On the Convergence of Swap Dynamics to Pareto-Optimal Matchings

Felix Brandt Anaëlle Wilczynski
Technische Universität München

brandtf@in.tum.de wilczyns@in.tum.de

Résumé

On étudie dans cet article la possibilité d'atteindre des couplages optimaux au sens de Pareto par le biais d'échanges entre paires d'agents dans des problèmes d'appariement en présence d'un couplage initial. On considère trois problèmes d'appariement particuliers : le problème d'allocation de maisons dans lequel à chaque agent doit être affecté exactement un objet, ainsi que le problème des mariages stables et le problème des colocataires, dans lesquels les agents sont couplés avec d'autres agents. Dans ce cadre, on examine trois différentes notions d'échanges améliorants entre paires d'agents. Tandis qu'il peut être vérifié en temps polynomial si un couplage Pareto-optimal peut être atteint lorsque les échanges sont améliorants relativement à des paires bloquantes d'agents, vérifier si toutes les séquences d'échanges mènent à un couplage Pareto-optimal est difficile. Alternativement, on prouve que ces deux problèmes sont difficiles lorsque l'on considère des échanges dans lesquels ce sont les agents qui échangent qui doivent améliorer leur situation. Ce résultat confirme et étend une conjecture faite par Damamme et al. [16] qui ont, de plus, montré que dans le problème particulier d'allocation de maisons, toute séquence d'échanges améliorants converge vers une allocation Pareto-optimale sous des préférences unimodales. On montre dans cet article que, dans les problèmes de mariages stables et de colocataires, l'unimodalité des préférences n'est pas suffisante pour obtenir un tel résultat de convergence. En revanche, la restriction encore plus forte des préférences 1-Euclidiennes le permet.

Abstract

We study whether Pareto-optimal stable matchings can be reached via pairwise swaps in one-to-one matching markets with initial assignments. We consider housing markets, marriage markets, and roommate markets as well as three different notions of swap rationality. Our main results are as follows. While it can be efficiently determined whether a Pareto-optimal stable matching can be reached when defining swaps via blocking pairs, checking whether this is the case for *all* such sequences is computationally intractable. When defining swaps such that all involved agents need to be better off, even deciding whether a Pareto-optimal stable matching can be reached via *some* sequence is intractable. This confirms and extends a conjecture made by Damamme et al. [16] who have furthermore shown that convergence to a Pareto-optimal matching is guaranteed in housing markets with single-peaked preferences. We show that in marriage and roommate markets, single-peakedness is not sufficient for this to hold, but the stronger restriction of one-dimensional Euclidean preferences is.

1 Introduction

One-to-one matchings, where individuals are matched with resources or other individuals, are omnipresent in everyday life. Examples include the job market, assigning offices to workers, pairing students in working groups, and online dating. The formal study of matching procedures leads to challenging algorithmic problems while being of immediate practical interest [25, 28]. One typically distinguishes between three different types of abstract one-to-one matching settings. In *housing markets* [32], each agent is matched with an object (usually referred to as a house). In *marriage markets* [21], agents are partitioned into two groups—say, males and females—and each member of one group is matched with an agent from the other group. Finally, in *roommate markets* [21], all agents belong to the same group and each agent is matched with another agent. By supposing that agents are rational and want to maximize their satisfaction, individual agreements may naturally occur among them and especially, for realistic reasons, between small groups of agents. An important question is then whether sequences of such individual agreements can lead to socially optimal outcomes. In many applications, it is reasonable to assume that there is an initial assignment because agents already live in a house, are engaged in a relationship, and are employed by a company [1, 29]. Under these assumptions, we focus on atomic agreements which require the least coordination: *pairwise swaps*.

We consider three different types of individual rationality for pairwise swaps. In housing markets, there is only one meaningful notion of swap rationality: two agents will only exchange objects if both of them are better off. By contrast, when matching agents with each other, one could require that all four agents involved in a swap or only two of them are better off. The latter requirement allows for two kinds of swap rationality: two agents who exchange their match are better off (e.g., a company and its subsidiary exchange employees without asking their consent) or two agents who decide to form a new pair are better off (e.g., two lovers leave their current partners to be together).

Social optimality in settings with ordinal preferences like that of matching markets is measured in terms of Pareto-optimality. We therefore study whether there exists a sequence of pairwise swaps that results in a Pareto-optimal matching that does not allow for further swaps (and hence is called stable). Whenever all sequences of pairwise swaps are of this kind, we say that the given type of swap dynamics converges.

It turns out that in all three types of matching markets and all three notions of swap rationality, it may not be possible to reach a Pareto-optimal stable matching from the initial assignment. We prove that deciding whether this is the case is NP-hard for two types of swap rationality while it can be solved in polynomial time for swaps based on blocking pairs. However, for all types of rationality, checking convergence is co-NP-hard. On the other hand, we show that when preferences are one-dimensional Euclidean—a natural but demanding restriction—swap dynamics for two types of swap rationality will always converge.

2 Related Work

Damamme et al. [16] investigated the dynamics of individually rational pairwise swaps in housing markets, where two agents are better off by exchanging their objects. Recently, variants of this problem that further restrict the agents' interactions using underlying graph structures have been examined [22, 23, 31].

In marriage and roommate markets, most of the literature focuses on deviations based on blocking pairs, where two agents decide to leave their old partners in order to be matched with each other. Blocking pairs are best known for their role in the definition of stability [21], but some papers also studied the dynamics of blocking pair swaps [2, 30]. The notion of exchange stability, where two agents agree to exchange their partners, has been investigated in both roommate markets [8, 13] and marriage markets [14]. We consider both types of swaps, i.e., blocking pair swaps and exchange rational swaps, but focus on the study of dynamics that reach Pareto-optimal matchings.

Perhaps closest to our work is a result by Damamme et al. [16] who proved that swap dynamics always converge to a Pareto-optimal matching in housing markets under single-peaked preferences. However, they left open the computational problem of deciding whether a Pareto-optimal stable matching can be reached for unrestricted preferences and conjectured this problem to be intractable. We solve this problem and extend it to marriage and roommate markets. Moreover, we prove that their convergence result for housing markets under single-peaked preferences does not extend to marriage and roommate markets, but can be restored when restricting preferences even further.

3 The Model

We are given a set N of agents $\{1, \ldots, n\}$ and a set O of objects $\{o_1, \ldots, o_n\}$ such that $|N| = |O| = n$. Each agent $i \in N$ has strict ordinal preferences, given by a linear order $\succ_i$, over a set A_i of alternatives to be matched with. In the matching markets we consider, A_i is either a subset of the set of agents N or the set of objects O. A tuple of preference relations $\succ = (\succ_1, \ldots, \succ_n)$ is called a *preference profile*.

3.1 Matching Markets

In this article, we are considering three different settings where the goal is to match the agents either with objects—like in housing markets—or with other agents—like in marriage or roommate markets. In all cases, we assume that there is an initial matching. More formally,

- a *housing market* consists of a preference profile where $A_i = O$ for all $i \in N$, and an initial endowment given as a bijection $\mu : N \to O$,
- a *marriage market* consists of a preference profile where $N = W \cup M$ with $W \cap M = \emptyset$, $A_i = M$ for all $i \in W$ and $A_i = W$ for all $i \in M$, and an initial matching given as a bijection $\mu : W \to M$, and
- a *roommate market* consists of a preference profile with even n and $A_i = N \setminus \{i\}$ for all $i \in N$, and an initial matching given as an involution $\mu : N \to N$ such that $\mu(i) \neq i$ for all $i \in N$.

When allowing for indifferences as well as unacceptabilities in the preferences, the three settings form a hierarchy: housing markets are marriage markets where the "objects" are indifferent between all agents, and marriage markets are roommate markets where all agents of the same type are considered unacceptable. In this paper, however, we do not make either assumption and therefore these inclusion relationships do not hold.

The key question studied in this paper is whether Pareto-optimal matchings can be reached from the initial matching via local modifications. A matching is *Pareto-optimal* if there is no other matching μ' such that for every agent i, $\mu'(i) \succeq_i \mu(i)$ and for at least one agent j, $\mu'(j) \succ_j \mu(j)$.

3.2 Preference Restrictions

We consider three restricted preference domains: single-peaked preferences [12], globally-ranked preferences [4, 6] and their common subdomain of one-dimensional Euclidean preferences [15]. A preference profile $\succ$ is *single-peaked* if there exists a linear order $>$ over the alternatives in $A := \bigcup_{i \in N} A_i$ such that for each agent i and each triple of alternatives $x, y, z \in A_i$ with $x > y > z$ or $z > y > x$, $x \succ_i y$ implies $y \succ_i z$. A preference profile $\succ$ is *globally-ranked* (we also speak about *correlated markets* [6]) if there exists a global order $>$ over all possible pairs in the matching market such that for every agent i and any two alternatives $x, y \in A_i$, $x \succ_i y$ iff $\{i, x\} > \{i, y\}$. Globally-ranked preferences impose no restriction in a housing market (the agents are matched with objects which do not express preference), but may capture in other markets the idea that each pair of agents generates an absolute profit and thus each agent prefers the agents with who she can get a better profit. A preference profile $\succ$ is *one-dimensional Euclidean (1-Euclidean)* if there exists an embedding $E : N \cup O \to \mathbb{R}$ on the real line such that for every agent i and any two alternatives $x, y \in A_i$, $x \succ_i y$ iff $|E(i) - E(x)| < |E(i) - E(y)|$.

One-dimensional Euclidean preferences form a subdomain of single-peaked preferences because every 1-Euclidean preference profile is singled-peaked for the linear order $>$ given by $x > y$ iff $E(x) > E(y)$. However, a single-peaked preference profile may not be 1-Euclidean, therefore the inclusion is strict. Moreover, one-dimensional Euclidean preferences form a subdomain of globally-ranked preferences: from a 1-Euclidean preference profile, a global ranking over all possible pairs can be extracted by sorting all pairs according to the Euclidean distance on the embedding E between the two partners. Reversely, a globally-ranked preference profile may not be 1-Euclidean, therefore the inclusion is strict. We know that 1-Euclidean preferences are both globally-ranked and single-peaked. However, the reverse is not true: a globally-ranked and single-peaked preference profile may not be 1-Euclidean, even in markets matching agents with each other. We omit the examples due to space restrictions.

While assuming that all agents have 1-Euclidean preferences certainly represents a strong restriction, there are nevertheless some applications where this assumption is not unreasonable. For example, in job markets, preferences could be 1-Euclidean because employees prefer one workplace to another if it is closer to their home, or when forming pairs of students for the realization of a project, a student could prefer to be matched with a student who is the most productive as the same hours as her.

Note that all considered preference restrictions are recognizable in polynomial time: there exist polynomial time algorithms for checking single-peakedness [10, 20] or the satisfaction of the 1-Euclidean property [18, 19, 26]. Moreover, checking whether a preference profile is globally-ranked boils down to checking the acyclicity of the directed graph defined on all possible pairs and where there is an arc from a pair $\{i, j\}$ to a pair $\{i, k\}$ if and only if $k \succ_i j$ [4]; this can be done in polynomial time.

3.3 Rational Swaps

We study sequences of matchings in which two pairs of the current matching are permuted. More formally, we assume that a swap w.r.t. two agents (i, j) transforms a matching μ into a matching μ' where agents i and j have exchanged their matches, i.e., $\mu'(i) = \mu(j)$ and $\mu'(j) = \mu(i)$, while the rest of the matching remains unchanged, i.e., $\mu'(k) = \mu(k)$ for every $k \notin \{i, j, \mu(i), \mu(j)\}$.

We furthermore require these swaps to be *rational* in the sense that they result from an agreement among agents, and thus make the agents involved in the agreement better off.

The most natural notion of rationality is exchange-rationality, which requires that the two agents who exchange their matches are better off [8]. A swap w.r.t. agents (i, j) from matching μ is *exchange rational (ER)* if the agents who exchange their matches are better off, i.e.,

$$\mu(j) \succ_i \mu(i) \text{ and } \mu(i) \succ_j \mu(j). \qquad (ER\text{-swap})$$

Exchange-rationality is the only meaningful notion of swap rationality in housing markets because only one side of the market has preferences. However, several notions of rationality emerge in marriage and roommate markets, where agents are matched with each other. One could demand that only two of the agents who agree to form a new pair need to be better off. This notion of rational swaps is based on the classic idea of *blocking pairs*, which forms the basis of the standard notion of stability [21]. A swap w.r.t. agents (i, j) from matching μ between agents is *blocking pair (BP) rational* if one of the new pairs in μ' forms a blocking pair, where both agents are better off, i.e.,

$$\left[\mu(j) \succ_i \mu(i) \text{ and } i \succ_{\mu(j)} j\right] \quad \text{or} \quad \left[\mu(i) \succ_j \mu(j) \text{ and } j \succ_{\mu(i)} i\right]. \qquad (BP\text{-swap})$$

We refer to a *BP*-swap by mentioning the associated blocking pair $((i, \mu(j))$ or $(j, \mu(i)))$. The old partners of the blocking pair are also assumed to be matched together. [1]

Finally, in marriage and roommate markets, a stronger notion of rationality is that of a *fully rational swap*, which makes all four involved agents better off. A swap w.r.t.

1. Once the old partners are alone, they have an incentive to form a new pair. Roth and Vande Vate [30] therefore decompose *BP*-swaps into two steps. We do not explicitly consider these steps in order to always maintain a perfect matching like, e.g., Knuth [27].

agents (i, j) from matching μ is *fully rational (FR)* if all four agents involved in the swap are better off, i.e.,

$$\mu(j) \succ_i \mu(i), \quad \mu(i) \succ_j \mu(j), \quad j \succ_{\mu(i)} i, \text{ and } \quad i \succ_{\mu(j)} j. \qquad (FR\text{-swap})$$

Note that for marriage and roommate markets, an *FR*-swap w.r.t. pair of agents (i, j) from a matching μ is an *ER*-swap w.r.t. pair (i, j) or $(\mu(i), \mu(j))$ and also a *BP*-swap w.r.t blocking pair $(i, \mu(j))$ or $(j, \mu(i))$. We thus obtain the following implications:

$$BP\text{-swap} \quad \Leftarrow \quad FR\text{-swap} \quad \Rightarrow \quad ER\text{-swap}$$

The different types of swap rationality are illustrated in the following example.

Example 1. *Consider a roommate market with six agents. The preferences of the agents are given below, where the initial assignment is marked with frames.*

$$\begin{array}{lccccccccc}
1: & 4 & \succ & \boxed{3} & \succ & 6 & \succ & 5 & \succ & 2 \\
2: & 3 & \succ & 1 & \succ & \boxed{4} & \succ & 6 & \succ & 5 \\
3: & 6 & \succ & 2 & \succ & \boxed{1} & \succ & 5 & \succ & 4 \\
4: & 5 & \succ & 1 & \succ & 3 & \succ & \boxed{2} & \succ & 6 \\
5: & 2 & \succ & \boxed{6} & \succ & 4 & \succ & 1 & \succ & 3 \\
6: & 4 & \succ & 3 & \succ & 1 & \succ & 2 & \succ & \boxed{5}
\end{array}$$

The swap w.r.t. pair $(1, 2)$*, which matches Agent* 1 *with Agent* 4 *and Agent* 2 *with Agent* 3*, is an FR-swap because every involved agent is better off. Hence, this is also an ER-swap for pair* $(1, 2)$ *or* $(3, 4)$ *because they both prefer to exchange their partner. It is also a BP-swap for blocking pair* $(2, 3)$ *or* $(1, 4)$ *because they both prefer to be together than with their current partner.*

The swap w.r.t. pair $(1, 6)$ *is a BP-swap for blocking pair* $(3, 6)$ *because Agent* 3*, the old partner of Agent* 1*, prefers to be with Agent* 6*, as well as Agent* 6 *who prefers* 3 *to her old partner* 5*. This is not an ER-swap (and hence not an FR-swap) because neither the agents in pair* $(1, 6)$ *nor in pair* $(3, 5)$ *want to exchange their partners.*

The swap w.r.t. pair $(4, 6)$ *is an ER-swap for* $(4, 6)$ *because Agent* 4 *prefers the current partner of* 6*, i.e., Agent* 5*, to her current partner and* 6 *prefers the current partner of* 4*, i.e., Agent* 2*, to her current partner. This is not a BP-swap (and hence not an FR-swap) because it matches Agent* 4 *with Agent* 5*, who prefers to stay with Agent* 6*, and Agent* 6 *with Agent* 2*, who prefers to stay with Agent* 4*.*

Stability can now be defined according to the different notions of rational swaps. A matching μ is *σ-stable*, for $\sigma \in \{FR, ER, BP\}$, if no σ-swap can be performed from matching μ. A sequence of σ-swaps, for $\sigma \in \{FR, ER, BP\}$, corresponds to a sequence of matchings $(\mu^0, \mu^1, \ldots, \mu^r)$ such that a σ-swap transforms each matching μ^t into matching μ^{t+1} for every $0 \leq t < r$. Then, matching μ is *σ-reachable* from initial matching μ^0 if there exists a sequence of σ-swaps $(\mu^0, \mu^1, \ldots, \mu^r)$ such that $\mu^r = \mu$. When the context is clear, we omit σ and the initial matching μ^0.

A *σ-dynamics* is defined according to initial matching μ^0 and a type σ of rational swaps. The σ-dynamics is *finite* if all associated sequences of σ-swaps terminate in a σ-stable matching, and it is said to *converge* if it is finite for every initial matching μ^0.

We consider the following decision problems related to the convergence of dynamics to a Pareto-optimal matching.

∃-σ-ParetoSequence / ∀-σ-ParetoSequence	
Input:	Matching market, type σ of rational swaps
Question:	Does there exist a sequence of σ-swaps terminating in a Pareto-optimal σ-stable matching? / Do all sequences of σ-swaps terminate in a Pareto-optimal σ-stable matching?

In order to tackle these two questions, we also study the stability and convergence properties of the considered dynamics in the three types of matching markets.

4 Exchange Rational Swaps

In housing markets, every *ER*-swap represents a Pareto improvement. Since the number of agents and objects is finite, *ER*-dynamics always converges and the existence of *ER*-stable matchings is guaranteed (simply because every Pareto-optimal matching happens to be *ER*-stable). However, it may be impossible to reach a Pareto-optimal matching from a given matching by only applying *ER*-swaps.

Proposition 1. *ER-dynamics may not converge to a Pareto-optimal matching in housing markets.*

Proof. Consider a housing market with n agents. The preferences of the agents are given below, where the initial assignment is marked with frames and [...] denotes an arbitrary order over the rest of objects.

$$\begin{array}{lccccc}
1: & o_1 & \succ & \boxed{o_2} & \succ & [\ldots] \\
2: & o_2 & \succ & \boxed{o_3} & \succ & [\ldots] \\
3: & o_3 & \succ & \boxed{o_4} & \succ & [\ldots] \\
 & & & \cdots & & \\
n: & o_n & \succ & \boxed{o_1} & \succ & [\ldots]
\end{array}$$

Observe that no *ER*-swap is possible in this instance, therefore the initial matching (framed objects) is the unique *ER*-reachable matching. However, there exists a unique Pareto-optimal matching (circled objects), and this matching is different from the initial one. Note that, in such an instance, even if exchanges involving up to $n-1$ agents are allowed, the Pareto-optimal matching will not be reached: the only *ER*-exchange would involve all the n agents. □

Nevertheless, it is known that *ER*-dynamics always converges to a Pareto-optimal matching in housing markets when the agents' preferences are single-peaked [16].

In marriage and roommate markets, an *ER*-stable matching may not exist, even for single-peaked preferences [8, 13]. However, it turns out that, for globally-ranked preferences, an *ER*-stable matching always exists, and, moreover, the convergence to such a matching is guaranteed.

Proposition 2. *ER-dynamics always converges in marriage / roommate markets for globally-ranked preferences.*

Proof. Denote by $>$ the global order over all possible pairs such that the preferences of the agents are globally-ranked with respect to this global order. Define as $f : \mu \to \mathbb{R}$ the potential function which assigns to each matching the sum of ranks in order $>$ of all the assigned pairs in the matching, i.e., $f(\mu) = \sum_{\{i,j\} s.t. \mu(i)=j} rank_>(\{i, j\}))$ with $rank_>$ the function which gives the rank of the pairs in order $>$. Now consider a sequence of *ER*-swaps given by the sequence of matchings $(\mu^0, \mu^1, \dots, \mu^r)$. Between each matchings μ^t and μ^{t+1}, with $0 \leq t < r$, an *ER*-swap is performed, say w.r.t. pair (i, j) of agents. That means, by definition of an *ER*-swap, that agents i and j prefer to exchange their partners in μ^t, and thus, $\mu^t(j) \succ_i \mu^t(i)$ and $\mu^t(i) \succ_j \mu^t(j)$. This implies, by correlation of the preferences, that $\{i, \mu^t(j)\} > \{i, \mu^t(i)\}$ and $\{j, \mu^t(i)\} > \{j, \mu^t(j)\}$. But agents i and $\mu^t(j)$ are matched in μ^{t+1}, as well as agents j and $\mu^t(i)$. Since the rest of the pairs remains unchanged between μ^t and μ^{t+1}, we get that $f(\mu^{t+1}) < f(\mu^t)$. Because the number of different matchings is finite, we can conclude that *ER*-dynamics always converges. □

In general, an *ER*-stable matching may not be Pareto-optimal, thus convergence to a Pareto-optimal matching is not guaranteed even when an *ER*-stable matching exists (note that deciding the existence of an *ER*-stable matching is NP-hard in marriage and roommate markets [13, 14]).

Proposition 3. *ER-dynamics may not converge to a Pareto-optimal matching, in marriage and roommate markets, even when an ER-stable matching exists and for globally-ranked preferences.*

Proof. Consider a marriage market with three women and three men. The preferences are given below and the initial assignment is marked with frames.

$$\begin{array}{lllll|lllll}
w_1: & m_1 & \succ & \boxed{m_2} & \succ m_3 & m_1: & w_1 & \succ & \boxed{w_3} & \succ w_2 \\
w_2: & m_2 & \succ & \boxed{m_3} & \succ m_1 & m_2: & w_2 & \succ & \boxed{w_1} & \succ w_3 \\
w_3: & m_3 & \succ & \boxed{m_1} & \succ m_2 & m_3: & w_3 & \succ & \boxed{w_2} & \succ w_1
\end{array}$$

No *ER*-swap is possible from initial matching μ^0 (framed agents), therefore μ^0 is the unique *ER*-reachable matching. However, there is another matching (circled agents) which is the unique Pareto-optimal matching. Note that this preference profile is globally-ranked with respect to, e.g., the global order $\{w_1, m_1\} > \{w_2, m_2\} > \{w_3, m_3\} > \{w_1, m_2\} > \{w_2, m_3\} > \{w_3, m_1\} > \{w_1, m_3\} > \{w_2, m_1\} > \{w_3, m_2\}$.

Now, consider a roommate market with six agents. Preferences of the agents are given below, where the initial partner of each agent is marked with frames and [. . .] denotes an arbitrary order over the rest of the agents.

$$\begin{array}{llllll llllll}
1: & 3 & \succ & \boxed{2} & \succ & [\dots] & 4: & 6 & \succ & \boxed{3} & \succ & [\dots] \\
2: & 5 & \succ & \boxed{1} & \succ & [\dots] & 5: & 2 & \succ & \boxed{6} & \succ & [\dots] \\
3: & 1 & \succ & \boxed{4} & \succ & [\dots] & 6: & 4 & \succ & \boxed{5} & \succ & [\dots]
\end{array}$$

No *ER*-swap is possible from initial matching μ^0 (framed agents), thus μ^0 is the unique *ER*-reachable matching. However, there is another matching (circled agents) which is the unique Pareto-optimal matching. This preference profile is globally-ranked w.r.t., e.g., the global order $\{4, 6\} > \{1, 3\} > \{3, 4\} > \{2, 5\} > \{1, 2\} > \{5, 6\} > [\dots]$. □

Note that the above preference profiles are not 1-Euclidean. In fact, they are not even single-peaked. Again, more positive results can be obtained by restricting the domain of admissible preferences.

Proposition 4. *Every ER-stable matching is Pareto-optimal when preferences are single-peaked in marriage and roommate markets.*

Proof. Let μ be an *ER*-stable matching. For any two agents i and j (in N for roommate markets, or both in either W or M for marriage markets) it holds that $\mu(i) \succ_i \mu(j)$ or $\mu(j) \succ_j \mu(i)$. Suppose there is another matching μ' such that $\mu'(i) \succeq_i \mu(i)$ for every $i \in N$ and there exists $j \in N$ such that $\mu'(j) \succ_j \mu(j)$. Then, there exists a Pareto improving cycle from μ to μ' along agents $(n_1, \dots, n_k)$ such that each agent n_i, $1 \leq i \leq k$, is matched in μ' with agent $\mu(n_{(i \bmod k)+1})$. For marriage markets, the agents in $(n_1, \dots, n_k)$ are restricted by definition to only one side of the market, but it impacts both sides since the agents exchange agents of the other side. But there is no problem of preferences of the matched agents because no agent is worse off in μ' compared to μ. The same holds for roommate markets. Since μ is *ER*-stable, it holds that $k > 2$. However, for single-peaked preferences, one can prove, by following the same proof by induction as Damamme et al. [16], that a Pareto improving cycle of any length cannot occur, contradicting the fact that μ is Pareto dominated. □

Propositions 2 and 4 allow us to conclude that sequences of *ER*-swaps will always terminate in Pareto-optimal matchings when preferences are both single-peaked and globally-ranked, like in 1-Euclidean preferences.

Corollary 1. *ER-dynamics always converges to a Pareto-optimal matching in marriage and roommate markets for 1-Euclidean preferences.*

For more general preferences, an interesting computational question is whether, given a preference profile and an initial assignment, a Pareto-optimal matching can be reached via *ER*-swaps. In the context of housing markets, the complexity of this question was mentioned as an open problem by Damamme et al. [16]. It turns out that this problem is computationally intractable for all kinds of matching markets, even for globally-ranked preferences.

Theorem 1. ∃-ER-ParetoSequence *is* NP*-hard in housing, marriage, and roommate markets even for globally-ranked preferences.*

Proof. For the case of housing markets, we perform a reduction from 2P1N-SAT, a variant of SAT known to be NP-complete [33], where the goal is to decide the satisfiability of a CNF propositional formula where each variable appears exactly twice as a positive literal and once as a negative literal. The idea of the proof is close to the one given by Gourvès et al. [22] for proving NP-hardness of determining whether a given object is reachable by a given agent. From an instance of 2P1N-SAT with formula φ on m clauses $C_1, \dots, C_m$ and p variables $x_1, \dots, x_p$, we build a housing market $(N, O, \succ, \mu^0)$ as follows.

For each clause C_j, with $1 \leq j \leq m$, we construct two clause-agents in N denoted by A_j and A'_j and two clause-objects in O denoted by a_j and a'_j such that $\mu^0(A_j) = a_j$ and $\mu^0(A'_j) = a'_j$. For each variable x_i, with $1 \leq i \leq p$, we construct six literal-agents in N corresponding to two copies of each literal, namely agents Y_i^ℓ and Z_i^ℓ who correspond to the ℓ^{th} ($\ell \in \{1, 2\}$) positive occurrence of variable x_i in formula φ, denoted by x_i^ℓ, and $\overline{Y_i}$ and $\overline{Z_i}$ who correspond to the negative occurrence of variable x_i in formula φ, denoted by $\overline{x_i}$; we also create their associated literal-objects y_i^ℓ, z_i^ℓ, $\overline{y_i}$ and $\overline{z_i}$ such that $\mu^0(Y_i^\ell) = y_i^\ell$, $\mu^0(Z_i^\ell) = z_i^\ell$, $\mu^0(\overline{Y_i}) = \overline{y_i}$ and $\mu^0(\overline{Z_i}) = \overline{z_i}$. The literal-agents are divided in two sets, denoted by Y and Z, which correspond to the original agents and their copy, respectively, i.e., $Y := \bigcup_{1 \leq i \leq p} \{Y_i^1, Y_i^2, \overline{Y_i}\}$ and $Z := \bigcup_{1 \leq i \leq p} \{Z_i^1, Z_i^2, \overline{Z_i}\}$. Three additional agents B, T and T' are created in N, with their initial assigned objects denoted by b, t and t', respectively.

The preferences are given below for each $1 \leq i \leq p$ and $1 \leq j < m$ ($[\dots]$ is an arbitrary order over the rest of the objects, $\{y_j\}$ is an arbitrary order over the literal-objects in $\bigcup_{1 \leq i \leq p} \{y_i^1, y_i^2, \overline{y_i}\}$ associated with literals of clause C_j and $cl(\ell_i)$ is the index of the clause in which literal ℓ_i appears).

$$
\begin{array}{rl|rl}
T: & t' \succ \{y_1\} \succ \boxed{t} \succ [\dots] & T': & a'_m \succ \{y_1\} \succ \boxed{t'} \succ [\dots] \\
A_j: & a'_j \succ \{y_{j+1}\} \succ t \succ & A'_j: & a_j \succ \{y_j\} \succ a'_m \succ \\
 & \{y_j\} \succ \boxed{a_j} \succ [\dots] & & \{y_{j+1}\} \succ \boxed{a'_j} \succ [\dots] \\
A_m: & b \succ t \succ \{y_m\} \succ \boxed{a_m} \succ [\dots] & A'_m: & a_m \succ \{y_m\} \succ \boxed{a'_m} \succ [\dots] \\
Y_i^1: & z_i^1 \succ a_{cl(\overline{x_i})} \succ a_{cl(x_i^1)} \succ & Z_i^1: & y_i^1 \succ \overline{y_i} \succ a_{cl(x_i^1)} \succ \\
 & \overline{y_i} \succ \boxed{y_i^1} \succ [\dots] & & a_{cl(\overline{x_i})} \succ \boxed{z_i^1} \succ [\dots] \\
Y_i^2: & z_i^2 \succ y_i^1 \succ a_{cl(x_i^2)} \succ & Z_i^2: & y_i^2 \succ \overline{y_i} \succ y_i^1 \succ \\
 & \overline{y_i} \succ \boxed{y_i^2} \succ [\dots] & & a_{cl(x_i^2)} \succ \boxed{z_i^2} \succ [\dots] \\
\overline{Y_i}: & \overline{z_i} \succ y_i^2 \succ \boxed{\overline{y_i}} \succ [\dots] & \overline{Z_i}: & \overline{y_i} \succ y_i^2 \succ \boxed{\overline{z_i}} \succ [\dots] \\
B: & t \succ \boxed{b} \succ [\dots] & &
\end{array}
$$

We claim that formula φ is satisfiable if and only if the matching assigning to each agent her best object is reachable (this is the only Pareto-optimal matching). The global idea is that the only way to reach this Pareto-optimal matching is to make object t reach agent A_m by first giving to each clause-agent A_j, via *ER*-swaps, a literal-object in $\{y_j\}$, objects associated with the literals of clause C_j. Once object t reaches clause-agent A_m, each agent except A'_m exchanges with her prime version agent (agents Z_i^1, Z_i^2 and $\overline{Z_i}$ are the prime versions of agents Y_i^1, Y_i^2 and $\overline{Y_i}$, respectively, and B is the prime version of A_m), and then the prime agents make among them the reverse sequence of swaps of the initial one where the goal was to make object t reach A_m, leading to the Pareto-optimal matching. By construction of the preferences among the literal-agents, once a literal-object associated with a positive (resp., negative) literal of a variable has been chosen to go with a clause-agent A_j, no literal-object associated with a negative (resp., positive) literal of this variable can reach a clause-agent. The details of the equivalence are omitted due to space restrictions.

To adapt the proof to the case of marriage and roommate markets, we now consider the objects in O as agents. More precisely, we build a marriage market $(N, \succ, \mu^0)$ where $N = M \cup W$ with $W = \{T, T', B, \{A_j, A'_j\}_{1 \leq j \leq m}, \{Y_i^1, Y_i^2, \overline{Y_i}, Z_i^1, Z_i^2, \overline{Z_i}\}_{1 \leq i \leq p}\}$ and $M = \{t, t', b, \{a_j, a'_j\}_{1 \leq j \leq m}, \{y_i^1, y_i^2, \overline{y_i}, z_i^1, z_i^2, \overline{z_i}\}_{1 \leq i \leq p}\}$. The preferences of women over men are the same as the preferences of agents over objects previously described, and the preferences of men over women are described below for $1 \leq j < m$ and $1 \leq i \leq p$. Notation $\{\mathcal{Y}_j\}$ (resp., $\{\mathcal{Z}_j\}$) denotes an arbitrary order over the literal-agents in Y (resp., Z) that are associated with a literal of clause C_j where each "negative" literal-agent $\overline{Y_i}$ (resp., $\overline{Z_i}$) is replaced by agent Y_i^1 (resp., Z_i^1), A_0 stands for T, A'_0 for T', and $[\dots]$ is an arbitrary order over the rest of the women.

$$
\begin{array}{rl|rl}
t: & B \succ A_m \succ \cdots \succ A_1 \succ \boxed{T} \succ [\dots] & t': & T \succ \boxed{T'} \succ [\dots] \\
a_j: & A'_j \succ \{\mathcal{Z}_j\} \succ \{\mathcal{Y}_j\} \succ \boxed{A_j} \succ [\dots] & a'_j: & A_j \succ \boxed{A'_j} \succ [\dots] \\
a_m: & A'_m \succ \{\mathcal{Z}_m\} \succ \{\mathcal{Y}_m\} \succ \boxed{A_m} \succ [\dots] & a'_m: & T' \succ A'_1 \succ \cdots \succ \\
y_i^1: & Z_i^1 \succ Z_i^2 \succ Y_i^2 \succ A'_{cl(x_i^1)} \succ A'_{cl(x_i^1)-1} \succ & & \boxed{A'_m} \succ [\dots] \\
 & A_{cl(x_i^1)-1} \succ A_{cl(x_i^1)} \succ \boxed{Y_i^1} \succ [\dots] & z_i^1: & Y_i^1 \succ \boxed{Z_i^1} \succ [\dots] \\
y_i^2: & Z_i^2 \succ \overline{Z_i} \succ \overline{Y_i} \succ A'_{cl(x_i^2)} \succ A'_{cl(x_i^2)-1} \succ & z_i^2: & Y_i^2 \succ \boxed{Z_i^2} \succ [\dots] \\
 & A_{cl(x_i^2)-1} \succ A_{cl(x_i^2)} \succ \boxed{Y_i^2} \succ [\dots] & \overline{z_i}: & \overline{Y_i} \succ \boxed{\overline{Z_i}} \succ [\dots] \\
\overline{y_i}: & \overline{Z_i} \succ Z_i^2 \succ Z_i^1 \succ A'_{cl(\overline{x_i})} \succ A'_{cl(\overline{x_i})-1} \succ & b: & A_m \succ \boxed{B} \succ [\dots] \\
 & A_{cl(\overline{x_i})-1} \succ A_{cl(\overline{x_i})} \succ Y_i^1 \succ Y_i^2 \succ \boxed{\overline{Y_i}} \succ [\dots] & &
\end{array}
$$

For roommate markets, we consider the same market but without distinguishing between men and women. The only difference in the preferences is that $[\dots]$ is an arbitrary order over all the rest of the agents. In such a way, there is no incentive to partner with an agent who belongs to the other side of the marriage market. Given these preferences, a swap is rational for one side of the market if and only if it is also rational for the other side (in the constructed roommate market, no *ER*-swap can occur between two agents who were from two different sides in the marriage market). In other words, the set of *ER*-swaps is identical to the set of *FR*-swaps. Hence, the sequences of swaps that may occur are exactly the same as in the proof for housing markets.

Note that one can exhibit a global order over pairs such that the preferences are globally ranked. □

Not surprisingly, for preferences more general than those restricted to the 1-Euclidean domain, recognizing the instances where *ER*-dynamics converges to a Pareto-optimal matching is intractable. The proof is omitted due to space restrictions but the idea is close to the proof of Theorem 1.

Theorem 2. ∀-ER-ParetoSequence *is* co-NP*-hard in housing, marriage and roommate markets even for globally-*

ranked preferences.

In housing markets, the size of a sequence of *ER*-swaps is bounded by $O(n^2)$ because every agent involved in a swap is strictly better off. Thus, since checking the Pareto-optimality of a matching in housing markets can be done in polynomial time [3], we get the following corollary.

Corollary 2. $\exists$-ER-ParetoSequence *is* NP*-complete and* $\forall$-ER-ParetoSequence *is* co-NP*-complete in housing markets even for globally-ranked preferences.*

5 Blocking Pair Swaps

BP-swaps cannot occur in housing markets because objects can never be better off. We thus focus in this section on matching markets that match agents with each other.

First, by definition of a blocking pair, any *BP*-stable matching is Pareto-optimal. Moreover, a *BP*-stable matching always exists in marriage markets by the Deferred Acceptance algorithm [21]. However, the convergence to such a state is not guaranteed, even for single-peaked preferences [27]. Nevertheless, there always exists a sequence of *BP*-swaps leading to a stable matching [30]. [2]

In roommate markets, even the existence of a *BP*-stable matching is not guaranteed [21], and actually this is the case even for single-peaked preferences. Nevertheless, checking the existence of a stable matching in a roommate market can be done in polynomial time [24], and there always exists a sequence of *BP*-swaps leading to a stable matching when there exists one [17]. Therefore, by combining these facts with the observation that every *BP*-stable matching is Pareto-optimal, we get the following corollary.

Corollary 3. $\exists$-BP-ParetoSequence *is solvable in polynomial time in marriage and roommate markets.*

However, in general, determining whether all sequences of *BP*-swaps terminate in a Pareto-optimal matching, i.e., checking convergence of *BP*-dynamics to a Pareto-optimal matching, is hard. This is due to the hardness of checking the existence of a cycle in *BP*-dynamics.

Theorem 3. *Determining whether BP-dynamics can cycle in marriage and roommate markets is* NP*-hard.*

Proof. We perform a reduction from (3,B2)-SAT, a variant of 3-SAT known to be NP-complete [11], where the goal is to decide the satisfiability of a CNF propositional formula with exactly three literals per clause and where each variable appears exactly twice as a positive literal and twice as a negative literal. From an instance of (3,B2)-SAT with formula φ on m clauses $C_1, \dots, C_m$ and p variables $x_1, \dots, x_p$, we build a marriage market $(N = W \cup M, \succ, \mu^0)$ as follows. For each clause C_j, with $1 \le j \le m$, we create four clause-agents A_j, B_j, Q_j and K_j, where $A_j, Q_j \in W$ and $B_j, K_j \in M$. For each occurrence of variable x_i, with $1 \le i \le p$, we create two literal-agents, i.e., agents $Z_i^\ell, D_i^\ell \in W$ and $Y_i^\ell, E_i^\ell \in M$ for the ℓ^{th} positive literal x_i^ℓ of x_i, with $\ell \in \{1,2\}$, and $\overline{Z}_i^\ell, \overline{D}_i^\ell \in W$ and $\overline{Y}_i^\ell, \overline{E}_i^\ell \in M$ for the ℓ^{th} negative literal $\overline{x}_i^\ell$ of x_i, with $\ell \in \{1,2\}$. Denote by A, B, Q, K, D, E, Y and Z the sets of agents associated with the same letter.

The preferences are given below, for $1 \le i \le p$, $1 \le j \le m$ and $\ell \in \{1,2\}$, with the initial assignment marked with frames. Notation $\{\mathcal{Y}_j\}$ (resp., $\{\mathcal{D}_j\}$, $\{\mathcal{E}_j\}$ and $\{\mathcal{Z}_j\}$) refers to an arbitrary order over the literal-agents in Y (resp., D, E and Z) corresponding to the literals of clause C_j, and $[\dots]$ is an arbitrary order over the rest of the agents of the other type. In general, when a set is given in the preferences, it refers to an arbitrary order over its elements minus the elements of the set already explicitly given in the rest of the preference ranking. Notation $cl(\ell_i)$ refers to the index of the clause in which literal ℓ_i appears. Note that A_0 (resp., B_0) stands for A_m (resp., B_m) and $\{\mathcal{Y}_{m+1}\}$ stands for $\{\mathcal{Y}_1\}$.

$$
\begin{array}{ll|ll}
A_j: & \{\mathcal{Y}_{j+1}\} \succ \{\mathcal{Y}_j\} \succ \boxed{B_j} \succ B \succ [\dots] & B_j: & \boxed{A_j} \succ A \succ \{\mathcal{Z}_{j+1}\} \succ \{\mathcal{Z}_j\} \succ [\dots] \\
 & & & \\
Z_i^1: & \overline{Y}_i^1 \succ \overline{Y}_i^2 \succ Y_i^1 \succ \boxed{E_i^1} \succ & Y_i^1: & \boxed{D_i^1} \succ \{\mathcal{D}_{cl(x_i^1)}\} \succ Q \succ \overline{Z}_i^1 \succ \overline{Z}_i^2 \succ \\
 & \quad Y \succ B_{cl(x_i^1)} \succ B_{cl(x_i^1)-1} \succ [\dots] & & \quad A_{cl(x_i^1)} \succ A_{cl(x_i^1)-1} \succ Z_i^1 \succ Q_{cl(x_i^1)} \succ [\dots] \\
Z_i^2: & \overline{Y}_i^2 \succ \overline{Y}_i^1 \succ Y_i^2 \succ \boxed{E_i^2} \succ & Y_i^2: & \boxed{D_i^2} \succ \{\mathcal{D}_{cl(x_i^2)}\} \succ Q \succ \overline{Z}_i^2 \succ \overline{Z}_i^1 \succ \\
 & \quad Y \succ B_{cl(x_i^2)} \succ B_{cl(x_i^2)-1} \succ [\dots] & & \quad A_{cl(x_i^2)} \succ A_{cl(x_i^2)-1} \succ Z_i^2 \succ Q_{cl(x_i^2)} \succ [\dots] \\
\overline{Z}_i^1: & Y_i^1 \succ Y_i^2 \succ \overline{Y}_i^1 \succ \boxed{\overline{E}_i^1} \succ & \overline{Y}_i^1: & \boxed{\overline{D}_i^1} \succ \{\mathcal{D}_{cl(\overline{x}_i^1)}\} \succ Q \succ Z_i^1 \succ Z_i^2 \succ \\
 & \quad Y \succ B_{cl(\overline{x}_i^1)} \succ B_{cl(\overline{x}_i^1)-1} \succ [\dots] & & \quad A_{cl(\overline{x}_i^1)} \succ A_{cl(\overline{x}_i^1)-1} \succ \overline{Z}_i^1 \succ Q_{cl(\overline{x}_i^1)} \succ [\dots] \\
\overline{Z}_i^2: & Y_i^2 \succ Y_i^1 \succ \overline{Y}_i^2 \succ \boxed{\overline{E}_i^2} \succ & \overline{Y}_i^2: & \boxed{\overline{D}_i^2} \succ \{\mathcal{D}_{cl(\overline{x}_i^2)}\} \succ Q \succ Z_i^2 \succ Z_i^1 \succ \\
 & \quad Y \succ B_{cl(\overline{x}_i^2)} \succ B_{cl(\overline{x}_i^2)-1} \succ [\dots] & & \quad A_{cl(\overline{x}_i^2)} \succ A_{cl(\overline{x}_i^2)-1} \succ \overline{Z}_i^2 \succ Q_{cl(\overline{x}_i^2)} \succ [\dots] \\
 & & & \\
D_i^\ell: & K_{cl(x_i^\ell)} \succ \boxed{Y_i^\ell} \succ Y \succ [\dots] & E_i^\ell: & Q_{cl(x_i^\ell)} \succ \boxed{Z_i^\ell} \succ Y \succ [\dots] \\
\overline{D}_i^\ell: & K_{cl(\overline{x}_i^\ell)} \succ \boxed{\overline{Y}_i^\ell} \succ Y \succ [\dots] & \overline{E}_i^\ell: & Q_{cl(\overline{x}_i^\ell)} \succ \boxed{\overline{Z}_i^\ell} \succ Y \succ [\dots] \\
 & & & \\
Q_j: & \{\mathcal{Y}_j\} \succ \boxed{K_j} \succ \{\mathcal{E}_j\} & K_j: & \{\mathcal{D}_j\} \succ \boxed{Q_j}
\end{array}
$$

We claim that *BP*-dynamics can cycle if and only if formula φ is satisfiable. The global idea of the reduction is the following. At the initial matching, the only possible *BP*-swaps involve blocking pairs with literal-agents in D and clause-agents in K associated with the same clause. By their swap, a literal-agent in D associated with clause C_j and clause-agent K_j can "unlock" exactly one literal-agent in Y associated with clause C_j who will not be matched with her most preferred agent anymore, and thus could have an incentive to form a blocking pair. By construction of the preferences, the only possibility to get a cycle in *BP*-dynamics is that, for each clause C_j, exactly one literal-agent $\mathtt{Y}_j$ in Y associated with C_j is unlocked and the cycle involves a sequence of blocking pairs $(A_j, \mathtt{Y}_j), (A_j, \mathtt{Y}_{j+1}), (A_{j+1}, \mathtt{Y}_{j+1}), \dots$ (with $j+1$ modulo m) all along the m clauses. For this cycle to occur, the unlocked literal-agents in Y must have been matched with their associated agent in Z. Therefore, two unlocked literal-agents in Y participating in the cycle cannot correspond to opposite literals, otherwise one of them would be matched at a moment of the cycle with an agent in Z corresponding to her opposite literal, and thus would not agree to form a

2. Assuming that the old partners also form a new pair does not alter this result.

blocking pair with a clause-agent. The details of the equivalence are omitted. This proof can be adapted to roommate markets by assuming that, in the preferences, [. . .] is an arbitrary order over the remaining agents where the agents of the same "type" in the marriage market are ranked last. □

Corollary 4. *∀-BP-ParetoSequence is co-NP-hard in marriage and roommate markets.*

Nevertheless, when the preferences are globally-ranked, we can always reach a stable matching thanks to *BP*-dynamics in both settings. Indeed, it has been proved that *BP*-dynamics always converges in marriage markets with globally-ranked preferences [7]. In roommate markets, there always exists a unique *BP*-stable matching under 1-Euclidean preferences [9]. We prove that convergence to this matching is guaranteed using a potential function argument, and further, this holds for more general preferences, namely globally-ranked preferences.

Proposition 5. *BP-dynamics always converges in roommate markets for globally-ranked preferences.*

Proof. Denote by $>$ the global order over all possible pairs such that the preferences of the agents are globally-ranked with respect to $>$. Let $d(\mu)$ be the $n/2$-vector of the ranks in $>$ of all the different pairs of μ, i.e., $d(\mu) = (rank_>(\{i, j\}))_{i,j \text{ s.t. } \mu(i)=j}$ with $rank_>$ the function which gives the rank of the pairs in order $>$. Consider a sequence of *BP*-swaps given by the following sequence of matchings $(\mu^0, \mu^1, \dots, \mu^r)$. Between each pair of matchings μ^t and μ^{t+1} with $0 \le t < r$, a *BP*-swap is performed, say w.r.t. blocking pair (i, j). By definition of a *BP*-swap, agents i and j prefer to be together than being with their partner in μ^t, so $j = \mu^{t+1}(i) \succ_i \mu^t(i)$ and $i = \mu^{t+1}(j) \succ_j \mu^t(j)$, which implies, by correlation of the preferences, that $\{i, j\} > \{i, \mu^t(i)\}$ and $\{i, j\} > \{j, \mu^t(j)\}$. Therefore, $(rank_>(\{i, j\}), rank_>(\{\mu^t(i), \mu^t(j)\}))$ is lexicographically strictly smaller than $(rank_>(\{i, \mu^t(i)\}), rank_>(\{j, \mu^t(j)\}))$. Since the rest of the pairs remains unchanged between μ^t and μ^{t+1}, it follows that $d(\mu^{t+1})$ is lexicographically strictly smaller than $d(\mu^t)$. The number of different matchings is finite, therefore *BP*-dynamics always converges. □

Since every *BP*-stable matching is Pareto-optimal, we obtain the following corollary.

Corollary 5. *BP-dynamics always converges to a Pareto-optimal matching in marriage and roommate markets when the preferences are globally-ranked.*

6 Fully Rational Swaps

Just as in the case of *ER*-swaps and housing markets, *FR*-swaps always represent Pareto improvements because all involved agents are strictly better off after the swap. Hence, *FR*-stable matchings are guaranteed to exist because every Pareto-optimal matching is *FR*-stable and *FR*-dynamics always converges because the number of agents is finite.

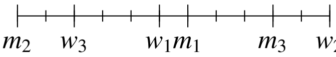
In Section 4, we have shown that *ER*-dynamics always converges to a Pareto-optimal matching when the preferences of the agents are 1-Euclidean. It turns out that this does not hold for *FR*-dynamics.

Proposition 6. *A sequence of FR-swaps may not converge to a Pareto-optimal matching in marriage and roommate markets, even for 1-Euclidean preferences.*

Proof. Consider a marriage market with three women and three men. The preferences are given below, where the initial assignment is marked with frames.

$$\begin{array}{lccccc|lccccc}
w_1: & m_1 & \succ & \boxed{m_3} & \succ & m_2 & m_1: & w_1 & \succ & \boxed{w_3} & \succ & w_2 \\
w_2: & m_3 & \succ & m_1 & \succ & \boxed{m_2} & m_2: & w_3 & \succ & w_1 & \succ & \boxed{w_2} \\
w_3: & m_2 & \succ & \boxed{m_1} & \succ & m_3 & m_3: & w_2 & \succ & \boxed{w_1} & \succ & w_3
\end{array}$$

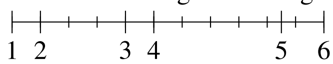
Initial matching μ^0 is the only reachable matching, because there is no *FR*-swap from μ^0. However, there is another matching (circled agents) which is not reachable but which Pareto dominates matching μ^0. The preferences are 1-Euclidean w.r.t. the following embedding on the real line.



Now, consider a roommate market with six agents. The preferences of the agents are given below, where the initial assignment is marked with frames.

$$\begin{array}{lccccccccc|lccccccccc}
1: & 2 & \succ & 3 & \succ & 4 & \succ & 5 & \succ & \boxed{6} & 4: & 3 & \succ & 2 & \succ & \boxed{5} & \succ & 1 & \succ & 6 \\
2: & 1 & \succ & \boxed{3} & \succ & 4 & \succ & 5 & \succ & 6 & 5: & 6 & \succ & \boxed{4} & \succ & 3 & \succ & 2 & \succ & 1 \\
3: & 4 & \succ & \boxed{2} & \succ & 1 & \succ & 5 & \succ & 6 & 6: & 5 & \succ & 4 & \succ & 3 & \succ & 2 & \succ & \boxed{1}
\end{array}$$

Initial matching μ^0 is the only reachable matching, because there is no *FR*-swap from μ^0. However, there is another matching (circled agents) which is not reachable but which Pareto dominates matching μ^0. The preferences are 1-Euclidean w.r.t. the following embedding on the real line.



□

The proofs of Theorems 1 and 2 only dealt with instances in which *FR*-swaps are identical to *ER*-swaps. We thus immediately obtain hardness of ∃-FR-ParetoSequence and ∀-FR-ParetoSequence. An *FR*-swap makes four agents strictly better off and no agent worse off, thus the size of a sequence of *FR*-swaps is bounded by $O(n^2)$. Moreover, the Pareto-optimality of a matching can be checked in polynomial time [5], therefore we get the membership of the problems to NP and co-NP, respectively.

Theorem 4. *∃-FR-ParetoSequence is NP-complete and ∀-FR-ParetoSequence is co-NP-complete in marriage and roommate markets even for globally-ranked preferences.*

7 Conclusion

We have studied the properties of different dynamics of rational swaps in matching markets with initial assignments and, in particular, the question of convergence to a Pareto-optimal matching. For all considered settings, the dynamics may not terminate in a Pareto-optimal matching because *(i)* there is no stable matching, *(ii)* the dynamics does not converge, or *(iii)* the stable matching that is eventually reached is not Pareto-optimal. An overview of our results is given in Table 1.

Market	Preferences	*ER*-Swaps	*BP*-Swaps	*FR*-Swaps
Housing	General / GR	Conv		
	SP	Pareto [16]		
	1-D	Pareto		
Marriage	General	–	Stable [21]	Conv
	GR	Conv (Prop. 2)	Pareto (Cor. 5)	Conv
	SP	– [13]	Stable	Conv
	1-D	Pareto (Cor. 1)	Pareto	Conv
Roommate	General	–	– [21]	Conv
	GR	Conv (Prop. 2)	Pareto (Cor. 5)	Conv
	SP	– [8]	–	Conv
	1-D	Pareto (Cor. 1)	Pareto	Conv

Pareto ⇒ Conv ⇒ Stable

Table 1 – Summary of the results on the existence of a stable matching (Stable), the guarantee of convergence (Conv) and the guarantee of convergence to a Pareto-optimal matching (Pareto) for the three different matching markets under study, according to different types of rational swaps and under different preference domains (General, globally-ranked (GR), single-peaked (SP), and 1-Euclidean (1-D)). Since Pareto ⇒ Convergence ⇒ Stable, we only mention the strongest result which is satisfied. The only meaningful type of rational swaps in housing markets are exchange-rational swaps; hence, the empty spaces.

Computationally, determining whether there *exists* a sequence of rational swaps terminating in a Pareto-optimal matching is NP-hard for fully rational swaps and exchange rational swaps in all matching markets even for globally-ranked preferences (Theorems 1 and 4). For swaps based on blocking pairs, this problem can be solved efficiently (Corollary 3). However, the convergence to a Pareto-optimal matching, that is whether *all* sequences of swaps terminate in a Pareto-optimal matching, is co-NP-hard to decide (Corollary 4). Not surprisingly, the same hardness result holds for fully rational and exchange rational swaps, even for globally-ranked preferences (Theorems 2 and 4). Our computational results are summarized in Table 2. Even if the existence of a sequence of swaps terminating in a Pareto-optimal matching is not guaranteed for single-peaked preferences in marriage and roommate markets, it would be interesting to know whether this preference restriction is nevertheless sufficient for efficiently solving our computational problems in these markets.

Market	Prefs	*ER*-Swaps ∃-ParSeq	*ER*-Swaps ∀-ParSeq	*BP*-Swaps ∃-ParSeq	*BP*-Swaps ∀-ParSeq	*FR*-Swaps ∃-ParSeq	*FR*-Swaps ∀-ParSeq
Housing	General / GR	NP-c. (Cor. 2)	co-NP-c. (Cor. 2)				
	SP	P [16]	P [16]				
Marriage / Roommate	General	NP-h. (Th. 1)	co-NP-h. (Th. 2)	P (Cor. 3)	co-NP-h. (Cor. 4)	NP-c. (Th. 4)	co-NP-c. (Th. 4)
	GR	NP-h. (Th. 1)	co-NP-h. (Th. 2)	P (Cor. 3)	P (Cor. 5)	NP-c. (Th. 4)	co-NP-c. (Th. 4)

Table 2 – Summary of the computational results on the existence (∃-ParSeq) or the guarantee (∀-ParSeq) of sequences of rational swaps terminating in a Pareto-optimal matching for the three different matching markets under study, according to different types of rational swaps and under different preference domains (General, globally-ranked (GR) and single-peaked (SP)). P means polynomial time solvable. The only meaningful type of rational swaps in housing markets are *ER*-swaps; hence, the empty spaces.

The convergence to a Pareto-optimal matching in housing markets for exchange rational dynamics and single-peaked preferences [16] does not hold for more general settings where the "objects" are agents with preferences. However, this convergence is guaranteed under 1-Euclidean preferences in marriage and roommate markets. Hence, the generalization of this convergence result to more general settings requires more structure in the preferences.

A natural extension of this work would be to study meaningful dynamics for hedonic games, where agents form groups consisting of more than two agents.

References

[1] Abdulkadiroğlu, A. et T. Sönmez: *House Allocation with Existing Tenants*. Journal of Economic Theory, 88(2) :233–260, 1999.

[2] Abeledo, H. et U. G. Rothblum: *Paths to marriage stability*. Discrete Applied Mathematics, 63(1) :1–12, 1995.

[3] Abraham, D. J., K. Cechlárová, D. F. Manlove et K. Mehlhorn: *Pareto Optimality in House Allocation Problems*. Dans *Proc. of 16th ISAAC*, tome 3341 de *LNCS*, pages 1163–1175, 2005.

[4] Abraham, D. J., A. Leravi, D. F. Manlove et G. O'Malley: *The stable roommates problem with globally-ranked pairs*. Dans *Proc. of 3rd WINE*, tome 4858 de *LNCS*, pages 431–444. Springer, 2007.

[5] Abraham, D. J. et D. F. Manlove: *Pareto optimality in the Roommates problem*. Rapport technique TR-2004-182, University of Glasgow, Department of Computing Science, 2004.

[6] Ackermann, H., P. W. Goldberg, V. S. Mirrokni, H. Röglin et B. Vöcking: *A Unified Approach to Congestion Games and Two-Sided Markets*. Dans *Proc. of 3rd WINE*, pages 30–41. Springer Berlin Heidelberg, 2007.

[7] Ackermann, H., P. W. Goldberg, V. S. Mirrokni, H. Röglin et B. Vöcking: *Uncoordinated two-sided matching markets*. SIAM Journal on Computing, 40(1) :92–106, 2011.

[8] Alcalde, J.: *Exchange-proofness or divorce-proofness ? Stability in one-sided matching markets*. Review of Economic Design, 1(1) :275–287, 1994.

[9] Arkin, E. M., S. W. Bae, A. Efrat, K. Okamoto, J. S. B. Mitchell et V. Polishchuk: *Geometric stable roommates*. Information Processing Letters, 109(4) :219–224, 2009.

[10] Bartholdi, III, J. et M. Trick: *Stable matching with preferences derived from a psychological model*. Operations Research Letters, 5(4) :165–169, 1986.

[11] Berman, P., M. Karpinski et A. Scott: *Approximation hardness of short symmetric instances of MAX-3SAT*. Rapport technique TR03-049, ECCC, 2003.

[12] Black, D.: *On the Rationale of Group Decision-making*. Journal of Political Economy, 56(1) :23–34, 1948.

[13] Cechlárová, K.: *On the complexity of exchange-stable roommates*. Discrete Applied Mathematics, 116(3) :279–287, 2002.

[14] Cechlárová, K. et D. F. Manlove: *The exchange-stable marriage problem*. Discrete Applied Mathematics, 152(1–3) :109–122, 2005.

[15] Coombs, C. H.: *Psychological scaling without a unit of measurement*. Psychological Review, 57(3) :145–158, 1950.

[16] Damamme, A., A. Beynier, Y. Chevaleyre et N. Maudet: *The Power of Swap Deals in Distributed Resource Allocation*. Dans *Proc. of 14th AAMAS Conference*, pages 625–633. IFAAMAS, 2015.

[17] Diamantoudi, E., E. Miyagawa et L. Xue: *Random paths to stability in the roommate problem*. Games and Economic Behavior, 48(1) :18–28, 2004.

[18] Doignon, J. P. et J. C. Falmagne: *A polynomial time algorithm for unidimensional unfolding representations*. Journal of Algorithms, 16(2) :218–233, 1994.

[19] Elkind, E. et P. Faliszewski: *Recognizing 1-Euclidean preferences : An alternative approach*. Dans *Proc. of 7th SAGT*, pages 146–157. Springer, 2014.

[20] Escoffier, B., J. Lang et M. Öztürk: *Single-Peaked Consistency and its Complexity*. Dans *Proc. of 18th European Conference on Artificial Intelligence (ECAI)*, pages 366–370. IOS Press, 2008.

[21] Gale, D. et L. S. Shapley: *College Admissions and the Stability of Marriage*. The American Mathematical Monthly, 69(1) :9–15, 1962.

[22] Gourvès, L., J. Lesca et A. Wilczynski: *Object Allocation via Swaps along a Social Network*. Dans *Proc. of 26th IJCAI*, pages 213–219. IJCAI, 2017.

[23] Huang, S. et M. Xiao: *Object reachability via swaps along a line*. Dans *Proc. of 33rd AAAI Conference*, pages 2037–2044. AAAI Press, 2019.

[24] Irving, R. W.: *An Efficient Algorithm for the "Stable Roommates" Problem*. Journal of Algorithms, 6(4) :577–595, 1985.

[25] Klaus, B., D. F. Manlove et F. Rossi: *Matching under Preferences*. Dans Brandt, F., V. Conitzer, U. Endriss, J. Lang et A. D. Procaccia (rédacteurs) : *Handbook of Computational Social Choice*, chapitre 14. Cambridge University Press, 2016.

[26] Knoblauch, V.: *Recognizing one-dimensional Euclidean preference profiles*. Journal of Mathematical Economics, 46(1) :1–5, 2010.

[27] Knuth, D. E.: *Mariages stables*. Les Presses de l'Université de Montréal, 1976.

[28] Manlove, D. F.: *Algorithmics of Matching Under Preferences*. World Scientific Publishing Company, 2013.

[29] Morrill, T.: *The roommates problem revisited*. Journal of Economic Theory, 145(5) :1739–1756, 2010.

[30] Roth, A. E. et J. H. Vande Vate: *Random Paths to Stability in Two-Sided Matching*. Econometrica, 58(6) :1475–1480, 1990.

[31] Saffidine, A. et A. Wilczynski: *Constrained Swap Dynamics over a Social Network in Distributed Resource Reallocation*. Dans *Proc. of 11th SAGT*, tome 11059 de *LNCS*, pages 213–225. Springer, 2018.

[32] Shapley, L. S. et H. Scarf: *On cores and indivisibility*. Journal of Mathematical Economics, 1(1) :23–37, 1974.

[33] Yoshinaka, R.: *Higher-Order Matching in the Linear Lambda Calculus*. Dans *Proc. of 16th RTA*, pages 235–249, 2005.