The Complexity of Rationalizing Matchings

Shankar Kalyanaraman[*]
Computer Science Department
California Institute of Technology
Pasadena, CA 91125
shankar@cs.caltech.edu

Christopher Umans[†]
Computer Science Department
California Institute of Technology
Pasadena, CA 91125
umans@cs.caltech.edu

March 2, 2008

Abstract

Given a set of observed economic choices, can one infer preferences and/or utility functions for the players that are consistent with the data? Questions of this type are called *rationalization* or *revealed preference* problems in the economic literature, and are the subject of a rich body of work.

From the computer science perspective, it is natural to study the complexity of rationalization in various scenarios. We consider a class of rationalization problems in which the economic data is expressed by a collection of matchings, and the question is whether there exist preference orderings for the nodes under which all the matchings are *stable*.

We show that the rationalization problem for one-one matchings is NP-complete. We propose two natural notions of approximation, and show that the problem is hard to approximate to within a constant factor, under both. On the positive side, we describe a simple algorithm that achieves a $3/4$ approximation ratio for one of these approximation notions. We also prove similar results for a version of many-one matching.

1

1 Introduction

Given a set of consumption choices in a market, it is natural to try to infer information about the players' preferences or utility functions. This branch of consumer demand theory is known as *revealed preference theory* because consumers, by dint of the choices they make, "reveal" their preferences for various outcomes [Afr67, Die73, Sam48, Ech06, FST04, Var82, Spr00]. It constitutes a major tool in econometric analysis used to estimate aggregate consumer demand [Afr67, Var06]. From the Computer Science perspective, this is a learning problem, and recent work initiated a study of its PAC-learnability [BV06].

Some classes of data cannot always be explained, or *rationalized* by simple (say, linear) utility functions, or even any reasonable utility function. Such settings are interesting to economists, because it becomes possible, in principle, to "test" various assumptions (e.g. that the players are maximizing a simple utility function). Several (classical and recent) results [Afr67, Var82, FST04, Ech06] in the economic literature establish criteria for when data is *always* rationalizable, thus delineating the limits of the "testable implications" of such data.

There is an important role for Computer Science in these questions, as the feasibility of performing such tests depends on being able to answer the rationalizability question *efficiently*. In other words, given a type of economic data, and a target form for an "explanation" (preference profile, a class of utility functions, etc...), we wish to understand the *complexity* of deciding whether the data can be rationalized by an explanation of the prescribed form. To our knowledge these sort of problems have not been studied before.

Among rationalization problems, one can identify at least two broad classes of problems. Some, such as inferring utility functions from consumption data, are rather easily solved efficiently using linear programming [Afr67, Var82]. Others are more combinatorial in nature, and their complexity is not at all obvious. One recent example is the problem of inferring costs from observations of spanning trees being formed to distribute some service, say power [Özs06].

Among the combinatorial-type rationalization problems, one of the most natural is the matchings problem that we study in this paper. Here we are given a set of bipartite matchings, and we wish to determine if there are preferences for the nodes under which all of the given matchings are stable. Matchings, or more precisely "two-sided matching markets," are a central abstraction in economics, investigated in relation to the similar "marriage models" in auction and labor markets [RS90, Fle03, EO04, EY07] and from the point of view of mechanism design [Sön96] and related strategic issues [STT01]. They are also a fundamental combinatorial abstraction from the computational perspective.

1.1 Our results

Given two sets of nodes, M ("men") and W ("women"), together with preferences for each node, the famous algorithm of Gale and Shapley [GS62] obtains a *stable matching*. We will be interested in the "reverse" question: given a set of matchings, are there preferences under which they are simultaneously stable? One may wonder why we should be given a collection of matchings instead of a single instance of a matching between the set of men and women. Indeed, we think of the men (and women) as representing instances of different *types* or populations that are matched differently in each matching and we are interested in determining the preference profiles that define these types based on the observed set of matchings. Before stating our results, we formalize the problem and introduce some terminology.

Definition 1.1. *Let M, W be disjoint sets of equal cardinality. A **one-one matching** μ is a bijection $\mu : M \cup W \rightarrow M \cup W$, such that for all $m \in M$, $\mu(m) \in W$, for all $w \in W$, $\mu(w) \in M$, and for all $m \in M, w \in W$, $\mu(m) = w \Leftrightarrow \mu(w) = m$.*

In the problems we consider, we will be seeking preferences for the elements of M and W, which are expressed as follows:

Definition 1.2. *A* **preference order** *for $m \in M$ (resp. $w \in W$) is a linear ordering of W (resp. M). We write $m : w > w'$ to mean that w occurs before w' in the preference order for m. A* **preference profile** *is a collection of preference orders for each $m \in M$ and $w \in W$.*

The "stability" of a matching with respect to a preference profile depends on the crucial notion of *blocking pair*:

Definition 1.3. *A* **blocking pair** *with respect to a matching μ and a preference profile $\mathcal{P}$ is a pair $(m, w) : m \in M, w \in W$ such that $\mu(m) \neq w$ and*

$$m : w > \mu(m) \text{ and } w : m > \mu(w).$$

Matching μ is **stable** *with respect to $\mathcal{P}$ if there is no blocking pair with respect to μ and $\mathcal{P}$.*

In other words, in a blocking pair (m, w) with respect to μ and $\mathcal{P}$, both people are "unhappy" with their current partner in μ and would instead prefer to be matched to each other.

Our first result is that rationalizing matchings is hard.

Theorem 1.4. *Given a collection of one-one matchings $\mathcal{H}$ on the sets M and W, it is NP-complete to determine if there exists a preference profile $\mathcal{P}$ such that every $\mu \in \mathcal{H}$ is stable with respect to $\mathcal{P}$.*

We call such a preference profile a *rationalization* of the matchings $\mathcal{H}$. The main gadget we use in the reduction is distilled from some fairly involved necessary and sufficient conditions for a preference profile to be a rationalization, discovered by Echenique [Ech06]. We describe the full conditions in Section 2. Our gadget is a configuration across two matchings, that looks like this:
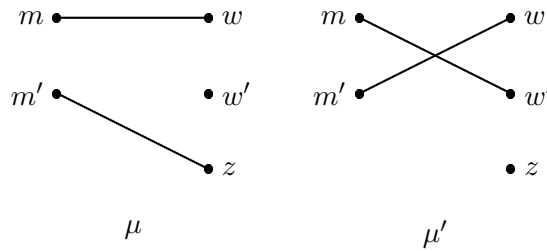
A preference profile $\mathcal{P}$ rationalizes the matchings containing this configuration only if either $m : w > w'$ and $m' : z > w$, or $m : w' > w$ and $m' : w > z$. Conversely, if these conditions hold (together with additional conditions concerning the remainder of the matchings) then $\mathcal{P}$ rationalizes the set of matchings. We use this gadget fundamentally as a Boolean choice gadget (either m prefers w over w' or w' over w), and as part of a scheme to ensure consistency (since the choice of m is tied to the choice of m').

Having ascertained that rationalizing a collection of matchings is NP-complete, we would next want to know how hard it is to solve the problem approximately. In this context, we first need to decide what exactly we mean by 'approximate' rationalization. Two notions are of particular interest: on the one hand, we can think of identifying a preference profile that rationalizes the maximum number of matchings.

Problem 1 (MAX-STABLE-MATCHINGS)**.** *Given a collection of matchings $\mathcal{H}$ on sets M, W, find a preference profile P that maximizes the number of matchings in $\mathcal{H}$ that are simultaneously rationalized by P.*

This problem is hard to approximate to within some constant factor:

Theorem 1.5. *There is a constant $\epsilon > 0$ for which it is NP-hard to approximate* MAX-STABLE-MATCHINGS *to within a factor of $(1 - \epsilon)$.*

A second natural notion of approximation attempts to maximize "stability" among the given set of matchings at a more fine-grained level, by maximizing the number of non-blocking pairs across all matchings.

Some effort is required to make this notion of approximation meaningful. In a typical instance there will be many pairs (m, w) for which m is not matched to w in *any* of the given matchings. We say such a pair is *non-active* and pairs that are matched in some matching are *active*. It is easy to ensure that all non-active pairs are non-blocking pairs with respect to any matching, by requiring the preference profile to be *valid*:

Definition 1.6. *A preference profile $\mathcal{P}$ is* valid *with respect to a collection of matchings $\mathcal{H}$ if for every $m \in M$, $m : w > w'$ if (m, w) is active and (m, w') is not active, and for every $w \in W$, $w : m > m'$ if (m, w) is active and (m', w) is not active.*

In other words, each man m prefers women that he is matched to in some matching over women that he is never matched to, and similarly for each women w. We argue that to have a meaningful notion of maximizing non-blocking pairs, one should consider only valid preference profiles, and therefore attempt to maximize the number of non-blocking pairs *among the active pairs* (since a valid preference profile automatically takes care of all of the non-active pairs). We are led to define the following optimization problem:

Problem 2 (MAX-STABILITY)**.** *Given a collection of matchings $\mathcal{H}$ on sets M, W, find a valid preference profile P for M, W that maximizes:*

$$|\{(m, w, \mu) : (m, w) \text{ is active}$$

$$\text{and is not a blocking pair with respect to } \mu, P\}|.$$

This problem is also hard to approximate to within some constant factor:

Theorem 1.7. *There is a constant $\epsilon > 0$ for which it is NP-hard to approximate* MAX-STABILITY *to within a factor of $(1 - \epsilon)$.*

Our proof uses the overall structure of the reduction used to prove Theorem 1.4 together with an explicit constant-degree expander to make aspects of the reduction robust enough to be gap-preserving.

An approximation of $3/4$ is achievable (in expectation) for this problem by a simple randomized assignment of preferences. Derandomizing via the method of conditional expectations yields:

Theorem 1.8. *There is a deterministic, polynomial-time approximation algorithm for* MAX-STABILITY *that achieves an approximation factor of 3/4.*

Finally, we turn to a generalization of the one-one matchings we have been considering:

Definition 1.9. *Let F, W be disjoint sets. A* **one-many matching** *is a pair of functions (μ, τ) with $\mu : F \to 2^W$, and $\tau : W \to F$ for which*

$$\forall w \in \mu(f), \tau(w) = f \text{ and } \forall w \in W, w \in \mu(\tau(w)).$$

Typically in economics literature, one-to-many matchings are spoken of in reference to firms and workers (or, similarly, hospitals and interns) and hence the notation of F, W is more prevalent. However, since this problem is so closely tied in with our discussion of one-to-one matchings we will continue to use the notation of "men" M and "women" W when we mention one-to-many matchings in the rest of the paper. One-many matching models have been widely studied [Rot82, Rot85].

In a one-many matching, preference order and preference profile are defined in the same way as for one-one matchings, except that each m has a linear ordering of 2^W instead of just W. Also analogous to the blocking pair for one-to-one matchings, we can define a *blocking set* and a notion of stability [EO04] for one-to-many matchings:

Definition 1.10. *A* **blocking set** *with respect to a one-many matching (μ, τ) and a preference profile $\mathcal{P}$ is a pair $(m, B) : m \in M, B \subseteq W$ such that $\mu(m) \cap B = \emptyset$ and*

$$\exists A \subseteq \mu(m) \text{ such that}$$

$$m : A \cup B > \mu(m) \text{ and } \forall w \in B \ \ w : m > \tau(w).$$

Matching (μ, τ) is **stable*** *with respect to $\mathcal{P}$ if there is no blocking set with respect to (μ, τ) and $\mathcal{P}$.*

The rationalization problem for one-many matchings is not likely to even be *in* NP, because a witness (preference profile) entails listing preference over 2^W, which is exponentially large. We are then led to consider a restricted version of the problem in which we only allow $m \in M$ to be matched to a set of cardinality at most some constant parameter ℓ. We call such matchings **one-ℓ** matchings.

The resulting rationalization problem is in NP and, we show, NP-complete:

Theorem 1.11. *For every fixed ℓ, given a collection of one-ℓ matchings $\mathcal{H}$ on the sets M and W, it is NP-complete to determine if there exists a preference profile $\mathcal{P}$ such that every $\mu \in \mathcal{H}$ is stable* with respect to $\mathcal{P}$.*

We can define the notion of an active pair (m, B) for one-ℓ matchings in analogy with active pairs, and also valid preference profiles as in Definition 1.6.

The two approximation problems arising with respect to one-ℓ matchings are hard to approximate to within some constant factor, just as in the one-one case:

Theorem 1.12. *There is a constant $\epsilon > 0$ for which it is NP-hard to approximate* MAX-STABLE-ONE-ℓ-MATCHINGS *to within a factor of $(1 - \epsilon)$.*

Theorem 1.13. *There is a constant $\epsilon > 0$ for which it is NP-hard to approximate* MAX-ONE-ℓ-STABILITY *to within a factor of $(1 - \epsilon)$.*

2 Preliminaries

In this section, we encapsulate the working of the result for one-one matchings due to Echenique [Ech06] and provide the necessary and sufficient conditions for the existence of a preference profile that rationalizes a given collection of matchings. We start with some definitions and notations.

Definition 2.1. *For any two matchings $\mu, \mu' \in \mathcal{H}$, a (μ, μ')-**pivot** is a $w \in W$ such that there exist some $m_k, m_\ell \in M$ such that $\mu(m_k) = \mu'(m_\ell) = w$.*

The key to proving Theorem 1.4 is a result due to Echenique [Ech06] which we encapsulate in Lemma 2.3 which sets down necessary and sufficient conditions for the existence of a preference profile that rationalizes a given collection of matchings. We first introduce some notation that will be necessary to describe Lemma 2.3. Consider the directed graph G_{ij} with M as vertex set and E_{ij} as edge-set where $(m, m') \in E_{ij}$ if $\mu_i(m) = \mu_j(m')$. Let $\mathbf{C}(\mu_i, \mu_j)$ denote the set of all connected components of G_{ij}. We will denote the analogous graph obtained by considering as vertex set W as H_{ij}. The following proposition now follows from our notation and establishes a correspondence between G_{ij} and H_{ij}.

Proposition 2.2. *(Echenique [Ech06]) C is a connected component of G_{ij} iff $\mu_i(C)$ is a connected component of H_{ij}. Furthermore, $\mu_i(C) = \mu_j(C)$.*

Echenique [Ech06] showed the following lemma to be true.

Lemma 2.3. *(Echenique [Ech06]) Let $\mathcal{H} = \{\mu_1, \ldots, \mu_\ell\}$ be rationalized by preference profile $\mathcal{P}$. Consider, for all $\mu_i, \mu_j \in \mathcal{H}$ the graph G_{ij} and all $C \in \mathbf{C}_{ij}$. Then, exactly one of (1) or (2) must be true:*

$$
\begin{aligned}
m : \mu_i(m) > \mu_j(m) \text{ for all } m \in C \text{ and} \\
w : \mu_j(w) > \mu_i(w) \text{ for all } w \in \mu_i(C)
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
m : \mu_i(m) < \mu_j(m) \text{ for all } m \in C \text{ and} \\
w : \mu_j(w) < \mu_i(w) \text{ for all } w \in \mu_i(C)
\end{aligned}
\tag{2}
$$

Conversely, if $\mathcal{P}$ is a preference profile such that for all $\mu_i, \mu_j \in \mathcal{H}$ and $C \in \mathbf{C}(\mu_i, \mu_j)$, exactly one of (1) or (2) holds, then $\mathcal{P}$ rationalizes $\mathcal{H}$.

3 Hardness of rationalizability of matchings

We are given two sets M, W with $|M| = |W| = N$ and a set $\mathcal{H}$ of s matchings $\mu_1, \ldots, \mu_s : M \to W$. We show that the problem of determining whether there exists a preference profile that rationalizes $\mathcal{H}$ is NP-complete by reducing from NAE-3SAT.

3.1 Proof outline

We give below a broad overview of the reduction used to prove Lemma 3.2. Our objective is to start with a set of clauses and construct matchings corresponding to them in such a way that the all-equal assignment to variables in a clause would lead to a conflicting preference relation for some element in the set of matchings. With this in mind, we build 'matching gadgets' corresponding to a given Boolean formula.

By way of example, consider a single clause $C_1 = (x_1 \vee \bar{x}_2 \vee \bar{x}_3)$. We associate with each variable x_i, the elements $m_{1i} \in M_1, w_{1i}, w'_{1i} \in W_1$. We will subsequently pad M_1 with dummy elements to ensure that $|M_1| = |W_1|$. For such a clause, we look up Table 4 (in Appendix **??**) to construct 10 partial matchings $\mu_1, \ldots, \mu_{10}$ involving $M_1 = \{m_{1i} | i = 1, 2, 3\} \cup \{u_1\}$ and $W_1 = \{w_{1i}, w'_{1i} | i = 1, 2, 3\} \cup \{y_1, z_1\}$. Our encoding of the truth assignment to a variable x_i in clause C_1 will then correspond to m_{1i} preferring w'_{1i} over w_{1i}, i.e. $m_{1i} : w'_{1i} > w_{1i}$ iff $x_i = 1$. The claim below gives a flavor of how the entire reduction works.

Claim 3.1. *There exists a rationalizable preference profile for M_1, W_1 for the matchings described in Table 4 iff there exists a not-all-equal satisfying assignment for C_1.*

Proof. (**Sketch**) Suppose there exists a not-all-equal satisfiable assignment to C_1. Then, in order to show that the corresponding preference profile obtained is rationalizable, we will show that it satisfies the conditions in Lemma 2.3. We fix the preference for each m_{1i} between w_{1i} and w'_{1i} based on the assignment to x_i for $i = 1, 2, 3$. We set $m_{1i} : w'_{1i} > w_{1i}$ if $x_i = 1$ and $m_{1i} : w_{1i} > w'_{1i}$ otherwise. Note that since an assignment $(0, 1, 1)$ or $(1, 0, 0)$ to (x_1, x_2, x_3) is ruled out, the matchings in Table 4 ensure that there will be no "cycles" in the preference orders of m_{11}, m_{12}, m_{13}. Furthermore, an assignment to x_1, x_2, x_3 only fixes a preference order for all $m \in M_1$ and so we can fix a preference order for $w \in W_1$ so that there is no conflict in the preference orders for all m, w and that the conditions in Lemma 2.3 are satisfied.

The converse is immediate because for a rationalizable preference profile for $m \in M_1, w \in W_1$, Lemma 2.3 holds and hence an all-equal assignment to C_1 is not allowed. For instance, suppose (x_1, x_2, x_3) were assigned $(0, 1, 1)$ then using Lemma 2.3 to draw up all the preference relations we would obtain a conflict, i.e. $m_{11} : w_{12} > w'_{11}$ (applying Lemma 2.3 to $\mu_{11}, \ldots, \mu_{18}$) and $m_{11} : w_{12} < w'_{11}$ (applying Lemma 2.3 to μ_{19}, μ_{110}). Therefore, setting each of the x_i to the values obtained depending on the preference relation for m_{1i} between w_{1i} and w'_{1i} as delineated above is a not-all-equal satisfying assignment. $\qed$

In a Boolean formula with m clauses, we repeat the exercise above but use disjoint sets M_ℓ, W_ℓ for each clause C_ℓ to avoid conflicting preference orders *across* clauses. This makes it necessary for us to enforce consistency between the preference relations for $m_{\ell i}$ and $w_{\ell i}, w'_{\ell i}$ for all $\ell = 1, \ldots, m$ and the assignment to x_i. To this end, we use additional matching gadgets from Table 5 and an auxiliary element v_i. Again applying Lemma 2.3, we see that for x_1 occurring in clauses C_1, C_2 say, we must have that $m_{11} : w'_{11} > w_{11} \iff m_{21} : w'_{21} > w_{21}$.

Note that in the manner our construction of matching gadgets is set up, it is necessary for our purposes to reduce from NAE-3SAT as opposed to 3SAT because, if an all-false assignment to a clause were to lead to a conflict in preference relation for some m, w, w', then by symmetry an all-true assignment would also lead to a contradictory preference relation.

3.2 Proof of Theorem 1.4.

The proof for Theorem 1.4 automatically follows from Lemma 3.2 which we formally state and prove below.

Lemma 3.2. *Let $\mathcal{Z}$ be an instance of* NAE-3SAT *over n variables $x_1, \ldots, x_n$ and m clauses $C_0, \ldots, C_{m-1}$. Then, there exists an instance $\mathcal{Z}'$ of $O(m)$ matchings between sets M and W, $|M| = |W| = O(m+n)$ such that there exists a rationalizable preference profile for all $m \in M, w \in W$ iff there exists a not-all-equal satisfiable assignment to $x_1, \ldots, x_n$. Furthermore, these matchings can be constructed in polynomial time.*

Proof. Consider a clause C_ℓ involving x_i, x_j, x_k. For C_ℓ, we consider the following sets of men and women: $\mathcal{M}_\ell = M_\ell \cup M'_\ell \cup B_\ell \cup U_\ell \cup V_\ell \cup T_\ell, \mathcal{W}_\ell = W_\ell \cup W'_\ell \cup G_\ell \cup Y_\ell \cup V'_\ell \cup Z_\ell$. Each of M_ℓ, W_ℓ comprises 3 men and women $\{m_{\ell i}, m_{\ell j}, m_{\ell k}\}$ and $\{w_{\ell i}, w_{\ell j}, w_{\ell k}\}$ respectively. The remaining sets are similarly constructed with each containing 3 elements. We then look up the corresponding table from Tables 1 through 4 and construct 10 partial matchings. In addition, we consider the singleton element v_ℓ which is used in matchings in Table 5. Note that each $m \in M_\ell$ corresponds to a variable occurring in C_ℓ. We will use $v_{\ell i}$ to match, say, $m_{\ell i} \in M_\ell$ for consistency in the assignment made to the variable x_i occurring in the first clause $C_r, r > \ell$. This gives rise to 4 matchings for each clause. Let $M = \cup_{\ell=1}^m \mathcal{M}_\ell, W = \cup_{\ell=1}^m \mathcal{W}_\ell$. Furthermore, we will denote $\mathcal{R}(C_\ell)$ to be the set of all matchings μ associated with clause C_ℓ as described above.

We now describe in detail the complete set of matchings between $\mathcal{M}_\ell$ and $\mathcal{W}_\ell$. The idea is to make sure that every element $m \in \mathcal{M}_\ell$ not already matched according to the tables is matched to some $w \in \mathcal{W}_\ell$. We use the following rules:

Table 1: For $C_\ell = (x_i + x_j + x_k), (\bar{x}_i + \bar{x}_j + \bar{x}_k)$:

$\mu_{\ell 1}$:	$(m_i, w_i')\ (m_j, w_i)$
$\mu_{\ell 2}$:	$(m_i, w_i)\ (m_j, y_\ell)$
$\mu_{\ell 3}$:	$(m_j, w_j')\ (m_k, w_j)$
$\mu_{\ell 4}$:	$(m_j, w_j)\ (m_k, z_\ell)$
$\mu_{\ell 5}$:	$(m_k, w_k')\ (u_\ell, w_k)$
$\mu_{\ell 6}$:	$(m_k, w_k)\ (u_\ell, w_j)$
$\mu_{\ell 7}$:	$(u_\ell, w_k)\ (m_i, w_j)$
$\mu_{\ell 8}$:	$(u_\ell, w_j)\ (m_i, w_i')$
$\mu_{\ell 9}$:	$(m_k, z_\ell)\ (m_i, w_j)$
$\mu_{\ell 10}$:	$(m_k, w_j)\ (m_i, w_i)$

Table 2: $C_\ell = (x_i + x_j + \bar{x}_k), (\bar{x}_i + \bar{x}_j + x_k)$

$\mu_{\ell 1}$:	$(m_i, w_i')\ (m_j, w_i)$
$\mu_{\ell 2}$:	$(m_i, w_i)\ (m_j, y_\ell)$
$\mu_{\ell 3}$:	$(m_j, w_j')\ (m_k, w_j)$
$\mu_{\ell 4}$:	$(m_j, w_j)\ (m_k, z_\ell)$
$\mu_{\ell 5}$:	$(m_k, w_k')\ (u_\ell, w_k)$
$\mu_{\ell 6}$:	$(m_k, w_k)\ (u_\ell, w_i')$
$\mu_{\ell 7}$:	$(u_\ell, w_k)\ (m_i, w_i')$
$\mu_{\ell 8}$:	$(u_\ell, w_i')\ (m_i, w_k)$
$\mu_{\ell 9}$:	$(m_k, z_\ell)\ (m_i, w_j)$
$\mu_{\ell 10}$:	$(m_k, w_j)\ (m_i, w_i)$

1. For $m_{\ell i}, \mu(m_{\ell i}) = \phi$, we match $m_{\ell i}$ to $g_{\ell i} \in G_\ell$ and $w_{\ell i}$ to $b_{\ell i} \in B_\ell$.

2. For $m_{\ell' i}, \ell' \neq \ell$ match $m_{\ell' i}$ to $g_{\ell' i} \in G_{\ell'}$ and $w_{\ell' i}$ to $b_{\ell' i} \in B_{\ell'}$. Match $m'_{\ell' i} \in \mathcal{M}'_{\ell'}$ to $w'_{\ell' i} \in \mathcal{W}'_{\ell'}$. Match $u_{\ell' i}$ to $y_{\ell' i}$, $v_{\ell' i}$ to $v'_{\ell' i}$ and $t_{\ell' i}$ to $z_{\ell' i}$.

3. Let $B'_\ell = \{b_{\ell k} | \mu(b_{\ell k}) = \phi\}, G'_\ell = \{g_{\ell r} | \mu(g_{\ell r}) = \phi\}$. Note that by the structure of our matching rules in Tables 1 through 4, $1 \leq |B'_\ell| \leq |G'_\ell| \leq 2$. For each $b_{\ell k} \in B_\ell$ we match to $g_{\ell r} \in G_\ell$ in ascending order of k, r.

4. If after (3), there is some $g_{\ell r} \in G_\ell, \mu(g_{\ell r}) = \phi$ match the first $m'_{\ell k} \in M'_\ell, \mu(m'_{\ell k}) = \phi$ to $g_{\ell r}$.

5. For all $m'_{\ell i} \in M'_\ell, \mu(m'_{\ell i}) = \phi$, match $m'_{\ell i}$ to the first $w'_{\ell j}, \mu(w'_{\ell j}) = \phi$. Similarly with $u_{\ell i}, t_{\ell i}$ and $z_{\ell i}, y_{\ell i}$.

6. Finally, for all $v_{\ell i}, \mu(v_{\ell i}) = \phi$ match $v_{\ell i}$ to $v'_{\ell i}$.

This specifies a complete matching $\mu : M \to W$. We have 10 such matchings for each clause, and at most 4 matchings for each variable in a clause to ensure consistency of assignment. Therefore, the total number of matchings is at most $22m$. The claims below demonstrate how our reduction works.

Table 3: For $C_\ell = (x_i + \bar{x}_j + x_k), (\bar{x}_i + x_j + \bar{x}_k)$

$\mu_{\ell 1}$:	$(m_i, w'_i)\, (m_j, w_i)$
$\mu_{\ell 2}$:	$(m_i, w_i)\, (m_j, y_\ell)$
$\mu_{\ell 3}$:	$(m_j, w'_j)\, (m_k, w_j)$
$\mu_{\ell 4}$:	$(m_j, w_j)\, (m_k, z_\ell)$
$\mu_{\ell 5}$:	$(m_k, w'_k)\, (u_\ell, w_k)$
$\mu_{\ell 6}$:	$(m_k, w_k)\, (u_\ell, w_i)$
$\mu_{\ell 7}$:	$(u_\ell, w_i)\, (m_i, w_k)$
$\mu_{\ell 8}$:	$(u_\ell, w_k)\, (m_i, w_i)$
$\mu_{\ell 9}$:	$(m_k, z_\ell)\, (m_i, w_j)$
$\mu_{\ell 10}$:	$(m_k, w_j)\, (m_i, w'_i)$

Table 4: For $C_\ell = (x_i + \bar{x}_j + \bar{x}_k), (\bar{x}_i + x_j + x_k)$:

$\mu_{\ell 1}$:	$(m_i, w'_i)\, (m_j, w_i)$
$\mu_{\ell 2}$:	$(m_i, w_i)\, (m_j, y_\ell)$
$\mu_{\ell 3}$:	$(m_j, w'_j)\, (m_k, w_j)$
$\mu_{\ell 4}$:	$(m_j, w_j)\, (m_k, z_\ell)$
$\mu_{\ell 5}$:	$(m_k, w'_k)\, (u_\ell, w_k)$
$\mu_{\ell 6}$:	$(m_k, w_k)\, (u_\ell, w_j)$
$\mu_{\ell 7}$:	$(u_\ell, w_k)\, (m_i, w_j)$
$\mu_{\ell 8}$:	$(u_\ell, w_j)\, (m_i, w_i)$
$\mu_{\ell 9}$:	$(m_k, z_\ell)\, (m_i, w_j)$
$\mu_{\ell 10}$:	$(m_k, w_j)\, (m_i, w'_i)$

Table 5: Consistency matching for x_p occurring in clauses C_i, C_j:

μ'_{p1}:	$(m_{ip}, w'_{ip})\, (v_{ip}, w_{ip})$
μ'_{p2}:	$(m_{ip}, w_{ip})\, (v_{ip}, w'_{jp})$
μ'_{p3}:	$(v_{ip}, w_{ip})\, (m_{jp}, w'_{jp})$
μ'_{p4}:	$(v_{ip}, w'_{jp})\, (m_{jp}, w_{jp})$

Claim 3.3. *Suppose there exists a not-all-equal satisfying assignment to an instance in m clauses $C_1, \ldots, C_m$ and n variables $x_1, \ldots, x_n$. Then, there exists a rationalizing preference profile $\mathcal{H}$ for the corresponding instance of matchings between M and W.*

Proof. We construct a valid preference profile and hence will only consider active pairs. Note that by the structure of our reduction setting up the matchings, each $m \in \mathcal{M}_\ell, w \in \mathcal{W}_\ell$ has at most five elements that it is matched to. In order to satisfy conditions in Lemma 2.3 we will construct these preference orders so that for every active pair, one of (1) or (2) holds.

Note that the only connected components possible in any graph $G_{\mu_1 \mu_2}$ constructed from matchings μ_1, μ_2 are either a cycle or a self-loop (when an element m is matched to the same w in both μ_1 and μ_2).

Consider the variable x_j and the set of matchings μ, μ' where $m_{\ell j}$ is matched to $w_{\ell j}$ and $w'_{\ell j}$ respectively. Note that by consequence of our construction of the matchings, for any element $m \in \mathcal{M}_\ell$ (resp. $w \in \mathcal{W}_\ell$) m (resp. w) occurs in a cycle in only those graphs involving at least one of μ, μ'. For all other such pairs of matchings, m occurs in a self-loop because m is connected to the same element in both such matchings. We look at the graph $G_{\mu \mu'}$.

For a cycle C in $G_{\mu \mu'}$ involving $m_{\ell j}$, the preference order is dictated by x_j's assignment: $x_j = 1 \Leftrightarrow m_{\ell j} : w'_{\ell j} > w_{\ell j}$. To satisfy Lemma 2.3, we will ensure in the preference order for all elements m occurring in C that $m : \mu'(m) > \mu(m)$ and similarly, for all elements w occurring in $\mu(C)$ in the graph $H_{\mu \mu'}$ that $w : \mu(w) > \mu'(w)$.

A preference order constructed as above will lead to a conflict in two possible ways. Firstly, there may exist a blocking pair (m, w) for some μ. Since our preference profile is a valid preference profile, there must exist some μ' such that $\mu'(m) = w$. Then, w is a (μ, μ')-pivot for m and $\mu(w) = m'$ say. But we ensured that for such a pair of matchings (μ, μ') either $m : w > \mu(m)$ and $w : \mu'(w) > m$ or $m : \mu(m) > w$ and $w : m > \mu'(w)$ and hence (m, w) cannot be a blocking pair.

Secondly, there may exist some $m \in \mathcal{M}_\ell$ (resp. $w \in \mathcal{W}_\ell$) for which some preference is contradictory, i.e. for instance when $m : w > w'$ and $m : w' > w$. For a not-all-equal satisfiable assignment to any clause C_ℓ containing x_j, it is easy to check given Tables 1 through 5 exhaustively amongst all w that m can be matched to that this is not the case. Furthermore, since each clause C_ℓ has a different set of $\mathcal{M}_\ell, \mathcal{W}_\ell$ from which the matchings are constructed, no contradictory preference order exists *across* any two clauses.

Finally, we remark since we wish to construct a valid preference profile, for all elements w for which (m, w) is not active, our preference order for m will have $m : w' > w$ for all w' such that (m, w') is active. This completes the proof of the claim. $\qquad \square$

Claim 3.4. *Let $\mathcal{H}$ be a rationalizing preference profile for the above instance of matchings. Then, the assignment*

$$x_i = \begin{cases} 1 & \forall \ell, m_{\ell i} : w'_{\ell i} > w_{\ell i} \\ 0 & otherwise. \end{cases}$$

for all i is a not-all-equal satisfying assignment.

Proof. We first point out that the consistency matchings involving v_ℓ and $m_{\ell i}, i = 1, \ldots, n$ ensure that any rationalizing preference profile $\mathcal{H}$ must satisfy either $(m_{\ell i} : w'_{\ell i} > w_{\ell i})$ or $(m_{\ell i} : w_{\ell i} > w'_{\ell i})$ *for all* $\ell = 1, \ldots, m$. This means that a truth assignment to $x_1, \ldots, x_n$ will be consistent in all clauses $C_1, \ldots, C_m$.

Consider an arbitrary clause C_ℓ. We show that if $\mathcal{H}$ is a rationalizing preference profile, then it is not possible to have an all-equal assignment made to variables in some C_ℓ. Suppose, by way of contradiction

that there existed such an assignment. Depending on the order and number of variables that appear negated in C_ℓ, we look up one of Tables 1 through 4. Then, as illustrated in Claim 3.1, we would obtain a conflict in preference orders for some m thereby giving a contradiction. $\qquad\square$

This completes the proof of the Lemma. $\qquad\square$

4 Hardness of approximate rationalizability of matchings

Our next step in exploring the computational aspects of rationalizability of matchings will be to look at the complexity of 'approximate' rationalizability.

4.1 Maximizing the number of rationalizable matchings

In the first setting, we wish to maximize the number of matchings that can be completely rationalized as stable by a preference profile. Theorem 1.5 states that this is hard to approximate within a constant factor.

Theorem 1.5 (restated). *There is a constant $\epsilon > 0$ for which it is NP-hard to approximate* MAX-STABLE-MATCHINGS *to within a factor of $(1 - \epsilon)$.*

To prove the theorem we show that it is NP-hard to rationalize any fixed set of matchings as captured in the lemma below.

Lemma 4.1. *Given a collection of matchings $\mathcal{H} = \{\mu_1, \ldots, \mu_k\}$ between M and W where k is some fixed constant, it is NP-hard to determine if there exists preferences for $m \in M, w \in W$ for which each of $\mu \in \mathcal{H}$ is a stable matching.*

In order to prove Lemma 4.1 we proceed as before by reducing from NAE-3SAT but we will use a special variant of the NAE-3SAT problem: NAE-3SAT(k') which has the property that every variable in the Boolean formula occurs in exactly k' clauses where $k' \geq 29$ is a constant. Lemma 4.2 below captures this reduction.

Lemma 4.2. *Let $\mathcal{Z}$ be an instance of NAE-3SAT(k') over n variables $x_1, \ldots, x_n$ and m clauses $C_0, \ldots, C_{m-1}$ where k' is some fixed constant. Then, there exists an instance $\mathcal{Z}'$ of $(10+k')$ matchings between sets M and W, $|M| = |W| = O(m+n)$ such that there exists a rationalizable preference profile for all $m \in M, w \in W$ iff there exists a not-all-equal satisfiable assignment to $x_1, \ldots, x_n$. Furthermore, these matchings can be constructed in polynomial time.*

The following claim is key to proving Lemma 4.2.

Claim 4.3. *Let $M_1, \ldots, M_k; W_1, \ldots, W_k$ be respectively k disjoint sets of men and women and $\mu_1, \ldots, \mu_k$ a collection of matchings with $\mu_i : M_i \cup W_i \to M_i \cup W_i$. There exists a set of preference orders $\mathcal{P}$ for M_i, W_i for $i = 1, \ldots, k$ that rationalizes $\mu_1, \ldots, \mu_k$ iff there exists a set of preference orders $\mathcal{P}'$ that rationalizes μ, where $\mu : M_1 \cup \ldots M_k \cup W_1 \cup \ldots W_k \to M_1 \cup \ldots M_k \cup W_1 \cup \ldots W_k$ is the matching obtained by setting $\mu(m) = \mu_i(m)$ for all $m \in M_i$; $i = 1, \ldots, k$.*

Proof. Suppose there exists a preference profile $\mathcal{P}$ for $\mu_1, \ldots, \mu_k$. Then, for μ we construct $\mathcal{P}'$ by assigning for $m \in M_i$ as sub-ordering over W_i, the corresponding preference order for m in $\mathcal{P}$. We complete the preference order for m by ranking all other $w \in W_1 \cup \ldots W_{i-1} \cup W_{i+1} \ldots W_k$ below the sub-ordering for $w \in W_i$. Conversely, for every preference order corresponding to some $m \in M_i$ in a rationalizing preference profile $\mathcal{P}'$ for μ, we obtain a preference order in $\mathcal{P}$ for μ_i by restricting the order over only W_i. $\qquad\square$

Proof. (Of Lemma 4.2) The proof essentially follows the same technique as that of Lemma 3.2 except that we need to be careful in our reduction to maintain the number of matchings at a constant. To this end, we revisit the matchings shown in Tables 1-4. Note that in our previous reduction, we required disjoint sets of M_ℓ, W_ℓ to correspond with each clause C_ℓ. Furthermore, each clause corresponds to 10 matchings (not counting those required to ensure consistency). The following proposition allows us to maintain the overall number of matchings at a constant by merging each of the 10 matchings *across* all the clauses $C_0, \ldots, C_{m-1}$.

Claim 4.3 tells us that we can merge all $\mu_{11}, \mu_{21}, \ldots, \mu_{m1}$ into one single matching μ'_1. We repeat this for all matchings $\mu_{\ell i}, \ell = 1, \ldots, m; i = 2, \ldots, 10$ to obtain 10 new matchings $\mu'_1, \ldots, \mu'_{10}$.

We now focus on the consistency matchings. We will exploit the fact that each variable x_i in Z occurs in at most k' clauses. Therefore, each x_i will correspond to at most $2k'$ matchings. Appealing once again to Claim 4.3, we can merge each of these matchings into a collection of $2k'$ matchings because each x_i is associated to a disjoint set of 'linking' elements $v_{i1}, \ldots, v_{ik'}$. Claims 3.3 and 3.4 go through with their proofs unchanged. This completes the proof for Lemma 4.2, and consequently Lemma 4.1. $\square$

From Lemma 4.1 it follows that it is NP-hard to approximate MAX-STABLE-MATCHINGS for $\mathcal{H}$ to within a factor of $(1 - \epsilon)$ where $\epsilon = 1/(k + 1)$.

Note that given a collection $\mathcal{H}$ of any two matchings, it is trivial to construct a (valid) preference profile that rationalizes $\mathcal{H}$ by arbitrarily assigning a preference for each element in M matched to W in one matching over the other and correspondingly assigning the reverse preference for elements in W.

4.2 Maximizing the number of non-blocking pairs

We look at the MAX-STABILITY problem. The motivation in considering this problem as a notion of approximate rationalizability is that we are now striving to ensure that given a collection of matchings between two sets M and W, there are optimally many different pairs (m, w) for which at least one of them is happy with their current partner and has no incentive to be matched to the other.

As a preliminary exercise, we ask how well would a simple randomized assignment of preferences to $m \in M, w \in W$ perform. It turns out that this would achieve a $3/4$-approximate solution. This is the content of Theorem 1.8.

Theorem 1.8 (restated). *There is a deterministic, polynomial-time approximation algorithm for* MAX-STABILITY *that achieves an approximation factor of 3/4.*

Proof. Note that since we are only interested in finding valid preference profiles, we will automatically accord the least preference for all $w \in W$ that $m \in M$ is not matched with in any of the matchings. Subsequently, each such (m, w) is by default a stable pair and is excluded from our estimation. Let P denote the total number of all remaining pairs for which no preference has been allocated as yet.

We start with an equivalent formulation of the problem. We are given sets M, W such that $|M| = |W| = n$, and a collection $\mathcal{H}$ of ℓ matchings $\mu_1, \ldots, \mu_\ell$. For some m (similarly, w), we associate a 'rank' function $r_m : W \to [n]$ (similarly, $r_w : M \to [n]$) which would completely describe m's (similarly w's) preference order with $r_m(w) < r_m(w')$ implying that $m : w > w'$. A pair (m, w) then is stable for some μ if either $r_m(w) > r_m(\mu(m))$ or $r_w(m) > r_w(\mu(w))$ is true. Let $S = \{(m, w, \mu)|(m, w) \text{ is a stable pair for } \mu\}$. Our objective then is to maximize $|S|$.

Consider the following scheme: for each $m \in M, w \in W$ we construct the rank function by assigning ranks uniformly at random to all $w' \in W$ and $m' \in W$ respectively. The probability that a pair (m, w) is stable for μ is $3/4$ and hence, the expected number of stable pairs denoted by $E[|S|]$ is $3P/4$. Furthermore,

we can easily derandomize this scheme by the method of conditional expectations. At every step when a value is to be assigned to $r_m(w)$, we can efficiently calculate the conditional expectation of the number of non-blocking pairs given the previous values assigned to all $w \in W$ and all other rank functions $r_{m'}$ (there are only a polynomial number of $w \in W$ and rank functions $r_{m'}$ to consider) and fix $r_m(w)$ to be the value that maximizes the conditional expectation. □

It suffices to mention here that a simple randomized preference order for all $m \in M, w \in W$ achieves the $3/4$-approximation factor in expectation and can subsequently be derandomized. How much better can we do than just a random assignment of preferences? Theorem 1.7 tells us that a constant-factor approximation is all we can hope for.

Theorem 1.7 (restated). *There is a constant $\epsilon > 0$ for which it is NP-hard to approximate* MAX-STABILITY *to within a factor of $(1 - \epsilon)$.*

To prove the theorem, we once again construct matchings corresponding to each clause in MAX-NAE-3SAT instance Z. Recall that in proving Lemma 3.2 we needed to construct auxiliary matchings to ensure consistency of assignment to the variables in accordance with the preferences of the corresponding elements in the matchings. To prove hardness of approximation, we will need to establish a gap-preserving reduction by boosting the *robustness* of these consistency gadgets. We do so by augmenting the number of matchings corresponding to the consistency and argue subsequently that if there exists a preference profile that achieves at least a $(1 - \epsilon')$ fraction of stable pairs, then there exists an assignment that would satisfy at least a $(1 - \epsilon)$ fraction of the clauses. Theorem 1.7 then follows from the following Lemma:

Lemma 4.4. *Let Z be an instance of* MAX-NAE-3SAT *over n variables $x_1, \ldots, x_n$ and m clauses $C_0, \ldots, C_{m-1}$ where k' is some fixed constant. Then, there exists a $\epsilon' < 1$ and a polynomial time reduction to an instance Z' of* MAX-STABILITY *of matchings between sets M and W, $|M| = |W| = O(m)$ such that the following is true:*

$$opt(Z) = 1 \implies opt(Z') = 1 \tag{3}$$
$$opt(Z) < 1 - \epsilon \implies opt(Z') < 1 - \epsilon' \tag{4}$$

Proof. The reduction is similar to what we used to prove Lemma 3.2. We set up matchings corresponding to the clauses $C_0, \ldots, C_{m-1}$ as before, but now we need to work harder to boost the robustness of the consistency gadgets. Previously, we used Table 5 to construct additional matchings using auxiliary elements to 'link' different copies of $m_{ji}; j = 1, \ldots, m$ corresponding to a single variable x_i. It will help to conceptualize this as a graph.

For a variable x_i which occurs in some t clauses $C_{j_1}, \ldots, C_{j_t}$, we associate elements from $M, m_{j_1 i}, \ldots, m_{j_t i}$ and define the consistency graph for x_i, G_i to comprise vertex set $V_i = \{m_{j_1 i}, \ldots, m_{j_t i}\}$. An edge exists between any two vertices $(m_{j_p i}, m_{j_q i})$ if they are 'linked' together by an auxiliary element.

Then, the consistency matchings described above in Lemma 3.2 correspond to a path in G_i. In order to boost the robustness, we will now replace the path in G_i by a constant-degree expander graph on t vertices. We make use of the edge expansion notion to define an expander graph: an (n, d, λ) expander graph is a d-regular graph on n vertices with the property that $|\partial(Y)|/|Y| \geq d(1 - \lambda)/2$ where $Y \subseteq V_i, |Y| \leq |V_i|/2$, $\partial(Y)$ is the set of all edges with exactly one end-point in Y and λ is the spectral expansion parameter of the graph. In particular, the following lemma will be useful (the proof can be found in [DH05]):

Lemma 4.5. *For a (t, d, λ) expander graph G and all $\delta \leq (1 - \lambda)/12$, upon removing $2\delta t$ vertices from G, there exists a connected component of size at least*

$$\left(1 - \frac{4\delta}{1 - \lambda}\right) t$$

Note that the total number of occurrences of variables in all the clauses is at most $3m$, and further, that in each clause a variable corresponds to an element m matched to at most an $O(1)$ elements in W. Therefore, the total number of pairs for which a matching exists is at most $O(m)$. Since we only consider valid preference profiles, this means that the number of active pairs under consideration is also $O(m)$ say. Additionally, the total number of auxiliary elements required to construct the expander graphs in the consistency gadgets is also at most $O(m)$ and hence $|M| = O(m)$.

Claim 3.3 from earlier goes through unchanged since our reduction is unchanged in how a satisfying assignment will correspond to a rationalizing preference profile (and hence, all stable pairs). It remains to show that (4) holds.

We shall show that if there is a valid preference profile for Z' such that there are at most an ϵ' fraction of blocking pairs, then there exists an assignment that fails to satisfy at most ϵ fraction of clauses in Z.

Suppose that there is a valid preference profile that allows at most $\epsilon'm$ blocking pairs. Note that if a pair (m, w) is a blocking pair for some matching μ, then Lemma 2.3 breaks down for μ. Since each matching in Z' can be identified with a clause, a blocking pair could result in the clause being unsatisfied.

For a blocking pair (m, w) for some matching μ in our reduction, we evaluate how many clauses are affected. Suppose μ corresponds to one of the matchings for clause C_ℓ. If $m \in M_\ell$ then m must be associated with some variable x_i occurring in C_ℓ, and we will label C_ℓ unsatisfiable. Otherwise, (m, w) has no effect on the satisfiability of C_ℓ.

Suppose μ corresponds to a matching constructed to ensure consistency. If $m \in M_\ell$ for some clause C_ℓ and x_i, then we delete the node $m_{\ell i}$ in G_i and as before label C_ℓ as unsatisfiable. However, now we also need to argue that (m, w) does not cause too many other clauses to be labeled unsatisfiable.

From Lemma 4.5 we know that deleting at most a constant fraction of vertices from G_i will result in a connected component of size at least $(1 - \frac{4\delta}{(1-\lambda)})t$. Taking the aggregate for every variable x_i and after deleting at most $\epsilon'm$ vertices from all the consistency graphs G_i together, the total sum of the largest connected components amongst all G_i will be some $(1 - \epsilon)m$ where ϵ is determined by ϵ', λ and the total number of occurrences of all variables in all the clauses. Therefore, at most ϵm of these occurrences in clauses will be discarded and the corresponding ϵm clauses labeled as unsatisfiable. $\qquad\square$

MAX-NAE-3SAT is known to be APX-complete [PY91] and not approximable to within 0.917 [Zwi98].

5   Rationalizing one-many matchings

For the generalized instance of rationalizing one-many matchings, the problem seems considerably harder. To begin with, since the preference order for any $m \in M$ is over 2^W, given sets of length n, expressing the preference order alone takes exponential time.

However, for a specific restriction of the problem where we allow $m \in M$ to be matched with at most ℓ elements $w \in W$ the problem is in NP and, in fact, NP-complete.

Theorem 1.11 (restated). *For every fixed ℓ, given a collection of one-ℓ matchings $\mathcal{H}$ on the sets M and W, it is NP-complete to determine if there exists a preference profile $\mathcal{P}$ such that every $\mu \in \mathcal{H}$ is stable* with respect to $\mathcal{P}$.*

Proof. Let Z be an instance of a collection $\mathcal{H}_Z = \{\mu_1, \ldots, \mu_r\}$ of one-to-one matchings between M_Z and W_Z. We need to construct an instance Z' of many-to-one matchings such that a stable preference profile for Z exists iff a stable* preference profile exists for Z'. Indeed, we show that $Z' = Z$ is itself such an instance. In other words, $M_{Z'} = M_Z; W_{Z'} = W_Z; \mathcal{H}_{Z'} = \mathcal{H}_Z$.

Claim 5.1. *Suppose there exists a stable preference profile for Z, then there exists a stable* preference profile for Z'.*

Proof. A stable preference profile for Z gives preference orders for all $m \in M_Z$ ($w \in W_Z$) over $w \in W_Z$ ($m \in M_Z$). Consider the following preference profile for Z': for each $m \in M_{Z'}$, we construct a preference order over all $B \subseteq W_{Z'}$ where $|B| \leq \ell$ as follows: we look at all singleton sets $B \subseteq W_{Z'}$ and affix preferences identical to the preference order for $m \in M_Z$ over $w \in W_Z$. Therefore, for $m \in Z'$, $m : w_1 > w_2 \Leftrightarrow$ for $m \in Z$, $m : \{w_1\} > \{w_2\}$. We fix preference for all other subsets $B \subseteq W_{Z'}$ below the singleton sets and in some consistent order (say lexicographic). It is not hard to see that by virtue of our construction, the preference profile outlined above for $m \in M_{Z'}$ is stable* if the corresponding preference profile for $m \in M_Z$ is stable. $\square$

Claim 5.2. *If there exists a stable* preference profile for Z', then there exists a stable preference profile for Z.*

Proof. We construct the preference order for $m \in M_Z$ as follows: we look at the preference order of the corresponding $m \in M_{Z'}$ and extract the partial order comprising m's preference for all $\{w\} \subseteq W_{Z'}$. Suppose that there is a blocking pair (m', w') in Z. Then, this would imply that $(m', \{w'\})$ is a blocking set in Z' which is a contradiction. $\square$

Claims 5.1 and 5.2 give us Theorem 1.11. $\square$

Given how the two problems of rationalizability are so naturally related, it is not surprising then to observe that the one-ℓ matchings problem would have a similar hardness of approximation performance with respect to both analogs of the optimization problem in the case of the one-one matchings.

Theorem 1.12 (restated). *There is a constant $\epsilon > 0$ for which it is NP-hard to approximate* MAX-STABLE-ONE-ℓ-MATCHINGS *to within a factor of $(1 - \epsilon)$.*

The proof follows immediately by combining Lemma 4.1 and Theorem 1.11.

Theorem 1.13 (restated). *There is a constant $\epsilon > 0$ for which it is NP-hard to approximate* MAX-ONE-ℓ-STABILITY *to within a factor of $(1 - \epsilon)$.*

The theorem follows from the lemma below.

Lemma 5.3. *Let Z be an instance of the* MAX-STABILITY *problem for a collection of matchings. Then, there exists an $\epsilon < 1$ and a polynomial-time reduction to an instance Z' of* MAX-ONE-ℓ-STABILITY *of one-ℓ matchings such that the following is true:*

$$opt(Z) = 1 \implies opt(Z') = 1$$
$$opt(Z) < 1 - \epsilon \implies opt(Z') < 1 - \epsilon$$

Proof. As in proving Theorem 1.11, we will use exactly Z as our instance for the MAX-ONE-ℓ-STABILITY problem. This means automatically that

$$opt(Z) = 1 \implies opt(Z') = 1$$

Note that we are looking at valid preference profiles. Since Z' matches all $m \in M$ exclusively to singleton elements in 2^W, these singleton elements are assigned preference over subsets $B \subseteq W, |B| \geq 2$. Hence, our estimate of the optimal number of stable sets will only include the pairs $(m, \{w\})$ which is the same as the optimal number of stable pairs in Z.

Suppose there exists a valid preference profile for Z' for which there are at most ϵ fraction of blocking sets. Then, each of these blocking sets also corresponds exactly to a blocking pair in Z and there cannot be any blocking pair in Z that does not have an equivalent blocking set in Z' for the same reasons as mentioned above in proving Theorem 1.11. Therefore, there are at most ϵ fraction of blocking pairs in Z hence giving us a contradiction and completing the proof to the lemma and the theorem. $\square$

6 Conclusions and Future work

There are many interesting opportunities for extensions to our work on the rationalization problem for matchings. It would be interesting to tighten the constant factor in Lemma 4.1: is it hard even to rationalize three matchings? It would also be satisfying to tighten the hardness of approximation result in Theorem 1.7. We can additionally look at other (restricted) variants of the matchings problem such as many-many matchings and pose the related complexity questions.

On a more general note, the question of rationalizability per se is very tantalizing because of the mutually interesting perspectives it offers within both economics and theoretical computer science.

Acknowledgments. We are indebted to Federico Echenique for numerous invaluable discussions and for getting us started on this work.

References

[Afr67] S.N. Afriat. The Construction of Utility Functions from Expenditure Data. *International Economic Review*, 8(1):67–77, 1967.

[BV06] E. Beigman and R. Vohra. Learning from revealed preference. In *ACM Conference on Electronic Commerce*, pages 36–42, 2006.

[DH05] C. Dwork and P. Harsha. Expanders in Computer Science (CS369E) – Lecture 5, 2005. Stanford University.

[Die73] E. Diewert. Afriat and Revealed Preference Theory. *Review of Economic Studies*, 40(3):419–425, 1973.

[Ech06] F. Echenique. What matchings can be stable? The testable implications of matching theory. Technical Report 1252, California Institute of Technology Social Science Working Papers, 2006.

[EO04] F. Echenique and J. Oviedo. Core many-to-one matchings by fixed-point methods. *Journal of Economic Theory*, 115(2):358–376, April 2004.

[EY07]   F. Echenique and M.B. Yenmez. A solution to matching with preferences over colleagues. *Games and Economic Behavior*, 59(1):46–71, 2007.

[Fle03]   T. Fleiner. A fixed-point approach to stable matchings and some applications. *Math. Oper. Res.*, 28(1):103–126, 2003.

[FST04]   A. Fostel, H.E. Scarf, and M.J. Todd. Two new proofs of Afriat's theorem. *Economic Theory*, 24(1):211–219, 2004.

[GS62]   D. Gale and L. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.

[Özs06]   H. Özsoy. A characterization of Bird's rule, 2006. Job market paper.

[PY91]   C.H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *Journal of Computing Systems and Sciences*, 43(3):425–440, 1991.

[Rot82]   A.E. Roth. The economics of matchings: stability and incentives. *Math Operations Research*, 7:617–628, 1982.

[Rot85]   A.E. Roth. The college admissions problem is not equivalent to the marriage problem. *Journal of Economic Theory*, 35:277–288, 1985.

[RS90]   A.E. Roth and M.A. Sotomayor. *Two-sided matching: A Study in Game-Theoretic Modeling and Analysis*, volume 18 of *Econometric Society Monographs*. Cambridge University Press, Cambridge, UK, second edition, 1990.

[Sam48]   P.A. Samuelson. Consumption Theory in terms of Revealed Preference. *Economica*, 15(60):243–253, 1948.

[Sön96]   T. Sönmez. Strategy-proofness in many-to-one matching problems. *Economic Design*, 3:365–380, 1996.

[Spr00]   Y. Sprumont. On the Testable Implications of Collective Choice Theories. *Journal of Economic Theory*, 93:205–232, 2000.

[STT01]   J. Sethuraman, C.-P. Teo, and W.-P. Tan. Gale-Shapley stable marriage revisited: strategic issues and applications. *Management Science*, 47(9):1252–1267, 2001.

[Var82]   H.R. Varian. The Nonparametric Approach to Demand Analysis. *Econometrica*, 50(4):945–973, 1982.

[Var06]   H.R. Varian. Revealed Preference. In M. Szenberg, L. Ramrattan, and A.A. Gottesman, editors, *Samuelson Economics and the Twenty-First Century*, chapter 6, pages 99–115. Oxford University Press, 2006.

[Zwi98]   U. Zwick. Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 201–210. ACM-SIAM, 1998.