MAGYAR TUDOMÁNYOS AKADÉMIA
Közgazdaság- és Regionális Tudományi Kutatóközpont

MTA KRTK KTI

Centre for Economic and Regional Studies
HUNGARIAN ACADEMY OF SCIENCES

MŰHELYTANULMÁNYOK   DISCUSSION PAPERS

MT-DP – 2017/26

The Stable Roommates problem with short lists

ÁGNES CSEH – ROBERT W. IRVING – DAVID F. MANLOVE

The Stable Roommates problem with short lists

Authors:

Ágnes Cseh,
research fellow
Institute of Economics, Research Centre for Economic and Regional Studies
Hungarian Academy of Sciences,
and Corvinus University of Budapest
E-mail: cseh.agnes@krtk.mta.hu

Robert W. Irving
School of Computing Science, University of Glasgow
E-mail: Rob.Irving@glasgow.ac.uk

David F. Manlove
School of Computing Science, Sir Alwyn Williams Building, University of Glasgow
E-mail: David.Manlove@glasgow.ac.uk

A stabil szobatársprobléma rövid preferencialistákkal

Cseh Ágnes - Robert W. Irving - David F. Manlove

Összefoglaló

A klasszikus stabil szobatársprobléma két fokszámkorlátozott változatát tanulmányozzuk. Az első változat az egalitáriánus d-SRI probléma, amely egalitáriánus stabil párosítás keresését írja elő olyan gráfon, amelyben minden csúcs foka legfeljebb d. Bebizonyítjuk, hogy a probléma már d=3-ra is NP-teljes. Pozitív eredményként adunk egy (2d+3)/7-közelítő algoritmust minden d=3,4,5-re, ami az eddig ismert legjobb 2-közelítő algoritmust javítja. A második tanulmányozott változat d-SRTI névre hallgat és az a sajátossága, hogy a legfeljebb d hosszú listák tartalmazhatnak döntetleneket is. Bebizonyítjuk, hogy a probléma már d=3-ra is NP-teljes. Ugyanezen problémának a „legstabilabb" változatát is megvizsgáljuk és egy erős közelíthetetlenségi korlátot adunk a d=3 esetre. Ugyanakkor azt is belátjuk, hogy a d=2 eset polinomidőben megoldható.

Kulcsszavak: stabil párosítás, korlátos preferencialisták, bonyolultság, közelítő algoritmus

JEL-kódok: C63, C78

The Stable Roommates problem with short lists

Ágnes Cseh · Robert W. Irving · David F. Manlove

Abstract We consider two variants of the classical Stable Roommates problem with Incomplete (but strictly ordered) preference lists (SRI) that are degree constrained, i.e., preference lists are of bounded length. The first variant, EGAL d-SRI, involves finding an egalitarian stable matching in solvable instances of SRI with preference lists of length at most d. We show that this problem is NP-hard even if $d = 3$. On the positive side we give a $\frac{2d+3}{7}$-approximation algorithm for $d \in \{3, 4, 5\}$ which improves on the known bound of 2 for the unbounded preference list case. In the second variant of SRI, called d-SRTI, preference lists can include ties and are of length at most d. We show that the problem of deciding whether an instance of d-SRTI admits a stable matching is NP-complete even if $d = 3$. We also consider the "most stable" version of this problem and prove a strong inapproximability bound for the $d = 3$ case. However for $d = 2$ we show that the latter problem can be solved in polynomial time.

Á. Cseh
Institute of Economics, Hungarian Academy of Sciences, and Department of Operations Research and Actuarial Sciences, Corvinus University of Budapest
Tel.:+361-309-2652
Fax: +361-309-2650
E-mail: cseh.agnes@krtk.mta.hu

R. W. Irving
School of Computing Science, University of Glasgow
E-mail: Rob.Irving@glasgow.ac.uk

D. F. Manlove
School of Computing Science, University of Glasgow
E-mail: David.Manlove@glasgow.ac.uk

1 Introduction

In the *Stable Roommates problem with Incomplete lists* (SRI), a graph $G = (A, E)$ and a set of preference lists $\mathcal{O}$ are given, where the vertices $A = \{a_1, \ldots, a_n\}$ correspond to *agents*, and $\mathcal{O} = \{\prec_1, \ldots, \prec_n\}$, where $\prec_i$ is a linear order on the vertices adjacent to a_i in G ($1 \leq i \leq n$). We refer to $\prec_i$ as a_i's *preference list*. The agents that are adjacent to a_i in G are said to be *acceptable* to a_i. If a_j and a_k are two acceptable agents for a_i where $a_j \prec_i a_k$ then we say that a_i *prefers* a_j to a_k.

Let M be a matching in G. If $a_i a_j \in M$ then we let $M(a_i)$ denote a_j. An edge $a_i a_j \notin M$ *blocks* M, or forms a *blocking edge* of M, if a_i is unmatched or prefers a_j to $M(a_i)$, and similarly a_j is unmatched or prefers a_i to $M(a_j)$. A matching is called *stable* if no edge blocks it. Denote by SR the special case of SRI in which $G = K_n$. Gale and Shapley [9] observed that an instance of SR need not admit a stable matching. Irving [15] gave a linear-time algorithm to find a stable matching or report that none exists, given an instance of SR. The straightforward modification of this algorithm to the SRI case is described in [12]. We call an SRI instance *solvable* if it admits a stable matching.

In practice agents may find it difficult to rank a large number of alternatives in strict order of preference. One natural assumption, therefore, is that preference lists are short, which corresponds to the graph being of bounded degree. Given an integer $d \geq 1$, we define *d*-SRI to be the restriction of SRI in which G is of bounded degree d. This special case of SRI problem has potential applications in organising tournaments. As already pointed out in a paper of Kujansuu et al. [18], SRI can model a pairing process similar to the Swiss system, which is used in large-scale chess competitions. The assumption on short lists is reasonable, because according to the Swiss system, players can be matched only to other players with approximately the same score.

A second variant of SRI, which can be motivated in a similar fashion, arises if we allow ties in the preference lists, i.e., $\prec_i$ ($1 \leq i \leq n$) is now a strict weak ordering. That is, $\prec_i$ is a strict partial order in which incomparability is transitive. We refer to this problem as the *Stable Roommates problem with Ties and Incomplete lists* (SRTI) [17]. As in the SRI case, define *d*-SRTI to be the restriction of SRTI in which G is of bounded degree d. Denote by SRT the special case of SRTI in which $G = K_n$. In the context of the motivating application of chess tournament construction as mentioned in the previous paragraph, *d*-SRTI is naturally obtained if a chess player has several potential partners of the same score and match history in the tournament.

In the SRTI context, ties correspond to indifference in the preference lists. In particular, if $a_i a_j \in E$ and $a_i a_k \in E$ where $a_j \not\prec_i a_k$ and $a_k \not\prec_i a_j$ then a_i is said to be *indifferent between* a_j and a_k. Thus preference in the SRI context corresponds to strict preference in the case of SRTI. Relative to the strict weak

orders in $\mathcal{O}$, we can define stability in SRTI instances in exactly the same way as for SRI. This means, for example, that if $a_i a_j \in M$ for some matching M, and a_i is indifferent between a_j and some agent a_k, then $a_i a_k$ cannot block M. The term *solvable* can be defined in the SRTI context in an analogous fashion to SRI. Using a highly technical reduction from a restriction of 3-SAT, Ronn [23] proved that the problem of deciding whether a given SRT instance is solvable is NP-complete. A simpler reduction was given by Irving and Manlove [17].

For solvable instances of SRI there can be many stable matchings. Often it is beneficial to work with a stable matching that is fair to all agents in a precise sense [11,16]. One such fairness concept can be defined as follows. Given two agents a_i, a_j in an instance $\mathcal{I}$ of SRI, where $a_i a_j \in E$, let $\mathrm{rank}(a_i, a_j)$ denote the rank of a_j in a_i's preference list (that is, 1 plus the number of agents that a_i prefers to a_j). Let A_M denote the set of agents who are matched in a given stable matching M. (Note that this set depends only on $\mathcal{I}$ and is independent of M by [12, Theorem 4.5.2].) Define $c(M) = \sum_{a_i \in A_M} \mathrm{rank}(a_i, M(a_i))$ to be the *cost* of M. An *egalitarian stable matching* is a stable matching M that minimises $c(M)$ over the set of stable matchings in $\mathcal{I}$. Finding an egalitarian stable matching in SR was shown to be NP-hard by Feder [7]. Feder [7,8] also gave a 2-approximation algorithm for this problem in the SRI setting. He also showed that an egalitarian stable matching in SR can be approximated within a factor of α of the optimum if and only if Minimum Vertex Cover can be approximated within the same factor α. It was proved later that, assuming the Unique Games Conjecture, Minimum Vertex Cover cannot be approximated within $2 - \varepsilon$ for any $\varepsilon > 0$ [19].

Given an unsolvable instance $\mathcal{I}$ of SRI or SRTI, a natural approximation to a stable matching is a *most-stable* matching [1]. Relative to a matching M in $\mathcal{I}$, define $bp(M)$ to be the set of blocking edges of M and let $bp(\mathcal{I})$ denote the minimum value of $|bp(M')|$, taken over all matchings M' in $\mathcal{I}$. Then M is a *most-stable* matching in $\mathcal{I}$ if $|bp(M)| = bp(\mathcal{I})$. The problem of finding a most-stable matching was shown to be NP-hard and not approximable within $n^{k-\varepsilon}$, for any $\varepsilon > 0$, unless $\mathsf{P} = \mathsf{NP}$, where $k = \frac{1}{2}$ if $\mathcal{I}$ is an instance of SR and $k = 1$ if $\mathcal{I}$ is an instance of SRT [1].

To the best of our knowledge, there has not been any previous work published on either the problem of finding an egalitarian stable matching in a solvable instance of SRI with bounded-length preference lists or the solvability of SRTI with bounded-length preference lists. This paper provides contributions in both of these directions, focusing on instances of d-SRI and d-SRTI for $d \geq 2$, with the aim of drawing the line between polynomial-time solvability and NP-hardness for the associated problems in terms of d.

Our contribution. In Section 2 we study the problem of finding an egalitarian stable matching in an instance of d-SRI. We show that this problem is NP-hard if $d = 3$, whilst there is a straightforward algorithm for the case that $d = 2$. We then consider the approximability of this problem for the case that $d \geq 3$. We give an approximation algorithm with a performance guarantee of $\frac{9}{7}$ for the case that $d = 3$, $\frac{11}{7}$ if $d = 4$ and $\frac{13}{7}$ if $d = 5$. These performance

	finding a stable matching	finding an egalitarian stable matching
d-SRI	in P [15,12]	in P for $d = 2$ ($*$) NP-hard even for $d = 3$ ($*$) $\frac{2d+3}{7}$-approximation for $d \in \{3, 4, 5\}$ ($*$) 2-approximation for $d \geq 6$ [7,8]
d-SRTI	in P for $d = 2$ ($*$) NP-hard even for $d = 3$ ($*$)	not well-defined (see Section 4)

Table 1 Summary of results for d-SRI and d-SRTI.

guarantees improve on Feder's 2-approximation algorithm for the general SRI case [7,8]. In Section 3 we turn to d-SRTI and prove that the problem of deciding whether an instance of 3-SRTI is solvable is NP-complete. We then show that the problem of finding a most-stable matching in an instance of d-SRTI is solvable in polynomial time if $d = 2$, whilst for $d = 3$ we show that this problem is NP-hard and not approximable within $n^{1-\varepsilon}$, for any $\varepsilon > 0$, unless P = NP. Due to various complications, as explained in Section 4, we do not attempt to define and study egalitarian stable matchings in instances of SRTI. Some open problems are presented in Section 5. A structured overview of previous results and our results (marked by $*$) for d-SRI and d-SRTI is contained in Table 1.

Related work. Degree-bounded graphs, most-stable matchings and egalitarian stable matchings are widely studied concepts in the literature on matching under preferences [21]. As already mentioned, the problem of finding a most-stable matching has been studied previously in the context of SRI [1]. In addition to the results surveyed already, the authors of [1] gave an $O(m^{k+1})$ algorithm to find a matching M with $|bp(M)| \leq k$ or report that no such matching exists, where $m = |E|$ and $k \geq 1$ is any integer. Most-stable matchings have also been considered in the context of d-SRI [4]. The authors showed that, if $d = 3$, there is some constant $c > 1$ such that the problem of finding a most-stable matching is not approximable within c unless P = NP. On the other hand, they proved that the problem is solvable in polynomial time for $d \leq 2$. The authors also gave a $(2d-3)$-approximation algorithm for the problem for fixed $d \geq 3$. This bound was improved to $2d - 4$ if the given instance satisfies an additional condition (namely the absence of a structure called an *elitist odd party*). Most-stable matchings have also been studied in the bipartite restriction of SRI called the *Stable Marriage problem with Incomplete lists* (SMI) [14,5]. Since every instance of SMI admits a stable matching M (and hence $bp(M) = \emptyset$), the focus in [14,5] was on finding maximum cardinality matchings with the minimum number of blocking edges.

Regarding the problem of finding an egalitarian stable matching in an instance of SRI, as already mentioned Feder [7,8] showed that this problem is NP-hard, though approximable within a factor of 2. A 2-approximation algorithm for this problem was also given independently by Gusfield and Pitt [13], and by Teo and Sethuraman [26]. These approximation algorithms can also be extended to the more general setting where we are given a weight function

on the edges, and we seek a stable matching of minimum weight. Feder's 2-approximation algorithm requires monotone, non-negative and integral edge weights, whereas with the help of LP techniques [25,26], the integrality constraint can be dropped, while the monotonicity constraint can be partially relaxed.

2 The Egalitarian Stable Roommates problem

In this section we consider the complexity and approximability of the problem of computing an egalitarian stable matching in instances of d-SRI. We begin by defining the following problems.

Problem 1. EGAL d-SRI
Input: A solvable instance $\mathcal{I} = \langle G, \mathcal{O} \rangle$ of d-SRI, where G is a graph and $\mathcal{O}$ is a set of preference lists, each of length at most d.
Output: An egalitarian stable matching M in $\mathcal{I}$.

The decision version of EGAL d-SRI is defined as follows:

Problem 2. EGAL d-SRI DEC
Input: $\mathcal{I} = \langle G, \mathcal{O}, K' \rangle$, where $\langle G, \mathcal{O} \rangle$ is a solvable instance $\mathcal{I}'$ of d-SRI and K' is an integer.
Question: Does $\mathcal{I}'$ admit a stable matching M with $c(M) \leq K'$?

In the following we give a reduction from the NP-complete decision version of Minimum Vertex Cover in cubic graphs to EGAL 3-SRI DEC, deriving the hardness of the latter problem.

Theorem 1 EGAL 3-SRI DEC *is* NP*-complete.*

Proof. Clearly EGAL 3-SRI DEC belongs to NP. To show NP-hardness, we begin by defining the NP-complete problem that we will reduce to EGAL 3-SRI DEC.

Problem 3. 3-VC
Input: $\mathcal{I} = \langle G, K \rangle$, where G is a cubic graph and K is an integer.
Question: Does G contain a vertex cover of size at most K?

3-VC is NP-complete [10,20].

Construction of the EGAL 3-SRI DEC **instance.** Let $\langle G, K \rangle$ be an instance of 3-VC, where $G = (V, E)$, $E = \{e_1, \ldots, e_m\}$ and $V = \{v_1, \ldots, v_n\}$. For each i $(1 \leq i \leq n)$, suppose that v_i is incident to edges e_{j_1}, e_{j_2} and e_{j_3} in G, where without loss of generality $j_1 < j_2 < j_3$. Define $e_{i,s} = e_{j_s}$ $(1 \leq s \leq 3)$. Similarly for each j $(1 \leq j \leq m)$, suppose that $e_j = v_{i_1} v_{i_2}$, where without loss of generality $i_1 < i_2$. Define $v_{j,r} = v_{i_r}$ $(1 \leq r \leq 2)$. The use of this notation is illustrated in Figure 1.

We now construct an instance $\mathcal{I}$ of 3-SRI as follows. We define the following sets of vertices.

Fig. 1 Notation derived from the 3-vc instance $\langle G, K \rangle$.

Fig. 2 Part of the constructed instance of EGAL 3-SRI DEC.

$$V' = \{v_i^r : 1 \le i \le n \ \land 1 \le r \le 4\}$$
$$E' = \{e_j^s : 1 \le j \le m \land 1 \le s \le 2\}$$
$$W = \{w_i^r : 1 \le i \le n \ \land 1 \le r \le 4\}$$
$$Z \ = \{z_i^r : 1 \le i \le n \ \land 1 \le r \le 4\}$$

Intuitively, $v_i^r \in V'$ corresponds to vertex v_i and its incident edge $e_{i,r}$, whilst $e_j^s \in E'$ corresponds to edge e_j and its incident vertex $v_{j,s}$. The set $V' \cup E' \cup W \cup Z$ constitutes the set of agents in $\mathcal{I}$, and the preference lists of the agents are as shown in Figure 2. In the preference list of an agent v_i^r ($1 \le i \le n$ and $1 \le r \le 3$), the symbol $e(v_i^r)$ denotes the agent $e_j^s \in E'$ such that $e_j = e_{i,r}$ and $v_i = v_{j,s}$ (that is, e_j is the rth edge incident to v_i and v_i is the sth end-vertex of e_j). Similarly in the preference list of an agent e_j^s ($1 \le i \le m$ and $1 \le s \le 2$), the symbol $v(e_j^s)$ denotes the agent $v_i^r \in V'$ such that $v_i = v_{j,s}$ and $e_j = e_{i,r}$ (that is, v_i is the sth end-vertex of e_j and e_j is the rth edge incident to v_i).

Finally we define some further notation in $\mathcal{I}$. Let $K' = 7m + 19n + K$. The following edge sets play a particular role in our proof. Addition is taken modulo 4 here.

$$
\begin{aligned}
V_i^c &= \{v_i^r w_i^r && : 1 \le i \le n \wedge 1 \le r \le 4\} \\
V_i^u &= \{v_i^r w_i^{r+1} && : 1 \le i \le n \wedge 1 \le r \le 4\} \\
E_j^1 &= \{e_j^1 e_j^2, e_j^3 e_j^4 && : 1 \le j \le m\} \\
E_j^2 &= \{e_j^1 e_j^4, e_j^2 e_j^3 && : 1 \le j \le m\} \\
M_Z &= \{z_i^1 z_i^2, z_i^3 z_i^4 && : 1 \le i \le n\}
\end{aligned}
$$

This finishes the construction of the EGAL 3-SRI DEC instance $\mathcal{I}$. In the remainder of the proof we show that G has a vertex cover C where $|C| \le K$ if and only if $\mathcal{I}$ has a stable matching M where $c(M) \le K'$.

Claim 2. *If G has a vertex cover C such that $|C| = k \le K$, then there is a stable matching M in $\mathcal{I}$ such that $c(M) \le K'$.*

Proof. Suppose that G has a vertex cover C such that $|C| = k \le K$. We construct a matching M in $\mathcal{I}$ as follows. For each i $(1 \le i \le n)$, if $v_i \in C$, add V_i^c to M, otherwise add V_i^u to M. For each j $(1 \le j \le m)$, if $v_{j,1} \in C$, add E_j^2 to M, otherwise add E_j^1 to M. Finally add the pairs in M_Z to M.

We now argue that M is stable. Suppose that $e_j^1 e_j^4 \in M$ for some j $(1 \le j \le m)$. Then $E_j^2 \subseteq M$, so $v_{j,1} \in C$. Let $v_i = v_{j,1}$. Then by construction, $V_i^c \subseteq M$, and hence v_i^r has his first choice for each r $(1 \le r \le 4)$. Thus e_j^1 does not form a blocking edge of M with $v(e_j^1)$. The argument is similar if $e_j^1 e_j^2 \in M$ for some j $(1 \le j \le m)$. Then $E_j^1 \subseteq M$, so $v_{j,2} \in C$. Let $v_i = v_{j,2}$. Then by construction, $V_i^c \subseteq M$, and hence v_i^r has his first choice for each r $(1 \le r \le 4)$. Thus e_j^2 does not form a blocking edge of M with $v(e_j^1)$. Now suppose that $v_i^r w_i^{r+1} \in M$ for some i $(1 \le i \le n)$ and r $(1 \le r \le 3)$. Then $V_i^u \subseteq M$, so $v_i \notin C$. Let $e_j^s = e(v_i^r)$. If $s = 1$ then $v_i = v_{j,1}$. Hence by construction of M, $E_j^1 \subseteq M$. Then e_j^1 has his first-choice partner, so v_i^r does not block M with $e(v_i^r)$. If $s = 2$ then $v_i = v_{j,2}$. As $v_{j,2} \notin C$, it follows that $v_{j,1} \in C$ as C is a vertex cover. Hence by construction of M, $E_j^2 \subseteq M$. Then e_j^2 has its first-choice partner, so v_i^r does not block M with $e(v_i^r)$. It is straightforward to verify that M cannot admit any other type of blocking edge, and thus M is stable in $\mathcal{I}$.

Clearly every agent in $\mathcal{I}$ is matched in M. We note that Theorem 4.5.2 of [12] implies that every stable matching in $\mathcal{I}$ matches every agent in $\mathcal{I}$ – we will use this fact in the next claim. We finally note that $c(M) = 4k + 12k + 9(n-k) + 2(n-k) + 4(n-k) + 7m + 4n = 7m + 19n + k \le K'$, considering the contributions from the agents matched in V_i^c, V_i^u $(1 \le i \le n)$, E_j^1, E_j^2 $(1 \le j \le m)$ and M_Z respectively. $\square$

Claim 3. *If there is a stable matching M in $\mathcal{I}$ such that $c(M) \le K'$ then G has a vertex cover C such that $|C| = k \le K$.*

Proof. Suppose that M is a stable matching in $\mathcal{I}$ such that $c(M) \le K'$. We construct a set of vertices C in G as follows. As M matches every agent in $\mathcal{I}$, then for each i $(1 \le i \le n)$, either $V_i^c \subseteq M$ or $V_i^u \subseteq M$. In the former case add v_i to C. Also, for each j $(1 \le j \le m)$, as M matches every agent in $\mathcal{I}$, either $E_j^1 \subseteq M$ or $E_j^2 \subseteq M$. Finally, it follows that $M_Z \subseteq M$.

We now argue that C is a vertex cover. Let j ($1 \leq j \leq m$) be given and suppose that $v_{j,1} \notin C$ and $v_{j,2} \notin C$. Suppose firstly that $E_j^1 \subseteq M$. Let $v_i = v_{j,2}$. Then $V_i^u \subseteq M$ by construction of C, so that e_j^2 blocks M with $v(e_j^2)$, a contradiction. Now suppose that $E_j^2 \subseteq M$. Let $v_i = v_{j,1}$. Then $V_i^u \subseteq M$ by construction of C, so that e_j^1 blocks M with $v(e_j^1)$, a contradiction. Hence C is a vertex cover in G.

Moreover if $k = |C|$ then given the composition of M, as noted in the previous claim, $c(M) = 7m + 19n + k$, and since $c(M) \leq K'$ it follows that $k \leq K$. □

Theorem 1 immediately implies the following result.

Corollary 4. EGAL 3-SRI *is* NP-*hard.*

We remark that EGAL 2-SRI is trivially solvable in polynomial time: the components of the graph are paths and cycles in this case, and the cost of a stable matching selected in one component is not affected by the matching edges chosen in another component. Therefore we can deal with each path and cycle separately, minimising the cost of a stable matching in each. Paths and odd cycles admit exactly one stable matching (recall that (i) the instance is assumed to be solvable, and (ii) the set of matched agents is the same in all stable matchings [12, Theorem 4.5.2]), whilst even cycles admit at most two stable matchings (to find them, test each of the two perfect matchings for stability) – we can just pick the stable matching with lower cost in such a case. The following result is therefore immediate.

Proposition 5. EGAL 2-SRI *admits a linear-time algorithm.*

Corollary 4 naturally leads to the question of the approximabilty of EGAL d-SRI. As mentioned in the Introduction, Feder [7,8] provided a 2-approximation algorithm for the problem of finding an egalitarian stable matching in an instance of SRI. As Theorems 6, 8 and 10 show, this bound can be improved for instances with bounded-length preference lists.

Theorem 6 EGAL 3-SRI *is approximable within* $9/7$.

Proof. Let $\mathcal{I}$ be an instance of 3-SRI and let M_{egal} denote an egalitarian stable matching in $\mathcal{I}$. First we show that any stable matching in $\mathcal{I}$ is a $4/3$-approximation to M_{egal}. We then focus on the worst-case scenario when this ratio $4/3$ is in fact realised. Then we design a weight function on the edges of the graph and apply Teo and Sethuraman's 2-approximation algorithm [25, 26] to find an approximate solution M' to a minimum weight stable matching M_{opt} for this weight function. This weight function helps M' to avoid the worst case for the $4/3$-approximation for a significant amount of the matching edges. We will ultimately show that M' is in fact a $9/7$-approximation to M_{egal}.

Claim 7. *In an instance of* EGAL 3-SRI*, any stable matching approximates* $c(M_{egal})$ *within a factor of* $4/3$.

Proof. Let M be an arbitrary stable matching in $\mathcal{I}$. Call an edge uv an (i,j)-*pair* $(i \leq j)$ if v is u's ith choice and u is v's jth choice. By Theorem 4.5.2 of [12], the set of agents matched in M_{egal} is identical to the set of agents matched in M. We will now study the worst approximation ratios in all cases of (i,j)-pairs, given that $1 \leq i \leq j \leq 3$ in 3-SRI.

- If $uv \in M_{\text{egal}}$ is a $(1,1)$-pair then u and v contribute 2 to $c(M_{\text{egal}})$ and also 2 to $c(M)$ since they must be also be matched in M (and in every stable matching).
- If $uv \in M_{\text{egal}}$ is a $(1,2)$-pair then u and v contribute 3 to $c(M_{\text{egal}})$ and at most 4 to $c(M)$. Since, if $uv \notin M$, then v must be matched to his 1st choice and u to his 2nd or 3rd, because one of u and v must be better off and the other must be worse off in M than in M_{egal}.
- If $uv \in M_{\text{egal}}$ is a $(1,3)$-pair then u and v contribute 4 to $c(M_{\text{egal}})$ and at most 5 to $c(M)$. Since, if $uv \notin M$, then v must be matched to his 1st or 2nd choice and u to his 2nd or 3rd.
- If $uv \in M_{\text{egal}}$ is a $(2,2)$-pair then u and v contribute 4 to $c(M_{\text{egal}})$ and at most 4 to $c(M)$. Since, if $uv \notin M$, then one must be matched to his 1st choice and the other to his 3rd.
- If $uv \in M_{\text{egal}}$ is a $(2,3)$-pair then u and v contribute 5 to $c(M_{\text{egal}})$ and at most 5 to $c(M)$. Since, if $uv \notin M$, then v must be matched to his 1st or 2nd choice and u to his 3rd.
- If $uv \in M_{\text{egal}}$ is a $(3,3)$-pair then u and v contribute 6 to $c(M_{\text{egal}})$ and also 6 to $c(M)$ since they must be also be matched in M (and in every stable matching – this follows by [12, Lemma 4.3.9]).

It follows that, for every pair $uv \in M_{\text{egal}}$,

$$\frac{\text{rank}(u, M(u)) + \text{rank}(v, M(v))}{\text{rank}(u, M_{\text{egal}}(u)) + \text{rank}(v, M_{\text{egal}}(v))} = \frac{\text{rank}(u, M(u)) + \text{rank}(v, M(v))}{\text{rank}(u,v) + \text{rank}(v,u)}$$
$$\leq 4/3.$$

Hence $c(M)/c(M_{\text{egal}}) \leq 4/3$ and Claim 7 is proved. $\qquad\square$

As shown in Claim 7, the only case when the approximation ratio $4/3$ is reached is where M_{egal} consists of (1,2)-pairs exclusively, while the stable matching output by the approximation algorithm contains (1,3)-pairs only. We will now present an algorithm that either delivers a stable solution M' containing at least a significant amount of the (1,2)-pairs in M_{egal} or a certificate that M_{egal} contains only a few (1,2)-pairs and thus any stable solution is a good approximation.

To simplify our proof, we execute some basic pre-processing of the input graph. If there are any (1,1)-pairs in G, then these can be fixed, because they occur in every stable matching and thus can only lower the approximation ratio. Similarly, if an arbitrary stable matching contains a (3,3)-pair, then this edge appears in all stable matchings and thus we can fix it. Those (3,3)-pairs that do not belong to the set of stable edges can be deleted from the graph.

From this point on, we assume that no edge is ranked first or last by both of its end vertices in G and prove the approximation ratio for such graphs.

Take the following weight function on all $uv \in E$:

$$w(uv) = \begin{cases} 0 & \text{if } uv \text{ is a (1,2)-pair,} \\ 1 & \text{otherwise.} \end{cases}$$

We designed $w(uv)$ to fit the necessary U-shaped condition of Teo and Sethuraman's 2-approximation algorithm [25, 26]. This condition on the weight function is as follows. We are given a function f_p on the neighbouring edges of a vertex p. Function f_p is *U-shaped* if it is non-negative and there is a neighbour q of p so that f_p is monotone decreasing on neighbours in order of p's preference until q, and f_p is monotone increasing on neighbours in order of p's preference after q. The approximation guarantee of Teo and Sethuraman's algorithm holds for an edge weight function $w(uv)$ if for every edge $uv \in E$, $w(uv)$ can be written as $w(uv) = f_u(uv) + f_v(uv)$, where f_u and f_v are U-shaped functions.

Our $w(uv)$ function is clearly U-shaped, because at each vertex the sequence of edges in order of preference is either monotone increasing or it is $(1, 0, 1)$. Since w itself is U-shaped, it is easy to decompose it into a sum of U-shaped f_v functions, for example by setting $f_v(uv) = f_u(uv) = \frac{w(uv)}{2}$ for every edge uv.

Let M denote an arbitrary stable matching and $M^{(1,2)}$ be the set of (1,2)-pairs in a matching M and M_{opt} be a minimum weight stable matching with respect to the weight function $w(uv)$. Since M_{opt} is by definition the stable matching with the largest number of (1,2)-pairs, $|M_{\mathrm{opt}}^{(1,2)}| \geq |M_{\mathrm{egal}}^{(1,2)}|$. We also know that $w(M) = |M| - |M^{(1,2)}|$ for every stable matching M.

Due to Teo and Sethuraman's approximation algorithm [25, 26], it is possible to find a stable matching M' whose weight approximates $w(M_{\mathrm{opt}})$ within a factor of 2. Formally,

$$|M| - |M'^{(1,2)}| = w(M') \leq 2w(M_{\mathrm{opt}}) = 2|M| - 2|M_{\mathrm{opt}}^{(1,2)}|.$$

This gives us a lower bound on $|M'^{(1,2)}|$.

$$|M'^{(1,2)}| \geq 2|M_{\mathrm{opt}}^{(1,2)}| - |M| \geq 2|M_{\mathrm{egal}}^{(1,2)}| - |M| \tag{1}$$

We distinguish two cases from here on, depending on the sign of the term on the right. In both cases, we establish a lower bound on $c(M_{\mathrm{egal}})$ and an upper bound on $c(M')$. These will give the desired upper bound of 9/7 on $\frac{c(M')}{c(M_{\mathrm{egal}})}$.

1) $2|M_{\mathrm{egal}}^{(1,2)}| - |M| \leq 0$

The derived lower bound for $|M'^{(1,2)}|$ is negative or zero in this case. Yet we know that at most half of the edges in M_{egal} are (1,2)-pairs, and $c(e) \geq 4$

for the rest of the edges in M_{egal}. Let us denote $|M| - 2|M_{\text{egal}}^{(1,2)}| \geq 0$ by x. Thus, $|M_{\text{egal}}^{(1,2)}| = \frac{|M|-x}{2}$.

$$c(M_{\text{egal}}) \geq \frac{|M|-x}{2} \cdot 3 + \frac{|M|+x}{2} \cdot 4 = 3.5|M| + 0.5x \qquad (2)$$

We use our arguments in the proof of Claim 7 to derive that an arbitrary stable matching approximates $c(M_{\text{egal}})$ on the $\frac{|M|-x}{2}$ (1,2)-edges within a ratio of $\frac{4}{3}$, while its cost on the remaining $\frac{|M|+x}{2}$ edges is at most 5. These imply the following inequalities for an arbitrary stable matching M.

$$c(M) \leq \frac{|M|-x}{2} \cdot 3 \cdot \frac{4}{3} + \frac{|M|+x}{2} \cdot 5 = 4.5|M| + 0.5x \qquad (3)$$

We now combine (2) and (3). The last inequality holds for all $x \geq 0$.

$$\frac{c(M)}{c(M_{\text{egal}})} \leq \frac{4.5|M| + 0.5x}{3.5|M| + 0.5x} \leq \frac{9}{7}$$

2) $2|M_{\text{egal}}^{(1,2)}| - |M| > 0$

Let us denote $2|M_{\text{egal}}^{(1,2)}| - |M|$ by $\hat{x}$. Notice that $|M_{\text{egal}}^{(1,2)}| = \frac{\hat{x}+|M|}{2}$. We can now express the number of edges with cost 3, and at least 4 in M_{egal}.

$$c(M_{\text{egal}}) \geq 3 \cdot \frac{\hat{x}+|M|}{2} + 4 \cdot \left(|M| - \frac{\hat{x}+|M|}{2}\right)$$
$$= 3.5|M| - 0.5\hat{x} \qquad (4)$$

Let $|M'^{(1,2)}| = z_1$. Then exactly z_1 edges in M' have cost 3. It follows from (1) that $z_1 \geq \hat{x}$. Suppose that $z_2 \leq z_1$ edges in $M'^{(1,2)}$ correspond to edges in $M_{\text{egal}}^{(1,2)}$. Recall that $|M_{\text{egal}}^{(1,2)}| = \frac{\hat{x}+|M|}{2}$. The remaining $\frac{|M|+\hat{x}}{2} - z_2$ edges in $M_{\text{egal}}^{(1,2)}$ have cost at most 4 in M'. This leaves $|M| - |M_{\text{egal}}^{(1,2)}| - (z_1 - z_2) = \frac{|M|-\hat{x}}{2} - z_1 + z_2$ edges in M_{egal} that are as yet unaccounted for; these have cost at most 5 in both M_{egal} and M'. We thus obtain:

$$c(M') \leq 3z_1 + 4\left(\frac{|M|+\hat{x}}{2} - z_2\right) + 5\left(\frac{|M|-\hat{x}}{2} - z_1 + z_2\right)$$
$$= 4.5|M| - 0.5\hat{x} - 2z_1 + z_2$$
$$\leq 4.5|M| - 1.5\hat{x} \qquad (5)$$

Combining (4) and (5) delivers the following bound.

$$\frac{c(M')}{c(M_{\text{egal}})} \leq \frac{4.5|M| - 1.5\hat{x}}{3.5|M| - 0.5\hat{x}} < \frac{9}{7}$$

The last inequality holds for every $\hat{x} > 0$.

We derived that M', the 2-approximate solution with respect to the weight function $w(uv)$ delivers a $\frac{9}{7}$-approximation in both cases. $\qquad \square$

uv	worst case cost at u	worst case cost at v	cost ratio
(1,2)	4	1	5/3
(1,3)	4	2	6/4
(1,4)	4	3	7/5
(2,2)	4	1	5/4
(2,3)	4	2	6/5
(2,4)	4	3	7/6
(3,3)	4	2	6/6
(3,4)	4	3	7/7

Table 2 uv edges and the corresponding costs in EGAL 4-SRI.

Using analogous techniques we can establish similar approximation bounds for EGAL 4-SRI and EGAL 5-SRI, as follows.

Theorem 8 EGAL 4-SRI *is approximable within* 11/7.

Proof. We start with a statement analogous to Claim 7.

Claim 9. *In an instance of* EGAL 4-SRI*, any stable matching approximates* $c(M_{egal})$ *within a factor of* 5/3.

Proof. As earlier, we can fix all (1,1)-pairs and eliminate all (4,4)-pairs from the instance. Table 2 contains all cases for uv edges in M_{egal} and the corresponding costs in an arbitrary stable matching. □

We define the same weight function $w(uv)$ as in the proof of Theorem 6. We remark here that $w(uv)$ remains U-shaped for preference lists of length 4, because at each vertex the sequence of edges in order of preference is either monotone increasing or it is (1,0,1,1). Since we derived Inequality (1) without using the bounded degree property, it holds for EGAL 4-SRI as well. We distinguish two cases based on the sign of $2|M_{\text{egal}}^{(1,2)}| - |M|$.

1) $2|M_{\text{egal}}^{(1,2)}| - |M| \leq 0$

   Let us denote $|M| - 2|M_{\text{egal}}^{(1,2)}| \geq 0$ by x. Thus, $|M_{\text{egal}}^{(1,2)}| = \frac{|M|-x}{2}$. Furthermore, let y denote the number of edges with cost at least 5 in M_{egal}.

   $$c(M_{\text{egal}}) \geq \frac{|M|-x}{2} \cdot 3 + \left(\frac{|M|+x}{2} - y\right) \cdot 4 + 5y$$
   $$= 3.5|M| + 0.5x + y$$

   $$c(M) \leq \frac{|M|-x}{2} \cdot 3 \cdot \frac{5}{3} + \left(\frac{|M|+x}{2} - y\right) \cdot 6 + 7y = 5.5|M| + 0.5x + y$$

   $$\frac{c(M)}{c(M_{\text{egal}})} \leq \frac{5.5|M| + 0.5x + y}{3.5|M| + 0.5x + y} \leq \frac{11}{7}$$

2) $2|M_{\text{egal}}^{(1,2)}| - |M| > 0$

Let $\hat{x}$ denote $2|M_{\text{egal}}^{(1,2)}| - |M|$ and y the number of edges with cost at least 5 in M_{egal}. Due to Inequality (1), we know that at least $\hat{x}$ $(1,2)$-pairs in M_{egal} correspond to edges of cost 3 in M'. The remaining $\frac{|M| - \hat{x}}{2}$ $(1,2)$-pairs in M_{egal} correspond to edges of cost at most 5 in M'.

$$c(M_{\text{egal}}) \geq \frac{\hat{x} + |M|}{2} \cdot 3 + 4 \cdot \left(\frac{|M| - \hat{x}}{2} - y\right) + 5y = 3.5|M| - 0.5\hat{x} + y$$

$$c(M') \leq 3\hat{x} + 5 \cdot \frac{|M| - \hat{x}}{2} + 6 \cdot \left(\frac{|M| - \hat{x}}{2} - y\right) + 7y = 5.5|M| - 2.5\hat{x} + y$$

$$\frac{c(M')}{c(M_{\text{egal}})} \leq \frac{5.5|M| - 2.5\hat{x} + y}{3.5|M| - 0.5\hat{x} + y} < \frac{11}{7} \qquad \square$$

Theorem 10 EGAL 5-SRI *is approximable within* 13/7.

Proof. Again we start with a statement analogous to Claim 7.

Claim 11. *In an instance of* EGAL 5-SRI*, any stable matching approximates* $c(M_{egal})$ *within a factor of* 2.

Proof. As earlier, we can fix all (1,1)-pairs and eliminate all (5,5)-pairs from the instance. Table 3 contains all cases for uv edges in M_{egal} and the corresponding costs in an arbitrary stable matching. $\qquad \square$

We remark that $w(uv)$ remains U-shaped for preference lists of length 5, because at each vertex the sequence of edges in order of preference is either monotone increasing or it is $(1,0,1,1,1)$. We observe that Inequality (1) holds for EGAL 5-SRI as well. Thus we distinguish two cases based on the sign of $2|M_{\text{egal}}^{(1,2)}| - |M|$.

1) $2|M_{\text{egal}}^{(1,2)}| - |M| \leq 0$

Let us denote $|M| - 2|M_{\text{egal}}^{(1,2)}| \geq 0$ by x. Thus, $|M_{\text{egal}}^{(1,2)}| = \frac{|M| - x}{2}$. Furthermore, let y be the number of edges with cost 5 and z the number of edges with cost at least 6 in M_{egal}.

$$c(M_{\text{egal}}) \geq \frac{|M| - x}{2} \cdot 3 + \left(\frac{|M| + x}{2} - y - z\right) \cdot 4 + 5y + 6z$$
$$= 3.5|M| + 0.5x + y + 2z$$

$$c(M) \leq \frac{|M| - x}{2} \cdot 3 \cdot \frac{6}{3} + \left(\frac{|M| + x}{2} - y - z\right) \cdot 7 + 8y + 9z$$
$$= 6.5|M| + 0.5x + y + 2z$$

uv	worst case cost at u	worst case cost at v	cost ratio
(1,2)	5	1	6/3
(1,3)	5	2	7/4
(1,4)	5	3	8/5
(1,5)	5	4	9/6
(2,2)	5	1	6/4
(2,3)	5	2	7/5
(2,4)	5	3	8/6
(2,5)	5	4	9/7
(3,3)	5	2	7/6
(3,4)	5	3	8/7
(3,5)	5	4	9/8
(4,4)	5	3	8/8
(4,5)	5	4	9/9

Table 3 uv edges and the corresponding costs in EGAL 5-SRI.

$$\frac{c(M)}{c(M_{\text{egal}})} \leq \frac{6.5|M| + 0.5x + y + 2z}{3.5|M| + 0.5x + y + 2z} \leq \frac{13}{7}$$

2) $2|M_{\text{egal}}^{(1,2)}| - |M| > 0$

Let $\hat{x}$ denote $2|M_{\text{egal}}^{(1,2)}| - |M|$, y the number of edges with cost 5 and z the number of edges with cost at least 6 in M_{egal}.

$$c(M_{\text{egal}}) \geq \frac{\hat{x} + |M|}{2} \cdot 3 + 4 \cdot \left(\frac{|M| - \hat{x}}{2} - y - z\right) + 5y + 6z = 3.5|M| - 0.5\hat{x} + y + 2z$$

$$c(M') \leq 3\hat{x} + 6 \cdot \frac{|M| - \hat{x}}{2} + 7 \cdot \left(\frac{|M| - \hat{x}}{2} - y - z\right) + 8y + 9z = 6.5|M| - 3.5\hat{x} + y + 2z$$

$$\frac{c(M')}{c(M_{\text{egal}})} \leq \frac{6.5|M| - 3.5\hat{x} + y + 2z}{3.5|M| - 0.5\hat{x} + y + 2z} < \frac{13}{7}$$

$\square$

Using a similar reasoning for each $d \geq 6$, our approach gives a c_d-approximation algorithm for EGAL d-SRI where $c_d > 2$. In these cases the 2-approximation algorithm of Feder [7,8] should be used instead.

3 Solvability and most-stable matchings in d-SRTI

In this section we study the complexity and approximability of the problem of deciding whether an instance of d-SRTI admits a stable matching, and the problem of finding a most-stable matching given an instance of d-SRTI.

We begin by defining two problems that we will be studying in this section from the point of view of complexity and approximability.

Problem 4. SOLVABLE d-SRTI

Input: $\mathcal{I} = \langle G, \mathcal{O} \rangle$, where G is a graph and $\mathcal{O}$ is a set of preference lists, each of length at most d, possibly involving ties.
Question: Is $\mathcal{I}$ solvable?

Problem 5. MIN BP d-SRTI

Input: An instance $\mathcal{I}$ of d-SRTI.
Output: A matching M in $\mathcal{I}$ such that $|bp(M)| = bp(\mathcal{I})$.

We will show that SOLVABLE 3-SRTI is NP-complete and MIN BP 3-SRTI is hard to approximate. In both cases we will use a reduction from the following satisfiability problem:

Problem 6. (2,2)-E3-SAT

Input: $\mathcal{I} = B$, where B is a Boolean formula in CNF, in which each clause comprises exactly 3 literals and each variable appears exactly twice in unnegated and exactly twice in negated form.
Question: Is there a truth assignment satisfying B?

(2,2)-E3-SAT is NP-complete, as shown by Berman et al. [2]. We begin with the hardness of SOLVABLE 3-SRTI.

Theorem 12 SOLVABLE 3-SRTI *is* NP-*complete.*

Proof. Clearly SOLVABLE 3-SRTI belongs to NP. To show NP-hardness, we reduce from (2,2)-E3-SAT as defined in Problem 6. Let B be a given instance of (2,2)-E3-SAT, where $X = \{x_1, x_2, \ldots, x_n\}$ is the set of variables and $C = \{c_1, c_2, \ldots, c_m\}$ is the set of clauses. We form an instance $\mathcal{I} = (G, \mathcal{O})$ of 3-SRTI as follows. Graph G consists of a *variable gadget* for each x_i ($1 \leq i \leq n$), a *clause gadget* for each c_j ($1 \leq j \leq m$) and a set of *interconnecting edges* between them; these different parts of the construction, together with the preference orderings that constitute $\mathcal{O}$, are shown in Figure 3 and will be described in more detail below.

When constructing G, we will keep track of the order of the three literals in each clause of B and the order of the two unnegated and two negated occurrences of each variable in B. Each of these four occurrences of each variable is represented by an interconnecting edge.

A variable gadget for a variable x_i ($1 \leq i \leq n$) of B comprises the 4-cycle $\langle v_i^1, v_i^2, v_i^3, v_i^4 \rangle$ with cyclic preferences. Each of these four vertices is incident to an interconnecting edge. These edges end at specific vertices of clause gadgets. The clause gadget for a clause c_j ($1 \leq j \leq m$) contains 20 vertices, three of which correspond to the literals in c_j; these vertices are also incident to an interconnecting edge.

Due to the properties of (2,2)-E3-SAT, x_i occurs twice in unnegated form, say in clauses c_j and c_k of B. Its first appearance, as the rth literal of c_j ($1 \leq r \leq 3$), is represented by the interconnecting edge between vertex v_i^1 in the variable gadget corresponding to x_i and vertex a_j^r in the clause gadget corresponding to c_j. Similarly the second occurrence of x_i, say as the sth literal of c_k ($1 \leq s \leq 3$) is represented by the interconnecting edge between v_i^3
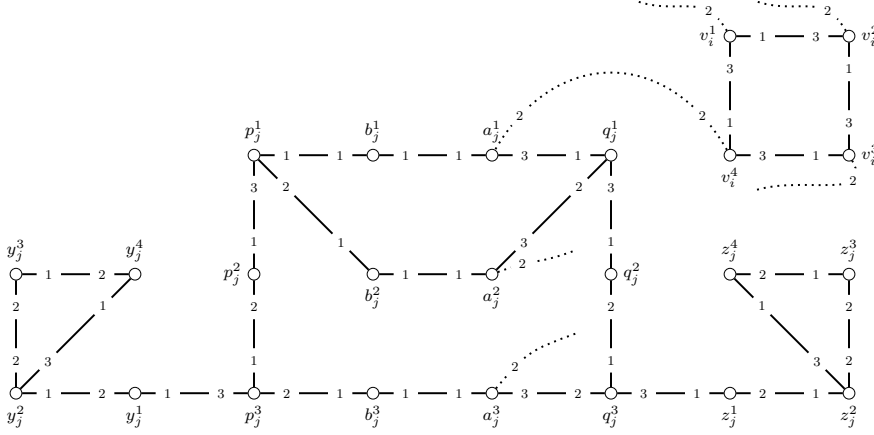
Fig. 3 Clause and variable gadgets for 3-SRTI. The dotted edges are the interconnecting edges. The notation used for edge $a_j^1 v_i^4$ implies that the first literal of the corresponding clause c_j is the second occurrence of the corresponding variable x_i in negated form.

$$M_i^T = \{v_i^1 v_i^2, v_i^3 v_i^4\}$$

$$M_i^F = \{v_i^1 v_i^4, v_i^2 v_i^3\}$$

$$M_j^1 = \{a_j^1 q_j^1, b_j^1 p_j^1, a_j^2 b_j^2, a_j^3 b_j^3, q_j^2 q_j^3, p_j^2 p_j^3, y_j^1 y_j^2, y_j^3 y_j^4, z_j^1 z_j^2, z_j^3 z_j^4\}$$

$$M_j^2 = \{a_j^2 q_j^1, b_j^2 p_j^1, a_j^1 b_j^1, a_j^3 b_j^3, q_j^2 q_j^3, p_j^2 p_j^3, y_j^1 y_j^2, y_j^3 y_j^4, z_j^1 z_j^2, z_j^3 z_j^4\}$$

$$M_j^3 = \{a_j^3 q_j^3, b_j^3 p_j^3, a_j^1 b_j^1, a_j^2 b_j^2, q_j^1 q_j^2, p_j^1 p_j^2, y_j^1 y_j^2, y_j^3 y_j^4, z_j^1 z_j^2, z_j^3 z_j^4\}$$

Fig. 4 The matchings corresponding to variable x_i if it is set to be true and false, respectively, and to the first, second or third literal being true in a fixed clause c_j.

and a_k^s. The same variable x_i also appears twice in negated form. Appropriate a-vertices in the gadgets representing those clauses are connected to v_i^2 and v_i^4. We remark that this construction involves a gadget similar to one presented by Biró et al. [4] in their proof of the NP-hardness of MIN BP 3-SRI.

Now we prove that there is a truth assignment satisfying B if and only if there is a stable matching M in $\mathcal{I}$.

Claim 13. *For any truth assignment satisfying B, a stable matching M can be constructed in $\mathcal{I}$.*

Proof. In Figure 4, we define two matchings, M_i^T and M_i^F, on the variable gadgets and three matchings, M_j^1, M_j^2 and M_j^3, on the clause gadgets.

If a variable x_i ($1 \le i \le n$) is assigned to be true, M_i^T is added to M, otherwise M_i^F is added. Similarly, since at least one literal in c_j ($1 \le j \le m$) is true, let r ($1 \le r \le 3$) be the minimum integer such that the literal at position r of c_j is true; add M_j^r to M. The intuition behind this choice is that if a literal is true, then the vertex representing it in the variable gadget is matched to its best choice. On the other hand, if some literals in a clause are true, then the vertex representing the appearance of one of them in that clause is matched to its last-choice vertex.

We claim that no edge blocks M. Checking the edges in the clause and variable gadgets is easy. The five special matchings were designed in such a way that no edge within the gadgets blocks them. More explanation is needed regarding the interconnecting edges. Suppose one of them, $a_j^r v_i^s$, ($r \in \{1,2,3\}$, $s \in \{1,2,3,4\}$) blocks M. Since M is a perfect matching, a_j^r needs to be matched to its last choice, a q-vertex. Similarly, v_i^s has to be matched to its worst partner. While the partner of a_j^r indicates that the literal represented by v_i^s (x_i or $\bar{x}_i$) is true in the clause, the partner of v_i^s means that the literal is false. $\qquad\square$

Claim 14. *For any stable matching M in $\mathcal{I}$, there is a truth assignment satisfying B.*

Proof. In the next three paragraphs we show that the restriction of M to any variable or clause gadget is one of the above listed special matchings, and no interconnecting edge is in M.

First of all, if a vertex u is the only first choice of another vertex, then u certainly needs to be matched in M. This property is fulfilled for all vertices of all clause gadgets except for y_j^3 and z_j^3 for each c_j ($1 \le j \le m$). Let us first study clause gadget c_j. If y_j^4 is matched to y_j^2, then $y_j^2 y_j^3$ blocks M. Thus, $y_j^3 y_j^4$, and similarly, $z_j^3 z_j^4$ are part of M for all clause gadgets.

Our proof for clause gadgets from this point involves considering matchings covering all twelve remaining vertices. We differentiate two possible cases, depending on the partner of p_j^3. In the first case, $p_j^3 b_j^3 \in M$. Therefore, $p_j^2 p_j^1 \in M$ too, because p_j^2 has to be matched. For similar reasons, $\{b_j^1 a_j^1, b_j^2 a_j^2, q_j^1 q_j^2, q_j^3 a_j^3\} \subseteq M$. This gives us matching M_j^3. In the second case, if p_j^3 is matched to p_j^2, then $\{b_j^3 a_j^3, q_j^3 q_j^2\} \subseteq M$. There are two possible matchings on the remaining six vertices: $\{p_j^1 b_j^1, a_j^1 q_j^1, b_j^2 a_j^2\}$ and $\{p_j^1 b_j^2, q_j^1 a_j^2, b_j^1 a_j^1\}$. These two matchings together with the lower part of the gadget form M_j^1 and M_j^2.

Since all a-vertices have a partner within their clause gadgets, no interconnecting edge can be a part of M. For the variable gadgets, it is straightforward to see that M_i^T and M_i^F are the only matchings covering all vertices of the 4-cycles.

The truth assignment to B is then defined in the following way. Each variable whose gadget has the edges of M_i^T in M is assigned to be true, while all other variables with M_i^F on their gadgets are false.

All that remains is to show that this is indeed a truth assignment. Suppose that there is an unsatisfied clause c_j in B. Since all three of c_j's literals are false, every vertex v_i^r ($1 \le i \le n$) such that $v_i^r a_j^s$ is an interconnecting edge prefers a_j^s to its partner in M ($1 \le s \le 3$). Hence a blocking edge can only be avoided if $a_j^1 b_j^1$, $a_j^2 b_j^2$ and $a_j^3 b_j^3$ are all in M, which never occurs in any stable matching as shown above. $\qquad\square$

This finishes the proof of Theorem 12. $\qquad\square$

Our construction shows that the complexity result holds even if the preference lists are either strictly ordered or consist of a single tie of length two. Moreover, Theorem 12 also immediately implies the following result.

Corollary 15. MIN BP 3-SRTI *is* NP-*hard.*

The following result strengthens Corollary 15.

Theorem 16 MIN BP 3-SRTI *is not approximable within* $n^{1-\varepsilon}$*, for any* $\varepsilon > 0$*, unless* P = NP*, where* n *is the number of agents.*

Proof. The core idea of the proof is to gather several copies of the 3-SRTI instance created in the proof of Theorem 12, together with a small unsolvable 3-SRTI instance. By doing so, we create a MIN BP 3-SRTI instance $\mathcal{I}$ in which $bp(\mathcal{I})$ is large if the Boolean formula B (originally given as an instance of (2,2)-E3-SAT) is not satisfiable, and $bp(\mathcal{I}) = 1$ otherwise. Therefore, finding a good approximation for $\mathcal{I}$ will imply a polynomial-time algorithm to decide the satisfiability of B. Our proof is similar to that of an analogous inapproximabilty result for the problem of finding a most-stable matching in an instance of the Hospitals / Residents problem with Couples [3].

The smallest unsolvable instance of 3-SRTI is a 3-cycle with cyclic strict preferences. Aside from this, we add k disjoint copies of 3-SRTI instance created in the proof of Theorem 12 (from the same Boolean formula B), for large enough k. In particular we let $c = \lceil 2/\varepsilon \rceil$ and $k = n_0^c$, where n_0 is the number of variables in B. We use m_0 to denote the number of clauses in B. Let $\mathcal{I}$ be the instance of 3-SRTI that has been constructed. Due to the proof of Theorem 12 above, if B is satisfiable then $bp(\mathcal{I}) = 1$, and if B is not satisfiable then $bp(\mathcal{I}) \geq k + 1$. Hence a k-approximation algorithm for MIN BP 3-SRTI could be used to solve (2,2)-E3-SAT in polynomial time.

In the remainder of the proof we show that $n^{1-\varepsilon} \leq k$, where n is the number of agents in $\mathcal{I}$, which will imply the statement of the theorem. With Inequalities (6)-(9) we give an upper bound for n. This is used in Inequalities (11)-(14) as we establish k as an upper bound for $n^{1-\varepsilon}$. Explanations for the steps are given as and when it is necessary after each set of inequalities.

$$n = k(4n_0 + 20m_0) + 3 \tag{6}$$

$$= k(4n_0 + 20\frac{4n_0}{3}) + 3 \tag{7}$$

$$\leq 32kn_0 \tag{8}$$

$$= 32n_0^{c+1} \tag{9}$$

In Equality (6) can be deduced by inspection of the 3-SRTI instance constructed in the proof of Theorem 12. In step (7) we substitute $m_0 = \frac{4n_0}{3}$, which follows from the structure of B. We can assume without loss of generality that $kn_0 \geq 3$, which we use in Inequality (8). Finally, in Equality (9) we substitute $k = n_0^c$.

Since $c = \lceil 2/\varepsilon \rceil$, the following inequality also holds.

$$\frac{c-1}{c+1} = 1 - \frac{2}{c+1} \geq 1 - \varepsilon \tag{10}$$

We can now establish the desired upper bound for $n^{1-\varepsilon}$.

$$n^{1-\varepsilon} \leq n^{\frac{c-1}{c+1}} \tag{11}$$

$$\leq 32^{\frac{c-1}{c+1}} n_0^{c-1} \tag{12}$$

$$\leq n_0^c \tag{13}$$

$$= k \tag{14}$$

Inequality (11) is obtained by raising n to the power of each side of Inequality (10). Inequality (12) follows from the bound for n established in Inequalities (6)-(9). Now in Inequality (13) we can assume without loss of generality that $n_0 \geq 32$ and use that $\frac{c-1}{c+1} < 1$. In the last step, we use the definition of k. $\qquad\square$

To complete the study of cases of MIN BP d-SRTI, we establish a positive result for instances with degree at most 2.

Theorem 17 MIN BP 2-SRTI *is solvable in* $\mathcal{O}(|V|)$ *time.*

Proof. For an instance $\mathcal{I}$ of MIN BP 2-SRTI, clearly every component of the underlying graph G is a path or cycle. We claim that $bp(\mathcal{I})$ equals the number of *odd parties* in G, where an *odd party* is a cycle $C = \langle v_1, v_2, ..., v_k \rangle$ of odd length, such that v_i strictly prefers v_{i+1} to v_{i-1} (addition and subtraction are taken modulo k).

Since an odd party never admits a stable matching, $bp(\mathcal{I})$ is bounded below by the number of odd parties [24]. This bound is tight: by taking an arbitrary maximum matching in an odd party component, a most-stable matching is already reached. Now we show that a stable matching M can be constructed in all other components.

Each component that is not an odd cycle is therefore a bipartite subgraph (indeed either a path or an even cycle). Such a subgraph therefore gives rise to the restriction of SRTI called the *Stable Marriage problem with Ties and Incomplete lists* (SMTI). An instance of SMTI always admits a stable solution and it can be found in linear time [22]. Thus these components contribute no blocking edge.

Regarding odd-length cycles that are not odd parties, we will show that there is at least one vertex not strictly preferred by either of its adjacent vertices. Leaving this vertex uncovered and adding a perfect matching in the rest of the cycle results in a stable matching.

Assume that every vertex along a cycle C_k (where k is an odd number) is strictly preferred by at least one of its neighbours. Since each of the k vertices is strictly preferred by at least one vertex, and a vertex v can prefer at most one other vertex strictly, every vertex along C_k has a strictly ordered preference

list. Now every vertex can point at its unique first-choice neighbour. To avoid an odd cycle, there must be a vertex pointed at by both of its neighbours. This implies that there is also a vertex v pointed at by no neighbour, and v is hence ranked second by both of its neighbours.   □

4 Egalitarian stable matchings in SRTI

In this section we outline the difficulties that would be encountered by attempting to define and study the concept of an egalitarian stable matching in instances of SRTI.

- When considering the approximability of EGAL d-SRI, we restricted attention to the case of solvable instances, in the knowledge that solvability can be determined in linear time [15]. However in the case of SRTI, we can no longer assume this, since SOLVABLE 3-SRTI is NP-complete as Theorem 12 shows.
- In instances of EGAL d-SRI, not all agents are necessarily matched in all stable matchings, but due to Theorem 4.5.2 of [12], which states that the same agents are matched in all stable matchings, we can discard unmatched agents and consider only the remaining agents when reasoning about approximation algorithms. There is no analogue of Theorem 4.5.2 in the case of d-SRTI (indeed, stable matchings can be of different sizes in a given instance of SRTI [17]). This means that any approximation algorithm for the problem of finding an egalitarian stable matching in an instance of SRTI would need to consider the cost of an unmatched agent in a given stable matching, and the choice of value for such a case is not universally agreed upon in the literature.
- Similarly in the case of SRTI, the choice of value for the rank of an agent a_j in a given agent a_i's preference list is again not universally agreed upon – for example if a_i has a tie of length 2 at the head of her preference list, followed strictly by a_j, then $\text{rank}(a_i, a_j)$ could reasonably be defined to be either 2 or 3 depending on the definition adopted.

5 Open questions

Theorems 6, 8 and 10 improve on the best known approximation factor for EGAL d-SRI for small d. It remains open to come up with an even better approximation or to establish an inapproximability bound matching our algorithm's guarantee. A more general direction is to investigate whether the problem of finding a minimum weight stable matching can be approximated within a factor less than 2 for instances of d-SRI for small d.

Chen et al. [6] consider different costs for an unmatched agent when finding egalitarian stable matchings in SRTI.

Acknowledgements

References

1. D. J. Abraham, P. Biró, and D. F. Manlove. "Almost stable" matchings in the roommates problem. In *Proceedings of WAOA '05*, vol. 3879 of *Lecture Notes in Computer Science*, pp. 1–14. Springer, 2006.
2. P. Berman, M. Karpinski, and A. D. Scott. Approximation hardness of short symmetric instances of MAX-3SAT. ECCC report, no. 49, 2003.
3. P. Biró, D. F. Manlove, and I. McBride. The Hospitals / Residents problem with couples: complexity and integer programming models. In *Proceedings of SEA '14*, vol. 8504 of *Lecture Notes in Computer Science*, pp. 10–21. Springer, 2014.
4. P. Biró, D. F. Manlove, and E. J. McDermid. "Almost stable" matchings in the roommates problem with bounded preference lists. *Theoretical Computer Science*, 432:10–20, 2012.
5. P. Biró, D. F. Manlove, and S. Mittal. Size versus stability in the marriage problem. *Theoretical Computer Science*, 411:1828–1841, 2010.
6. J. Chen, D. Hermelin, M. Sorge and H. Yedidsion. How hard is it to satisfy (almost) all roommates? ArXiv preprint 1707.04316, 2017.
7. T. Feder. A new fixed point approach for stable networks and stable marriages. *Journal of Computer and System Sciences*, 45:233–284, 1992.
8. T. Feder. Network flow and 2-satisfiability. *Algorithmica*, 11:291–319, 1994.
9. D. Gale and L. S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69:9–15, 1962.
10. M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1:237–267, 1976.
11. D. Gusfield. Three fast algorithms for four problems in stable marriage. *SIAM Journal on Computing*, 16(1):111–128, 1987.
12. D. Gusfield and R. W. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, 1989.
13. D. Gusfield and L. Pitt. A bounded approximation for the minimum cost 2-SAT problem. *Algorithmica*, 8:103–117, 1992.
14. K. Hamada, K. Iwama, and S. Miyazaki. An improved approximation lower bound for finding almost stable maximum matchings. *Information Processing Letters*, 109:1036–1040, 2009.
15. R. W. Irving. An efficient algorithm for the "stable roommates" problem. *Journal of Algorithms*, 6:577–595, 1985.
16. R. W. Irving, P. Leather, and D. Gusfield. An efficient algorithm for the "optimal" stable marriage. *Journal of the ACM*, 34:532–543, 1987.
17. R. W. Irving and D. F. Manlove. The stable roommates problem with ties. *Journal of Algorithms*, 43:85–105, 2002.
18. E. Kujansuu, T. Lindberg and E. Mäkinen. The stable roommates problem and chess tournament pairings. *Divulgaciones Matemáticas*, 7:19–28, 1999.
19. S. Khot and O. Regev. Vertex cover might be hard to approximate to within 2- ε. *Journal of Computer and System Sciences*, 74:335–349, 2008.
20. D. Maier and J. Storer. A note on the complexity of the superstring problem. Technical Report 233, Princeton University, Department of Electrical Engineering and Computer Science, Princeton, NJ, October 1977.
21. D. F. Manlove. *Algorithmics of Matching Under Preferences*. World Scientific, 2013.
22. D. F. Manlove, R. W. Irving, and K. Iwama and S. Miyazaki and Y. Morita Hard variants of stable marriage. *Theoretical Computer Science*, 276:261–279, 2002.
23. E. Ronn. NP-complete stable matching problems. *Journal of Algorithms*, 11:285–304, 1990.

24. J. J. M. Tan. A necessary and sufficient condition for the existence of a complete stable matching. *Journal of Algorithms*, 12:154–178, 1991.
25. C.-P. Teo and J. Sethuraman. LP based approach to optimal stable matchings. In *Proceedings of SODA '97*, pp. 710–719. ACM-SIAM, 1997.
26. C.-P. Teo and J. Sethuraman. The geometry of fractional stable matchings and its applications. *Mathematics of Operations Research*, 23:874–891, 1998.