Von Neumann-Morgenstern Stability and Internal Closedness in Matching Theory

Yuri Faenza[1], Clifford S. Stein[1], and Jia Wan[2](✉)

[1] Columbia University, New York, NY 10027, USA
yf2414@columbia.edu, cliff@ieor.columbia.edu
[2] Massachusetts Institute of Technology, Cambridge, MA 02139, USA
jiawan@mit.edu

Abstract. Gale and Shapley's stability criterion enjoys a rich mathematical structure, which propelled its application in various settings. Although immensely popular, the approach by Gale and Shapley cannot encompass all the different features that arise in applications, motivating the search for alternative solution concepts. We investigate alternatives that rely on the concept of internal stability, a notion introduced for abstract games by von Neumann and Morgenstern and motivated by the need of finding a set of mutually compatible solutions. The set of stable matchings is internally stable (IS). However, the class of IS sets is much richer, for an IS set of matchings may also include unstable matchings and/or exclude stable ones. In this paper, we focus on two families of IS sets of matchings: *von Neumann-Morgenstern* (vNM) stable and *internally closed*. We study algorithmic questions around those concepts in both the marriage and the roommate models. One of our results imply that, in the marriage model, internally closed sets are an alternative to stable matchings that is as tractable as stable matchings themselves, a fairly rare occurrence in the area. Both our positive and negative results rely on new structural insights and extensions of classical algebraic structures associated with sets of matchings, which we believe to be of independent interest.

Keywords: Stable matching · rotation · poset · distributive lattice · vNM stability

1 Introduction

The *marriage model* was introduced by Gale and Shapley in their classical work [12] to address the problem of fairly allocating college seats to students. In (a slight generalization of) their setting, we are given a two-sided matching market, with each agent listing a subset of the agents from the opposite side of the market in a strict preference order. The goal is to find a matching that respects a fairness property called *stability*. Stable matchings enjoy a rich

J. Vygen and J. Byrka (Eds.): IPCO 2024, LNCS 14679, pp. 168–181, 2024.
https://doi.org/10.1007/978-3-031-59835-7_13

mathematical structure that has been leveraged on to design various algorithms (see, e.g., [15,23,25]) to assign students to schools, doctors to hospitals, workers to firms, partners in online dating, agents in ride-sharing platforms, and more, see, e.g., [1,16,24,30]. Although immensely popular, stability is not the right notion for some applications, where we may want, e.g., a matching of larger size or more favorable to one side of the market. Concepts alternative to stability, such as popularity [8,9,13,17,19,21] and Pareto-optimality [2,26] have therefore become an important area of research. However, such alternative concepts often lack many of the attractive structural and algorithmic properties of stable matchings. Much research has therefore been devoted to defining alternative solution concepts that are computationally tractable [3,5,11,23].

Internal Stability, von Neumann-Morgenstern Stability, and Internal Closedness. In this paper, we study two solution concepts alternative to stability, in both the marriage and the roommate (i.e., non-bipartite) model. These concepts rely on the fundamental game theory notion of *internal stability* [28]. A set of matchings $\mathcal{M}'$ of a (marriage or roommate) instance is *internally stable* if there are no matchings $M, M' \in \mathcal{M}'$ such that an edge of M blocks M', i.e., there is no pair of agents matched in M that strictly prefer each other to their respective partners in M'. Note that while stability is a property of a matching, internal stability is a property of a set of matchings, and that the set of stable matchings is internally stable. As discussed by von Neumann and Morgenstern [28], a family of internally stable solutions $\mathcal{M}'$ to a game (a family of matchings, in our case) can be thought of as the family of "standard behaviours" within an organization: solutions in $\mathcal{M}'$ are all and only those that are compatible with predefined rules. Also, as argued in [7,28], one can question why other solutions are excluded from being members of $\mathcal{M}'$. Hence, in order for an internally stable set $\mathcal{M}'$ to be deemed an acceptable standard, it is required that $\mathcal{M}'$ satisfies further conditions. Typically, one requires $\mathcal{M}'$ to be *externally stable*: for every matching $M \notin \mathcal{M}'$, there is a matching $M' \in \mathcal{M}'$ containing an edge that blocks M. A set that is both internally and externally stable is called *von Neumann-Morgenstern* (*vNM*) stable. Von Neumann and Morgenstern proposed vNM stability as the main solution concept for cooperative games [28]. Shubik [27] reports more than 100 works that investigate this concept until 1973 (see also [22]). Later research in game theory shifted the focus from vNM stability to the core of games, which enjoys stronger properties, and is often easier to work with, than vNM stable sets [6]. However, vNM stable sets in the marriage model have been shown to enjoy strong algorithmic and structural properties [6,10,29].

To define internal closedness, we relax the external stability condition to inclusionwise maximality. A set is *internally closed* if it is an inclusionwise maximal set of internally stable matchings. Our definition is motivated by two considerations. First, a vNM stable set may not always exist in a roommate instance, while an internally closed set of matchings always exists. Second, a central planner may want a specific set of internally stable matchings $\mathcal{M}'$ to be part of the family of feasible solutions. Consider for instance the problem the planner faces when given a family of internally stable matchings $\mathcal{M}'$, representing the currently accepted matchings. The planner's goal is to look for a more comprehensive set of solutions, that are possibly an improvement over the status quo, but are also compatible with it, i.e., matchings that neither block nor are blocked by matchings in $\mathcal{M}'$. Formally, can we efficiently obtain an internally closed set $\mathcal{M}''$ so that $\mathcal{M}'' \supseteq \mathcal{M}'$? Clearly, $\mathcal{M}''$ always exists, as one can start from $\mathcal{M}'$ and iteratively enlarge it as to achieve inclusionwise maximality while preserving internal stability. As we will see, our analysis of internally closed sets of matching also leads to an algorithmic understanding of vNM stable sets.

1.1 Overview of Contributions and Techniques

Motivated by the discussion above, in this paper we investigate structural and algorithmic properties of vNM stable and internally closed sets of matching. Our results have implications for the theory of stable matchings more generally. We present our contributions from the end, i.e., from their algorithmic implications.

Algorithmic and Complexity Results. We show that, in a marriage instance, one can find an internally closed set of matchings containing any given internally stable set of matchings in polynomial time (Theorem 1). Conversely, in a roommate instance, even deciding if a set of matchings is internally closed, or vNM stable, is co-NP-hard, and the problem of finding a vNM stable set of matchings is also co-NP-hard (Theorem 6).

From Matchings to Edges. The algorithmic statements from the previous paragraph glossed over the complexity issue of representing the input and the output to our problems, since a family of internally stable matchings may have size exponential in the number of agents. We bypass this concern by showing that every internally closed set of matchings coincides with the set $\mathcal{S}'$ of stable matchings in a *subinstance* of the original instance, i.e., an instance obtained from it by removing certain entries from preference lists. Hence, denoting an instance by a pair $(G, >)$ (with G being the graph with agents as nodes and matchable pairs as edges, and $>$ the agents' preferences), our input is given by a set $E_0 \subseteq E(G)$ implicitly describing the set $\mathcal{S}'$ of stable matchings in $(G[E_0], >)$. Similarly, an internally closed or vNM stable set of matchings is described compactly by $E_C \subseteq E(G)$. See Sect. 2 for details. This fact allows us to work with (polynomially-sized) sets of edges, rather than (possibly exponentially-sized) sets of matchings. It also implies that any question on an internally closed set of matchings, once the set E_C had been determined, reduces to the analogous question on stable matchings, for which algorithms are often known.

Our next step lies in understanding how to enlarge the input set of edges E_0, in order to obtain the set E_C. The challenge here lies in the fact that adding *any single edge* to E_0 may not lead to a strictly larger internally stable set of matchings, while adding certain edges may prevent the possibility of finding an internally closed set of matching. Hence, any algorithm will need to iteratively add *sets of edges* rather than single edges, possibly leading to a search space that is exponentially large. However, (extensions of) classical algebraic properties of stable matchings come to our rescue.

T =

Agent	Preference List
1	Ⓐ Ⓑ Ⓒ E
2	Ⓑ A Ⓒ Ⓓ
3	Ⓒ Ⓓ Ⓐ E
4	Ⓓ C Ⓐ Ⓑ
5	Ⓔ C
A	③ ④ 2 ①
B	④ ① ②
C	5 ① ② 4 ③
D	② ③ ④
E	1 ⑤ 3

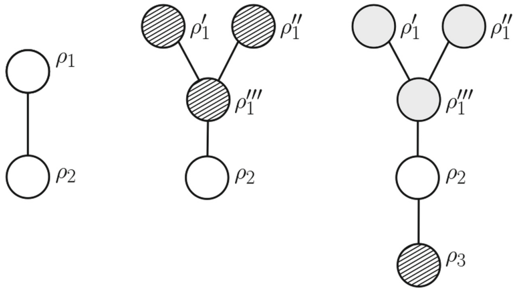


Fig. 1. Left: A marriage instance described by its preference table T with, circled, the entries of subtable T'. The set $\mathcal{S}'$ of stable matchings of T' is internally stable but not internally closed: matching $\{1B, 2A, 3D, 4C, 5E\}$ neither blocks nor is blocked by any matching from $\mathcal{S}'$. Right: We iteratively expand the poset of rotations of $\mathcal{S}'$ first by dissecting ρ_1 into $\rho_1', \rho_1'', \rho_1'''$ (leading to the poset of rotations of the stable matchings of $T \backslash \{1E, 5C, 3E\}$) and then by vertically expanding the poset via ρ_3 (poset associated to $T \backslash \{3E\}$). The set of stable matchings of $T \backslash \{3E\}$ is internally closed: no further enlargement of the poset is possible.

From Edges to Rotations: The Marriage Case. We first focus on the marriage case, investigated in Sect. 3. Since vNM stable sets in this model are well understood [6,10,29], we consider only internally closed sets of matchings. A classical result gives a bijection between the set of stable matchings of a marriage instance and the family of closed sets of the associated poset of *rotations* $(R, \sqsupseteq$ $)$ [20]. Rotations are certain cycles in the marriage instance (see Sect. 3). We first give a characterization of internally closed sets of matchings that relies on a certain "maximality" property of the poset of rotations. Roughly speaking, a family of matchings is internally closed if and only if, in the poset of rotations associated to it, no rotation can be *dissected*, i.e., replaced with a poset of new rotations, and we cannot *vertically expand* the poset by adding rotations that are maximal or minimal wrt $\sqsupseteq$ (see Theorem 4). On the way to this result, we introduce the novel concept of *generalized rotations*, that may be useful for other questions in the area. Our characterization leads to a polynomial-time algorithm that iteratively enlarges a set of internally stable matchings by trying to dissect rotations or vertically expand the associated poset (see Fig. 1 for an example). This algorithms shows that an internally closed set of matchings containing a given internally stable set of matchings can be found in polynomial time (see Theorem 1).

From Edges to Rotations: The Roommate Case. We investigate the roommate case in Sect. 4. The set of stable matchings in the roommate case does not seem to have a relevant lattice structure, and may be empty. However, when it is non-empty, it can be described via a poset of *singular* and *non-singular* rotations (first introduced in [14], defined differently from the marriage case). Our main structural contribution here is to show that the set of stable matchings is internally closed if and only if the poset of rotations associated to it cannot be augmented via what we call a *stitched rotation* (see Definition 6 and Theorem 5). We then show how to construct, from any instance ϕ of 3-SAT, a roommate instance that has a stitched rotation if and only if ϕ is satisfiable, thus proving the claimed hardness result. Internally closed sets give us a new tool to study vNM stability, also leading to a proof of the co-NP-hardness of deciding if a set of matchings is vNM stable, and of finding a vNM stable set (see Theorem 6).

Note. All proofs are deferred to the journal version of the paper.

2 Preliminaries

Basic Notation. For $n \in \mathbb{N}$, we let $[n] = \{1, \ldots, n\}$ and $[n]_0 = \{0\} \cup [n]$. On top of the classical graph representation for marriage and roommate instances, we represent instances via preference tables (see, e.g., [15]). A (*roommate*) *instance* is therefore described as a set A of agents and, for each $z \in A$, a *preference list* consisting of a subset $A(z) \subseteq A \backslash \{z\}$ and a strict ordering (i.e., without ties) of elements from this list. For each $z \in A$, $A(z)$ contains therefore all and only the agents that are acceptable to z. The collection of all preference lists is then represented by the *(preference) table* $T(A, >)$ (or T in short), where $>$ collects, for $z \in A$, the strict ordering within the preference list of agent z, denoted as $>_z$. We assume that preferences are symmetric, i.e., z_1 is on z_0's preference list if and only if z_0 is on z_1's. See Fig. 1 and Fig. 2 for examples.

Preferences. For $z, z_1, z_2 \in A$ with $z_1, z_2 \in A(z)$, we say that z *strictly prefers* z_1 to z_2 if $z_1 >_z z_2$. We say that z *(weakly) prefers* z_1 to z_2, and write $z_1 \geq_z z_2$, if $z_1 >_z z_2$ or $z_1 = z_2$. For $z \in A$, we extend $>_z$ to $\emptyset$ by letting $z_1 >_z \emptyset$ for all $z_1 \in A(z)$. That is, all agents strictly prefer being matched to some agent in their preference list than being left unmatched. Because of the symmetry assumption, we say that $z_0 z_1 \in T$ if z_1 is on z_0's preference list and call $z_0 z_1$ an *edge (of T)*. Let $r_T(z_0, z_1)$ denote the *ranking* (i.e., the position, counting from left to right) of z_1 in the preference list of z_0 in preference table T: $r_T(z_0, z_1) < r_T(z_0, z_2)$ if and only if agent z_0 strictly prefers z_1 to z_2 within preference table T. We let $r_T(z, \emptyset) = +\infty$. For a preference table T and agent z, let $f_T(z)$, $s_T(z)$, $\ell_T(z)$ denote the first, second and last agent on z's preference list.

Consistency, Subtables. Two roommate instances T, T' have *consistent* preference lists if, for any pair $z_0 z_1, z_0 z_2 \in T$ and $z_0 z_1, z_0 z_2 \in T'$, we have $r_T(z_0, z_1) < r_T(z_0, z_2)$ if and only if $r_{T'}(z_0, z_1) < r_{T'}(z_0, z_2)$. We write $T' \subseteq T$ when T', T have the following properties: (a) $z_0 z_1 \in T$ for all $z_0 z_1 \in T'$ and (b) T, T' have consistent preference lists. In this case, we call T' a *subtable* of T. For subtables T_1, T_2 of T, we let $T_1 \cup T_2$ be the subtable of T that contains all edges that are in T_1, T_2, or both.

Matching Basics. Fix a roommate instance $T(A, >)$. A *matching* M of T is a collection of disjoint pairs of agents from A, with the property that if $zz' \in M$, then z appears in z' preference list. For $z_0 \in A$, we let $M(z_0)$ be the partner of z_0 in matching M. If $z_0 z_1 \notin M$ for every $z_1 \in A$, we write $M(z_0) = \emptyset$. If $M(z_0) \neq \emptyset$, we say that z_0 is *matched* (*in* M). For a matching $M \subseteq T$, we say that $ab \in T$ is a *blocking pair* for M (and M is blocked by ab) if $b >_a M(a)$ and $a >_b M(b)$. M is *stable* if it is not blocked by any pair in T, and *unstable* if it is blocked by some pair in T. We say that a matching M' *blocks* a matching M if M' contains a blocking pair ab for M. Any matching M can be interpreted as a preference table T, where $z_0 z_1 \in T$ if and only if $z_0 z_1 \in M$. For a roommate instance T, we let $\mathcal{M}(T)$ denote the set of matchings of T, and $\mathcal{S}(T)$ denote the set of stable matchings of T. If $xy \in T$ is contained in some stable matching of T, then it is called a *stable edge* or *stable pair*. Let $E_S(T)$ denote the subtable of T containing all and only the stable edges. If $E_S(T) = T$, then T is called a *stable table*. When T is clear from the context, we abbreviate $\mathcal{M}(T), \mathcal{S}(T), r_T(\cdot,\cdot), E_S(T), \ldots$ by $\mathcal{M}$, $\mathcal{S}$, $r(\cdot,\cdot), E_S, \ldots$. We say an instance T is *solvable* if it admits at least one stable matching.

Internal Stability and Related Concepts. Let T be a roommate instance. We say that a set of matchings $\mathcal{M}' \subseteq \mathcal{M}(T)$ is *internally stable* if given any two matchings $M, M' \in \mathcal{M}'$, M does not block M' and M' does not block M. For an internally stable set of matchings $\mathcal{M}' \subseteq \mathcal{M}(T)$, we define its closure $\overline{\mathcal{M}'} = \{M \in \mathcal{M}(T) : \{M\} \cup \mathcal{M}' \text{ is internally stable}\}$. Note that $\overline{\mathcal{M}'}$ may not be internally stable. If $\overline{\mathcal{M}'} = \mathcal{M}'$, we say that $\mathcal{M}'$ is *internally closed*. Note that internally closed sets of matchings are exactly the inclusionwise maximal internally stable sets. For an internally stable set $\mathcal{M}'' \supseteq \mathcal{M}'$, we say that $\mathcal{M}''$ is an *internal closure* of $\mathcal{M}'$ if $\mathcal{M}''$ is internally closed. Clearly, every internally stable set of matching admits an internal closure, which may not be unique. The following lemma gives more basic structural results.

Lemma 1. *Let T be an instance of the roommate problem, $\mathcal{M}' \subseteq \mathcal{M}(T)$, and $\widetilde{T} = \cup\{M | M \in \mathcal{M}'\}$. (a) $\mathcal{M}'$ is internally stable if and only if $\mathcal{M}' \subseteq \mathcal{S}(\widetilde{T})$; (b) If $\mathcal{M}'$ is internally closed, then $\mathcal{M}' = \mathcal{S}(\widetilde{T})$.*

By Lemma 1, part (b), we can succinctly represent any internally closed set of matchings $\mathcal{M}'$ via $\widetilde{T} \subseteq T$ such that $\mathcal{M}' = \mathcal{S}(\widetilde{T})$. A set $\mathcal{M}' \subseteq \mathcal{M}(T)$ is called *externally stable* (in T) if for each $M \in \mathcal{M}(T) \setminus \mathcal{M}'$, there exists $M' \in \mathcal{M}'$ that blocks M. A set that is both internally and externally stable is called *vNM stable* (in T). Note that any vNM stable set $\mathcal{M}'$ is necessarily internally closed, therefore by Lemma 1, part (b), we have $\mathcal{M}' = \mathcal{S}(\widetilde{T})$, where $\widetilde{T} = \cup\{M : M \in \mathcal{M}'\}$.

The Problems of Interest. We now present the algorithmic questions investigated in this paper. Marriage instances are a special case of roommate instances (see Sect. 3 for a definition). As previously discussed, the input to the first three problems below consists of a subtable $\widetilde{T} \subseteq T$, which implicitly describes the associated set of internally stable matchings $\mathcal{S}(\widetilde{T})$, see Lemma 1.

Find an internal closure of a given internally stable set, Marriage Case($\widetilde{T}, T$) `(IStoIC-MC)`

Given: A marriage instance T and a stable table $\widetilde{T}$ such that $\widetilde{T} \subseteq T$.
Find: $T' \subseteq T$ such that $\mathcal{S}(T')$ is an internal closure of $\mathcal{S}(\widetilde{T})$.

Check Internal Closedness($\widetilde{T}, T$) `(CIC)`

Given: A roommate instance T and a stable table $\widetilde{T}$ such that $\widetilde{T} \subseteq T$.
Decide: If $\mathcal{S}(\widetilde{T})$ is internally closed.

Check vNM Stability($\widetilde{T}, T$) `(CvNMS)`

Given: A roommate instance T and a stable table $\widetilde{T}$ such that $\widetilde{T} \subseteq T$.
Decide: If $\mathcal{S}(\widetilde{T})$ is vNM stable.

Find a vNM Stable Set(T) `(FvNMS)`

Given: A solvable roommate instance T.
Find: $T' \subseteq T$ such that $\mathcal{S}(T')$ is a vNM stable set, or conclude that no vNM stable set exists.

3 Internally Closed Sets: The Marriage Case

In a *marriage* instance [12], the set of agents A can be partitioned into two disjoint sets $X = \{x_1, x_2, \ldots, x_p\}$ and $Y = \{y_1, y_2, \ldots, y_\ell\}$, where the preference list of any agent in X consists only of a subset of agents in Y, and vice versa. All marriage instances are solvable [12]. An agent $x \in X$ is called an X*-agent*, and similarly an agent $y \in Y$ is called a Y*-agent*. In this section, we give a characterization of internally closed sets of matchings in the marriage case that relies on a generalization of the classical concept of rotations (see Theorem 4). We then use the characterization to show the following result.

Theorem 1. *Given a marriage instance with n agents,* `IStoIC-MC` *can be solved in time $O(n^4)$.*

Throughout this section, fix a marriage instance T.

3.1 Algebraic Structures Associated to Stable Matchings

Partial Order. We start by discussing known features of the poset of stable matchings obtained via an associated partial order (the poset is in fact a distributive lattice, but we will not use this fact explicitly in the exposition). For an extensive treatment of this topic, see [15]. Define the following domination relationship between matchings:

$$M \succeq_X M' \text{ if, for every } x \in X,\ M(x) \geq_x M'(x).$$

If, in addition, there exists at least one agent $x \in X$ such that $M(x) >_x M'(x)$ (or, equivalently, if $M' \neq M$), then we write $M \succ_X M'$. We symmetrically define the relations $M \succeq_Y M'$ and $M \succ_Y M'$. It is well-known that, for $Z \in \{X, Y\}$, there exists a matching $M_Z^T \in \mathcal{S}$ (or simply M_Z) such that $M_Z^T \succeq_Z M'$ for all $M' \in \mathcal{S}$. M_Z^T is called *Z-optimal.* Note that M_Z^T is, by definition, stable.

Classical Rotations and Properties. We sum up here definitions and known facts about the classical concept of rotations (first introduced in [20]).

Definition 1. *Let $M \in \mathcal{S}(T)$. Following [15], given distinct X-agents $x_0, \ldots, x_{r-1} \in X$ and Y-agents $y_0, \ldots, y_{r-1} \in Y$, we call the finite ordered list of ordered pairs*

$$\rho = (x_0, y_0), (x_1, y_1), \ldots, (x_{r-1}, y_{r-1}) \tag{1}$$

*a (*classical*)* X-rotation[1] exposed in *M if, for every $i \in [r-1]_0$:*

a) $x_i y_i \in M$; *b) $x_i >_{y_{i+1}} x_{i+1}$;* *c) $y_i >_{x_i} y_{i+1}$;*
d) $M(y) >_y x_i$ for all $y \in Y$ such that $x_i y \in T$ and $y_i >_{x_i} y >_{x_i} y_{i+1}$.

We abuse notation and write $x \in \rho$ (resp., $y \in \rho$) if $(x, y) \in \rho$ for some y (resp., for some x), and similarly we write $x \notin \rho$ (resp., $y \notin \rho$) if such y (resp., x) does not exist. The *elimination* of an X-rotation ρ exposed in a stable matching M maps M to $M' := M/\rho$ where $M(x) = M'(x)$ for $x \in X \setminus \rho$ and $M'(x_i) = y_{i+1}$ for all $i \in [r-1]_0$. Note that $M \succ_X M'$ and M' differs from M by a cyclic shift of each X-agent in ρ to the partner in M of the next X-agent in ρ. Rotations can be used to describe the set of stable matchings, as shown next.

Theorem 2. *There is exactly one set of X-rotations $R_X = \{\rho_1, \rho_2, \ldots, \rho_h\}$ such that, for $i \in [h]$, ρ_i is exposed in $((M_X/\rho_1)/\rho_2)\ldots/\rho_{i-1}$, and $M_Y = M_X/R_X = (((M_X/\rho_1)/\rho_2)/\ldots)/\rho_h$. Moreover, R_X is exactly the set of all X-rotations exposed in some stable matching of T.*

Extending the definition from the previous theorem, for $R = \{\rho_1, \rho_2, \ldots, \rho_k\} \subseteq R_X$, such that, for $i \in [k]$, ρ_i is exposed in $((M/\rho_1)/\rho_2)\ldots/\rho_{i-1}$, we let $M/R := (((M/\rho_1)/\rho_2)/\ldots)/\rho_k$. Define the poset $(R_X, \sqsupseteq)$ as follows: $\rho \sqsupseteq \rho'$ if

[1] We often omit "X-" and call ρ simply a rotation. Note that an X-rotation (and a generalized X-rotation, defined later) with r elements is equivalent up to a constant shift (modulo r) of all indices of its pairs. Hence, we will always assume that indices in entries of a (generalized) X-rotation are taken modulo r.

for any sequence of X-rotation eliminations $M_X/\rho_1/\dots/\rho_k$ with $\rho = \rho_k$, we have $\rho' \in \{\rho_0, \rho_1, \dots, \rho_k\}$. If moreover $\rho' \neq \rho$, we write $\rho \sqsupset \rho'$. When we want to stress the instance T we use to build the (po)set of X-rotations, we denote it by $R_X(T)$.

A set $R \subseteq R_X$ is called *closed* if $\rho \in R, \rho' \in R_X : \rho \sqsupseteq \rho' \Rightarrow \rho' \in R$. For $\rho \in R_X$, we let $R(\rho) = \{\rho' \in R_X : \rho \sqsupset \rho'\}$. Note that $R(\rho)$ is closed and does not include ρ. The following extension of Theorem 2 can be seen as a specialization of Birkhoff's representation theorem [4] to the lattice of stable matchings.

Theorem 3. *The following map defines a bijection between closed sets of rotations and stable matchings:* $R \subseteq R_X,\ R$ *closed* $\ \rightarrow\ M_X/R$.

Generalized X-Rotations. We now introduce an extension of the classical concept of X-rotation, which we call *generalized X-rotation*, and define its associated digraph.

Definition 2. *Let $M \in \mathcal{M}(T)$. Given distinct X-agents $x_0, \dots, x_{r-1}$ and Y-agents $y_0, \dots, y_{r-1}$, we call the ordered set of ordered pairs*

$$\rho^g = (x_0, y_0), (x_1, y_1), \dots, (x_{r-1}, y_{r-1}) \tag{2}$$

a generalized X-rotation exposed in M *if, for every $i \in [r-1]_0$, it satisfies properties a)-b)-c) from Definition 1 (but not necessarily d). Often, we omit "X-" when clear from the context.*

Note that a classical X-rotation exposed in a stable matching M is also a generalized rotation exposed in M. We again write $x \in \rho^g$ (resp., $y \in \rho^g$) if $(x, y) \in \rho^g$ for some y (resp., for some x). The *elimination* of a generalized X-rotation, or of a set of generalized X-rotations, is defined analogously to the classical rotation case. For a (generalized) X-rotation ρ^g exposed at a matching M, we still have $M \succ_X M/\rho^g$. For a generalized X-rotation ρ^g as in (2), we let $E(\rho^g) = \{x_i y_i\}_{i \in [r-1]_0} \cup \{x_i y_{i+1}\}_{i \in [r-1]_0}$, and also interpret $E(\rho^g)$ as a subtable of T. Similarly, we interpret ρ^g as a subtable with edges $x_i y_i$ for $i \in [r-1]_0$.

The Generalized Rotation Digraphs. We define the following generalized X-rotation digraph for a matching M of T, denoted as $D_X(M, T)$ or simply $D_X(M)$ when T is clear from the context. The set of nodes is given by $X \cup Y$. For any agents $x \in X, y \in Y$, add arc (x, y) if $x >_y M(y)$ and $M(x) >_x y$; and add arc (y, x) if $M(y) = x$. Note that the outdegree of each X-agent can be larger than 1, but the outdegree of every Y-agent is at most 1, see Fig. 2 and Example 1. The next lemma follows directly from the definition of $D_X(M)$ and it is similar to a known statement for classical rotations, see, e.g., [15,20].

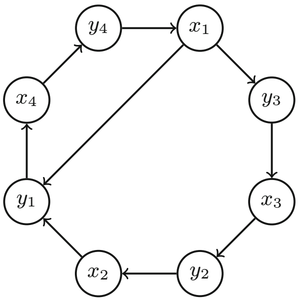
Lemma 2. *Let $M \in \mathcal{M}(T)$. $x_0 \rightarrow y_1 \rightarrow x_1 \rightarrow \cdots \rightarrow y_0 \rightarrow x_0$ is a cycle in $D_X(M, T)$ if and only if $\rho^g = (x_0, y_0), (x_1, y_1), \dots, (x_{r-1}, y_{r-1})$ is a generalized X-rotation exposed in M. We say that ρ^g and the cycle* correspond *to each other.*

Classical and generalized Y-rotations are defined similarly to classical and generalized X-rotations, with the role of agents in X and Y swapped. By symmetry, all definitions and properties carry over.

Example 1. Consider the instance T in Fig. 2, left, and its stable matching $M = \{x_1y_4, x_2y_2, x_3y_3, x_4y_1\}$. The generalized rotation digraph $D_X(M)$ is given in Fig. 2, center. $P = x_1 \to y_3 \to x_3 \to y_2 \to x_2 \to y_1 \to x_4 \to y_4 \to x_1$ is a cycle in $D_X(M)$. Cycle P corresponds to a generalized X-rotation $\rho = (x_1, y_4), (x_3, y_3), (x_2, y_2), (x_4, y_1)$, which is exposed in M (wrt table T). One can check that ρ satisfies properties a)-b)-c) but not d) from Definition 1. On the other hand, $\rho_1 = (x_1, y_4), (x_4, y_1)$ satisfies properties a)-b)-c)-d) from Definition 1, hence it is an X-rotation exposed in M, corresponding to cycle $x_1 \to y_1 \to x_4 \to y_4 \to x_1$ of $D_X(M)$.

$T =$

Agent	Preference list
x_1	$y_4\ y_1\ y_3$
x_2	$y_2\ y_1$
x_3	$y_3\ y_2$
x_4	$y_1\ y_4$
y_1	$x_2\ x_1\ x_4$
y_2	$x_3\ x_2$
y_3	$x_1\ x_3$
y_4	$x_4\ x_1$



$T' =$

Agent	Preference list
x_1	$y_4\ y_3$
x_2	$y_2\ y_1$
x_3	$y_3\ y_2$
x_4	$y_1\ y_4$
y_1	$x_2\ x_4$
y_2	$x_3\ x_2$
y_3	$x_1\ x_3$
y_4	$x_4\ x_1$

Fig. 2. Illustrations from Example 1. From left to right: T, $D_X(M)$, T'.

Dissection of a Rotation. A new, key concept for proving our characterization of internally closed sets is that of dissecting set for a rotation.

Dissecting a Rotation(T, T', ρ) (DR)

Given: Marriage instances $T' \subseteq T$ with T' stable, and $\rho \in R_X(T')$.
Find: A set $R = \{\rho_1, \rho_2, \dots, \rho_k\}$ satisfying a) $R \subseteq R_X(T^*) \setminus R_X(T')$ with $T^* = T' \cup_{j=1}^{k} E(\rho_j)$; b) $M_X^{T'}/R(\rho)/\rho = M_X^{T'}/R(\rho)/R$; or output $\{\rho\}$ if R as above does not exist.

If DR(T, T', ρ) outputs a set $R \neq \{\rho\}$, then R is called a *dissecting set* for (T, T', ρ). Note that a dissecting set has at least two elements.

Example 2. Consider again the instance from Example 1. T has 3 inclusionwise maximal matchings: M, $M_1 = \{x_1y_1, x_2y_2, x_3y_3, x_4y_4\}$, $M_2 = \{x_1y_3, x_2y_1, x_3y_2, x_4y_4\}$. Let $T' = M \cup E(\rho)$. M is the X-optimal matching within T' and ρ is the only (classical) X-rotation exposed in M in T'. Note that we have $M_2 = M/\rho$. Let $T^* = T = T' \cup E(\rho_1) \cup E(\rho_2)$, with ρ_1 as in Example 1 and $\rho_2 = (x_1, y_1), (x_3, y_3), (x_2, y_2)$. M is also the X-optimal matching within T. Note that $\rho_1 \in R_X(T^*)$ is exposed in M, $\rho_2 \in R_X(T^*)$ is exposed in M/ρ_1, and $M_2 = M/\rho_1/\rho_2$. Yet, $\rho_1, \rho_2 \notin R_X(T')$. Hence, $\{\rho_1, \rho_2\}$ is a dissecting set for (T, T', ρ). When enlarging T' to T^*, the set of stable matchings becomes larger, and ρ is dissected into ρ_1 and ρ_2, increasing the size of the poset of rotations.

In Example 2, the poset of rotations in T^* is an "expansion" of the poset in T'. We next show that whenever a rotation is dissected, a similar phenomenon happens.

Lemma 3 (Internal expansion of the rotation poset via rotation dissection). *Let $T' \subseteq T$ with T' stable. Let $\rho \in R_X(T')$ and $R = \{\rho_1, \ldots, \rho_k\}$ be a dissecting set for (T, T', ρ). Then a) $T^* = T' \cup_{j=1}^{k} \rho_j$ is a stable table and b) $R_X(T^*) = (R_X(T') \backslash \{\rho\}) \cup \{\rho_1, \ldots, \rho_k\}$.*

Generalized Rotations Exposed in the Y-Optimal Stable Matching. We next show that an internally stable set of matchings can be "vertically expanded" by adding to a stable table the edges from a generalized X- (resp., Y-)rotation exposed in the Y- (resp., X-)optimal stable matching. The next lemma is stated for X-rotations, but by symmetry extends to Y-rotations.

Lemma 4 (Vertical expansion of the rotation poset). *Let T be a marriage instance and $T' \subseteq T$, with T' stable. Let $M_Y = M_Y^{T'}$ and suppose ρ^g is a generalized X-rotation corresponding to a cycle of $D_X(M_Y, T)$. Let $M^* = M_Y/\rho^g$ and $\overline{T} = T' \cup E(\rho^g) = T' \cup M^*$. Then a) $M^* \notin \mathcal{S}(T')$, b) the set $\{M^*\} \cup \mathcal{S}(T')$ is internally stable, c) $\overline{T}$ is a stable table, d) M^* is the Y-optimal stable matching in $\mathcal{S}(\overline{T})$, and e) $R_X(\overline{T}) = R_X(T') \cup \{\rho^g\}$.*

3.2 Characterization of Internally Closed Sets of Matchings

We now have all ingredients to state a characterization of internally closed sets of matchings based on generalized rotations and on rotation dissections.

Theorem 4. *Let T be a marriage instance and $\mathcal{M}' \subseteq \mathcal{M}(T)$. $\mathcal{M}'$ is internally closed if and only if $\mathcal{M}' = \mathcal{S}(T')$ for a stable subtable $T' \subseteq T$ such that:*

a) (no rotation can be dissected) $DR(T, T', \rho)$ returns $\{\rho\}$ for all $\rho \in R_X(T')$.
b) (no vertical expansion) $D_X(M_Y^{T'}, T)$ and $D_Y(M_X^{T'}, T)$ have no cycles.

We can now give a sketch of the proof of how to solve `(IStoIC-MC)` and prove Theorem 1. Starting from input $\widetilde{T}$, we iteratively check if conditions a) and b) from Theorem 4 are verified and, if not, rely on Lemma 3 and Lemma 4 to create internally stable tables strictly containing $\widetilde{T}$. While Lemma 3 does not explain how to enlarge the current subtable of T and the rotation poset, we can perform all operations efficiently by using properties of the stable matchings lattice and of (generalized) rotations.

4 Internally Closed and vNM Stable Sets: The Roommate Case

In this section, we deal with roommate instances. Our major structural contribution is an extension of the concept of rotations which we call *stitched rotations*. We show that the set of stable matchings of a solvable roommate instance is internally closed if and only if a stitched rotation does not exist (Theorem 5). The complexity of finding stitched rotations then allow us to deduce our hardness results, see Theorem 6.

Rotations. A major difference between the marriage and the roommate problem is the absence of a (known) relevant lattice structure in the latter. This change calls for a different concept of rotation (first defined in [18]) for analyzing the problem. Throughout the section, assume that T is a solvable roommate instance.

Definition 3. *Let $T' \subseteq T$. A sequence $\rho = (x_0, y_0), (x_1, y_1), ..., (x_{r-1}, y_{r-1})$ is called a* rotation exposed in T' *if $y_i = f_{T'}(x_i)$ and $y_{i+1} = s_{T'}(x_i)$ for all $i \in [r-1]_0$, where indices are taken modulo r. T'/ρ denotes the table obtained from T' by deleting all pairs $y_i z$ such that y_i strictly prefers x_{i-1} to z. We refer to this process as to the* elimination of ρ in T. *Recall that $f_{T'}(x)$ gives the first preference of x within table T', and $s_{T'}(x)$ gives the second preference.*

Let $\mathcal{Z}(T)$ denote the set of all rotations exposed in some table obtained from T by iteratively eliminating exposed rotations. Throughout the rest of the section, we fix a rotation $\rho = (x_0, y_0), (x_1, y_1), ..., (x_{r-1}, y_{r-1}) \in \mathcal{Z}(T)$ and again we also interpret ρ as a table with entries $x_i y_i$ for $i \in [r-1]_0$. Rotations can be further classified into two categories: *singular* and *non-singular*.

Definition 4. *ρ is called a* non-singular *rotation if $\overline{\rho} \in \mathcal{Z}(T)$, where*

$$\overline{\rho} = (y_1, x_0), (y_2, x_1), ..., (y_i, x_{i-1}), ..., (y_0, x_{r-1}), \tag{3}$$

and it is called singular *otherwise. The subset of $\mathcal{Z}(T)$ containing all singular (resp., non-singular) rotations is denoted by $\mathcal{Z}_s(T)$ (resp., $\mathcal{Z}_{ns}(T)$).*

Antipodal Edges and Stitched Rotations. Let T^* be the subtable of T containing all and only its stable edges, i.e., $T^* = E_S(T)$. We first argue that to "expand" $\mathcal{S}(T^*)$ to a strictly larger internally stable set of matchings, i.e., to find a stable table $T' \supsetneq T^*$ such that $\mathcal{S}(T') \supsetneq \mathcal{S}(T^*)$, any edge $e \in T' \setminus T^*$ must satisfy what we call the *antipodal* condition, defined below. We then introduce stitched rotations, a new object that allows us to assemble antipodal edges and expand the set of stable matchings to a larger internally stable set.

Definition 5. *We say that an edge $e = xy \in T \setminus T^*$ satisfies the antipodal condition (wrt T^*) if exactly one of the following is true:*

$$y >_x f_{T^*}(x) \, and \, x <_y \ell_{T^*}(y), \quad or \quad x >_y f_{T^*}(y) \, and \, y <_x \ell_{T^*}(y).$$

Lemma 5. *Let $e \in T \backslash T^*$. Assume that e is not an antipodal edge. Then every $M \in \mathcal{M}(T)$ with $e \in M$ is blocked by some edge from T^*.*

Definition 6. *Suppose $\rho' \in \mathcal{Z}(T^* \cup \rho')$ is exposed in $T^* \cup E(\rho')$ and, for all $i \in [r-1]_0$, $x_i y_i$ satisfies the antipodal condition with respect to T^*. If $\rho' \in \mathcal{Z}_{ns}(T^* \cup \rho')$, we call ρ' a* stitched rotation *with respect to T^*.*

The next theorem shows that the stable subtable T^* of a solvable instance is internally closed if and only if there is no stitched rotation w.r.t. T^*.

Theorem 5 (Expansion of internally stable sets via stitched rotations). *1. Let $M \in \mathcal{M}(T) \backslash \mathcal{S}(T^*)$ and assume that $\{M\} \cup \mathcal{S}(T^*)$ is internally stable. Then there exists a stitched rotation w.r.t. T^*. 2. Conversely, if ρ' is a stitched rotation w.r.t. T^*, there exists $M \in \mathcal{M}(T) \backslash \mathcal{S}(T^*)$ with $\rho' \subseteq M$ so that $\{M\} \cup \mathcal{S}(T^*)$ is internally stable.*

Hardness Results. For each instance ϕ of 3-SAT, we create an instance T with stable subtable $\widetilde{T}$ such that $\mathcal{S}(T) = \mathcal{S}(\widetilde{T})$. Moreover, there exists a satisfiable assignment for ϕ if and only if there exists a rotation that is stitched w.r.t. $\widetilde{T}$. Together with Theorem 5, this reduction and variations of it lead to the following (the details of these constructions are postponed to the journal version).

Theorem 6. *`CIC`, `CvNMS`, and `FvNMS` are co-NP-hard.*

Acknowledgments. Yuri Faenza acknowledges the support of the NSF Grant 2046146 *CAREER: An Algorithmic Theory of Matching Markets*. Cliff Stein is supported in part by NSF grant CCF-2218677, ONR grant ONR-13533312, and by the Wai T. Chang Chair in Industrial Engineering and Operations Research.

References

1. Abdulkadiroğlu, A., Sönmez, T.: School choice: a mechanism design approach. Am. Econ. Rev. **93**(3), 729–747 (2003)
2. Abraham, D.J., Cechlárová, K., Manlove, D.F., Mehlhorn, K.: Pareto optimality in house allocation problems. In: Fleischer, R., Trippen, G. (eds.) ISAAC 2004. LNCS, vol. 3341, pp. 3–15. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30551-4_3
3. Abraham, D.J., Levavi, A., Manlove, D.F., O'Malley, G.: The stable roommates problem with globally-ranked pairs. In: Deng, X., Graham, F.C. (eds.) WINE 2007. LNCS, vol. 4858, pp. 431–444. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-77105-0_48
4. Birkhoff, G.: Rings of sets. Duke Math. J. **3**(3), 443–454 (1937)
5. Cseh, Á., Faenza, Y., Kavitha, T., Powers, V.: Understanding popular matchings via stable matchings. SIAM J. Discrete Math. **36**(1), 188–213 (2022)
6. Ehlers, L.: Von Neumann-Morgenstern stable sets in matching problems. Journal of Economic Theory **134**(1), 537–547 (2007)
7. Ehlers, L., Morrill, T.: (Il) legal assignments in school choice. Rev. Econ. Stud. **87**(4), 1837–1875 (2020)

8. Faenza, Y., Kavitha, T.: Quasi-popular matchings, optimality, and extended formulations. Math. Oper. Res. **47**(1), 427–457 (2022)
9. Faenza, Y., Kavitha, T., Powers, V., Zhang, X.: Popular matchings and limits to tractability. In: Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 2790–2809. SIAM (2019)
10. Faenza, Y., Zhang, X.: Legal assignments and fast EADAM with consent via classical theory of stable matchings. Oper. Res. **70**(3), 1873–1890 (2022)
11. Faenza, Y., Zhang, X.: Affinely representable lattices, stable matchings, and choice functions. Math. Program. **197**(2), 721–760 (2023)
12. Gale, D., Shapley, L.S.: College admissions and the stability of marriage. Am. Math. Mon. **69**(1), 9–15 (1962)
13. Gärdenfors, P.: Match making: assignments based on bilateral preferences. Behav. Sci. **20**(3), 166–173 (1975)
14. Gusfield, D.: The structure of the stable roommate problem: efficient representation and enumeration of all stable assignments. SIAM J. Comput. **17**(4), 742–769 (1988)
15. Gusfield, D., Irving, R.W.: The Stable Marriage Problem: Structure and Algorithms. MIT Press, Cambridge (1989)
16. Hitsch, G.J., Hortaçsu, A., Ariely, D.: Matching and sorting in online dating. Am. Econ. Rev. **100**(1), 130–63 (2010)
17. Huang, C.C., Kavitha, T.: Popularity, mixed matchings, and self-duality. Math. Oper. Res. **46**(2), 405–427 (2021)
18. Irving, R.W.: An efficient algorithm for the "stable roommates" problem. J. Algorithms **6**(4), 577–595 (1985)
19. Irving, R.W., Kavitha, T., Mehlhorn, K., Michail, D., Paluch, K.E.: Rank-maximal matchings. ACM Trans. Algorithms (TALG) **2**(4), 602–610 (2006)
20. Irving, R.W., Leather, P.: The complexity of counting stable marriages. SIAM J. Comput. **15**(3), 655–667 (1986)
21. Kavitha, T.: A size-popularity tradeoff in the stable marriage problem. SIAM J. Comput. **43**(1), 52–71 (2014)
22. Lucas, W.F.: von Neumann-Morgenstern stable sets. In: Handbook of Game Theory with Economic Applications, vol. 1, pp. 543–590 (1992)
23. Manlove, D.: Algorithmics of Matching Under Preferences, vol. 2. World Scientific (2013)
24. Roth, A.E.: Stability and polarization of interests in job matching. Econometrica J. Econometric Soc. **52**(1), 47–57 (1984)
25. Roth, A.E., Sotomayor, M.: Two-sided matching. In: Handbook of Game Theory with Economic Applications, vol. 1, pp. 485–541 (1992)
26. Shapley, L., Scarf, H.: On cores and indivisibility. J. Math. Econ. **1**(1), 23–37 (1974)
27. Shubik, M.: Game Theory in the Social Sciences: Concepts and Solutions (1982)
28. Von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (2007)
29. Wako, J.: A polynomial-time algorithm to find von Neumann-Morgenstern stable matchings in marriage games. Algorithmica **58**(1), 188–220 (2010)
30. Wang, X., Agatz, N., Erera, A.: Stable matching for dynamic ride-sharing systems. Transp. Sci. **52**(4), 850–867 (2018)