Reaching Individually Stable Coalition Structures in Hedonic Games

Felix Brandt,[1] **Martin Bullinger,** [1] **Anaëlle Wilczynski** [2]
[1] Institut für Informatik, Technische Universität München
[2] MICS, CentraleSupélec, Université Paris-Saclay, France
brandtf@in.tum.de, bullinge@in.tum.de, anaelle.wilczynski@centralesupelec.fr

Abstract

The formal study of coalition formation in multiagent systems is typically realized using so-called hedonic games, which originate from economic theory. The main focus of this branch of research has been on the existence and the computational complexity of deciding the existence of coalition structures that satisfy various stability criteria. The actual process of forming coalitions based on individual behavior has received little attention. In this paper, we study the convergence of simple dynamics leading to stable partitions in a variety of classes of hedonic games, including anonymous, dichotomous, fractional, and hedonic diversity games. The dynamics we consider is based on individual stability: an agent will join another coalition if she is better off and no member of the welcoming coalition is worse off. We identify conditions for convergence, provide elaborate counterexamples of existence of individually stable partitions, and study the computational complexity of problems related to the coalition formation dynamics. In particular, we settle open problems suggested by Bogomolnaia and Jackson (2002), Brandl, Brandt, and Strobel (2015), and Boehmer and Elkind (2020).

1 Introduction

Coalitions and coalition formation are central concerns in the study of multiagent systems as well as cooperative game theory. Typical real-world examples include individuals joining clubs or societies such as orchestras, choirs, or sport teams, countries organizing themselves in international bodies like the European Union (EU) or the North Atlantic Treaty Organization (NATO), students living together in shared flats, or employees forming unions. The formal study of coalition formation is often realized using so-called hedonic games, which originate from economic theory and focus on coalition structures (henceforth partitions) that satisfy various stability criteria based on the agents' preferences over coalitions. A partition is defined to be stable if single agents or groups of agents cannot gain by deviating from the current partition by means of leaving their current coalition and joining another coalition or forming a new one. Which kinds of deviations are permitted depends on the underlying notion of stability. Two important and well-studied questions in this context concern the existence of stable partitions in restricted classes of hedonic games and the computational complexity of finding a stable partition. However, stability is only concerned with the end-state of the coalition formation process and ignores how these desirable partitions can actually be reached. Essentially, an underlying assumption in most of the existing work is that there is a central authority that receives the preferences of all agents, computes a stable partition, and has the means to enforce this partition on the agents. By contrast, our work focuses on simple dynamics, where starting with some partition (e.g., the partition of singletons), agents deliberately decide to join and leave coalitions based on their individual preferences. We study the convergence of such a process and the stable partitions that can arise from it. For example, in some cases the only partition satisfying a certain stability criterion is the grand coalition consisting of all agents, while the dynamics based on the agents' individual decisions can never reach this partition and is doomed to cycle.

The dynamics we consider is based on *individual stability*, a natural notion of stability going back to Drèze and Greenberg (1980): an agent will join another coalition if she is better off and no member of the welcoming coalition is worse off. Individual stability is suitable to model the situations mentioned above. For instance, by Article 49 of the Treaty on European Union, admitting new members to the EU requires the unanimous approval of the current members. Similarly, by Article 10 of their founding treaty, unanimous agreement of all parties is necessary to become a member of the NATO. Also, for joining a choir or orchestra it is often necessary to audition successfully, and joining a shared flat requires the consent of all current residents. This distinguishes individual stability from Nash stability, which ignores the consent of members of the welcoming coalition.

The analysis of coalition formation processes provides more insight in the natural behavior of agents and the conditions that are required to guarantee that desirable social outcomes can be reached without a central authority. Similar dynamic processes have been studied in the special domain of matching, which only allows coalitions of size 2 (e.g., Roth and Vande Vate 1990; Abeledo and Rothblum 1995; Brandt and Wilczynski 2019). More recently, the dynamics of coalition formation have also come under scrutiny in the context of hedonic games (Bilò et al. 2018; Hoefer, Vaz, and Wagner 2018; Carosi, Monaco, and Moscardelli 2019).

While coalition formation dynamics are an object of study worthy for itself, they can also be used as a means to design algorithms that compute stable outcomes, and have been implicitly used for this purpose before. For example, the algorithm by Boehmer and Elkind (2020) for finding an individually stable partition in hedonic diversity games predefines a promising partition and then reaches an individually stable partition by running the dynamics from there. Similarly, the algorithm by Bogomolnaia and Jackson (2002) for finding an individually stable partition on games with ordered characteristics, a generalization of anonymous hedonic games, runs the dynamics using a specific sequence of deviations starting from the singleton partition.

In many cases, the convergence of the dynamics of deviations follows from the existence of potential functions, whose local optima form individually stable states. Generalizing a result by Bogomolnaia and Jackson (2002), Suksompong (2015) has shown via a potential function argument that an individually stable—and even a Nash stable—partition always exists in subset-neutral hedonic games, a generalization of symmetric additively-separable hedonic games. Using the same potential function, it can straightforwardly be shown that the dynamics converge.[1]

Another example are hedonic games with the common ranking property, a class of hedonic games where preferences are induced by a common global order (Farrell and Scotchmer 1988). The dynamics associated with core-stable deviations is known to converge to a core-stable partition that is also Pareto-optimal, thanks to a potential function argument (Caskurlu and Kizilkaya 2019). The same potential function implies convergence of the dynamics based on individual stability.

In this paper, we study the coalition formation dynamics based on individual stability for a variety of classes of hedonic games, including anonymous hedonic games (AHGs), hedonic diversity games (HDGs), fractional hedonic games (FHGs), and dichotomous hedonic games (DHGs). Whether we obtain positive or negative results often depends on the initial partition and on restrictions imposed on the agents' preferences. Computational questions related to the dynamics are investigated in two ways: the existence of a *path to stability*, that is the existence of a sequence of deviations that leads to a stable state, and the *guarantee of convergence* where every sequence of deviations should lead to a stable state. The former gives an optimistic view on the behavior of the dynamics and may be used to motivate the choice of reachable stable partitions (we can exclude "artificial" stable partitions that may never naturally form). If such a sequence can be computed efficiently, it enables a central authority to coordinate the deviations towards a stable partition. However, since this approach does not give any guarantee on the outcome of the dynamics, we also study the latter, more pessimistic, problem. Our main results are as follows.

[1]By inclusion, convergence also holds for symmetric additively-separable hedonic games. Symmetry is essential for this result to hold since an individually stable partition may not exist in additively-separable hedonic games, even under additional restrictions (Bogomolnaia and Jackson 2002).

- In AHGs, the dynamics converges for (naturally) single-peaked strict preferences. We provide a 15-agent example showing the non-existence of individually stable partitions in general AHGs. The previous known counterexample by Bogomolnaia and Jackson (2002) requires 63 agents and the existence of smaller examples was an acknowledged open problem (see Ballester 2004; Boehmer and Elkind 2020).
- In HDGs, the dynamics converges for strict and naturally singled-peaked preferences when starting from the singleton partition. In contrast to empirical evidence reported by Boehmer and Elkind (2020), we show that these preference restrictions are not sufficient to guarantee convergence from an arbitrary initial partition.
- In FHGs, the dynamics converges for simple symmetric preferences when starting from the singleton partition or when preferences form an acyclic digraph. We show that individually stable partitions need not exist in general symmetric FHGs, which was left as an open problem by Brandl, Brandt, and Strobel (2015).
- For each of these four classes, including DHGs, we show that deciding whether there is a sequence of deviations leading to an individually stable partition is NP-hard while deciding whether all sequences of deviations lead to an individually stable partition is co-NP-hard. Some of these results hold under preference restrictions and even when starting from the singleton partition.

2 Preliminaries

Let $N = [n] = \{1, \dots, n\}$ be a set of n agents. The goal of a coalition formation problem is to partition the agents into different disjoint coalitions according to their preferences. A solution is then a partition $\pi : N \to 2^N$ such that $i \in \pi(i)$ for every agent $i \in N$ and either $\pi(i) = \pi(j)$ or $\pi(i) \cap \pi(j) = \emptyset$ holds for every agents i and j, where $\pi(i)$ denotes the coalition to which agent i belongs. Two prominent partitions are the *singleton partition* π given by $\pi(i) = \{i\}$ for every agent $i \in N$, and the *grand coalition* π given by $\pi = \{N\}$.

Since we focus on dynamics of deviations, we assume that there exists an initial partition π_0, which could be a natural initial state (such as the singleton partition) or the outcome of a previous coalition formation process.

2.1 Classes of Hedonic Games

In a hedonic game, the agents only express preferences over the coalitions to which they belong, i.e., there are no externalities. Let $\mathcal{N}_i$ denote all possible coalitions containing agent i, i.e., $\mathcal{N}_i = \{C \subseteq N : i \in C\}$. A hedonic game is defined by a tuple $(N, (\succsim_i)_{i \in N})$ where $\succsim_i$ is a weak order over $\mathcal{N}_i$ which represents the preferences of agent i. Since $|\mathcal{N}_i| = 2^{n-1}$, the preferences are rarely given explicitly, but rather in some concise representation. These representations give rise to several classes of hedonic games:

- *Anonymous hedonic games (AHGs)* (Bogomolnaia and Jackson 2002): The agents only care about the size of the coalition they belong to, i.e., for each agent $i \in N$, there exists a weak order $\succsim_i$ over integers in $[n]$ such that $\pi(i) \succsim_i \pi'(i)$ iff $|\pi(i)| \succsim_i |\pi'(i)|$.

- *Hedonic diversity games (HDGs)* (Bredereck, Elkind, and Igarashi 2019): The agents are divided into two different types, red and blue agents, represented by the subsets R and B, respectively, such that $N = R \cup B$ and $R \cap B = \emptyset$. Each agent only cares about the proportion of red agents present in her own coalition, i.e., for each agent $i \in N$, there exists a weak order $\succsim_i$ over $\{\frac{p}{q} : p \in [|R|] \cup \{0\}, q \in [n]\}$ such that $\pi(i) \succsim_i \pi'(i)$ iff $\frac{|R \cap \pi(i)|}{|\pi(i)|} \succsim_i \frac{|R \cap \pi'(i)|}{|\pi'(i)|}$.
- *Fractional Hedonic Games (FHGs)* (Aziz et al. 2019): The agents evaluate a coalition according to how much they like each of its members on average, i.e., for each agent i, there exists a utility function $v_i : N \to \mathbb{R}$ where $v_i(i) = 0$ such that $\pi(i) \succsim_i \pi'(i)$ iff $\frac{\sum_{j \in \pi(i)} v_i(j)}{|\pi(i)|} \geq \frac{\sum_{j \in \pi'(i)} v_i(j)}{|\pi'(i)|}$. An FHG can be represented by a weighted complete directed graph $G = (N, E)$ where the weight of arc (i, j) is equal to $v_i(j)$. An FHG is *symmetric* if $v_i(j) = v_j(i)$ for every pair of agents i and j, i.e., it can be represented by a weighted complete undirected graph with weights $v(i, j)$ on each edge $\{i, j\}$. An FHG is simple if $v_i : N \to \{0, 1\}$ for every agent i, i.e., it can be represented by an unweighted directed graph where $(i, j) \in E$ iff $v_i(j) = 1$. We say that a simple FHG is *asymmetric* if, for every pair of agents i and j, $v_i(j) = 1$ implies $v_j(i) = 0$, i.e., it can be represented by an asymmetric directed graph.
- *Dichotomous hedonic games (DHGs)*: The agents only approve or disapprove coalitions, i.e., for each agent i there exists a utility function $v_i : \mathcal{N}_i \to \{0, 1\}$ such that $\pi(i) \succsim_i \pi'(i)$ iff $v_i(\pi(i)) \geq v_i(\pi'(i))$. When the preferences are represented by a propositional formula, such games are called *Boolean hedonic games* (Aziz et al. 2016).

An anonymous game (resp., hedonic diversity game) is *generally single-peaked* if there exists a linear order $>$ over integers in $[n]$ (resp., over ratios in $\{\frac{p}{q} : p \in [|R|] \cup \{0\}, q \in [n]\}$) such that for each agent $i \in N$ and each triple of integers $x, y, z \in [n]$ (resp., $x, y, z \in \{\frac{p}{q} : p \in |R| \cup \{0\}, q \in [n]\}$) with $x > y > z$ or $z > y > x$, $x \succsim_i y$ implies $y \succsim_i z$. The obvious linear order $>$ that comes to mind is, of course, the natural order over integers (resp., over rational numbers). We refer to such games as *naturally single-peaked*. Clearly, a naturally single-peaked preference profile is generally single-peaked but the converse is not true.

2.2 Dynamics of Individually Stable Deviations

Starting from the initial partition, agents can leave and join coalitions in order to improve their well-being. We focus on unilateral deviations, which occur when a single agent decides to move from one coalition to another. A *unilateral deviation* performed by agent i transforms a partition π into a partition π' where $\pi(i) \neq \pi'(i)$ and, for all agents $j \neq i$,

$$\pi'(j) = \begin{cases} \pi(j) \setminus \{i\} & \text{if } j \in \pi(i) \\ \pi(j) \cup \{i\} & \text{if } j \in \pi'(i) \\ \pi(j) & \text{otherwise} \end{cases}.$$

Since agents are assumed to be rational, agents only engage in a unilateral deviation if it makes them better off, i.e., $\pi'(i) \succ_i \pi(i)$. Any partition in which no such deviation is possible is called *Nash stable (NS)*.

This type of deviation can be refined by additionally requiring that no agent in the welcoming coalition is worse off when agent i joins. A partition in which no such deviation is possible is called *individually stable (IS)*. Formally, a unilateral deviation performed by agent i who moves from coalition $\pi(i)$ to $\pi'(i)$ is an IS-deviation if $\pi'(i) \succ_i \pi(i)$ and $\pi'(i) \succsim_j \pi(j)$ for all agents $j \in \pi'(i)$. Clearly, an NS partition is also IS.[2] In this article, we focus on dynamics based on IS-deviations. By definition, all terminal states of the dynamics have to be IS partitions.

We are mainly concerned with whether sequences of IS-deviations can reach or always reach an IS partition. If there exists a sequence of IS-deviations leading to an IS partition, i.e., a path to stability, then agents can coordinate (or can be coordinated) to reach a stable partition. The corresponding decision problem is described as follows.

∃-IS-SEQUENCE-[HG]	
Input:	Instance of a particular class of hedonic games [HG], initial partition π_0
Question:	Does there exist a sequence of IS-deviations starting from π_0 leading to an IS partition?

In order to provide some guarantee, we also examine whether *all* sequences of IS-deviations terminate. Whenever this is the case, we say that the dynamics *converges*. The corresponding decision problem is described below.

∀-IS-SEQUENCE-[HG]	
Input:	Instance of a particular class of hedonic games [HG], initial partition π_0
Question:	Does every sequence of IS-deviations starting from π_0 reach an IS partition?

We mainly investigate this problem via the study of its complement: given a hedonic game and an initial partition, does there exist a sequence of IS-deviations that cycles?

A common idea behind hardness reductions concerning these two problems is to exploit prohibitive subconfigurations that evolve from instances without an IS partition or instances which allow for cycling starting from a certain partition.

3 Anonymous Hedonic Games (AHGs)

Bogomolnaia and Jackson (2002) showed that IS partitions always exist in AHGs under naturally single-peaked preferences, and proved that this does not hold under general preferences, by means of a 63-agent counterexample. Here, we provide a counterexample that only requires 15 agents and additionally satisfies general single-peakedness.

Due to space restrictions, we omit some of the proofs or provide only proof sketches.

[2] It is possible to weaken the notion of individual stability even further by also requiring that no member of the *former* coalition of agent i is worse off. The resulting stability notion is called contractual individual stability and guarantees convergence of our dynamics.

Proposition 3.1. *There may not exist an IS partition in AHGs even when $n = 15$ and the agents have strict and generally single-peaked preferences.*

Sketch of proof. Let us consider an AHG with 15 agents with the following (incompletely specified) preferences.

$$\begin{array}{rccccccccccccc} 1: & 2 & \succ & 3 & \succ & 13 & \succ & 12 & \succ & 1 & \succ & [\dots] & & \\ 2: & 13 & \succ & 3 & \succ & 2 & \succ & 1 & \succ & 12 & \succ & [\dots] & & \\ 3,\ 4: & 3 & \succ & 2 & \succ & 1 & \succ & [\dots] & & & & & & \\ 5,\dots,15: & 13 & \succ & 12 & \succ & 15 & \succ & 14 & \succ & 11 & \succ & [\dots] & \succ & 1 \end{array}$$

They can be completed to be generally single-peaked w.r.t. axis $1 < 2 < 3 < 13 < 12 < 15 < 14 < 11 < \dots < 4$.

One can prove that in an IS partition,

(i) agents 3 and 4 are in a coalition of size at most 3;
(ii) agents 5 to 15 are in the same coalition;
(iii) agents 3 and 4 are in the same coalition;
(iv) agents 1 and 2 cannot be both alone.

Therefore, agents 3 and 4 must be together, as well as agents 5 to 15, but not in the same coalition. It remains to identify the coalitions of agents 1 and 2. By *(i)*, they cannot be both with agents 3 and 4. If one agent among them is alone and the other one with agents 5 to 15, then the alone agent can deviate to join them, a contradiction. The remaining possible partitions are present in the cycle of IS-deviations below (the deviating agent is written on top of the arrows).

$$\begin{array}{ccccc} \{\{1\},\{2,3,4\},\{5,\dots,15\}\} & \xrightarrow{1} & \{\{2,3,4\},\{1,5,\dots,15\}\} & \xrightarrow{2} & \{\{3,4\},\{1,2,5,\dots,15\}\} \\ 2\uparrow & & & & \downarrow 1 \\ \{\{1,2\},\{3,4\},\{5,\dots,15\}\} & \xleftarrow{1} & \{\{2\},\{1,3,4\},\{5,\dots,15\}\} & \xleftarrow{2} & \{\{1,3,4\},\{2,5,\dots,15\}\} \end{array}$$

Hence, there is no IS partition in this instance. □

However, even in smaller examples where IS partitions do exist, there may still be cycles in the dynamics.

Proposition 3.2. *The dynamics of IS-deviations may cycle in AHGs even when starting from the singleton partition or grand coalition, for strict generally single-peaked preferences, and for $n < 15$.*

Proof. Let us consider an AHG with 7 agents with the following (incompletely specified) preferences.

$$\begin{array}{rccccccccccc} 1: & 2 & \succ & 3 & \succ & 5 & \succ & 4 & \succ & 1 & \succ & [\dots] \\ 2: & 5 & \succ & 3 & \succ & 2 & \succ & 1 & \succ & 4 & \succ & [\dots] \\ 3,4: & 3 & \succ & 2 & \succ & 1 & \succ & [\dots] & & & & \\ 5,6,7: & 5 & \succ & 4 & \succ & 3 & \succ & 2 & \succ & 1 & \succ & [\dots] \end{array}$$

They can be completed to be generally single-peaked w.r.t. axis $1 < 2 < 3 < 5 < 4 < 6 < 7$. Note that $\{\{1\},\{3,5,6\},\{2,4,7\}\}$ is an IS partition. We represent below a cycle in IS-deviations that can be reached from the singleton partition or the grand coalition.

$$\begin{array}{ccccc} \{1,2\},\{3,4\},\{5,6,7\} & \xrightarrow{2} & \{1\},\{2,3,4\},\{5,6,7\} & \xrightarrow{1} & \{2,3,4\},\{1,5,6,7\} \\ 1\uparrow & & & & \downarrow 2 \\ \{2\},\{1,3,4\},\{5,6,7\} & \xleftarrow{2} & \{1,3,4\},\{2,5,6,7\} & \xleftarrow{1} & \{3,4\},\{1,2,5,6,7\} \end{array}$$

□

We know that it is NP-complete to recognize instances for which an IS partition exists in AHGs, even for strict preferences (Ballester 2004). We prove that both checking the existence of a sequence of IS-deviations ending in an IS partition and checking convergence are hard.

Theorem 3.3. $\exists$-IS-SEQUENCE-AHG *is* NP*-hard and* $\forall$-IS-SEQUENCE-AHG *is* co-NP*-hard, even for strict preferences.*

However, this hardness result does not hold under strict naturally single-peaked preferences, since we show in the next proposition that *every* sequence of IS-deviations is finite under such a restriction.

Proposition 3.4. *The dynamics of IS-deviations always converges to an IS partition in AHGs for strict naturally single-peaked preferences.*

Proof. Assume for contradiction that there exists a cycle of IS-deviations. The key idea is to construct an infinite sequence of agents $(a_k)_{k\geq 1}$ that perform deviations from coalitions $(C_k)_{k\geq 1}$, which are strictly increasing in size. Let a_1 be an agent that deviates within this cycle towards a larger coalition by an IS-deviation. This transforms, say, partition π_1 into partition π_1^1. Set $C_1 = \pi_1(a_1)$ and $\hat{C}_1 = \pi_1^1(a_1)$. One can for instance take an agent that performs a deviation from a coalition of minimum size amongst all coalitions from which any deviation is performed. We will now observe how the coalition $\hat{C}_1$ evolves during the cycle. After possibly some agents outside $\hat{C}_1$ joined it or some left it, some agent b originally in $\hat{C}_1$ must deviate from the coalition evolved from $\hat{C}_1$. Otherwise, we cannot reach partition π_1 again in the cycle. If $b \neq a_1$, we assume that the deviation transforms partition π_2 into partition π_2^1 and we set $a_2 = b$, $C_2 = \pi_2(b)$, and $\hat{C}_2 = \pi_2^1(b)$. Note that $|\hat{C}_2| > |C_2| \geq |\hat{C}_1|$, by single-peakedness and the fact that $|\hat{C}_2| \succ_b |C_2| \succ_b |C_2| - 1 \succ_b \cdots \succ_b |\hat{C}_1| \succ_b |\hat{C}_1| - 1$ (where all preferences but the first follow from the assumption of strictness when some other agent joined the coalition of b). In particular, $|C_2| > |C_1|$.

If $b = a_1$, assume that the deviation transforms partition π_1^2 into π_1^3, where possibly $\pi_1^2 = \pi_1^1$. We update $\hat{C}_1 = \pi_1^3(a_1)$. Note that still $|\hat{C}_1| > |C_1|$ by single-peakedness, because the original deviation of a_1 performed in partition π_1 was towards a larger coalition and $|\pi_1^2(a_1)| \succeq_{a_1} |\pi_1^1(a_1)|$ (equality if the partitions are the same). We consider again the next deviation from $\hat{C}_1$ until it is from an agent $b \neq a_1$, in which case we proceed as in the first case. This must eventually happen, because every time the deviation is again performed by agent a_1 she gets closer to her peak. We proceed in the same manner. In step k, we are given a coalition $\hat{C}_k$ with $|\hat{C}_k| > |C_k|$ which was just joined by an agent. When the next agent originally in $\hat{C}_k$ deviates from the coalition evolved from $\hat{C}_k$, it is either an agent different from a_k and we call it a_{k+1}, and find coalitions C_{k+1} and $\hat{C}_{k+1}$ with $|\hat{C}_{k+1}| > |C_{k+1}| \geq |\hat{C}_k|$; or this agent is a_k, she moves towards an updated coalition $\hat{C}_k$ which maintains $|\hat{C}_k| > |C_k|$.

We have thus constructed an infinite sequence of coalitions $(C_k)_{k\geq 1}$ occurring in the cycle with $|C_{k+1}| > |C_k|$ for all $k \geq 1$, a contradiction. □

An interesting open question is whether this convergence result still holds under naturally single-peaked preferences with indifference. However, convergence is also guaranteed under other constrained anonymous games, called *neutral anonymous games*, which are subset-neutral, as defined by Suksompong (2015), thanks to the use of the same potential function argument.

4 Hedonic Diversity Games (HDGs)

Hedonic diversity games take into account more information about the identity of the agents, changing the focus from coalition sizes to proportions of given types of agents. We obtain more positive results regarding the existence of IS partitions. Indeed, there always exists an IS partition in a hedonic diversity game, even with preferences that are not single-peaked (Boehmer and Elkind 2020). However, we prove that there may exist cycles in IS-deviations, even under some strong restrictions. This stands in contrast to empirical evidence for convergence based on extensive computer simulations by Boehmer and Elkind (2020).

Proposition 4.1. *The dynamics of IS-deviations may cycle in HDGs even*

1. *when preferences are strict and naturally single-peaked,*
2. *when preferences are strict and the initial partition is the singleton partition or the grand coalition, or*
3. *when preferences are naturally single-peaked and the initial partition is the singleton partition.*

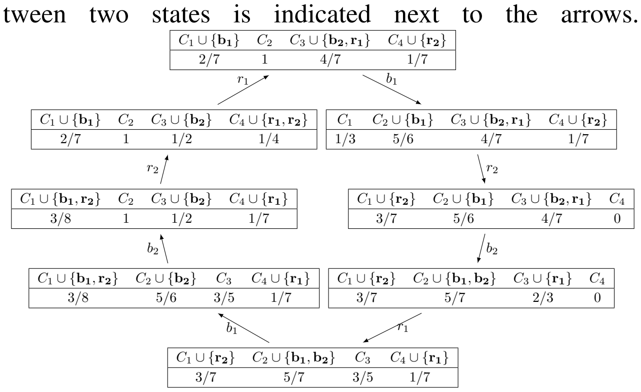
Sketch of proof. We only provide the counterexample for an HDG with strict and naturally single-peaked preferences (restriction 1). Let us consider an HDG with 26 agents: 12 red agents and 14 blue agents. There are four deviating agents: red agents r_1 and r_2 and blue agents b_1 and b_2, and four fixed coalitions C_1, C_2, C_3 and C_4 such that:

- C_1 contains 2 red agents and 4 blue agents;
- C_2 contains 5 red agents;
- C_3 contains 3 red agents and 2 blue agents;
- C_4 contains 6 blue agents.

The relevant part of the preferences is given below.

$$\begin{array}{llll} b_1: & \frac{3}{8} \succ \frac{5}{7} \succ \frac{5}{6} \succ \frac{2}{7} & C_1: & \frac{3}{8} \succ \frac{3}{7} \succ \frac{1}{3} \\ b_2: & \frac{5}{7} \succ \frac{4}{7} \succ \frac{1}{2} \succ \frac{5}{6} & C_2: & \frac{5}{7} \succ \frac{5}{6} \succ 1 \\ r_1: & \frac{4}{7} \succ \frac{1}{4} \succ \frac{1}{7} \succ \frac{2}{3} & C_3: & \frac{4}{7} \succ \frac{1}{2} \succ \frac{3}{5} \\ r_2: & \frac{1}{4} \succ \frac{3}{8} \succ \frac{3}{7} \succ \frac{1}{7} & C_4: & \frac{1}{4} \succ \frac{1}{7} \succ 0 \end{array}$$

Consider the following sequence of IS-deviations that describe a cycle in the dynamics. The four deviating agents of the cycle b_1, b_2, r_1 and r_2 are marked in bold and the specific deviating agent between two states is indicated next to the arrows.



□

This example does not show the impossibility to reach an IS partition since the IS partition $\{C_1 \cup \{b_1, r_2\}, C_2, C_3 \cup \{r_1, b_2\}, C_4\}$ can be reached via IS-deviations from some partitions in the cycle. Thus, starting in these partitions, a path to stability may still exist. Nevertheless, it may be possible that every sequence of IS-deviations cycles, even for strict or naturally single-peaked preferences (with indifference), as the next proposition shows. An interesting open question is whether strict and single-peaked preferences allow for the existence of a path to stability.

Proposition 4.2. *The dynamics of IS-deviations may never reach an IS partition in HDGs, whatever the chosen path of deviations, even for (1) strict preferences or (2) naturally single-peaked preferences with indifference.*

However, convergence is guaranteed by combining all previous restrictions, as stated in the next proposition.

Proposition 4.3. *The dynamics of IS-deviations starting from the singleton partition always converges to an IS partition in HDGs for strict naturally single-peaked preferences.*

Sketch of proof. One can easily prove that at any step of the dynamics, a coalition is necessarily of the form $\{r_1, b_1, \ldots, b_k\}$ or $\{b_1, r_1, \ldots, r_{k'}\}$ or $\{b_1\}$ or $\{r_1\}$ where $r_i \in R$ and $b_j \in B$ for every $i \in [k']$, $j \in [k]$ and $k \leq |B|$ and $k' \leq |R|$. Therefore, the ratio of a coalition can only be equal to $\frac{1}{k+1}$, $\frac{k'}{k'+1}$, 0 or 1. Let us define as $\rho(C)$ the modified ratio of a valid coalition C formed by the dynamics where $\rho(C) = \begin{cases} \frac{|R \cap C|}{|C|} & \text{if } C = \{b_1, r_1, \ldots, r_{k'}\} \text{ for } k' \geq 1 \\ 1 - \frac{|R \cap C|}{|C|} & \text{if } C = \{r_1, b_1, \ldots, b_k\} \text{ for } k \geq 2 \\ 0 & \text{otherwise, i.e., } C = \{r_1\} \text{ or } C = \{b_1\} \end{cases}$.
For each partition in a sequence of IS-deviations, we consider the vector composed of the modified ratios $\rho(C)$ for all coalitions C in the partition. One can prove that for each sequence of IS-deviations, either this vector strictly increases lexicographically at each deviation or there is an equivalent sequence of IS-deviations where it does. □

Under strict preferences, checking the existence of a path to stability and convergence are hard.

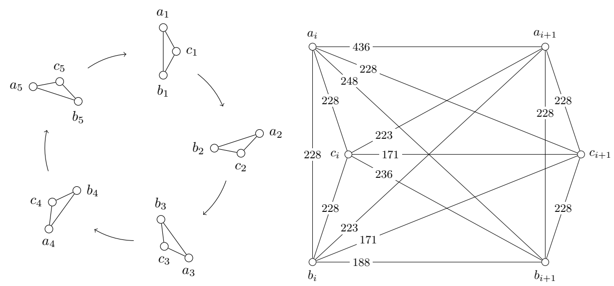
Theorem 4.4. $\exists$-IS-SEQUENCE-HDG *is* NP*-hard and* $\forall$-IS-SEQUENCE-HDG *is* co-NP*-hard, even for strict preferences.*

5 Fractional Hedonic Games (FHGs)

Next, we study fractional hedonic games, which are closely related to hedonic diversity games, but instead of agent types, utilities rely on a cardinal valuation function of the other agents. The first part of the section deals with symmetric games, the second part with simple games.

An open problem for symmetric FHGs was whether they always admit an IS partition (Brandl, Brandt, and Strobel 2015). Here, we provide a counterexample using 15 agents.

Theorem 5.1. *There exists a symmetric FHG without an IS partition.*



(a) Five triangles are ordered in a cycle such that there is a tendency of agents in N_i to deviate to coalitions in N_{i+1}.

(b) The transition weights between the triangles allow for infinite loops of deviations.

Figure 1: Description of the graph associated with the constructed symmetric FHG without an IS partition.

Sketch of proof. Define the sets of agents $N_i = \{a_i, b_i, c_i\}$ for $i \in \{1, \dots, 5\}$ and consider the FHG on the agent set $N = \bigcup_{i=1}^{5} N_i$ where symmetric weights are given as in Figure 1b. All weights not specified in this figure are set to -2251. The FHG consists of five triangles that form a cycle. Its structure is illustrated in Figure 1a.

There is an infinite cycle of deviations starting with the partition $\{N_5 \cup N_1, N_2, N_3, N_4\}$. First, a_1 deviates by joining N_2. Then, b_1 joins this new coalition, then c_1. After this step, we are in an isomorphic state as in the initial partition. It can be shown that there exists no IS partition in this instance. □

Employing this counterexample, the methods of Brandl, Brandt, and Strobel (2015), which originate from hardness constructions of Sung and Dimitrov (2010), can be used to show that it is NP-hard to decide about the existence of IS partitions in symmetric FHGs.

Corollary 5.2. *Deciding whether there exists an individually stable partition in symmetric FHGs is* NP*-hard.*

If we consider symmetric, non-negative utilities, the grand coalition forms an NS, and therefore IS, partition of the agents. However, deciding about the convergence of the IS dynamics starting with the singleton partition is NP-hard. The reduction is similar to the one in the previous statement and avoids negative weights by the fact that, due to symmetry of the weights, in a dynamics starting with the singleton partition, all coalitions that can be obtained in the process must have strictly positive mutual utility for all pairs of agents in the coalition.

Theorem 5.3. $\exists$-IS-SEQUENCE-FHG *is* NP*-hard and* $\forall$-IS-SEQUENCE-FHG *is* co-NP*-hard, even in symmetric FHGs with non-negative weights. The former is even true if the initial partition is the singleton partition.*

From now on, we consider simple FHGs. We start with the additional assumption of symmetry.

Proposition 5.4. *The dynamics of IS-deviations starting from the singleton partition always converges to an IS partition in simple symmetric FHGs in at most $\mathcal{O}(n^2)$ steps. The dynamics may take $\Omega(n\sqrt{n})$ steps.*

Sketch of proof. We only prove the upper bound. Note that all coalitions formed through the deviation dynamics are cliques. Hence, every deviation step will increase the total number of edges in all coalitions. More precisely, the dynamics will increase the potential $\Lambda(\pi) = \sum_{C \in \pi} |C|(|C| - 1)/2$ in every step by at least 1. Since the total number of edges is quadratic, this proves the upper bound. □

Note that there is a simple way to converge in a linear number of steps starting with the singleton partition by forming largest cliques and removing them from consideration.

If we allow for asymmetries, the dynamics is not guaranteed to converge anymore. For instance, the IS dynamics on an FHG induced by a directed triangle will not converge for any initial partition except for the grand coalition. We can, however, characterize convergence on asymmetric FHGs. Tractability highly depends on the initial partition. First, we assume that we start from the singleton partition.

The key insight is that throughout the dynamic process on an asymmetric FHG starting from the singleton partition, the subgraphs induced by coalitions are always transitive and complete. Convergence is then shown by a potential function argument.

Proposition 5.5. *The dynamics of IS-deviations starting from the singleton partition converges in asymmetric FHGs if and only if the underlying graph is acyclic. Moreover, under acyclicity, it converges in $\mathcal{O}(n^4)$ steps.*

The previous statement shows convergence of the dynamics for asymmetric, acyclic FHGs. In addition, it is easy to see that there is always a sequence converging after n steps, starting with the singleton partition. One can use a topological order of the agents and allow agents to deviate in decreasing topological order towards a best possible coalition.

There are two interesting further directions. One can weaken either the restriction on the initial partition or on asymmetry. If we allow for general initial partitions, we immediately obtain hardness results that apply in particular to the broader class of simple FHGs.

Theorem 5.6. $\exists$-IS-SEQUENCE-FHG *is* NP*-hard and* $\forall$-IS-SEQUENCE-FHG *is* co-NP*-hard, even in asymmetric FHGs.*

On the other hand, if we transition to simple FHGs while maintaining the initial partition, the problem of deciding whether a path to stability exists becomes hard.

Theorem 5.7. $\exists$-IS-SEQUENCE-FHG *is* NP*-hard even in simple FHGs when starting from the singleton partition.*

6 Dichotomous Hedonic Games (DHGs)

By taking into account the identity of other agents in the preferences of agents over coalitions, it can be more complicated to get positive results regarding individual stability (see, e.g., Theorem 5.1). However, by restricting the evaluation of coalitions to dichotomous preferences, the existence of an IS partition is guaranteed (Peters 2016), as well as convergence of the dynamics of IS-deviations, when starting from the grand coalition (Boehmer and Elkind 2020).

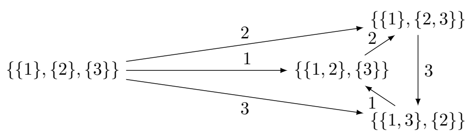
Nevertheless, the convergence of the dynamics is not guaranteed for an arbitrary initial partition and no sequence of IS-deviations may ever reach an IS partition.

Proposition 6.1. *The dynamics of IS-deviations may never reach an IS partition in DHGs, whatever the chosen path of deviations, even when starting from the singleton partition.*

Proof. Let us consider an instance of a DHG with three agents. Their preferences are described in the table below.

Agent	1	2	3
Approvals	$\{1,2\}$	$\{2,3\}$	$\{1,3\}$
Disapprovals	$\{1\},\{1,3\},\{1,2,3\}$	$\{2\},\{1,2\},\{1,2,3\}$	$\{3\},\{2,3\},\{1,2,3\}$

There is a unique IS partition which consists of the grand coalition $\{1,2,3\}$. We represent below all possible IS-deviations between all the other possible partitions. An IS-deviation between two partitions is indicated by an arrow mentioning the name of the deviating agent.



One can check that the described deviations are IS-deviations. A cycle is necessarily reached when starting from a partition different from the unique IS partition, which can be reached only if it is the initial partition. □

Moreover, it is hard to decide on the existence of a sequence of IS-deviations ending in an IS partition, even when starting from the singleton partition, as well as checking convergence.

Theorem 6.2. $\exists$-IS-SEQUENCE-DHG *is* NP*-hard even when starting from the singleton partition, and* $\forall$-IS-SEQUENCE-DHG *is* co-NP*-hard.*

Note that the counterexample provided in the proof of Proposition 6.1 exhibits a global cycle in the preferences of the agents: $\{1,2\} \rhd \{1,3\} \rhd \{2,3\} \rhd \{1,2\}$. However, by considering dichotomous preferences with *common ranking property*, that is, each agent has a threshold for acceptance in a given global order, we obtain convergence thanks to the same potential function argument used by Caskurlu and Kizilkaya (2019), for proving the existence of a core-stable partition in hedonic games with common ranking property.

Note that when assuming that if a coalition is approved by one agent, then it must be approved by all the members of the coalition (so-called *symmetric dichotomous preferences*), we obtain a special case of preferences with common ranking property where all the approved coalitions are at the top of the global order. Therefore, convergence is also guaranteed under symmetric dichotomous preferences.

7 Conclusion

We have investigated dynamics of deviations based on individual stability in hedonic games. The two main questions we considered were whether there exists *some* sequence of deviations terminating in an IS partition, and whether *all* sequences of deviations terminate in an IS partition, i.e., the dynamics converges. Our results are mostly negative with examples of cycles in dynamics or even non-existence of IS partitions under rather strong preference restrictions. In particular, we have answered a number of open problems proposed in the literature. On the other hand, we have identified natural conditions for convergence that are mostly based on preferences relying on a common scale for the agents, like the common ranking property, single-peakedness or symmetry. An overview of our results can be found in Table 1.

Class	Convergence	Hardness
AHGs	✓ strict & nat. SP (single-peaked) (Prop. 3.4) ✓ neutral (derived from Suksompong (2015)) ○ strict & gen. SP; singletons / grand coalition (Prop. 3.2)	$\exists$ strict (Th. 3.3) $\forall$ strict (Th. 3.3)
HDGs	✓ strict & nat. SP; singletons (Prop. 4.3) ○ strict & nat. SP (Prop. 4.1) ○ strict; singletons / grand coalition (Prop. 4.1) ○ nat. SP; singletons (Prop. 4.1)	$\exists$ strict (Th. 4.4) $\forall$ strict (Th. 4.4)
FHGs	✓ simple & sym.; singletons (Prop. 5.4) ✓ acyclic digraph (Th. 5.5) ○ sym. (Th. 5.1)	$\exists$ sym. (Th. 5.3) $\exists$ simple; singletons (Th. 5.7) $\exists$ asym. (Th. 5.6) $\forall$ sym. (Th. 5.3) $\forall$ asym. (Th. 5.6)
DHGs	✓ grand coalition (Boehmer and Elkind) ✓ common ranking property or symmetric (derived from Caskurlu and Kizilkaya (2019)) ○ singletons (Prop. 6.1)	$\exists$ singletons (Th. 6.2) $\forall$ general (Th. 6.2)

Table 1: Convergence and hardness results for the dynamics of IS-deviations in various classes of hedonic games. Symbol ✓ marks convergence under the given preference restrictions and initial partition (if applicable) while ○ marks non-convergence, i.e., cycling dynamics. Symbol $\exists$ (resp., $\forall$) denotes that problem $\exists$-IS-SEQUENCE-HG (resp., $\forall$-IS-SEQUENCE-HG) is NP-hard (resp., co-NP-hard).

For all hedonic games under study, it turned out that the existence of cycles for IS-deviations is sufficient to prove the hardness of recognizing instances for which there exists a finite sequence of deviations or whether all sequences of deviations are finite, i.e., the dynamics converges. While our results cover a broad range of hedonic games considered in the literature, there are still promising directions for further research. First, even though our hardness results hold under strong restrictions, the complexity of these questions remains open for other interesting preference restrictions, some of which do not guarantee convergence. Following our work, the most intriguing cases are AHGs under single-peaked weak preferences, simple symmetric FHGs with arbitrary initial partitions, and HDGs under single-peaked preferences. Secondly, one could investigate more specific rules of IS-deviations that quickly terminate in IS partitions, even in classes of hedonic games that allow for cyclic IS-deviations. For instance, for simple symmetric FHGs, there is the possibility of convergence such that each agent deviates at most once, but the selection of the deviating agents in this approach requires to solve a maximum clique problem (cf. the discussion after Proposition 5.4). Finally, the dynamics we consider only guarantee individual stability. One could also aim at reaching outcomes that satisfy Pareto optimality or other desirable properties on top of individual stability.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft under grant BR 2312/12-1. We thank the anonymous reviewers for helpful comments.

References

Abeledo, H.; and Rothblum, U. G. 1995. Paths to marriage stability. *Discrete Applied Mathematics* 63(1): 1–12.

Aziz, H.; Brandl, F.; Brandt, F.; Harrenstein, P.; Olsen, M.; and Peters, D. 2019. Fractional Hedonic Games. *ACM Transactions on Economics and Computation* 7(2): 1–29.

Aziz, H.; Harrenstein, P.; Lang, J.; and Wooldridge, M. 2016. Boolean Hedonic Games. In *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR)*, 166–175.

Ballester, C. 2004. NP-completeness in hedonic games. *Games and Economic Behavior* 49(1): 1–30.

Bilò, V.; Fanelli, A.; Flammini, M.; Monaco, G.; and Moscardelli, L. 2018. Nash Stable Outcomes in Fractional Hedonic Games: Existence, Efficiency and Computation. *Journal of Artificial Intelligence Research* 62: 315–371.

Boehmer, N.; and Elkind, E. 2020. Individual-Based Stability in Hedonic Diversity Games. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*, 1822–1829.

Bogomolnaia, A.; and Jackson, M. O. 2002. The Stability of Hedonic Coalition Structures. *Games and Economic Behavior* 38(2): 201–230.

Brandl, F.; Brandt, F.; and Strobel, M. 2015. Fractional Hedonic Games: Individual and Group Stability. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 1219–1227.

Brandt, F.; and Wilczynski, A. 2019. On the Convergence of Swap Dynamics to Pareto-Optimal Matchings. In *Proceedings of the 15th International Conference on Web and Internet Economics (WINE)*, 100–113.

Bredereck, R.; Elkind, E.; and Igarashi, A. 2019. Hedonic Diversity Games. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 565–573.

Carosi, R.; Monaco, G.; and Moscardelli, L. 2019. Local Core Stability in Simple Symmetric Fractional Hedonic Games. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 574–582.

Caskurlu, B.; and Kizilkaya, F. E. 2019. On Hedonic Games with Common Ranking Property. In *Proceedings of the 11th International Conference on Algorithms and Complexity (CIAC)*, 137–148. Springer International Publishing.

Drèze, J. H.; and Greenberg, J. 1980. Hedonic Coalitions: Optimality and Stability. *Econometrica* 48(4): 987–1003.

Farrell, J.; and Scotchmer, S. 1988. Partnerships. *Quarterly Journal of Economics* 103: 279–297.

Hoefer, M.; Vaz, D.; and Wagner, L. 2018. Dynamics in matching and coalition formation games with structural constraints. *Artificial Intelligence* 262: 222–247.

Peters, D. 2016. Complexity of Hedonic Games with Dichotomous Preferences. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*, 579–585.

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

Suksompong, W. 2015. Individual and Group Stability in Neutral Restrictions of Hedonic Games. *Mathematical Social Sciences* 78: 1–5.

Sung, S. C.; and Dimitrov, D. 2010. Computational Complexity in Additive Hedonic Games. *European Journal of Operational Research* 203(3): 635–639.