I)$.

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 exactmm 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 [17, 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 $P=N P$.

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 $G^{\prime}$ ) 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 $K \leq \beta(G) \leq \min \{|U|,|W|\}$. Suppose that $U=\left\{u_{1}, u_{2}, \ldots, u_{n_{1}}\right\}$ and $W=\left\{w_{1}, w_{2}, \ldots, w_{n_{2}}\right\}$. 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 $B=\left\lceil\frac{3}{\varepsilon}\right\rceil$ and let $C=\left(n_{1}+n_{2}\right)^{B+1}-\left(n_{1}+n_{2}\right)+1$. 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 $|\mathcal{U}|=|\mathcal{W}|=3 n_{1}+4 n_{2}+2 m C-K$. Let $U^{1}=\left\{u_{i}^{1}: 1 \leq i \leq n_{1}\right\}$ and let $W^{1}=\left\{w_{j}^{1}: 1 \leq j \leq n_{2}\right\}$.

For each $u_{i} \in U$ and $w_{j} \in W$ such that $\left\{u_{i}, w_{j}\right\} \in E$, 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 $u_{i}=u_{t_{j}}$.

```
\(\mathcal{U}=\left(\cup_{i=1}^{n_{1}} U_{i}\right) \cup\left(\cup_{\left\{u_{i}, w_{j}\right\} \in E} G_{i, j}\right) \cup\left(\cup_{i=1}^{n_{2}} V_{i}\right) \cup X\)
\(\mathcal{W}=\left(\cup_{j=1}^{n_{2}} W_{j}\right) \cup\left(\cup_{\left\{u_{i}, w_{j}\right\} \in E} H_{i, j}\right) \cup\left(\cup_{j=1}^{n_{1}} Z_{j}\right) \cup Y\)
\(G_{i, j}=G_{i, j}^{1} \cup G_{i, j}^{2} \quad\left(\left\{u_{i}, w_{j}\right\} \in E\right)\)
\(G_{i, j}^{d}=\left\{g_{i, j}^{c, d}: 1 \leq c \leq C\right\} \quad\left(\left\{u_{i}, w_{j}\right\} \in E \wedge 1 \leq d \leq 2\right)\)
\(H_{i, j}=H_{i, j}^{1} \cup H_{i, j}^{2}\)
\(H_{i, j}^{d}=\left\{h_{i, j}^{c, d}: 1 \leq c \leq C\right\}\)
\(\left(\left\{u_{i}, w_{j}\right\} \in E \wedge 1 \leq d \leq 2\right)\)
\(U_{i}=\left\{u_{i}^{1}, u_{i}^{2}, u_{i}^{3}\right\}\)
\(V_{i}=\left\{v_{i}^{1}, v_{i}^{2}, v_{i}^{3}\right\} \quad\left(1 \leq i \leq n_{2}\right)\)
\(W_{j}=\left\{w_{j}^{1}, w_{j}^{2}, w_{j}^{3}, w_{j}^{4}\right\}\)
\(\left(1 \leq j \leq n_{2}\right)\)
\(X=\left\{x_{i}: 1 \leq i \leq n_{2}-K\right\}\)
\(Y=\left\{y_{j}: 1 \leq j \leq n_{1}-K\right\}\)
\(Z_{j}=\left\{z_{j}^{1}, z_{j}^{2}\right\}\)
\(\left(1 \leq j \leq n_{1}\right)\).
```

Fig. 1. Men and women in the constructed instance of max size min bp smi.

Preference lists for the men and women in $I$ are as shown in Figure 2. In a given person'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 $\left\{u_{i}, w_{j}\right\} \in M$, the relevant pair in $U_{i} \times W_{j}$ (who rank each other in second place) will be added to a matching $M^{\prime}$ 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 $M^{\prime}$ 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 $M^{\prime}$ 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 $M^{\prime}$. It is then possible to extend $M^{\prime}$ 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 $\left\{u_{i}, w_{j}\right\} \in E$. Hence $\left|b p\left(M^{\prime}\right)\right|=n_{1}+n_{2}$. Conversely, from a perfect matching $M^{\prime}$ in $I$, it is straightforward to extract a matching $M$ in $G$ of size $K$. If $M$ is not maximal then there is some $u_{i} \in U$ and $w_{j} \in W$, both unmatched in $M$, such that $\left\{u_{i}, w_{j}\right\} \in E$. In this case, for each $c(1 \leq c \leq C)$, either $\left(u_{i}^{1}, h_{i, j}^{c, 1}\right) \in b p\left(M^{\prime}\right)$ or $\left(g_{i, j}^{c, 1}, w_{j}^{1}\right) \in b p\left(M^{\prime}\right)$, and hence $\left|b p\left(M^{\prime}\right)\right| \geq C$. This introduces the required 'gap' for the inapproximability result.

The formal proof of correctness of the reduction is based on a number of claims as follows, each of which is proved in [5].

Claim 1: I admits a perfect matching. Claim 2: if $G$ admits a maximal matching of size $K$, then $b p(I) \leq n_{1}+n_{2}$. Claim 3: if $G$ admits no maximal matching of size $K$ then $b p(I)>\left(n_{1}+n_{2}\right)^{B+1}$.

$$
\begin{aligned}
& u_{i}^{1}: z_{i}^{1} w_{p_{i}}^{\tau_{i, p_{i}}}\left[H_{i, p_{i}}^{1}\right]\left[H_{i, q_{i}}^{1}\right][[Y]] \quad\left(1 \leq i \leq n_{1}\right) \\
& u_{i}^{2}: z_{i}^{2} w_{q_{i}}^{\tau_{i, q_{i}}} \quad\left(1 \leq i \leq n_{1}\right) \\
& u_{i}^{3}: z_{i}^{1} z_{i}^{2} \quad\left(1 \leq i \leq n_{1}\right) \\
& g_{i, j}^{c, 1}: h_{i, j}^{c, 1} w_{j}^{1} h_{i, j}^{c, 2} \quad\left(\left\{u_{i}, w_{j}\right\} \in E \wedge 1 \leq c \leq C\right) \\
& g_{i, j}^{c, 2}: h_{i, j}^{c, 2} \quad h_{i, j}^{c, 1} \\
& v_{i}^{1}: w_{i}^{1} w_{i}^{4} \\
& v_{i}^{2}: w_{i}^{2} w_{i}^{4} \\
& v_{i}^{3}: w_{i}^{3} w_{i}^{4} \\
& x_{i}: \quad\left[\left[W^{1}\right]\right] \\
& w_{j}^{1}: v_{j}^{1} u_{r_{j}}^{\sigma_{j, r_{j}}}\left[G_{j, r_{j}}^{1}\right]\left[G_{j, s_{j}}^{1}\right]\left[G_{j, t_{j}}^{1}\right][[X]] \quad\left(1 \leq j \leq n_{2}\right) \\
& w_{j}^{2}: v_{j}^{2} u_{s_{j}}^{\sigma_{j, s_{j}}} \quad\left(1 \leq j \leq n_{2}\right) \\
& w_{j}^{3}: v_{j}^{3} u_{t_{j}}^{\sigma_{j, t_{j}}} \quad\left(1 \leq j \leq n_{2}\right) \\
& w_{j}^{4}: v_{j}^{1} v_{j}^{2} v_{j}^{3} \quad\left(1 \leq j \leq n_{2}\right) \\
& h_{i, j}^{c, 1}: g_{i, j}^{c, 2} \quad u_{i}^{1} \quad g_{i, j}^{c, 1} \\
& \left(\left\{u_{i}, w_{j}\right\} \in E \wedge 1 \leq c \leq C\right) \\
& h_{i, j}^{c, 2}: g_{i, j}^{c, 1} \quad g_{i, j}^{c, 2} \\
& z_{j}^{1}: u_{j}^{1} u_{j}^{3} \\
& z_{j}^{2}: u_{j}^{2} u_{j}^{3} \\
& y_{j}: \quad\left[\left[U^{1}\right]\right]
\end{aligned}
$$

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

Hence the existence of a $\left(n_{1}+n_{2}\right)^{B}$-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 $\mathrm{P}=$ NP. Claim 4: $\left(n_{1}+n_{2}\right)^{B} \geq$ $n^{1-\varepsilon}$, completing the proof.

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

Corollary 1. max size exact bp smi is $N P$-complete.
Proof. We use the same reduction as in the proof of Theorem 1 and set $K^{\prime}=$ $n_{1}+n_{2}$ 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 $M^{\prime}$ such that $\left|b p\left(M^{\prime}\right)\right| \leq K^{\prime}$. However it is straightforward to verify that any perfect matching $M^{\prime}$ in $I$ satisfies $\left|b p\left(M^{\prime}\right)\right| \geq K^{\prime}$, 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 2 ([2]). MAX SIZE EXACT BP SMI is solvable in polynomial time when $K^{\prime}$ is fixed.

## 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 denote the restriction of MAX SIZE MIN BP SMI in which each man's preference list is of length at most $p$, and each woman's list is of length at most $q$. We use $p=\infty$ or $q=\infty$ to denote the possibility that the men's lists or women's lists 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 $\mathrm{P}=\mathrm{NP}$. 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 3. Given any $\varepsilon\left(0<\varepsilon<\frac{1}{2032}\right)$, MAX SIZE MIN BP (3,3)-SMI is not approximable within $\frac{3557}{3556+2032 \varepsilon}$ unless $P=N P$.
Proof. Let $\varepsilon\left(0<\varepsilon<\frac{1}{2032}\right)$ be given. Let $B$ be an instance of max (2,2)-E3-SAT. Let $V=\left\{v_{0}, v_{1}, \ldots, v_{n-1}\right\}$ and $C=\left\{c_{1}, c_{2}, \ldots, c_{m}\right\}$ be the set of variables and clauses in $B$ respectively. Then for each $v_{i} \in V$, each of literals $v_{i}$ and $\overline{v_{i}}$ appears exactly twice in $B$. Also $\left|c_{j}\right|=3$ 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 $X \cup P \cup Q$ and the set of women in $I$ is $Y \cup C^{\prime} \cup Z$, where $X=\cup_{i=0}^{n-1} X_{i}, X_{i}=\left\{x_{4 i+r}: 0 \leq r \leq 3\right\}(0 \leq i \leq n-1)$, $P=\cup_{j=1}^{m} P_{j}, P_{j}=\left\{p_{j}^{1}, p_{j}^{2}, p_{j}^{3}\right\}(1 \leq j \leq m), Q=\left\{q_{j}: c_{j} \in C\right\}, Y=\cup_{i=0}^{n-1} Y_{i}$, $Y_{i}=\left\{y_{4 i+r}: 0 \leq r \leq 3\right\}(0 \leq i \leq n-1), C^{\prime}=\left\{c_{j}^{r}: c_{j} \in C \wedge 1 \leq r \leq 3\right\}$ and $Z=\left\{z_{j}: c_{j} \in C\right\}$.

The preference lists of the men and women in $I$ are shown in Figure 3. In the preference list of an agent $x_{4 i+r} \in X(0 \leq i \leq n-1$ and $r \in\{0,1\})$, the symbol $c\left(x_{4 i+r}\right)$ denotes the woman $c_{j}^{s} \in C^{\prime}$ 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 $c\left(x_{4 i+r}\right)$ denotes the woman $c_{j}^{s} \in C^{\prime}$ 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 $c_{j}^{s} \in C^{\prime}$, if literal $v_{i}$ appears at position $s$ of clause $c_{j} \in C$, the symbol $x\left(c_{j}^{s}\right)$ denotes the man $x_{4 i+r-1}$ 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 $x\left(c_{j}^{s}\right)$ denotes the man $x_{4 i+r+1}$ where $r=1,2$ according as this is the first or second occurrence of literal $\overline{v_{i}}$ in $B$. Clearly each preference list is of length at most 3 .

For each $i(0 \leq i \leq n-1)$, let $T_{i}=\left\{\left(x_{4 i+r}, y_{4 i+r}\right): 0 \leq r \leq 3\right\}$ and $\left.F_{i}=\left\{\left(x_{4 i+r}, y_{4 i+r+1}\right)\right\}: 0 \leq r \leq 3\right\}$, where addition is taken module $4 i$. We firstly note that $M$ is a perfect matching of the men and women in $I$, where

$$
M=\bigcup_{i=0}^{n-1} T_{i} \cup\left\{\left(p_{j}^{1}, c_{j}^{1}\right),\left(p_{j}^{2}, c_{j}^{2}\right),\left(p_{j}^{3}, z_{j}\right),\left(q_{j}, c_{j}^{3}\right): 1 \leq j \leq m\right\} .
$$

$$
\begin{aligned}
& x_{4 i}: y_{4 i} c\left(x_{4 i}\right) \quad y_{4 i+1} \\
& x_{4 i+1}: y_{4 i+1} c\left(x_{4 i+1}\right) \quad y_{4 i+2} \\
& x_{4 i+2}: y_{4 i+3} c\left(x_{4 i+2}\right) \quad y_{4 i+2} \\
& x_{4 i+3}: y_{4 i} c\left(x_{4 i+3}\right) \quad y_{4 i+3} \\
& p_{j}^{r}: c_{j}^{r} z_{j} \\
& q_{j}: c_{j}^{1} \quad c_{j}^{2} \quad c_{j}^{3} \\
& y_{4 i}: x_{4 i} \quad x_{4 i+3} \\
& y_{4 i+1}: x_{4 i} \quad x_{4 i+1} \\
& y_{4 i+2}: x_{4 i+1} \quad x_{4 i+2} \\
& y_{4 i+3}: x_{4 i+2} \quad x_{4 i+3} \\
& c_{j}^{r}: p_{j}^{r} x\left(c_{j}^{r}\right) q_{j} \\
& z_{j}: p_{j}^{1} p_{j}^{2} p_{j}^{3}
\end{aligned}
$$

```
\((0 \leq i \leq n-1)\)
\((0 \leq i \leq n-1)\)
( \(0 \leq i \leq n-1\) )
\((0 \leq i \leq n-1)\)
\((1 \leq j \leq m \wedge 1 \leq r \leq 3)\)
\((1 \leq j \leq m)\)
\((0 \leq i \leq n-1)\)
\((0 \leq i \leq n-1)\)
\((0 \leq i \leq n-1)\)
\((0 \leq i \leq n-1)\)
\((1 \leq j \leq m \wedge 1 \leq r \leq 3)\)
\((1 \leq j \leq m)\)
```

Fig. 3. Preference lists in the constructed instance of max size min bp (3, 3)-SMI.

We now give some intuition behind this construction. The people in $X_{i} \cup Y_{i}$ correspond to variable $v_{i} \in V$, whilst the people in $P_{j} \cup\left\{q_{j}, c_{j}^{1}, c_{j}^{2}, c_{j}^{3}, z_{j}\right\}$ 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 $x_{4 i}$ and $x_{4 i+1}$ (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 $x_{4 i+2}$ and $x_{4 i+3}$ (corresponding to the first and second occurrences of literal $\overline{v_{i}}$ ) has his third choice in $M$. Hence if any clause $c_{j}$ is false under $f$, then since $\left(q_{j}, c_{j}^{s}\right) \in M$ for some $s \in\{1,2,3\}$, it follows that $\left(x\left(c_{j}^{s}\right), c_{j}^{s}\right) \in b p(M)$. 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 $P_{j} \times C^{\prime}$ for each $c_{j} \in C$.

For the formal argument showing the correctness of the reduction, we claim (see [5] for the proof) that $t(B)+b p^{+}(I)=n+2 m=\frac{11}{4} m$, since $3 m=4 n$.

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), $b p^{+}(I) \leq\left(\frac{3556}{2032}+\varepsilon\right) m$, whilst in case (ii), $b p^{+}(I) \geq\left(\frac{3558}{2032}-\varepsilon\right) m$. 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 $\mathrm{P}=\mathrm{NP}$.

## 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. Define the underlying graph of $I$ to be a bipartite graph
$G=(V, E)$, where $V=\mathcal{U} \cup \mathcal{W}$ and $E$ is the set of mutually acceptable pairs. Let $n=|V(G)|$ and $m=|E(G)|$. Note that $m \leq 2 \cdot|\mathcal{U}|<2 n$.

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

The algorithm is quite simple; the description can be found in [5]. 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 2. MEN-COVER MIN BP $(2, \infty)$-SMI is solvable in $O\left(n^{2}\right)$ 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.

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 3. WOMEN-COVER MIN BP $(2, \infty)$-SMI is solvable in $O\left(n^{3}\right)$ 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 $b p(M)$ 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 $b p_{\text {int }}(M)$ denote the set of internal blocking pairs for $M$, those blocking pairs that are covered by $M$. Furthermore, let $b p_{\text {ext }}(M)$ denote the external blocking pairs, where the men are uncovered by $M$. Note that $b p(M)=b p_{\text {int }}(M) \cup b p_{\text {ext }}(M)$.

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 $M, b p_{\text {ext }}(M) \neq \emptyset$. We can construct a matching $M^{*}$ such that $b p_{\text {int }}(M) \supseteq b p_{\text {int }}\left(M^{*}\right)=b p\left(M^{*}\right)$.

Suppose that $\left(u_{i}, w_{j}\right) \in b p_{\text {ext }}(M)$, and if $\left(u_{i}, w_{k}\right)$ is also in $b p_{\text {ext }}(M)$ then $u_{i}$ prefers $w_{j}$ to $w_{k}$. Let $M^{\prime}=M \backslash\left(M\left(w_{j}\right), w_{j}\right) \cup\left(u_{i}, w_{j}\right)$. We get $b p_{\text {int }}\left(M^{\prime}\right) \subseteq$ $b p_{\text {int }}(M)$ since only $u_{i}$ and $w_{j}$ could be part of a new internal blocking pair. This is because $\left(u_{i}, w_{k}\right)$ cannot be blocking since either $u_{i}$ prefers $w_{j}$ if $\left(u_{i}, w_{k}\right)$ is blocking for $M$ or $\left(u_{i}, w_{k}\right)$ 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 $b p_{\text {int }}\left(M^{*}\right)=b p\left(M^{*}\right)$. 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.

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 $P=\left\{\left(u_{0}, w_{1}\right),\left(w_{1}, u_{1}\right),\left(u_{1}, w_{2}\right), \ldots,\left(u_{k-1}, w_{k}\right),\left(w_{k}, u_{k}\right)\right\}$ is an alternating path if $\left(w_{i}, u_{i}\right) \in M$ and $\left(u_{i-1}, w_{i}\right) \notin M$ 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 $\left(u_{i}, w_{i}\right)$ from $M$ and adding $\left(u_{i-1}, w_{i}\right)$ 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 $P_{\mathcal{U}}=\left\{u_{1}, u_{2}, \ldots, u_{k}\right\}=M\left(P_{\mathcal{W}}\right)$ and $P_{\mathcal{U}}^{0}=\left\{u_{0}, u_{1}, \ldots, u_{k-1}\right\}=(M \oplus P)\left(P_{\mathcal{W}}\right)$. Finally, let $D(S)$ denote the set of edges incident with the set of vertices $S$.

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

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 $\left(u_{i}, w_{j}\right)$ is a new internal blocking pair. If $u_{i} \notin P_{\mathcal{U}}^{0}$ then $u_{i}$ is either uncovered by $M \oplus P$ or has the same partner as in $M$, so $\left(u_{i}, w_{j}\right)$ cannot be a new internal blocking pair. If $u_{i} \in P_{\mathcal{U}}^{0} \cap P_{\mathcal{U}}$ then $\left(u_{i}, w_{j}\right) \neq E(G)$ since $u_{i}$ has only two women in his list and both of them are in $P_{\mathcal{W}}$. Finally, if $u_{i}=u_{0}=P_{\mathcal{U}}^{0} \backslash P_{\mathcal{U}}$ then $\left(u_{0}, w_{j}\right)$ cannot be blocking since $u_{0}$ was uncovered by $M$ and we supposed that no external blocking pair exists for $M$, a contradiction.

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

Claim 3: Suppose that $b p_{\text {int }}(M)=b p(M)$ and there is a matching $M^{\text {opt }}$ such that $\left|b p\left(M^{o p t}\right)\right|<|b p(M)|$ then there must be either an alternating path $P$ such
that $\left|b p_{\text {int }}(M \oplus P) \cap D\left(P_{\mathcal{W}}\right)\right|<\left|b p_{\text {int }}(M) \cap D\left(P_{\mathcal{W}}\right)\right|$ or an alternating cycle $C$ such that $\left|b p_{\text {int }}(M \oplus C) \cap D\left(C_{\mathcal{W}}\right)\right|<\left|b p_{\text {int }}(M) \cap D\left(C_{\mathcal{W}}\right)\right|$.

By Claim 1 we can suppose that $b p_{\text {int }}\left(M^{o p t}\right)=b p\left(M^{o p t}\right)$. Considering the symmetric difference of $M$ and $M^{o p t}$ we get some alternating paths, some alternating cycles and some pairs that remain matched in $M^{\text {opt }}$ 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 $M^{o p t}$. Furthermore, let $\mathcal{P}_{\mathcal{U}}=M\left(\mathcal{P}_{\mathcal{W}}\right), \mathcal{P}_{\mathcal{U}}^{0}=M^{\text {opt }}\left(\mathcal{P}_{\mathcal{W}}\right)$, $\mathcal{C}_{\mathcal{U}}=M\left(\mathcal{C}_{\mathcal{W}}\right)$ and $\mathcal{R}_{\mathcal{U}}=M\left(\mathcal{R}_{\mathcal{W}}\right)$. Finally, let $\mathcal{D} \mathcal{I} \mathcal{F}=\mathcal{C}_{\mathcal{U}} \cup\left(\mathcal{P}_{\mathcal{U}} \cap \mathcal{P}_{\mathcal{U}}^{0}\right)$ denote the set of men who are matched with different partners in $M$ and $M^{o p t}$.

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 $M^{o p t}$. Let us consider a pair $\left(u_{i}, w_{j}\right)$. If $u_{i} \in \mathcal{R}_{\mathcal{U}}$ then $\left(u_{i}, w_{j}\right)$ is blocking for $M$ if and only if it is blocking for $M^{o p t}$ too, obviously. If $u_{i} \in \mathcal{D} \mathcal{I} \mathcal{F}$ then $\left(u_{i}, w_{j}\right) \notin E(G)$ since $u_{i}$ has only two women in his list: $M\left(u_{i}\right)$ and $M^{\text {opt }}\left(u_{i}\right)$, who are in $\mathcal{P}_{\mathcal{W}} \cup \mathcal{C}_{\mathcal{W}}$. Finally, if $u_{i} \in \mathcal{P}_{\mathcal{U}}^{0} \backslash \mathcal{P}_{\mathcal{U}}$ then $u_{i}$ is uncovered by $M$, so $\left(u_{i}, w_{j}\right)$ cannot be blocking since there is no external blocking pair for $M$. Similarly, if $u_{i} \in \mathcal{P}_{\mathcal{U}} \backslash \mathcal{P}_{\mathcal{U}}^{0}$ then $u_{i}$ is uncovered by $M^{o p t}$, so ( $u_{i}, w_{j}$ ) cannot be blocking since there is no external blocking pair for $M^{o p t}$.

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 $M^{o p t}$, then we get either an alternating path $P$ or an alternating cycle $C$ such that either $\left|b p\left(M^{o p t}\right) \cap D\left(P_{\mathcal{W}}\right)\right|<\left|b p(M) \cap D\left(P_{\mathcal{W}}\right)\right|$ or $\left|b p\left(M^{o p t}\right) \cap D\left(C_{\mathcal{W}}\right)\right|<$ $\left|b p(M) \cap D\left(C_{\mathcal{W}}\right)\right|$.

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

Similarly, if for an alternating cycle $C,\left|b p\left(M^{\text {opt }}\right) \cap D\left(C_{\mathcal{W}}\right)\right|<\mid b p(M) \cap$ $D\left(C_{\mathcal{W}}\right) \mid$ then we can prove in the same way that $\left\{b p_{\text {int }}(M \oplus C) \cap D\left(C_{\mathcal{W}}\right)\right\} \subseteq$ $\left\{b p\left(M^{o p t}\right) \cap D\left(C_{\mathcal{W}}\right)\right\}$ which implies $\left|b p_{\text {int }}(M \oplus C) \cap D\left(C_{\mathcal{W}}\right)\right|<\mid b p_{\text {int }}(M) \cap$ $D\left(C_{\mathcal{W}}\right) \mid$.

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 cycle the number of internal blocking pairs reduces. Finally, the overall algorithm has complexity $O\left(n^{3}\right)$ (see [5] for full details).

Theorem 4. MAX size min Bp $(2, \infty)$-Smi is solvable in $O\left(n^{3}\right)$ 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 $\mathcal{W}_{2}=M\left(\mathcal{U}_{2}\right)$. Furthermore, let $\mathcal{U}_{3}=X \backslash \mathcal{U}_{2}$, $\mathcal{U}_{1}=\mathcal{U} \backslash X, \mathcal{W}_{1}=Y$ and $\mathcal{W}_{3}=\mathcal{W} \backslash\left(\mathcal{W}_{1} \cup \mathcal{W}_{2}\right)$. 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 $\left(\mathcal{W}_{1} \cup \mathcal{W}_{2}\right) \times \mathcal{U}_{3}$ 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 $u_{i}^{\prime} s$ list must be in $\mathcal{W}_{1} \cup \mathcal{W}_{2}$ by the definition of $\mathcal{U}_{2}$, so no woman in $u_{i}^{\prime}$ '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 $\mathcal{U}_{1} \cup \mathcal{U}_{2} \cup \mathcal{W}_{1} \cup \mathcal{W}_{2}$.

## 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, 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. So far all preference lists in this paper have been strictly ordered, however ties arise in practice: for example a large hospital with many applicants may be indifferent between those in certain groups.

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; 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 polynomialtime solvable, whilst 'NPc' denotes the NP-completeness of the related decision problem. Furthermore, ' $=0$ ' denotes the fact that an optimal solution admits 0 blocking pairs, whilst ' $(*)$ ' indicates that the complexity result is established in this paper. Indeed, the complexity result in the last row shown in boldface implies the result immediately to its right.

Table 1 already indicates that the hardness results of Sections 2 and 3 also apply to the extension of SMI to the case where preference lists may include ties. However it remains to extend the algorithms of Section 4 to this setting. Similar

| The problem is to find a matching $M$ | where <br> $M$ is | SMI instances |  | SR instances |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | strict | with ties | strict | with ties |
| such that M | arbitrary | $\mathrm{P}[7]$ | $\mathrm{P}[7,9]$ | P [10] | NPc [19, 13] |
| is | maximum | $\mathrm{P}[7,8]$ | NPc [16] | $\mathrm{P}[10,9]$ | NPc [19, 13] |
| such that $M$ has min | arbitrary | $\mathrm{P}(=0)[7]$ | $\mathrm{P}(=0)[7,9]$ | NPc [2] | NPc [2] |
| no. of blocking pairs | maximum | NPc (*) | NPc (*) | NPc [2] | NPc [2] |

Table 1. Complexity results for problems involving finding stable matchings and finding matchings with the minimum number of blocking pairs.
remarks apply if we are to consider the extension of the results to HR (and its generalisation HRT, where preference lists may include ties).

The inapproximability result established by Theorem 3 leaves open the question as to whether there is a $c$-approximation algorithm for MAX SIZE MIN BP $(3,3)$-SMI, for some constant $c>1$.

We conclude by mentioning an alternative way to minimise the instability of a maximum matching $M$ in an instance of Smi. As described by Eriksson and Häggström [6], rather than trying to minimise $|b p(M)|$, one could try to minimise the number of people who are involved in blocking pairs of $M$. We can modify the proofs of the results in Sections 2 and 3 so that they hold for this variant of mAX SIZE MIN BP SMI (the details are omitted for space reasons), however again it remains to extend the results of Section 4 to this case.

## Acknowledgement

We would like to thank Rob Irving and the 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.
3. D.J. Abraham, R.W. Irving, and D.F. Manlove. Two algorithms for the StudentProject 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. P. Biró, D.F. Manlove, and S.Mittal. Size versus stability in the marriage problem. Technical Report TR-2008-283, University of Glasgow, Department of Computing Science, 2008.
6. K. Eriksson and O. Häggström. Instability of matchings in decentralized markets with various preference structures. International Journal of Game Theory, to appear, 2008.
7. D. Gale and L.S. Shapley. College admissions and the stability of marriage. American Mathematical Monthly, 69:9-15, 1962.
8. D. Gale and M. Sotomayor. Some remarks on the stable matching problem. Discrete Applied Mathematics, 11:223-232, 1985.
9. D. Gusfield and R.W. Irving. The Stable Marriage Problem: Structure and Algorithms. MIT Press, 1989.
10. R.W. Irving. An efficient algorithm for the "stable roommates" problem. Journal of Algorithms, 6:577-595, 1985.
11. 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.
12. R.W. Irving and P. Leather. The complexity of counting stable marriages. SIAM Journal on Computing, 15(3):655-667, 1986.
13. R.W. Irving and D.F. Manlove. The Stable Roommates Problem with Ties. Journal of Algorithms, 43:85-105, 2002.
14. S. Khuller, S.G. Mitchell, and V.V. Vazirani. On-line algorithms for weighted bipartite matching and stable marriages. Theoretical Computer Science, 127:255267, 1994.
15. E. Kujansuu, T. Lindberg, and E. Mäkinen. The stable roommates problem and chess tournament pairings. Divulgaciones Matemáticas, 7(1):19-28, 1999.
16. 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.
17. G. O'Malley. Algorithmic Aspects of Stable Matching Problems. PhD thesis, University of Glasgow, Department of Computing Science, 2007.
18. 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.
19. E. Ronn. NP-complete stable matching problems. Journal of Algorithms, 11:285304, 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, T. Sönmez, and M. Utku Ünver. Kidney exchange. Quarterly Journal of Economics, 119:457-488, 2004.
22. A.E. Roth, T. Sönmez, and M. Utku Ünver. Pairwise kidney exchange. Journal of Economic Theory, 125:151-188, 2005.
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.nrmp.org (National Resident Matching Program website).
25. http://www.carms.ca (Canadian Resident Matching Service 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).

[^0]:    * 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.

