Reaching Individually Stable Coalition Structures

FELIX BRANDT, Department of Informatics, Technical University of Munich, Germany
MARTIN BULLINGER, Department of Informatics, Technical University of Munich, Germany
ANAËLLE WILCZYNSKI, MICS, CentraleSupélec, Université Paris-Saclay, France

The formal study of coalition formation in multi-agent systems is typically realized in the framework of 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 established 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.

Our results are threefold. First, we identify conditions for the (fast) convergence of our dynamics. To this end, we develop new techniques based on the simultaneous usage of multiple intertwined potential functions and establish a reduction uncovering a close relationship between anonymous hedonic games and hedonic diversity games. Second, we provide elaborate counterexamples determining tight boundaries for the existence of individually stable partitions. Third, we study the computational complexity of problems related to the coalition formation dynamics. In particular, we settle open problems suggested by Bogomolnaia and Jackson [9], Brandl et al. [10], and Boehmer and Elkind [8].

CCS Concepts: • **Computing methodologies → Multi-agent systems**; • **Theory of computation → Design and analysis of algorithms**.

Additional Key Words and Phrases: coalition formation, hedonic games, individual stability, game dynamics

1 INTRODUCTION

Coalitions and coalition formation are central concerns in the study of multi-agent 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. The permitted deviations can be constrained by other agents, for instance by means of contracts with an existing coalition or by the necessity of consent when admitting a new member. These constraints lead to a large body of stability concepts [4]. 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 as an end-state, 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

Authors' addresses: Felix Brandt, brandtf@in.tum.de, Department of Informatics, Technical University of Munich, Germany; Martin Bullinger, bullinge@in.tum.de, Department of Informatics, Technical University of Munich, Germany; Anaëlle Wilczynski, anaelle.wilczynski@centralesupelec.fr, MICS, CentraleSupélec, Université Paris-Saclay, France.

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 [16]: 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 for matchings, which can be seen as a special domain of coalition formation where only coalitions of size 2 are allowed [e.g., 1, 12, 21]. More recently, dynamics of coalition formation have also come under scrutiny in the context of hedonic games [7, 14, 19]. While coalition formation dynamics are an object of study worthy for itself, they can also be used as a means to design algorithms for the computation of stable outcomes, and have been implicitly used for this purpose before. For example, the algorithm by Boehmer and Elkind [8] 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 [9] 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 [9], Suksompong [22] 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 [17]. Here, 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 [15]. 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

[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 [9].

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. On the other hand, guaranteed convergence allows agents to reach a stable outcome without further coordination. This provides strong stability guarantees under more pessimistic assumptions on the agents' behavior. Our results indicate clear boundaries for possibilities towards various dimensions concerning in particular the structure of preferences, the initial partition, or the selection of deviations. Our main results are summarized as follows.

- In AHGs, the dynamics converges for (naturally) single-peaked 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 [9] requires 63 agents and the existence of smaller examples was an acknowledged open problem [see 5, 8].
- We provide an elaborate reduction for HDGs that eventually establishes a close relationship to AHGs and show convergence of the dynamics for strict and naturally singled-peaked preferences when starting from the singleton partition and agents' deviations satisfy a weak constraint. In contrast to empirical evidence reported by Boehmer and Elkind [8], we show that all of the above assumptions are essential for the convergence result. In particular, cycling of the dynamics is possible if the starting partition is the singleton partition and preferences are restricted to be strict and naturally single-peaked.
- 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 et al. [10].
- For each of the above classes and DHGs, we identify computational boundaries. In particular, 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 AND MODEL

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. Formally, a solution is a *partition* of N, i.e., a subset $\pi \subseteq 2^N$ such that $\bigcup_{C\in\pi} C = N$, and for every pair $C, D \in \pi$, it holds that $C = D$ or $C \cap D = \emptyset$. An element of a partition is called *coalition*, and given a partition π, we denote by $\pi(i)$ the coalition containing agent i. 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)* [9]: 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)* [13]: 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)* [2]: The agents evaluate a coalition by assessing 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* [3].

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 $z > y > x$, $x \succ_i y$ 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)*.

Draft – April 19, 2022

This type of deviation can be refined by additionally requiring that no agent in the welcoming coalition is worse off when agent i joins. 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 $j \in \pi'(i)$. A partition in which no IS-deviation is possible is called *individually stable (IS)*. 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 although the agents perform myopic deviations, they can optimistically reach (or can be guided towards) 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 the existence of instances without an IS partition or instances which allow for cycling starting from a certain partition. These can be used to create prohibitive subconfigurations in reduced instances.

3 ANONYMOUS HEDONIC GAMES

Bogomolnaia and Jackson [9] 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.

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.*

Proof. Let us consider an AHG with 15 agents with the following (incompletely specified) preferences ([...] denotes an arbitrary order over the remaining coalition sizes).

$$\begin{array}{rccccccccccccccc}
1: & 2 & \succ & 3 & \succ & 13 & \succ & 12 & \succ & 1 & \succ & [\dots] \\
2: & 13 & \succ & 3 & \succ & 2 & \succ & 1 & \succ & 12 & \succ & [\dots] \\
3,\ 4: & 3 & \succ & 2 & \succ & 1 & \succ & [\dots] \\
5,\dots,15: & 13 & \succ & 12 & \succ & 15 & \succ & 14 & \succ & 11 & \succ & 10 & \succ & \dots & \succ & 1
\end{array}$$

They can be completed to be generally single-peaked with respect to axis $1 > 2 > 3 > 13 > 12 > 15 > 14 > 11 > 10 > \cdots > 4$.

[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.

Draft – April 19, 2022

Note that in an IS partition,

(i) agents 3 and 4 are in a coalition of size at most 3: Otherwise, they prefer to deviate to be alone.

(ii) agents 5 to 15 are in the same coalition: Suppose, for the sake of contradiction, that agents 5 to 15 are not in the same coalition. By (i), at most two of them are with agent 3 and at most two of them with agent 4. Agents 1 and 2 cannot be in a coalition of size 12 or 13 if all these 11 agents $5, \ldots, 15$ are not with them (agents 3 and 4 cannot be in such a big coalition by (i)). Therefore, agents 1 and 2 cannot be in a coalition of size larger than 3, otherwise they would deviate to be alone. It follows that at most two agents from $\{5, \ldots, 15\}$ are with agent 1 and at most two of them with agent 2. Then, in the worst case, it remains only three agents within $\{5, \ldots, 15\}$ who are not in a coalition with agents 1, 2, 3 or 4. These three agents cannot enter into the other coalitions but they prefer to group together, forming a coalition of size three. Afterwards, all the agents from $\{5, \ldots, 15\}$ that are in coalitions with agents 1, 2, 3 or 4 will deviate to join them because they prefer to be in bigger coalitions, and they can benefit from a coalition of size at least four by joining these remaining agents, whereas they are blocked in a coalition of size at most three, a contradiction.

(iii) agents 3 and 4 are in the same coalition: Suppose for the sake of contradiction that agents 3 and 4 are not in the same coalition. By (i), none of them can belong to the big coalition containing the agents $5, \ldots, 15$ (ii). Moreover, if they are both alone, then they have incentive to group together, contradicting the stability. Therefore, they must form coalitions with agents 1 and 2. If agents 1 and 2 are both with agent 3 (resp., 4) and agent 4 (resp., 3) is alone, then agent 1 has incentive to leave the coalition $\{1, 2, 3\}$ (resp., $\{1, 2, 4\}$) to join agent 4 (resp., 3), contradicting the stability. Therefore, one agent between 1 and 2 must be with agent 3 or agent 4. But, in such a case, the agent between 3 and 4, say 3, who is with agent 1 will move to the coalition with agent 2 and agent 4, contradicting the stability. Therefore, agents 3 and 4 must be in the same coalition.

(iv) agents 1 and 2 cannot be both alone: Otherwise, they would deviate to group together.

From the previous observations, we get that agents 3 and 4 must be together, as well as agents $5, \ldots, 15$, but not in the same coalition. The remaining question concerns the coalitions to which agents 1 and 2 belong. It is not possible that both agents 1 and 2 are in a coalition with agents 3 and 4, otherwise it would contradict condition (i). If one agent between agents 1 and 2 is alone and the other one is with agents 5 to 15, then the alone agent can deviate to join them, contradicting the stability. The remaining possible partitions are present in the following cycle of IS-deviations (the deviating agent is written on top of the arrows).

$$\begin{tikzcd}
\{\{1\},\{2,3,4\},\{5,\ldots,15\}\} \arrow[r, "1"] & \{\{2,3,4\},\{1,5,\ldots,15\}\} \arrow[r, "2"] & \{\{3,4\},\{1,2,5,\ldots,15\}\} \arrow[d, "1"] \\
\{\{1,2\},\{3,4\},\{5,\ldots,15\}\} \arrow[u, "2"] & \{\{2\},\{1,3,4\},\{5,\ldots,15\}\} \arrow[l, "1"'] & \{\{1,3,4\},\{2,5,\ldots,15\}\} \arrow[l, "2"']
\end{tikzcd}$$

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

The construction in Proposition 3.1 does not seem to leave room for improvements, and we conjecture that the counterexample may even be minimal, that is, an IS partition always exists when $n < 15$. However, even when $n < 15$ and 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, preferences are strictly generally single-peaked, and* $n = 7$.

Draft – April 19, 2022

Proof. Let us consider an AHG with 7 agents with the following (incompletely specified) preferences ([…] denotes an arbitrary order over the remaining coalition sizes).

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

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

$$\begin{tikzcd}
\{1,2\},\{3,4\},\{5,6,7\} \arrow[r, "2"] & \{1\},\{2,3,4\},\{5,6,7\} \arrow[r, "1"] & \{2,3,4\},\{1,5,6,7\} \arrow[d, "2"] \\
\{2\},\{1,3,4\},\{5,6,7\} \arrow[u, "1"] & \{1,3,4\},\{2,5,6,7\} \arrow[l, "2"'] & \{3,4\},\{1,2,5,6,7\} \arrow[l, "1"']
\end{tikzcd}$$

□

Note that $\{\{1\},\{3,5,6\},\{2,4,7\}\}$ is an IS partition in the example of the previous proposition.

We know that it is NP-complete to recognize instances for which an IS partition exists in AHGs, even for strict preferences [5]. 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. *∃-IS-Sequence-AHG is* NP*-hard and ∀-IS-Sequence-AHG is* co-NP*-hard, even for strict preferences.*

However, these hardness results do not hold under naturally single-peaked preferences, even if preferences may be weak. Indeed, we show in the next proposition that *every* sequence of IS-deviations is finite under such a restriction. We thus generalize a result of Bogomolnaia and Jackson [9] who have proved that there always exists an IS partition if the preferences are naturally single-peaked, with a constructive proof involving specific IS-deviations starting from the singleton partition.

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

Proof. Let an AHG be given with naturally single-peaked preferences. Assume for contradiction that there exists a cycle of IS-deviations. We will identify a specific deviation within this cycle that cannot be repeated throughout the execution of the dynamics, obtaining a contradiction. Let C be a coalition of smallest cardinality such that there exists an agent $d \in C$ that performs a deviation leaving C. Such a coalition exists, because there are only finitely many different deviations performed in the cycle. We will argue that it is impossible to alter the coalition $C \setminus \{d\}$ within the execution of the cycle. Hence, the coalition C will never be reached again.

First, our assumption of minimality implies that $C \setminus \{d\}$ cannot be altered by having an agent leave this coalition. Therefore, we have to show that it cannot happen that an agent ever joins $C \setminus \{d\}$. Assume for contradiction that there exists an agent x that joins $C \setminus \{d\}$. Using our minimality assumption again, the deviation of agent x when joining $C \setminus \{d\}$ must originate from a coalition C' with $|C'| > |C|$. Hence, agent x deviates towards a smaller coalition. Hence, single-peakedness implies that the peak p_x of agent x must satisfy $p_x < |C'|$. In particular, it follows by single-peakedness that it cannot be the case that $y \succ_x |C'|$ for any $y \geq |C'|$.

We claim that it is impossible that agent x ever reaches a coalition of size at least $|C'|$ again. To see this, let C_k be the k-th coalition that x is part of after leaving C', i.e., $C_1 = (C \setminus \{d\}) \cup \{x\}$ and C_{k+1} evolves from C_k by having some agent join or leave C_k, or C_{k+1} is the new coalition of x if x

performs a deviation to leave C_k. We will show by induction over k that, for every k, $|C_k| < |C|$ or $C_k \succ_x C'$. Since $C_1 \succ_x C'$, the claim is true for $k = 1$. Now, let $k \geq 1$ and assume that $|C_k| < |C|$ or $|C_k| \succ_x |C'|$. Consider first the case that $|C_k| < |C|$. By our minimality assumption, x is not allowed to perform a deviation and therefore C_{k+1} evolves by having an agent leave or join C_k. For the same reason, no other agent may leave the coalition. If an agent joins the coalition, then the size remains to be smaller than $|C|$, or is exactly $|C|$ and we already know that $|C| \succ_x |C'|$.

It remains to consider the case that $C_k \succ_x C'$. If C_{k+1} forms via a deviation of agent x, then $C_{k+1} \succ_x C_k \succ_x C'$ and the claim is true. If some agent joins C_k to form C_{k+1} then x has to approve this, and we can conclude that $C_{k+1} \succsim_x C_k \succ_x C'$. Hence, it remains to consider the case that some agent leaves C_k and the remaining coalition is C_{k+1}. As we have argued above, $|C| \succ_x |C'|$ implies that $|C_k| < |C'|$. If $|C_{k+1}| < |C|$, then the assertion is true. If $|C_{k+1}| = |C|$, then $C_{k+1} \succ_x C'$, and the assertion is true. Finally, it remains the case that $|C| < |C_{k+1}| = |C_k| - 1 < |C_k| < |C'|$, and we know that $|C| \succ_x |C'|$ and $|C_k| \succ_x |C'|$. Hence, single-peakedness implies that $C_{k+1} \succ_x C'$. This completes the proof of the induction hypothesis. It follows that agent x cannot reach a coalition of size at least $|C'|$ again, a contradiction. Hence, there can be no agent joining $C \setminus \{d\}$. This shows that $C \setminus \{d\}$ can never be altered again, our final contradiction. □

Our final goal in this section is to provide a polynomial bound on the running time of the dynamics. Unfortunately, our proof relies on strict preferences, leaving the case of weak preferences as an interesting open problem. Even the proof under strict preferences needs far more sophisticated methods than the proof of existence. The key idea is to distinguish deviations towards a smaller and a larger coalition, respectively, and to make use of a potential function that aggregates values for agents and coalitions to bound the number of deviations towards a larger coalition by n^2. Using a second, much simpler potential function yields an overall polynomial running time.

Theorem 3.5. *The dynamics of IS-deviations always converges to an IS partition in AHGs for strict naturally single-peaked preferences in $O(n^3)$ steps.*

Proof. Let an AHG be given with strict and naturally single-peaked preferences where the peak of agent j is at position p_j. Consider a sequence of IS-deviations starting at some initial partition π_0. Assume that the deviations lead to the sequence $(\pi_k)_{k=0}^m$ where, for $k = 0, \dots, m-1$, π_{k+1} evolves from π_k through an IS-deviation of agent d_k. We call a deviation an R-move (resp., L-move) if it is towards a larger (resp., smaller) partition. The main part of the proof provides a bound of n^2 for the number of R-moves. As we will see, this implies at most n^3 L-moves.

The idea is to define a potential that is based on a value v_j^k for each agent $j \in N$ and a value v_C^k for each coalition $C \in \pi_k$. This potential will not only depend on the partition π_k, but also on the starting partition and the specific sequence of deviations to derive π_k. It will be increased strictly during an R-move and will not decrease during an L-move. We need also to keep track of the last agent l_C^k that entered a coalition $C \in \pi_k$ if this agents plays a 'special role' within her coalition.

Initially, define $v_j^0 = 0$ for all agents $j \in N$, and $v_C^0 = 0$ for all coalitions $C \in \pi_0$. Also, there is no last agent that entered a coalition so far, so we initiate $l_C^0 = \bot$. Now, assume that we transition from partition π_k to partition π_{k+1} through an IS-deviation of agent d_k. Denote $D_k = \pi_k(d_k) \setminus \{d_k\}$ and $E_k = \pi_{k+1}(d_k)$, i.e., the new coalitions in π_{k+1} compared to π_k, and denote $l_k = l^k_{\pi_k(d_k)}$, which will be the last agent that entered $\pi_k(d_k)$ (unless $l^k_{\pi_k(d_k)} = \bot$, in which case such an agent does either not exist or its identity is unimportant for the updates of the values). We specify first the updates that are done independently of the kind of deviation.

We set $v_j^{k+1} = v_j^k$ for all $j \in N \setminus (D_k \cup E_k)$, i.e., the value of agents not involved in the deviation does not change. Similarly, we set $v_C^{k+1} = v_C^k$ and $l_C^{k+1} = l_C^k$ for all coalitions $C \in \pi_k \cap \pi_{k+1} = \pi_{k+1} \setminus \{D_k, E_k\}$.

Draft – April 19, 2022

The updates for the last agents do not depend on the kind of move. Set $l_{E_k}^{k+1} = d_k$, and set $l_{D_k}^{k+1} = \perp$ if $l_k = d_k$ and $l_{D_k}^{k+1} = l_{\pi_k(d_k)}^k$, otherwise.

It remains to specify new values for the agents which are part of one of the coalitions involved in the deviation, and the values of these coalitions. The intuition for defining the agent and coalition values is as follows. The value of an agent is always strictly smaller than her peak. In addition, it represents a size that is preferred at most as much as an agents current coalition size. It is easy to achieve this for agents which are joined because strictness and single-peakedness imply that they move towards their peak.

Hence, we set $v_j^{k+1} = |E_k| - 1$ for all $j \in E_k \setminus \{d_k\}$. If the deviating agent performs an R-move, then an appropriate bound is the size of the left coalition. In the case of an L-move, the deviating agent can only be the last agent who joined the coalition, unless she abandons a coalition that was never joined by an agent (this intuition follows from the third invariant in the formal proof below). Hence, we assign the 'right' value by maintaining the current value.

Formally, we set $v_{d_k}^{k+1} = |D_k| + 1 = |\pi_{d_k}(\pi_k)|$ if it was an R-move and $v_{d_k}^{k+1} = v_{d_k}^k$ if the deviation was an L-move. In either case, we update afterwards $v_{E_k}^{k+1} = v_{d_k}^{k+1}$.

The role of the coalitional value is mainly important to negate changes in agent values for the case that the abandoned coalition reaches the size corresponding to the value of the last joined agent. If the deviation was an L-move or $d_k = l_k$, we update $v_j^{k+1} = v_j^k$ for all $j \in D_k$ and $v_{D_k}^{k+1} = 0$. If the agents in D_k did not change their value since the initialization, i.e., if $v_j^k = 0$ for all $j \in D_k$, we also leave $v_j^{k+1} = v_j^k$ for all $j \in D_k$ and $v_{D_k}^{k+1} = 0$. It remains the case that $d_k \neq l_k$ while $v_{d_k}^k \neq 0$ (within every coalition, either the value of none or all agents will always be 0). If $v_{l_k}^k = |D_k|$, then set $v_j^{k+1} = |D_k|$ for all $j \in D_k$, $v_{D_k}^{k+1} = 0$, and $l_{D_k}^{k+1} = \perp$ (we update this last agent *again*, because all agents play the same role now within D_k). Otherwise, set $v_j^{k+1} = |D_k| - 1$ for all $j \in D_k \setminus \{l_k\}$, $v_{l_k}^{k+1} = v_{l_k}^k$, and $v_{D_k}^{k+1} = v_{D_k \cup \{d_k\}}^k$ (the last agent of D_k does not need a second update since it still plays a special role).

Given a partition π_k that occurs in the dynamics, define its potential $\Lambda(\pi_k) = \sum_{j \in N} v_j^k + \sum_{C \in \pi_k} v_C^k$. Again, note that this potential can depend both on the starting partition and the specific sequence of deviations. We will prove the following claim.

Claim 3.1. *For all $k = 1, \dots, m$, $\Lambda(\pi_k) \geq \Lambda(\pi_{k-1})$. If π_{k+1} evolved from π_k by an R-move, then $\Lambda(\pi_{k+1}) > \Lambda(\pi_k)$.*

Proof. We will prove this main claim along with the following useful invariants by induction over $k = 0, \dots, m$:

(1) For all $C \in \pi_k$, $v_C^k \leq |C| - 1$.
(2) For all $C \in \pi_k$ and $j \in C$, if $v_C^k > 0$, then $v_j^k \leq |C| - 1$.
(3) For all agents $j \in N$ with $v_j^k > 0$, $v_j^k \preceq_j |\pi_k(j)|$ and $p_j > v_j^k$.
(4) For all $C \in \pi_k$ with $l^k(C) \neq \perp$, $v^k(C) = v^k(l^k(C))$ or $v^k(C) = 0$.

For $k = 0$, the main claim is vacant, and the four invariants hold by our initialization of the agent and coalition values. So assume that all of them are true for some $0 \leq k < m$. We use the notation for the agents d_k and l_k, and the coalitions D_k and E_k as in the description of the updates of the values.

(1) The first invariant follows from the update rules and by induction. Let us provide the details for the affected coalitions. Specifically, $v_{D_k}^{k+1} = 0$, unless we are in the last case of the update rule, where $v_{D_k}^{k+1} = v_{D_k \cup \{d_k\}}^k = v_{l_k}^k \leq |D_k| - 1$. Now, let us consider coalition E_k. If d_k

Draft – April 19, 2022

performed an R-move, then $v^{k+1}_{E_k} = |D_k| + 1 \leq |E_k| - 1$. On the other hand, assume that d_k performed an L-move. Recall that then $v^k_{E_k} = v^k_{d_k}$. It holds that $v^k_{d_k} = 0$ (and we are done), or d_k was the last agent who joined $\pi_k(d_k)$ (other agents in $\pi_k(d_k)$ can only perform R-moves due to strict preferences and single-peakedness). Assume for contradiction that $v^k_{E_k} \geq |E_k|$. If $v^k_{E_k} = |E_k|$, then the third invariant yields $|E_k| = v^k_{d_k} \preceq_{d_k} |\pi_k(d_k)|$, contradicting the fact that d_k must improve her coalition after her deviation. Hence, using the third invariant again, we have that $p_{d_k} > v^k_{d_k} > |E_k|$ and $v^k_{d_k} \prec_{d_k} p_k$. Hence, single-peakedness implies that $|E_k| \prec_{d_k} v^k_{d_k}$, and the deviation was again not improving, a contradiction.

(2) The second invariant follows by induction and the update rules. In particular, it holds for the agent $d_k \in E_k$ if d_k performed an L-move, because then $v^{k+1}_{d_k} = v^{k+1}_{E_k}$ and the invariant follows from the first invariant. The value of the coalition D_k will either be set to 0 or the invariant will be maintained for all agents in D_k.

(3) The third invariant holds by induction for the agents who do not change their value. For the agents in $E_k \setminus \{d_k\}$, it holds by definition of an IS-deviation, because they welcome agent d_k. The same is true for d_k if she performed an R-move. Otherwise, if d_k performs an L-move, we can apply induction, because she improved her utility.
Finally, we have to consider the agents in D_k. The invariant holds if no agent ever joined them. Otherwise, all agents $j \in D_k \setminus \{l_k\}$ have approved an agent to join and therefore $p_j > |D_k| \geq v^k_j$. In particular, single-peakedness implies that $v^k_j \preceq_j |D_k| = |\pi_k(j)|$. Finally, for l_k, the argument is the same if $v_{l_k} = |D_k|$ and we can apply induction (and single-peakedness) if $v_{l_k} < |D_k|$. Note that the case $v_{l_k} > |D_k|$ is excluded by induction for the second invariant.

(4) The forth invariant is immediate for all coalitions in $\pi_{k+1} \setminus \{D_k\}$. The only case where $v^{k+1}(D_k) \neq 0$ is if $v^{k+1}_{l_k} = v^k_{l_k}$, and $v^{k+1}_{D_k} = v^k_{D_k \cup \{d_k\}}$. If $v^{k+1}_{l_k} \neq 0$, then $v^k_{l_k} \neq 0$, and induction yields for this case that $v^{k+1}_{D_k} = v^k_{D_k \cup \{d_k\}} = v^k(l_k) = v^{k+1}(l_k)$.

Finally, we can show our main claim. We can restrict attention to coalitions D_k and E_k and agents within these. Note that either $v^k_{E_k} = 0$, or $v^k_{E_k} \leq |E_k| - 1$ while $v^{k+1}_j - v^k_j \geq 1$ for all $j \in E_k \setminus \{d_k\}$ (using the second invariant). Hence,

$$\sum_{j \in E_k \setminus \{d_k\}} v^{k+1}_j - v^k_j - v^k_{E_k} \geq 0. \tag{$*$}$$

If an L-move was performed, then $d_k = l_k$ or the agents in $D_k \cup \{l_k\}$ did not participate in a deviation, yet. In any case, $v^{k+1}_{E_k} \geq v^k_{D_k}$ (using the forth invariant). Hence,

$$\Lambda(\pi_{k+1}) - \Lambda(\pi_k) \overset{(*)}{\geq} v^{k+1}_{D_k} + \sum_{j \in D_k \cup \{d_k\}} v^{k+1}_j - v^k_j = 0.$$

Assume now that an R-move was performed. If the agents in D_k had value 0, then $\Lambda(\pi_{k+1}) - \Lambda(\pi_k) \geq v^{k+1}_{d_k} + v^{k+1}_{E_k} = 2(|E_k| - 1) > 0$.

Next, assume otherwise and that $d_k = l_k$. Then,

$$\Lambda(\pi_{k+1}) - \Lambda(\pi_k) \overset{(*)}{\geq} \underbrace{v^{k+1}_{d_k} - v^k_{d_k}}_{>0} + \underbrace{v^{k+1}_{E_k}}_{\geq 0} - \underbrace{v^k_{D_k}}_{=0} > 0.$$

Finally, consider the case that $d_k \neq l_k$. If $v^k_{l_k} = |D_k|$, then $v^{k+1}_j = v^k_j$ for all $j \in D_k$. Hence, as in the previous case, $\Lambda(\pi_{k+1}) - \Lambda(\pi_k) \geq (v^{k+1}_{d_k} - v^k_{d_k}) + v^{k+1}_{E_k} - v^k_{D_k} > 0$.

Draft – April 19, 2022

If $v^k_{l_k} \neq |D_k|$, then by the second invariant, $v^k_{l_k} < |D_k|$. Hence, $\sum_{j\in D_k} v^k_j - v^{k+1}_j \leq |D_k| - 1$ and $v^{k+1}_{E_k} = |E_k| - 1 \geq |D_k|$. Hence, $\Lambda(\pi_{k+1}) - \Lambda(\pi_k) > v^{k+1}_{d_k} - v^k_{d_k} > 0$. This completes the induction. $\lhd$

Now, note that for all $0 \leq k \leq m$, $\Lambda(\pi_k) \geq 0$ and the potential is integer-valued. Moreover, $\Lambda(\pi_m) = \sum_{j\in N} v^m_j + \sum_{C\in\pi_k} v^m_C = \sum_{C\in\pi_k}(v^m_C + \sum_{j\in C} v^m_j) \leq \sum_{C\in\pi} |C|^2 \leq n^2$. We use that if $v^m_C = 0$, then $v^m_j \leq |C|$ for all $j \in C$, and the first and second invariant. Hence, there are at most n^2 R-moves.

To get a global bound on the number of moves, we consider the simple potential that counts the pairs of agents that form common coalitions. Given a coalition C, this value is precisely $|C|(|C|-1)/2$. Define therefore the potential $\Gamma(\pi) = \sum_{C\in\pi} |C|(|C|-1)/2$. Note that $0 \leq \Gamma(\pi) \leq n(n-1)/2$. Now, every R-move raises the potential Γ by at most $n-1$, and every L-move diminishes it by at least 1. Hence, there can be at most $\Gamma(\pi_0) + n^2(n-1) \leq n^3$ L-moves. This bounds the total length of the execution of the dynamics by $n^2 + n^3 = O(n^3)$. $\square$

4 HEDONIC DIVERSITY GAMES

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. Following the definition of HDGs, we assume throughout this section that the agent set is always partitioned into sets R and B of red and blue agents, respectively. It is known that IS partitions always exist in HDGs, even without restrictions such as single-peakedness of preferences [8]. However, we prove that the dynamics of IS-deviations may start to cycle, even under very strong restrictions. This stands in contrast to empirical evidence for the convergence of dynamics based on extensive computer simulations by Boehmer and Elkind [8]. To this end, we consider natural restrictions of the preferences, of the starting partitions, and specific selection rules for the performed deviations. We show that most combinations of them still allow for infinite dynamics. Most surprisingly, we can show that the dynamics may cycle even if we start from the singleton partition and the preferences are strict and single-peaked.[3] However, if we add an arguably weak selection rule, we obtain convergence of the dynamics. To define this rule, we call a coalition $C \subseteq N$ *homogeneous* if it consists only of agents of one type, i.e., $C \subseteq R$ or $C \subseteq B$. We say that a deviation satisfies *solitary homogeneity* if, whenever the target coalition of the deviator is homogeneous, then it is a singleton coalition. Note that whenever an agent can perform an IS-deviation, then she can perform a deviation satisfying solitary homogeneity, simply by forming the homogeneous singleton coalition instead of joining existing homogeneous coalitions. Hence, assuming solitary homogeneity of deviations yields valid selection rules, i.e., whenever a deviation is possible, then solitary homogeneity does not prohibit all possible deviations.

In the second part of this section, we show that combining all considered restrictions leads to convergence of the dynamics. In other words, the IS-dynamics may cycle if and only if any of the four properties of Theorem 4.2 is violated. For the proof of the theorem, we make use of a lemma which already highlights the special role of homogeneous coalitions.

LEMMA 4.1. *Given a set R_a of red (respectively, set B_a of blue) agents whose preferences satisfy $\frac{2}{3} \succ 1 \succ \frac{1}{2}$ (respectively, $\frac{1}{3} \succ 0 \succ \frac{1}{2}$), it is possible to create the homogeneous coalition R_a (respectively, B_a) by means of an individual dynamics starting from singleton coalitions.*

We are ready to prove the theorem.

THEOREM 4.2. *The dynamics of IS-deviations may cycle in HDGs even if any three of the following restrictions apply:*

(1) preferences are naturally single-peaked,

[3] This corrects a statement in the conference version of this paper [11].

Draft – April 19, 2022

(2) *preferences are strict,*
(3) *the starting partition is the singleton partition, or*
(4) *all deviations satisfy solitary homogeneity.*

Proof. We provide examples for any triple of the four restrictions. The example where all properties except the condition on the starting partition, and where all properties except deviation selection according to solitary homogeneity are satisfied are closely related. First we show how to deal with the former case. Then, we show how to reach a configuration within the cycle of this case by starting from the singleton coalition. However, for reaching this cycle, some of the performed deviations violate solitary homogeneity. We defer the other two examples to the appendix.

(¬3) We start with an example of an HDG where all preferences are strict and naturally single-peaked and all agents' deviations satisfy solitary homogeneity. Therefore, 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 of the agents is given below.

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

Consider the sequence of IS-deviations in Figure 1 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.
Note that all deviations result in non-homogeneous target coalitions, and therefore satisfy solitary homogeneity.

(¬4) Our next goal is to provide an example of cycling under strict and single-peaked preferences while the starting partition is the singleton partition. Our example makes use of the previous example. As a first step, we show, how we can create the coalitions C_i for $i \in [4]$ by starting from the singleton coalition. In a second step, we show how to add the deviators r_1, r_2, b_1, and b_2 of the previous example to these coalitions to reach a partition from the cycle. As a consequence, cycling can occur by following the cycle from the first part of the proof. For highlighting the relationship of the two examples, we will use bold face for the relevant part of the preferences of the constructed coalitions.
To create the desired coalitions, we use Lemma 4.1 to create homogeneous auxiliary coalitions. As in the proof of the lemma, we assume that we take *new* (auxiliary) agents for every step of every construction.
Now, we show how to create the coalitions C_i for $i \in [4]$ one by one.

Creating C_2 and C_4. Given the insights gained in Lemma 4.1, it is straightforward to manufacture the homogeneous coalitions C_2 and C_4. Therefore, define the coalitions $C_2 = \{r_{2,i} : i \in [5]\}$ and $C_4 = \{b_{4,i} : i \in [6]\}$ together with the following single-peaked preferences.

$$
\begin{array}{ll}
C_2: & \frac{2}{3} \succ \boldsymbol{\frac{5}{7} \succ \frac{5}{6} \succ 1} \succ \frac{1}{2} \\
C_4: & \frac{1}{3} \succ \boldsymbol{\frac{1}{4} \succ \frac{1}{7} \succ 0} \succ \frac{1}{2}
\end{array}
$$

Draft – April 19, 2022

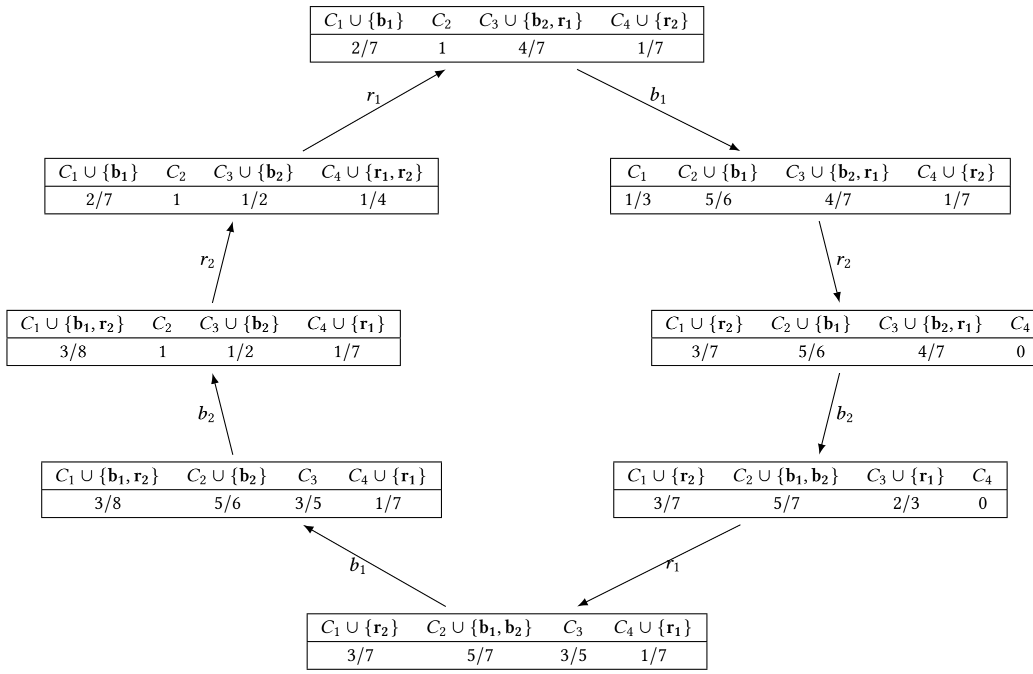


Fig. 1. Possibility of cycling of IS-dynamics in HDG under strict and single-peaked preferences. All deviations satisfy solitary homogeneity.

Since the preferences satisfy the assumptions of Lemma 4.1, we can apply it to create C_2 and C_4.

Creating C_1. Creating the coalition C_1 is also not very difficult. We can simply apply Lemma 4.1 to form a coalition of all the blue agents, and let the red agents join this coalition. More formally, consider the coalition $C_1 = \{b_{1,1}, b_{1,2}, b_{1,3}, b_{1,4}, r_{1,1}, r_{1,2}\}$ with the following single-peaked preferences.

$$C_1: \quad \tfrac{3}{8} \succ \tfrac{3}{7} \succ \tfrac{1}{3} \succ \tfrac{1}{5} \succ 0 \succ \tfrac{1}{2} \succ 1$$

We can apply Lemma 4.1 to form $B_a = \{b_{1,1}, b_{1,2}, b_{1,3}, b_{1,4}\}$. Then, $r_{1,1}$ and $r_{1,2}$ can perform IS-deviations to join B_a one after another. This results in the coalition C_1, as desired.

Creating C_3. The by far most difficult coalition to create is C_3, where we have to combine several steps. The central idea is to apply Lemma 4.1 to create a sufficiently large homogeneous coalition of auxiliary blue agents. Then, the red agents of the future coalition C_3 will join. This is followed by having the blue agents of the former homogeneous coalition abandon the so created coalition. An essential step is to create further coalitions to incentivize them to perform the necessary deviations. Finally, the two blue agents from C_3 can join.
To this end, consider the coalition $C_3 = \{b_{3,1}, b_{3,2}, r_{3,1}, r_{3,2}, r_{3,3}\}$ with the following strict and naturally single-peaked preferences.

$$C_3: \quad \tfrac{4}{7} \succ \tfrac{1}{2} \succ \tfrac{3}{5} \succ \tfrac{3}{4} \succ \tfrac{3}{10} \succ \tfrac{2}{9} \succ \tfrac{1}{8} \succ 1 \succ 0$$

For creating C_3, we consider a set of auxiliary agents A containing agents with single-peaked preferences such that her peak is at $\frac{4}{13}$ and the preferences satisfy $\frac{1}{3} \succ \frac{3}{11} \succ 0 \succ \frac{1}{2} \succ 1$.

Draft – April 19, 2022

Note that these agents can be of any color and we can just ignore the preference for 0 or 1 if we consider a red or blue agent from the set, respectively.
Now, let $i \in [7]$ and consider a set of *blue* agents $C_{3,1} = \{b_{a,i} \colon i \in [7]\} \subseteq A$. By Lemma 4.1, we can create the coalition $C_{3,1}$. Now, since we move the coalition ratio towards the peak of the blue agents, we can have the red agents from C_3 join one by one to form the coalition $C_{3,2} = C_{3,1} \cup \{r_{3,1}, r_{3,2}, r_{3,3}\}$. For this, note that $\frac{3}{10} < \frac{4}{13}$. The next goal is to get rid of the agents in $C_{3,1}$. To make this happen, we create auxiliary coalitions such that the agents in $C_{3,1}$ can move there and get into a most preferred coalition. Therefore, we create 7 identical coalitions as follows. By Lemma 4.1, we can create a homogeneous coalition consisting of 8 blue agents from A. Then, we let 4 red agents from A join one after another. Note that this only consists of IS-deviations, even though we cross the peak of these agents, because all agents satisfy $\frac{4}{12} = \frac{1}{3} \succ \frac{3}{11}$. Also, all agents in the resulting coalition would allow another blue agent to join, because this deviation would lead to reaching the peak of $\frac{4}{13}$. Hence, we let the agents from $C_{3,1}$, one after another, deviate to distinct auxiliary coalitions. All of these steps are a strict improvement for the deviators leaving $C_{3,2}$. The first of the deviators leaves a coalition of ratio $\frac{3}{10}$ and reaches her peak. Afterwards, the ratio of the abandoned coalitions is at least $\frac{3}{9} = \frac{1}{3} > \frac{4}{13}$ and therefore all other deviators improve strictly.
We obtain the coalition $\{r_{3,1}, r_{3,2}, r_{3,3}\}$ and the blue agents from C_3 can join one after another to form coalition C_3.

Starting cycling. The final step for this example is to show how to start the cycle constructed in the first HDG. Therefore, we have to add the deviator agents r_1, r_2, b_1, and b_2 with the following preferences.

$$\begin{aligned}
b_1 :&\quad \tfrac{3}{8} \succ \tfrac{5}{7} \succ \tfrac{5}{6} \succ \tfrac{2}{7} \succ 0\\
b_2 :&\quad \tfrac{5}{7} \succ \tfrac{4}{7} \succ \tfrac{1}{2} \succ \tfrac{5}{6} \succ 0\\
r_1 :&\quad \tfrac{4}{7} \succ \tfrac{1}{4} \succ \tfrac{1}{7} \succ \tfrac{2}{3} \succ 1\\
r_2 :&\quad \tfrac{1}{4} \succ \tfrac{3}{8} \succ \tfrac{3}{7} \succ \tfrac{1}{7} \succ 1
\end{aligned}$$

Given the constructed coalitions C_i, $i \in [4]$, we perform the following IS-deviations:

- Agent r_2 joins coalition C_1.
- Agent b_1 joins coalition C_2.
- Agent b_2 joins coalition C_3.
- Agent r_1 joins coalition $C_3 \cup \{b_2\}$.

This results in a partition that occurs in the cycle of the first example. Hence, the IS dynamics can start to cycle as shown before.

□

The previous examples do not show the impossibility to reach an IS partition since, e.g., in the first case, 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 4.3. *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.*

Draft – April 19, 2022

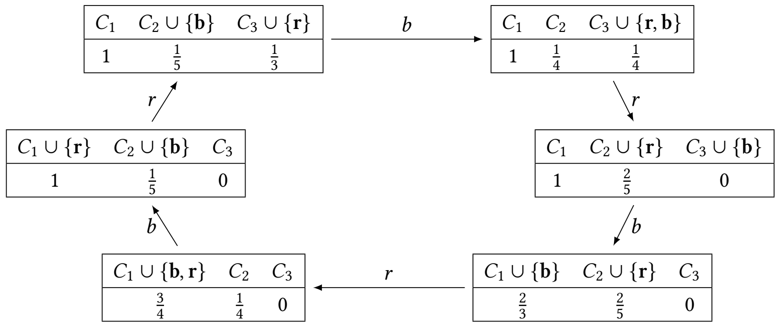


Fig. 2. Impossibility of convergence of IS-dynamics.

Proof. Let us consider an HDG with 10 agents: 4 red agents and 6 blue agents. There are two deviating agents: red agent r and blue agent b, and three fixed coalitions C_1, C_2 and C_3 such that:

- C_1 contains 2 red agents;
- C_2 contains 1 red agent and 3 blue agents;
- C_3 contains 2 blue agents.

The relevant part of the preferences of the agents is given below, with on the left the preferences for the case of (1) strict preferences and on the right the preferences for the case of (2) naturally single-peaked preferences with indifference.

$$\begin{array}{rl}
& (1) \\
r: & \frac{3}{4} \succ \frac{2}{5} \succ \frac{1}{4} \succ \frac{1}{3} \succ 1 \\
b: & \frac{1}{4} \succ \frac{1}{5} \succ \frac{3}{4} \succ \frac{2}{3} \succ 0 \\
C_1: & \frac{3}{4} \succ \frac{2}{3} \succ 1 \\
C_2: & \frac{2}{5} \succ \frac{1}{5} \succ \frac{1}{4} \\
C_3: & \frac{1}{4} \succ \frac{1}{3} \succ 0
\end{array}
\qquad
\begin{array}{rl}
& (2) \\
r: & \frac{3}{4} \succ \frac{2}{5} \succ \frac{1}{4} \sim \frac{1}{3} \succ 1 \\
b: & \frac{1}{4} \succ \frac{1}{5} \succ \frac{3}{4} \sim \frac{2}{3} \succ 0 \\
C_1: & \frac{3}{4} \succ \frac{2}{3} \succ 1 \\
C_2: & \frac{2}{5} \succ \frac{1}{5} \sim \frac{1}{4} \\
C_3: & \frac{1}{4} \succ \frac{1}{3} \succ 0
\end{array}$$

Consider the sequence of IS-deviations in Figure 2 that describe a cycle in the dynamics. The two deviating agents of the cycle r and b are marked in bold and the specific deviating agent between two states is indicated next to the arrows.

Note that at each state, the deviation performed by agent r or b is the only possible one that they can do. Moreover, for the other agents, by assuming that in case (1) all the omitted fractions are ranked after the mentioned partial preferences, and in case (2) that they are all indifferent, none of them has incentive to deviate at any state. Therefore, the cycle is the only possible sequence of IS-deviations, and the cycle cannot be avoided in this instance. □

We show now that convergence is guaranteed by combining all restrictions of Theorem 4.2. Interestingly, part of the proof is a reduction to the positive result about AHGs, and reveals a close relationship of the two classes of hedonic games if HDGs are sufficiently restricted. However, this reduction requires careful preprocessing of the initial HDG. The first key insight is to show that, for every coalition occurring during the dynamics, there exists an agent type of which at most one agent is present. Consequently, given an instance with b blue and r red agents, respectively, the only important ratios (apart from 0 and 1) are $\frac{k}{k+1}$ and $\frac{1}{k'+1}$, where $1 \leq k \leq r$ and $1 \leq k' \leq b$. The next step is to show how to transition to an HDG where every agent will only end up in coalitions of one of these ratio types. This transition is essentially performed by omitting certain steps in the

Draft – April 19, 2022

dynamics. From there, we can observe the structure of an AHG by identifying the ratios $\frac{k}{k+1}$ and $\frac{1}{k+1}$ with a coalition size of $k+1$. This correspondence is reasonable. For instance, the former ratio corresponds to a coalition with one blue and k red agents, i.e., a total number of $k+1$ agents. We can apply Theorem 3.4 to bound the length of the transformed HDG. Interestingly, the identification with an AHG requires some auxiliary agents, and the transformed dynamics is not starting from the singleton partition anymore. In this respect, we even need the full power of Theorem 3.4.

Theorem 4.4. *The dynamics of IS-deviations satisfying solitary homogeneity always converges in $O(n^5)$ steps when starting from the singleton partition in an HDG where agents entertain strict and naturally single-peaked preferences.*

Proof. Consider an HDG with agent set $N = R \cup B$, where agents have strict and naturally single-peaked preferences. Let $(\pi_k)_{k=0}^{K}$ be a sequence of partitions of an execution of the dynamics of IS-deviations satisfying solitary homogeneity, where π_0 is the singleton partition and, for every $1 \leq k \leq K$, π_k evolves from π_{k-1} by an IS-deviation of agent d_k.

The first step of the proof is to show the specific structure of the attained coalitions.

Claim 4.1. *For every $k \geq 0$, it holds that every coalition in π_k is of the form $\{r_1\}$, $\{b_1\}$, $\{r_1, b_1, \dots, b_k\}$, or $\{b_1, r_1, \dots, r_{k'}\}$, where $1 \leq k \leq |B|$ and $1 \leq k' \leq |R|$ and $r_i \in R$ and $b_j \in B$ for every $i \in [k']$, $j \in [k]$. Moreover, the following statements hold:*

(1) If $\{r_1, b_1, \dots, b_m\} \in \pi_k$ for $m \geq 2$, then $\frac{1}{m+1} \succ_{r_1} \frac{1}{m} \succ_{r_1} \cdots \succ_{r_1} \frac{1}{2} \succ_{r_1} 1$.

(2) If $\{b_1, r_1, \dots, r_{m'}\} \in \pi_k$ for $m' \geq 2$, then $\frac{m'}{m'+1} \succ_{b_1} \frac{m'-1}{m'} \succ_{b_1} \cdots \succ_{b_1} \frac{1}{2} \succ_{b_1} 0$.

Proof. We will show by induction over k for $0 \leq k \leq K$ that every coalition in π_k is of the form $\{r_1\}$, $\{b_1\}$, $\{r_1, b_1, \dots, b_m\}$, or $\{b_1, r_1, \dots, r_{m'}\}$, where $b_i \in B$ and $r_j \in R$ for every $i \in [m]$, $j \in [m']$, $1 \leq m \leq |B|$, and $1 \leq m' \leq |R|$. Simultaneously to this main claim, we will prove the additional statements as auxiliary claims.

Clearly, the singleton partition satisfies the main and auxiliary claims. Now, assume that the assertion is true for some fixed $0 \leq k < K$. Assume without loss of generality that d_{k+1} is a red agent (the case for a blue agent is symmetric and uses the second auxiliary claim where we use the first auxiliary claim). We have to consider the two coalitions affected by d_{k+1} to show that π_{k+1} satisfies the claims.

First, assume for contradiction that the coalition $\pi_{k+1}(d_{k+1})$ breaks the main claim. Then, $\pi_{k+1}(d_{k+1}) \setminus \{d_{k+1}\}$ is of the form $\{r_1\}$ or $\{r_1, b_1, \dots, b_m\}$ with $2 \leq m \leq |B|$. The former case is excluded as the deviation satisfies solitary homogeneity. In the latter case, $\frac{2}{2+m} > \frac{1}{m} > \frac{1}{m+1}$, and we know by the first auxiliary claim in step k that $\frac{1}{m+1} \succ_{r_1} \frac{1}{m}$. Hence, single-peakedness implies $\frac{1}{m} \succ_{r_1} \frac{2}{2+m}$, and, by transitivity of the preferences, we obtain that $\frac{1}{m+1} \succ_{r_1} \frac{2}{2+m}$. This contradicts the fact that the deviation by d_{k+1} was approved by agent r_1. Hence, $\pi_{k+1}(d_{k+1})$ satisfies the main claim and must be of the form $\{b_1, r_1, \dots, r_{m'}\}$, where $m' \geq 1$.

We proceed with the auxiliary claims for this coalition. As the coalition contains only one blue agent, the first auxiliary claim is vacant. Further, since b_1 gave her consent to letting d_{k+1} join, it satisfies the second auxiliary claim (extending the second auxiliary claim for b_1 at step k) if $m' \geq 3$. If $m' = 2$, then the consent of b_1 implies $\frac{2}{3} \succ_{b_1} \frac{1}{2}$, and single-peakedness implies $\frac{2}{3} \succ_{b_1} \frac{1}{2} \succ_{b_1} 0$. If $m' = 1$, this claim is also vacant.

Second, assume for contradiction that the coalition abandoned by agent d_{k+1} violates the main claim. Then, $\pi_k(d_{k+1})$ was of the form $\{r, b_1, \dots, b_m\}$ with $m \geq 2$ and $r = d_{k+1}$. We already know that $\pi_{k+1}(d_{k+1})$ is of the desired form. It cannot be the coalition $\{r\}$, because of the first auxiliary claim for r. Also, it cannot be of the form $\{r, \hat{b}_1, \dots, \hat{b}_{\hat{m}}\}$ with $\hat{m} \geq 2$, because then $\pi_k(\hat{b}_1)$ violates the main claim in step k. Hence, we know that $\pi_{k+1}(d_{k+1})$ is of the form $\{b, r_1, \dots, r_{m'}\}$ with $m' \geq 1$.

Draft – April 19, 2022

Using $m \geq 2$, we have that $\frac{1}{m+1} < \frac{m'}{m'+1} < 1$ and, since the deviation was performed by d_{k+1}, also $\frac{m'}{m'+1} \succ_{d_{k+1}} \frac{1}{m+1}$. Hence, single-peakedness implies $1 \succ_{d_{k+1}} \frac{m'}{m'+1}$, and therefore, by transitivity, $1 \succ_{d_{k+1}} \frac{1}{m+1}$. However, this contradicts the first auxiliary claim for agent $d_{k+1} = r$ in π_k.

It remains to prove the auxiliary claims for the abandoned coalition. The first auxiliary claim is vacant. The second auxiliary claim follows directly by induction, whenever it is not vacant. ◁

In the sequel, we use the notation $f_i(\pi) = \frac{|R \cap \pi(i)|}{|\pi(i)|}$, which specifies the fraction of red agents in the coalition of agent i with respect to partition π. Also, given an agent i, denote her peak by p_i. We distinguish agents according to their peaks. To this end, define the agent sets

- $R_S = \{r \in R \colon 0 < p_r < 1/2\}$,
- $R_L = \{r \in R \colon 1/2 \leq p_r \leq 1\}$,
- $B_S = \{b \in B \colon 0 \leq p_b \leq 1/2\}$, and
- $B_L = \{b \in B \colon 1/2 < p_b < 1\}$.

The subscripts indicate whether the peak is large (L) or small (S). We would like to analyze a dynamics where $f_i(\pi)$ is always close to the peak of an agent. This is achieved by agents in R_L and B_S.

Claim 4.2. *Let $k \geq 0$. Then, the following statements hold:*

(1) If $r \in R_L$, then $f_r(\pi_k) \geq \frac{1}{2}$.
(2) If $b \in B_S$, then $f_b(\pi_k) \leq \frac{1}{2}$.

Proof. We show the statement by induction over k for $0 \leq k \leq K$. Clearly, the statement is true for $k = 0$, because π_0 is the singleton partition. Now, assume that the assertion is true for some fixed $0 \leq k < K$. We assume without loss of generality that agent d_{k+1} is red (the case of a blue agent follows from a symmetric argument). Clearly, all agents not affected by the deviation maintain the two invariants claimed in the lemma. Therefore, we have to consider the abandoned and joined coalitions.

By Claim 4.1, $\pi_k(d_{k+1})$ is of the form $\{d_{k+1}\}$ or $\{b_1, r_1, \dots, r_m, d_{k+1}\}$ for some $m \geq 0$, where $b_1 \in B$ and $r_1, \dots, r_m \in R$. The former case is vacant. In the latter case, $f_{b_1}(\pi_{k+1}) < f_{b_1}(\pi_k)$, and the second invariant follows by induction if $b \in B_S$. Furthermore, if $m \geq 1$, then $f_{b_1}(\pi_{k+1}) \geq \frac{1}{2}$, and the first invariant is true if $r_i \in R_L$ for $1 \leq i \leq m$.

Applying Claim 4.1 again, $\pi_{k+1}(d_{k+1})$ is also of the form $\{d_{k+1}\}$ or $\{b_1, r_1, \dots, r_{m'}, d_{k+1}\}$ for some $m' \geq 0$, where $b_1 \in B$ and $r_1, \dots, r_{m'} \in R$. Hence, the first invariant is satisfied for all red agents and for b_1 if $m' = 0$. It remains the case $m' \geq 1$. Then, since the deviation was approved by agent b_1, it holds $\frac{m'+1}{m'+2} \succ_{b_1} \frac{m'}{m'+1}$. Then, single-peakedness implies that $p_{b_1} > \frac{m'}{m'+1} \geq \frac{1}{2}$ and therefore $b_1 \notin B_S$. Hence, the second invariant is vacant in this case.

Altogether, we have shown that both invariants are satisfied for partition π_{k+1}, which completes the induction step. ◁

A similar statement is not true for agents in R_S or B_L. Therefore, the next step will *modify* the dynamics such that agents are only contained in coalitions close to their peaks, unless they are in a singleton coalition. Given a partition π, define the subset of agents $F_\pi \subseteq N$ as $F_\pi = \{r \in R_S \colon \frac{1}{2} < f_r(\pi) < 1\} \cup \{b \in B_L \colon \frac{1}{2} > f_b(\pi) > 0\}$, i.e., the set of agents whose ratio is *far* from their peak. By definition of F_π and Claim 4.1, a red agent $r \in R \cap F_\pi$ is in a coalition of the form $\{b_1, r_1, \dots, r_m\}$ in partition π and, symmetrically, a blue agent $b \in B \cap F_\pi$ is in a coalition of the form $\{r_1, b_1, \dots, b_m\}$ in partition π. Moreover, by Claim 4.1 and strict single-peakedness of the preferences, an agent in F_π is the last agent who entered her coalition in π.

Now, consider a modified dynamics $(\sigma_k)_{k=0}^K$, where, for every $0 \leq k \leq K$, $\sigma_k = \bigcup_{C \in \pi_k} \{C \setminus F_{\pi_k}\} \cup \bigcup_{i \in F_{\pi_k}} \{\{i\}\}$. This modification has essentially the following effects: We omit deviations of agents

Draft – April 19, 2022

where they land in the set F_π while keeping them in a singleton coalition. Sometimes, it can happen that the deviator satisfies $d_k \in F_{\pi_{k-1}} \setminus F_{\pi_k}$. In this case, the modified dynamics *sees* the deviator join her new coalition from a singleton coalition. The second effect that plays a role is the case where a non-deviator is in $F_{\pi_{k-1}} \setminus F_{\pi_k}$. This happens exactly if an agent in $F_{\pi_{k-1}}$ is abandoned, and thereby left in a coalition of size 2, which consists of one agent of each type. Hence, we insert a suitable deviation to obtain a valid modified dynamics of IS-deviations. This requires two lemmas. The first lemma gives more structural insight and establishes that every coalition can contain at most one agent from F_{π_k}.

Claim 4.3. *Let $0 \leq k \leq K$ and $C \subseteq \pi_k$. Then, $|C \cap F_{\pi_k}| \leq 1$. Moreover, the following statements hold:*

(1) If $\{b_1, r_1, \ldots, r_m\} \in \pi_k$, then, for all $1 \leq i \leq m$ with $r_i \notin F_{\pi_k}$, $f_{r_i}(\pi_k) \succeq_{r_i} f_{r_i}(\sigma_k)$.
(2) If $\{r_1, b_1, \ldots, b_m\} \in \pi_k$, then, for all $1 \leq i \leq m$ with $b_i \notin F_{\pi_k}$, $f_{b_i}(\pi_k) \succeq_{b_i} f_{b_i}(\sigma_k)$.

Proof. We show all statements simultaneously by induction over k for $0 \leq k \leq K$. Clearly, all statements are true for $k = 0$. Now, assume that the statements are true for some fixed $0 \leq k < K$. Clearly, all coalitions in π_{k+1} except possibly $\pi_k(d_{k+1}) \setminus \{d_{k+1}\}$ and $\pi_{k+1}(d_{k+1})$ satisfy the claim. Assume without loss of generality that d_{k+1} is a red agent (the case of a blue agent is symmetric).

We start with the abandoned coalition. By Claim 4.1, $\pi_k(d_{k+1}) \setminus \{d_{k+1}\}$ is the empty set (if d_{k+1} was in a singleton coalition) or of the type $\{b_1, r_1, \ldots, r_m\}$ where $b_1 \in B$ and $r_1, \ldots, r_m \in R$ for some $m \geq 0$. In the former case, all claims are vacant, so assume the latter case. If $m = 0$, the abandoned coalition is a singleton coalition, and the assertions are true (the additional statements are then vacant). If $m = 1$, then $f_{b_1}(\pi_{k+1}) = \frac{1}{2}$, and therefore $\pi_{k+1}(b_1) \cap F_{\pi_{k+1}} = \emptyset$. In particular, $\pi_{k+1}(b_1) = \sigma_{k+1}(b_1)$, and all claims are true. If $m \geq 2$, it follows from $f_{b_1}(\pi_{k+1}) \geq \frac{2}{3}$ that $b_1 \notin F_{\pi_{k+1}}$. Moreover, $\{r_1, \ldots, r_m\} \cap F_{\pi_{k+1}} \subseteq \{r_1, \ldots, r_m\} \cap F_{\pi_k}$. Therefore, $|\pi_{k+1}(b_1) \cap F_{\pi_{k+1}}| \leq 1$ follows from induction for step k, which implies that $|\{r_1, \ldots, r_m\} \cap F_{\pi_k}| \leq |\pi_k(b_1) \cap F_{\pi_k}| \leq 1$. Additionally, the additional statement is trivially true unless $\{r_1, \ldots, r_m\} \cap F_{\pi_{k+1}} \neq \emptyset$, say $r_1 \in F_{\pi_{k+1}}$. Then, $r_1 \in F_{\pi_k}$ and we have that $f_{b_1}(\sigma_{k+1}) < f_{b_1}(\pi_{k+1}) = f_{b_1}(\sigma_k) < f_{b_1}(\pi_k)$. Let $2 \leq i \leq m$. Then, induction implies $f_{r_i}(\pi_k) \succ_{r_i} f_{r_i}(\sigma_k)$. Hence, single-peakedness implies $f_{r_i}(\pi_{k+1}) \succ_{r_i} f_{r_i}(\sigma_{k+1})$.

The proof for the joined coalition is similar. Using Claim 4.1 again, we know that $\pi_{k+1}(d_{k+1})$ is a singleton coalition or of the type $\{b_1, r_1, \ldots, r_m, d_{k+1}\}$ where $b_1 \in B$ and $r_1, \ldots, r_m \in R$ for some $m \geq 0$. A singleton coalition fulfills the claim. Therefore, assume the latter case. Since $f_{b_1}(\pi_{k+1}) \geq \frac{1}{2}$, it holds that $b_1 \notin F_{\pi_{k+1}}$. Moreover, if $m \geq 1$ and $1 \leq i \leq m$, then $\frac{m+1}{m+2} \succ_{r_i} \frac{m}{m+1}$ and single-peakedness implies that $p_{r_i} > \frac{m}{m+1} \geq \frac{1}{2}$. Hence, $r_i \notin R_S$, and therefore $r_i \notin F_{\pi_{k+1}}$. Together, $\pi_{k+1}(d_{k+1}) \cap F_{\pi_{k+1}} \subseteq \{d_{k+1}\}$, and the first statement is true. The additional statement is clear if $d_{k+1} \notin F_{\pi_{k+1}}$, in which case $\pi_{k+1}(b_1) = \sigma_{k+1}(b_1)$. If $d_{k+1} \in F_{\pi_{k+1}}$, it follows for the other red agents, because they approve that d_{k+1} joins. ◁

We are ready to show how to obtain the valid dynamics.

Claim 4.4. *Let $1 \leq k \leq K$. If $\sigma_k \neq \sigma_{k-1}$, then σ_k evolves from σ_{k-1} by performing at most 2 IS-deviations. If two deviations have to be performed, then the intermediate partition evolves from σ_{k-1} by merging two agents from singleton coalitions.*

Proof. Let $1 \leq k \leq K$ with $\sigma_k \neq \sigma_{k-1}$. The only agents that matter to us are in $\pi_k(d_k) \cup \pi_{k-1}(d_k)$. Other agents did not change their coalition in the original dynamics, and therefore, their membership in F_{π_k} is also not affected. Without loss of generality, we assume that d_k is a red agent (the case of a blue agent is again symmetric).

First, we show that $(\pi_k(d_k) \setminus \{d_k\}) \cap F_{\pi_{k-1}} = \emptyset$ and $\pi_k(d_k) \cap F_{\pi_k} = \emptyset$. Assume for contradiction that this is not the case. By Claim 4.1, this can only be the case if $\pi_k(d_k)$ is of the form $\{b_1, r_1, \ldots, r_m, d_k\}$

Draft – April 19, 2022

where $b_1 \in B$ and $r_1, \dots, r_m \in R$ for some $m \geq 1$. As $f_{b_1}(\pi_k) \geq \frac{2}{3}$ and $f_{b_1}(\pi_{k-1}) \geq \frac{1}{2}$, $b_1 \notin F_{\pi_k}$ and $b_1 \notin F_{\pi_{k-1}}$, respectively. For $1 \leq i \leq m$, it holds that $\frac{m+1}{m+2} \succ_{r_i} \frac{m}{m+1}$, and therefore, by single-peakedness, $p_i > \frac{m}{m+1} \geq \frac{1}{2}$. Hence, $r_i \in R_L$, and therefore $r_i \notin F_{\pi_k}$ and $r_i \notin F_{\pi_{k-1}}$. This shows already that $(\pi_k(d_k) \setminus \{d_k\}) \cap F_{\pi_{k-1}} = \emptyset$.

Finally, it remains to exclude that $d_k \in F_{\pi_k}$. Assume for contradiction that $d_k \in F_{\pi_k}$. Then, our considerations about $\pi_k(d_k)$ imply that $\sigma_k(b_1) = \sigma_{k-1}(b_1)$. Moreover, by Claim 4.1, $f_{d_k}(\pi_{k-1}) \geq \frac{1}{2}$. Since $d_k \in F_{\pi_k}$, it holds $p_{d_k} < \frac{1}{2}$. As we already know that $f_{d_k}(\pi_k) > \frac{1}{2}$, single-peakedness and the fact that d_k has improved her ratio imply that $f_{d_k}(\pi_{k-1}) \geq f_{d_k}(\pi_k) > \frac{1}{2}$. Therefore, $d_k \in F_{\pi_{k-1}}$ and Claim 4.2 implies that $(\pi_{k-1}(d_k) \setminus \{d_k\}) \cap F_{\pi_{k-1}} = \emptyset$. Additionally, $f_{d_k}(\pi_{k-1}) > \frac{1}{2}$ also implies that $(\pi_{k-1}(d_k) \setminus \{d_k\}) \cap F_{\pi_k} = \emptyset$. Hence, the coalition abandoned by d_k has not changed from σ_{k-1} to σ_k. Together, this contradicts that $\sigma_k \neq \sigma_{k-1}$. Hence, our assumption that $d_k \in F_{\pi_k}$ was wrong, and therefore $\pi_k(d_k) \cap F_{\pi_k} = \emptyset$. In particular, we have shown so far that

$$\sigma_{k-1}(b_1) = \pi_{k-1}(b_1) \text{ and } \sigma_k(b_1) = \pi_k(b_1) = \pi_{k-1} \cup \{d_k\}. \tag{1}$$

Next, we consider the abandoned coalition. By Claim 4.1, the coalition $\pi_{k-1}(d_k)$ is of the form $\{d_k\}$ or $\{b'_1, r'_1, \dots, r'_{m'}, d_k\}$ where $b'_1 \in B$ and $r'_1, \dots, r'_{m'} \in R$ for some $m' \geq 0$. In the first case, we know that $\{d_k\} \in \sigma_{k-1}$, and, together with Equation 1, $\sigma_{k-1} = \pi_{k-1}$ and $\sigma_k = \pi_k$. Therefore, σ_k evolves from σ_{k-1} by an IS-deviation of agent d_k.

Next, we consider the case that $\pi_{k-1}(d_k)$ is of the form $\{b'_1, r'_1, \dots, r'_{m'}, d_k\}$. Note that $f_{b'_1}(\pi_{k-1}) \geq \frac{1}{2}$ and $f_{b'_1}(\pi_k) \geq \frac{1}{2}$ or $f_{b'_1}(\pi_k) = 0$ and therefore $b'_1 \notin F_{\pi_{k-1}}$ and $b'_1 \notin F_{\pi_k}$. Moreover, it holds that $\{r'_1, \dots, r'_{m'}\} \cap F_{\pi_k} \subseteq \{r'_1, \dots, r'_{m'}\} \cap F_{\pi_{k-1}}$. We are ready to consider the final cases.

First assume that $d_k \in F_{\pi_{k-1}}$. Then, Claim 4.2 and the previous fact imply that $\{b'_1, r'_1, \dots, r'_{m'}\} \cap F_{\pi_k} = \emptyset$ and $\{b'_1, r'_1, \dots, r'_{m'}\} \cap F_{\pi_{k-1}} = \emptyset$. Hence, $\{b'_1, r'_1, \dots, r'_{m'}\} \in \sigma_k$ and $\{b'_1, r'_1, \dots, r'_{m'}\} \in \sigma_{k-1}$. This, together with Equation 1 and the definition of σ_{k-1} implies that σ_k evolves from σ_{k-1} by a unilateral deviation of agent d_k from a singleton coalition to coalition $\sigma_k(d_k)$. Since $\sigma_k \neq \sigma_{k-1}$, we know that $\sigma_k(d_k) \neq \{d_k\}$. Hence, $\frac{1}{2} \leq f_{d_k}(\sigma_k) < 1$. This, together with $d_k \in R_S$ implies that $p_{d_k} \leq f_{d_k}(\sigma_k) < 1$. Hence, single-peakedness implies that the deviation was a Nash deviation. By Equation 1, the deviation was also approved by all agents in the joined coalition. Hence, σ_k evolves from σ_{k-1} through an IS-deviation of agent d_k.

It remains the case that $d_k \notin F_{\pi_{k-1}}$. If $\{r'_1, \dots, r'_{m'}\} \cap F_{\pi_{k-1}} = \emptyset$, then $\sigma_{k-1} = \pi_{k-1}$ and, together with Equation 1, σ_k evolves from σ_{k-1} by an IS-deviation of agent d_k.

Assume therefore that there exists $1 \leq i \leq m'$ with $r'_i \in F_{\pi_{k-1}}$. If $m' \geq 2$, then $\{b'_1, r'_1, \dots, r'_{m'}\} \cap F_{\pi_{k-1}} = \{b'_1, r'_1, \dots, r'_{m'}\} \cap F_{\pi_k}$. In this case, σ_k evolves from σ_{k-1} through a unilateral deviation of d_k. Since $d_k \notin F_{\pi_{k-1}}$, the first additional statement of Claim 4.2 implies that $f_{d_k}(\pi_{k-1}) \succeq_{d_k} f_{d_k}(\sigma_{k-1})$. Therefore, d_k performs a Nash deviation because she performed a Nash deviation from π_{k-1} to π_k. The consent of the joined coalition follows again from Equation 1.

Finally, assume that $\{r'_1, \dots, r'_{m'}\} \cap F_{\pi_{k-1}} \neq \emptyset$ and $m' = 1$. Then, $\pi_{k-1}(b'_1) = \{b'_1, r'_1, d_k\}$ and $r'_1 \in F_{\pi_{k-1}}$. Hence, σ_k evolves from σ_{k-1} by transforming $\{r'_1\}$, $\{b'_1, d_k\}$, and $\pi_k(d_k) \setminus \{d_k\}$ into $\{b'_1, r'_1\}$ and $\pi_k(d_k)$. These changes can be achieved by two unilateral deviations. First, d_k joins $\pi_k(d_k) \setminus \{d_k\}$ and then r'_1 joins b'_1. The first deviation is an IS-deviation as in the previous case. The second deviation is also an IS-deviation. The approval of b'_1 follows from the second auxiliary statement in Claim 4.1 applied to $\pi_{k-1}(b'_1) = \{b'_1, r'_1, d_k\}$. Also, the deviation is improving for r'_1, because $r'_1 \in F_{\pi_{k-1}}$. Therefore $r'_1 \in R_S$. Hence, $1 > \frac{1}{2} \geq p_{r'_1}$, and therefore $\frac{1}{2} \succ_{r'_1} 1$. $\lhd$

Using the insights gained in the previous claim, we can define a valid modified dynamics based on $(\sigma_k)_{k=1}^K$. First, we insert the partitions identified in the proof of Claim 4.4 where two agents are merged. This yields a dynamics $(\tau_l)_{l=1}^L$ such that τ_l evolves from τ_{l-1} through a deviation of agent $\hat{d}_l$

whenever $\tau_l \neq \tau_{l-1}$. Then, we remove all steps where $\tau_l = \tau_{l-1}$ to obtain a dynamics $(\rho_p)_{p=1}^P$, which is a dynamics where every step corresponds to an IS-deviation. Define $C = \{1 \leq k \leq K \colon \sigma_k = \sigma_{k-1}\}$, that is, the set of steps where the modified dynamics remains unchanged. Then, $K \leq L = |C| + P$.

Hence, we would like to obtain bounds on $|C|$ and P, respectively. The next claim allows us to bound $|C|$ by replacing it with an appropriate bound with respect to P. The key insight for proving the next claim follows from the observation says that essentially every n-th deviation of an agents has to correspond to a deviation of the modified dynamics.

Claim 4.5. *It holds that $L \leq n^2 + nP$.*

Proof. We first show that if $\sigma_k = \sigma_{k-1}$, then $d_k \in F_{\pi_k}$. We prove this fact by contraposition. Assume that $d_k \notin F_{\pi_k}$. Note that, by the definition of F_π, it holds for every partition π and every coalition $C \in \pi$ with $C \cap F_\pi \neq \emptyset$ that $|C| \geq 3$. Hence, if $\sigma_k(d_k) \neq \{d_k\}$, then there exists an agent $x \in \sigma_k(d_k) \setminus \{d_k\}$. Since d_k was the deviator, it holds that $\sigma_k(x) \neq \sigma_{k-1}(x)$ and therefore $\sigma_k \neq \sigma_{k-1}$. It remains the case that $\sigma_k(d_k) = \{d_k\}$. Then, as $d_k \notin F_{\pi_k}$, $\pi_k(d_k) = \{d_k\}$. This implies that $d_k \notin F_{\pi_{k-1}}$. To see this, we assume without loss of generality that $d_k \in R$. Indeed, if $d_k \in F_{\pi_{k-1}}$, then $d_k \in R_S$ and $\frac{1}{2} \leq f_{d_k}(\pi_{k-1})$. Then, single-peakedness implies that $f_{d_k}(\pi_{k-1}) \succ_{d_k} 1$, contradicting that d_k performed an IS-deviation to form a singleton coalition. Hence, $d_k \notin F_{\pi_{k-1}}$. As in the first case, we find an agent $x \in \sigma_{k-1}(d_k) \setminus \{d_k\}$, for which it holds that $\sigma_k(x) \neq \sigma_{k-1}(x)$ and therefore $\sigma_k \neq \sigma_{k-1}$.

The key insight for this claim is that every agent can only perform few successive deviations corresponding to steps in C. Indeed, the first part of the proof implies that $d_k \in F_{\pi_k}$ whenever $k \in C$. Consider an arbitrary agent $r \in R_S$. We define a potential function

$$\lambda_r(\pi) = \begin{cases} |R| + 1 & f_r(\pi) \leq \frac{1}{2} \text{ or } f_r(\pi) = 1 \\ m & f_r(\pi) = \frac{m}{m+1} \text{ for } 2 \leq m \leq |R| \end{cases}.$$

Note that λ_r is integer-valued and $2 \leq \lambda_r \leq |R| + 1$. We will show that λ_r decreases whenever r performs a deviation at step k where she lands in F_{π_k}, and can only increase through a deviation of r in $(\tau_l)_{l=1}^L$. In particular, we will show that the potential does not increase if another agent performs a deviation, unless when $r \in F_{\pi_{k-1}} \setminus F_{\pi_k}$ which corresponds to the case of inserting a deviation by r, which also corresponds to a deviation in $(\tau_l)_{l=1}^L$.

Consider a step k in the dynamics where $d_k = r$ and $r \in F_{\pi_k}$. Then, $\frac{1}{2} < f_r(\pi_k) < 1$. By Claim 4.1, $f_r(\pi_k) = \frac{m}{m+1}$ for some $m \in [|R|]$ and therefore $\lambda_r(\pi_k) = m$. Also by Claim 4.1, r is not allowed to perform a deviation if $f_r(\pi_{k-1}) < \frac{1}{2}$. Hence, single-peakedness implies that $f_r(\pi_{k-1}) > \frac{m}{m+1}$, and therefore $\lambda_r(\pi_k) < \lambda_r(\pi_{k-1})$.

If $d_k = r$ and $r \notin F_{\pi_k}$, then a deviation happens where $\sigma_k \neq \sigma_{k-1}$, and therefore this corresponds to a deviation of r in $(\tau_l)_{l=1}^L$. Next, we want to inspect how r is affected from a deviator if $d_k \neq r$. In this case, d_k cannot join $\pi_{k-1}(r)$ if $\lambda_r(\pi_{k-1}) \leq m$ (that is, in the case where r is in a coalition of 'large' ratio). Indeed, since $r \in R_S$, r would block any red agent to join, and a blue agent cannot join due to Claim 4.1. Hence, $\pi_{k-1}(r)$ is only affected if $d_k \in \pi_{k-1}(r)$. If $|\pi_{k-1}(r)| = 2$, then $f_r(\pi) = \frac{1}{2}$, and the potential cannot go up. Otherwise, Claim 4.1 implies that d_k is red. Since $r \in R_s$, it holds in addition that $r \notin F_{\pi_{k-1}}$. Hence, $\lambda_r(\pi_k) = \lambda_r(\pi_{k-1}) - 1$ or $\pi_k(r)$ is of size 2 and $\lambda_r(\pi_k) = |R| + 1$. As then $r \in F_{\pi_{k-1}} \setminus F_{\pi_k}$, this case corresponds exactly to inserting the deviation of r to form a coalition of size 2 in $(\tau_l)_{l=1}^L$.

Together, there can be at most $|R| - 1 \leq n - 1$ successive deviations by r corresponding to steps in C until there is a deviation by r in $(\tau_l)_{l=1}^L$. We obtain a bound for the deviations by r which matter. To make this formal, we consider the following quantities. Given an agent x, define

Draft – April 19, 2022

$C_x = |\{k \in C\colon d_k = x\}|$ and $L_x = |\{1 \le l \le L\colon \hat{d}_l = x\}|$. Since at least every n-th deviation counts towards L_r but not towards C_r, we can conclude that $L_r - C_r \ge \lfloor \frac{L_r}{n} \rfloor \ge \frac{L_r}{n} - 1$.

By an analogous argument where we consider an analogous potential function for blue agents, we obtain that $L_b - C_b \ge \frac{L_b}{n} - 1$ for every $b \in B_L$. Additionally, the definition of F_π implies that $k \notin C$ if $d_k \in R_L$ or $d_k \in B_S$. Hence, for $x \in R_L \cup B_S$, it holds that $L_x - C_x = L_x \ge \frac{L_x}{n} - 1$ (where the latter inequality is of course a strong estimate, but it is all we need).

Summing up the inequalities for all agents, we obtain

$$P = L - |C| = \sum_{x \in N} L_x - C_x \ge \sum_{x \in N} \frac{L_x}{n} - 1 = \frac{L}{n} - n.$$

Solving for L yields the desired inequality. $\lhd$

It remains to analyze the dynamics $(\rho_p)_{p=1}^P$. To this end, we will show that this dynamics essentially behaves like a specific AHG, where we have to replace some agents by multiple copies. This yields an AHG with at most n^2 agents. Hence, Theorem 3.5 would provide a running time of $O(n^6)$. However, we can do better. By exploiting structural properties of the AHG and a close inspection of the potentials in the proof of Theorem 3.5, we can reduce the running time of our transformed dynamics on the AHG to $O(n^4)$.

Claim 4.6. *It holds that $P \in O(n^4)$.*

Proof. First, let us note that, by construction of $(\rho_p)_{p=0}^P$, $F_{\rho_p} = \emptyset$ for all $0 \le p \le P$. Hence, all agents in R_S (respectively, B_L) only perform deviations towards singletons or coalitions of ratio at most $\frac{1}{2}$ (respectively, at least $\frac{1}{2}$). Moreover, we may assume that an agent $r \in R_S$ never forms a coalition of size 2 with an agent in $b \in B_L$. Due to single-peakedness and their respective peaks, this can only happen if both of them come from singleton coalitions. However, then no further agent can join $\{r, b\}$. Further red agents would be blocked by r and further blue agents by b. Hence, this coalition can only be altered if one of these agents leaves. But this deviation can be performed right away from the singleton coalition. Similarly, we can exclude the formation of coalitions of size 2 by agents in R_L and B_S.

As this shortcutting can only remove every second step (and an initial $n/2$ steps for forming a first set of pairs), it leaves us with a dynamics $(\rho'_p)_{p=0}^{P'}$ with $P' \ge \frac{P - n/2}{2}$ such that, for $1 \le p \le P'$, ρ'_p evolves from ρ'_{p-1} through an IS-deviation of some agent d'_p.

Even more, we may assume that agents in R_S (respectively, B_L) never perform deviations. First, according to Claim 4.1, the only coalition that such an agent can leave is a coalition of size 2 with ratio $\frac{1}{2}$. Hence, single-peakedness implies that forming the singleton coalition is not beneficial. Furthermore, Claim 4.1 implies that they could only form coalitions of size 2. By the first part of the proof, their partner has to be from B_S (respectively, R_L). Since the preferences are strict, we may assume that their partner performs the deviation.

Now, we define an AHG $(N^A, (\succ_x^A)_{x \in N^A})$ as follows. In principle, the only part of the preferences are on ratios $\frac{1}{m+1}$ or $\frac{m}{m+1}$, and we want to identify these ratios with coalitions of size $m+1$, because all coalitions of these ratios have exactly this size (using Claim 4.1). This part of the preferences will also inherit single-peakedness from the HDG. However, we have to deal with the preference over 1 for agents in R_L (respectively, over 0 for agents in B_S). To maintain single-peakedness, we should identify these ratios with coalition sizes $|R| + 1$ and $|B| + 1$, respectively. To achieve this goal, we introduce some auxiliary agents. Let the agent set of the AHG therefore be $N^A = R_S \cup B_L \cup \{r_0, \dots, r_{|R|}\colon r \in R_L\} \cup \{b_0, \dots, b_{|B|}\colon b \in B_S\}$ and define strict and single-peaked preferences as follows (where we present only the relevant part of the preferences).

Draft – April 19, 2022

- If $r \in R_S$ and $2 \leq i, j \leq |B| + 1$, then
  - $i \succ_r^A j$ if and only if $\frac{1}{i} \succ_r \frac{1}{j}$,
  - $1 \succ_r^A i$ if and only if $1 \succ_r \frac{1}{i}$, and
  - $i \succ_r^A 1$ if and only if $\frac{1}{i} \succ_r 1$.
- If $b \in B_L$ and $2 \leq i, j \leq |R| + 1$, then
  - $i \succ_b^A j$ if and only if $\frac{i-1}{i} \succ_b \frac{j-1}{j}$,
  - $1 \succ_b^A i$ if and only if $0 \succ_b \frac{i-1}{i}$, and
  - $i \succ_b^A 1$ if and only if $\frac{i-1}{i} \succ_b 0$.
- If $r \in R_L$ and $2 \leq i, j \leq |R|$, then
  - $i \succ_{r_0}^A j$ if and only if $\frac{i-1}{i} \succ_r \frac{j-1}{j}$,
  - $|R| + 1 \succ_{r_0}^A i$ if and only if $1 \succ_r \frac{i-1}{i}$,
  - $i \succ_{r_0}^A |R| + 1$ if and only if $\frac{i-1}{i} \succ_r 1$, and
  - $|R| + 1 \succ_{r_l}^A |R|$ if $l \in [|R|]$.
- If $b \in B_S$ and $2 \leq i, j \leq |B|$, then
  - $i \succ_{b_0}^A j$ if and only if $\frac{1}{i} \succ_b \frac{1}{j}$,
  - $|B| + 1 \succ_{b_0}^A i$ if and only if $0 \succ_b \frac{1}{i}$,
  - $i \succ_{b_0}^A |B| + 1$ if and only if $\frac{1}{i} \succ_b 0$, and
  - $|B| + 1 \succ_{b_l}^A |B|$ if $l \in [|B|]$.

Next, we define the modified dynamics. Therefore, let $0 \leq p \leq P'$ and define the partition $\omega_p = \{\{r_0, \dots, r_{|R|}\} \colon r \in R_L, \{r\} \in \rho'_p\} \cup \{\{r_1, \dots, r_{|R|}\} \colon r \in R_L, \{r\} \notin \rho'_p\} \cup \{\{b_0, \dots, b_{|B|}\} \colon b \in B_S, \{b\} \in \rho'_p\} \cup \{\{b_1, \dots, b_{|B|}\} \colon b \in B_S, \{b\} \notin \rho'_p\} \cup \{\{b, r_0^1, \dots, r_0^m\} \colon b \in B_S, \rho'_p(b) = \{b, r^1, \dots, r^m\}$ for $m \geq 0\} \cup \{\{r, b_0^1, \dots, b_0^m\} \colon r \in R_L, \rho'_p(r) = \{r, b^1, \dots, b^m\}$ for $m \geq 0\}$.

Note that ω_p is well-defined, because every agent in R_L (respectively, B_S), which is not in a singleton coalition is part of a coalition solely consisting of agents in R_L (respectively, B_S), and a unique agent in B_L (respectively, R_S).

Next, let $1 \leq p \leq P'$. Then, ω_p evolves from ω_{p-1} through an IS-deviation of agent d'_0. This follows directly from the preferences in the AHG, where a fraction of 1 (respectively, 0) plays the role of the coalition size $|R| + 1$ (respectively, $|B| + 1$) for agents in R_L (respectively, B_S). Hence, $(\omega_p)_{p=0}^{P'}$ is an execution of an individual dynamics in AHG $(N^A, (\succ_x^A)_{x \in N^A})$.

To bound its running time, we have to inspect the potentials in the proof of Theorem 3.5. First, $v_j^{P'} \leq n$ for all agents $j \in N^A$. Second, $v_C^{P'} \leq n$ for all $C \in \omega_{P'}$, and $|\omega_{P'}| \leq 2n$. The latter bounds hold, because the copies of every original agent are only part of at most 2 coalitions. Hence, $\Lambda(\omega_{P'}) \leq n^3 + 2n^2$, and there can be at most that many R-moves. Moreover, $\Lambda(\omega_k) - \Lambda(\omega_{k-1}) \leq n - 1$ for every R-move. Hence, as in the proof of Theorem 3.5, we obtain a bound of $n^4 + 2n^3$ L-moves. Hence, the dynamics on the AHG runs for $P' \in O(n^4)$ steps. Therefore, as $P \leq 2P' + \frac{n}{2}$, we obtain $P \in O(n^4)$. $\lhd$

Finally, we can combine Claim 4.5 and Claim 4.6 to obtain

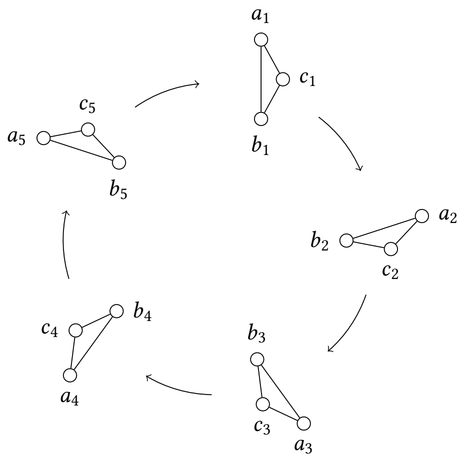
$$K \leq L \leq n^2 + n(L - |C|) = n^2 + nP \in O(n^5).$$

$\square$

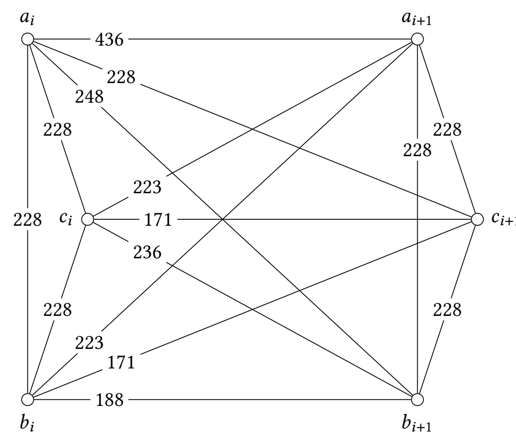
Under strict preferences, checking the existence of a path to stability and convergence are hard.

THEOREM 4.5. *$\exists$-IS-SEQUENCE-HDG is* NP*-hard and $\forall$-IS-SEQUENCE-HDG is* co-NP*-hard, even for strict preferences.*

Draft – April 19, 2022



(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.

Fig. 3. Description of the graph associated with the constructed symmetric FHG without IS partition.

5 FRACTIONAL HEDONIC GAMES

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 [10]. Here, we provide a counterexample using 15 agents. The weights were found with the help of a computer.

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

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 by

- $v(a_i, b_i) = v(b_i, c_i) = v(a_i, c_i) = 228, i \in \{1, \dots, 5\}$,
- $v(a_i, a_{i+1}) = 436, v(a_i, b_{i+1}) = 228, v(a_i, c_{i+1}) = 248, i \in \{1, \dots, 5\}$,
- $v(b_i, a_{i+1}) = 223, v(b_i, b_{i+1}) = 171, v(b_i, c_{i+1}) = 236, i \in \{1, \dots, 5\}$,
- $v(c_i, a_{i+1}) = 223, v(c_i, b_{i+1}) = 171, v(c_i, c_{i+1}) = 188, i \in \{1, \dots, 5\}$, and
- $v(x, y) = -2251$ for all agents $x, y \in N$ such that the weight is not defined yet.

In the above definition, all indices are to be read modulo 5 (where the modulo function is assumed to map to $\{1, \dots, 5\}$). Note that the large negative weight exceeds the sum of positive weights incident to any agents. Hence, agents linked by a negative weight, can never be in a common coalition in any IS partition. The FHG consists of five triangles that form a cycle. The structure of the game is illustrated in Figure 3. There is an infinite loop 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 join this new coalition, then c_1. After this step, we are in an isomorphic state as in the initial partition.

Let π be any partition of the agents and assume that π is IS. In particular, no agent receives negative utility. Therefore there exists an $i \in \{1, \dots, 5\}$ such that $\pi(a_i) \cap \{a_1, \dots, a_5\} = \{a_i\}$. We may assume, w.l.o.g., that a_1 is such an agent. In the following, we will exclude all possible coalitions for agent a_1, deriving a contradiction.

Draft – April 19, 2022

- Goal 1: $c_2 \notin \pi(a_1)$.
  First, assume for contradiction that $c_2 \in \pi(a_1)$. If $\pi(a_2) \subseteq \{a_2, b_2, b_1, c_1\}$, then $v_{a_2}(\pi) \leq 168.5$ while $v_{a_2}(\{a_2\} \cup \pi(a_1)) \geq 221$ and a_2 has an incentive to deviate (making no agent in $\pi(a_1)$ worse). Hence, $\pi(a_2) \subseteq \{a_2, b_2, a_3, b_3, c_3\}$. In addition, by the same potential deviation, $a_3 \in \pi(a_2)$ and $|\pi(a_2)| \geq 3$.
  Next, consider the case that $b_2 \in \pi(a_2)$. Then, $v_{c_2}(\pi) \leq 142.5$, while $v_{c_2}(\{c_2\} \cup \pi(a_2)) \geq 169.75$ and c_2 would deviate. Hence, $b_2 \notin \pi(a_2)$. If $b_2 \notin \pi(a_1)$, then $v_{b_2}(\pi) \leq \max\{141.34, 119.67\} = 141.34$ (this is if b_2 forms a coalition with $N_1 \backslash \{a_1\}$ and $N_3 \backslash \{a_3\}$, respectively), while $v_{b_2}(\{b_2\} \cup \pi(a_1)) \geq 158.6$ and it is easily seen that b_2 can only improve agents in $\pi(a_1)$. It follows that $b_2 \in \pi(a_1)$.
  If $b_3 \notin \pi(a_2)$, then $\pi(a_2) = \{a_2, a_3, c_3\}$ and b_3 would deviate by joining $\pi(a_2)$. Hence, $\{a_2, a_3, b_3\} \subseteq \pi(a_2)$. But then $v_{c_2}(\pi) \leq 159.6$ while $v_{c_2}(\{c_2\} \cup \pi(a_2)) \geq 171.6$ and joining with c_2 makes no agent worse. In conclusion, the initial assumption was wrong and $c_2 \notin \pi(a_1)$.
- Goal 2: $b_2 \notin \pi(a_1)$.
  Second, assume for contradiction that $b_2 \in \pi(a_1)$. As in the previous case, it is easily seen that $\pi(a_2) \subseteq \{a_2, b_2\} \cup N_3$, $a_3 \in \pi(a_2)$, and $|\pi(a_2)| \geq 3$. If $c_2 \notin \pi(a_2)$, then $v_{c_2}(\pi) \leq 118$, while $v_{c_2}(\{c_2\} \cup \pi(a_2)) \geq 155.5$. Hence, $c_2 \in \pi(a_2)$. But then $v_{b_2}(\pi) \leq 168$ while $v_{b_2}(\{b_2\} \cup \pi(a_2)) \geq 169.75$ and b_2 would join $\pi(a_2)$ making no agent worse. We conclude that $b_2 \notin \pi(a_1)$ and can therefore assume that $\pi(a_1) \subseteq N_1 \cup N_5$.
- Goal 3: $b_1 \notin \pi(a_1)$.
  Third, assume for contradiction that $b_1 \in \pi(a_1)$. Then, $v_{a_5}(\pi) \leq \max\{223, 171\} = 223$ (where the first utility in the maximum refers to the coalition $N_4 \cup N_5$ and the second utility to $N_5 \cup \{c_1\}$). However, $v_{a_5}(\{a_5\} \cup \pi(a_1)) \geq 228$. Since joining $\pi(a_1)$ with a_5 makes no agent worse, this is not possible. Hence, $b_1 \notin \pi(a_1)$.
- Goal 4: $c_1 \notin \pi(a_1)$.
  Forth, assume for contradiction that $c_1 \in \pi(a_1)$. Then, $v_{b_1}(\{b_1\} \cup \pi(a_1)) \geq 152$ and adding b_1 to $\pi(a_1)$ leaves no agent worse off. Since, $v_{b_1}(\{b_1\} \cup N_2) = 145.5$, it must hold that $\pi(b_1) \subseteq \{b_1\} \cup N_5$ and even $\{a_5, c_5\} \subseteq \pi(b_1)$ since otherwise $v_{b_1}(\pi) \leq 145.4$. But then $v_{a_1}(\pi) \leq 150.4$ while $v_{a_1}(\{a_1\} \cup \pi(b_1)) \geq 221.75$ and a_1 would deviate making no agent worse. It follows that $c_1 \notin \pi(a_1)$.
- Goal 5: $\pi(a_1) \nsubseteq \{a_1, b_5, c_5\}$.
  It remains the case that $\pi(a_1) \subseteq \{a_1, b_5, c_5\}$. If $|\pi(b_1)| \geq 2$, then a_1 would deviate by joining $\pi(b_1)$, making no agent worse. If, however, b_1 is in a singleton coalition, then b_1 would join $\pi(a_1)$, making no agent worse and improving her utility.

□

Employing this counterexample, the methods of Brandl et al. [10], which originate from hardness constructions of Sung and Dimitrov [23], can be used to show that it is NP-hard to decide about the existence of IS partitions in symmetric FHGs.

For the following corollary, we only give a brief proof sketch, because the main method will also be applied in the proof of Theorem C.1, which considers convergence of the IS dynamics in the case that the FHGs even have non-negative weights.

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

Sketch of proof. In the reduction by Brandl et al. [10, Theorem 5], we replace the non-symmetric gadget by a vertex-minimal symmetric FHG that admits no IS partition. Such an FHG

Draft – April 19, 2022

exists according to Theorem 5.1. The weights in the symmetric part of their reduced instances must be large enough to incentivize the agent in the gadget to stay in a coalition outside the gadget. □

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. *∃-IS-Sequence-FHG is* NP*-hard and ∀-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 converges in simple symmetric FHGs in at most $O(n^2)$ steps. The dynamics may take $\Omega(n\sqrt{n})$ steps.*

Proof. We start with the lower bound. Consider the FHG induced by the complete graph on $n = k(k+1)/2$ agents for some non-negative integer $k \geq 1$. We partition the agents arbitrarily into sets $C_1, \dots, C_k$ where $|C_j| = j$ for $j = 1, \dots, k$. Now, we perform two phases of IS-deviations. In the first phase, we form the coalitions C_j by having agents join one by one. In the second phase, there are $k-1$ steps. In step j, the agents of coalition C_j join coalitions $C_{j+1}, \dots, C_k$ performing $k-j$ deviations each. The total number of deviations in the second phase is therefore $\sum_{j=1}^{k-1} k \cdot (k-j) = \frac{1}{6}(k-1)k(k+1) = \Theta(k^3) = \Theta(k^2\sqrt{k^2}) = \Theta(n\sqrt{n})$. In particular, there can be $\Omega(n\sqrt{n})$ IS-deviation steps starting from the singleton partition.

For the upper bound, let an FHG be given. 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 bounded by $n(n-1)/2$, 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 $O(n^4)$ steps.*

Proof. Let $G = (V, A)$ be an asymmetric graph on $n = |V|$ vertices. If the graph contains a cycle, it is easy to find a non-converging series of deviations. There exists a cycle of length at least 3. Let an edge coalition propagate along the cycle.

Assume that the graph is acyclic. Our first observation is that, in every step of the dynamics, the subgraphs induced by coalitions are always transitive tournaments, i.e., linear orders (on their vertices). Indeed, by induction, in a deviation, the coalition that is left still induces a transitive tournament and the new coalition induced a transitive tournament before the deviation. Hence, every agent except one has at least one outgoing edge and will only accept the new agent if she likes her. Since the deviating agent must have non-negative utility after the deviation, she needs to approve the single agent without outgoing edge. Hence, the newly formed coalition still induces a transitive tournament.

We will now define two potentials based on the agents that receive 0 utility in a partition, and based on the coalition sizes. The first potential is monotonically decreasing and bounded. The second potential is strictly increasing whenever the first potential is not strictly decreasing, and bounded. Hence, we establish convergence of the dynamics.

First, fix a topological order of the agents, i.e., a bijection $\sigma : V \longrightarrow [n]$ such that for all $(v, w) \in A$, $\sigma(v) < \sigma(w)$. For a given partition π of the agents, we define the vector $v^\sigma(\pi)$ of length $|\pi|$ that sorts the numbers $\max_{i\in C} \sigma(i)$ for $C \in \pi$ in decreasing order, that is it sorts the coalitions in decreasing topological score of the agent with the highest number due to the topological order. This is exactly the unique agent in every coalition receiving 0 utility. In addition, we define the vector $w(\pi)$ of length $|\pi|$ that sorts the coalition sizes in increasing order. Note that this vector does not depend on the underlying topological order.

For two vectors $v = (v_i)_{i=1}^k$ and $w = (w_i)_{i=1}^l$, not necessarily of the same length, we say

$$v >_{lex} w \iff \text{there is } i < \max\{k, l\} \text{ with } v_j = w_j\ \forall 1 \le j \le i \text{ and } v_{i+1} > w_{i+1}, \text{ or } k > l \text{ and } v_j = w_j\ \forall 1 \le j \le k$$

In other words, $v >_{lex} w$ if v is lexicographically greater than w.

The key insight is that, for π' formed from π by an IS deviation, $v^\sigma(\pi') <_{lex} v^\sigma(\pi)$, or $v^\sigma(\pi') =_{lex} v^\sigma(\pi)$ and $w(\pi') >_{lex} w(\pi)$. For a proof, assume that π' is formed from π by an IS deviation of agent i. Note that $\max_{j\in\pi'(i)} \sigma(j) = \max_{j\in\pi'(i)\setminus\{i\}} \sigma(j)$. We distinguish two cases. Either $i = \arg\max_{j\in\pi(i)} \sigma(j)$ and it follows $v^\sigma(\pi') <_{lex} v^\sigma(\pi)$. Otherwise, $v_i(\pi(i)) \ge \frac{1}{|\pi(i)|}$, and because i is improving her utility, $\frac{1}{|\pi'(i)|} = v_i(\pi') > \frac{1}{|\pi(i)|}$. It follows that $|\pi(i)| > |\pi'(i)|$. Hence, $|\pi'(i)| - 1 < \min\{|\pi'(i)|, |\pi(i)|\}$, and therefore $w(\pi') >_{lex} w(\pi)$.

We estimate the running time in two steps. First, we bound the number of times that the lexicographic score of $v^\sigma(\pi)$ can decrease. Then, we estimate the number of deviations that can happen while this score does not change. We call the first kind of deviations *primal* and the second type *secondary*.

The idea to bound the number of primal deviations is to associate with every agent $v \in V$ a set D_v that stores a certain amount of deviating agents, so that at every step in the algorithm $\sum_{v\in V} |D_v|$ is the number of primal deviations so far. We ensure that we can always add the agent v performing a deviation to a set D_w such that $\sigma(v) > \sigma(x)$ for all $x \in D_w$. Hence, at the end of the sequence of deviations, $\sum_{v\in V} |D_v| \le n^2$.

Initially, set $D_v^{\pi_0} = \emptyset$ for all $v \in V$ and the starting partition π_0 of the dynamics. Assume first that agent v performs a primary deviation that changes partition π into partition π'. If v was in a singleton coalition, update $D_v^{\pi'} = \{v\}$ and leave all other sets the same, i.e., $D_x^{\pi'} = D_x^\pi$ for all $x \neq v$. Otherwise, let $w = \arg\max_{x\in\pi(v)\setminus\{v\}} \sigma(x)$ be the agent in $\pi(v)$ different from v of highest topological score, i.e., the agent in $\pi(v)$ of second-highest topological score. We update $D_v^{\pi'} = D_w^\pi \cup \{v\}$, $D_w^{\pi'} = \emptyset$, and $D_x^{\pi'} = D_x^\pi$ for all $x \neq v, w$. If a secondary deviation is performed from π to π', leave all sets the same, i.e., $D_x^{\pi'} = D_x^\pi$ for all $x \in V$.

Given a set of agents $W \subseteq V$, let $m_W = \arg\max_{x \in W} \sigma(x)$ be the agent in W maximizing the topological score. We have the following invariants for every partition π during the heuristics and for every agent $v \in V$:

- If $v = m_{\pi(v)}$, then $D_v^\pi = \emptyset$.
- If $v \neq m_{\pi(v)}$, then $\sigma(x) < \sigma(m_{\pi(v)})$ for all $x \in D_v^\pi$.
- The number of primal deviations of the dynamics until partition π is $\sum_{v \in V} |D_v^\pi|$

The first and third invariants follow directly from the update rules, where we use that the agent newly added to a set has not been in this set due to the second invariant. The second invariant follows by induction, because if v performs a primal deviation from π to π', then for all $x \in D_v^{\pi'}$, $\sigma(x) \leq \sigma(v) < \sigma(m_{\pi'(v)})$, where the first inequality follows by induction and the second by the fact that the agent in $\pi'(v)$ which gives positive utility to v has a higher topological score than v. Hence, there can be at most n^2 primal deviations, because for the terminal partition π^* of the dynamics, $\sum_{v \in V} |D_v^{\pi^*}| \leq n^2$.

While the topological score is the same, there can be at most n^2 secondary deviations, which follows from the same reasoning as in the proof of Proposition 5.4. Hence, together there are at most n^4 deviations. □

In the previous proposition, it seems that there is still space for improvement of the bound on the running time, in particular due to the interplay of the two nested potentials.

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

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 [20], as well as convergence of the dynamics of IS-deviations, when starting from the grand coalition [8]. 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.

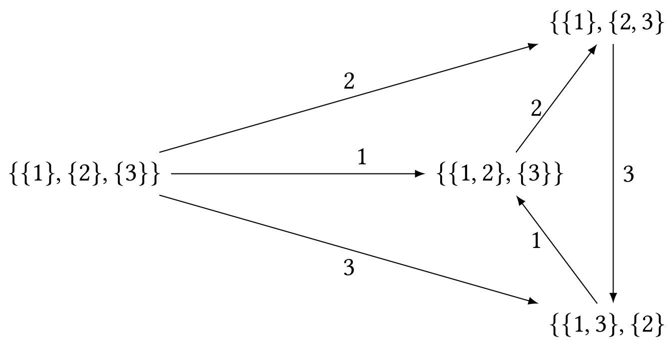
Proposition 6.1. *The dynamics of IS-deviations may never reach an IS partition in DHGs, whatever the chosen path of deviations, even when starting from the singleton partition.*

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

Draft – April 19, 2022

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: $\{1,2\} \succ_1 \{1,3\} \succ_3 \{2,3\} \succ_2 \{1,2\}$. 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 [15], for proving the existence of a core-stable partition in hedonic games with common ranking property.

Also 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. Many of our results are negative, that is, examples of cycles in dynamics or even non-existence of IS partitions under strong preference restrictions. In particular, we have answered a number of open problems proposed in the literature leading to boundaries of dynamics. For all hedonic games under study, it turned out that the existence of cycles for IS-deviations is

Draft – April 19, 2022

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 ∃ (resp., ∀) denotes that problem ∃-IS-Sequence-HG (resp., ∀-IS-Sequence-HG) is NP-hard (resp., co-NP-hard).

Class	Convergence	Hardness
AHGs	✓ natural SP (single-peaked) (Th. 3.4) ✓ neutral (Suksompong [22]) ∘ strict & general SP; singletons / grand coalition (Prop. 3.2)	∃ strict (Th. 3.3) ∀ strict (Th. 3.3)
HDGs	✓ strict & natural SP; singletons; solitary homogeneity (Th. 4.4) ∘ any three of: strict, natural SP, singletons, and solitary homogeneity (Th. 4.2)	∃ strict (Th. 4.5) ∀ strict (Th. 4.5)
FHGs	✓ simple & sym.; singletons (Prop. 5.4) ✓ acyclic digraph (Th. 5.5) ∘ symmetric (Th. 5.1)	∃ symmetric (Th. 5.3) ∃ simple; singletons (Th. 5.7) ∃ asymmetric (Th. 5.6) ∀ symmetric (Th. 5.3) ∀ asymmetric (Th. 5.6)
DHGs	✓ grand coalition (Boehmer and Elkind [8]) ✓ common ranking property or symmetric (Caskurlu and Kizilkaya [15]) ∘ singletons (Prop. 6.1)	∃ singletons (Th. 6.2) ∀ general (Th. 6.2)

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. On the other hand, we have identified natural conditions for convergence that are based on *i*) initial conditions, that is, the starting partition, *ii*) selection rules for the performed deviation, and *iii*) preference restrictions such as a common scale for the agents (e.g., the common ranking property), single-peakedness, or symmetry. An overview of our results can be found in Table 1. In particular, we have made sophisticated use of potential functions to show the polynomial running time of the dynamics for anonymous hedonic games and hedonic diversity games with restrictions at the boundary of convergence.

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 some interesting preference restrictions, some of which do not guarantee convergence. Following our work, the most intriguing cases are simple symmetric FHGs with arbitrary initial partitions and HDGs under single-peaked preferences.

Second, one could investigate specific selection rules for the performed deviations. With the exception of Theorem 4.5, we do not have to pose any assumptions on the performed deviations. However, selecting appropriate deviations may lead to a quick termination of the dynamics 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).

Third, one can define dynamics also based on other stability concepts like Nash stability or contractual individual stability. For the latter, cycling is not possible, and therefore an analysis within the complexity class PLS as local search algorithms is a natural approach that measures the complexity of convergence. Such an analysis is also appropriate for dynamics guaranteed

Draft – April 19, 2022

to converge based on a potential function argument as it was already done for Nash stability in additively separable hedonic games [9, 18].

Finally, the final states reached in the dynamics we consider do not provide information beyond individual stability. One could therefore additionally aim to reach efficient outcomes, potentially measured amongst stable outcomes. The notion of Pareto efficiency seems natural here, because it gives also rise to a natural improvement dynamics. Note that the compatibility of fairness and strategyproofness is already problematic for simple instances of hedonic games that are often symmetric amongst the agents [9].

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft under grant BR 2312/12-1. We would like to thank Abheek Ghosh for sharing insights on how to extend our convergence result for AHGs to weak preferences. We also thank Leo Tappe and the anonymous reviewers for helpful discussions. A preliminary version of this article appeared in the Proceedings of the 35th AAAI Conference on Artificial Intelligence (February, 2021). Results from this article were presented at the 3rd Games, Agents, and Incentives Workshop (May, 2021) and the 15th Journées d'Intelligence Artificielle Fondamentale (July, 2021).

REFERENCES

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

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

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

[4] H. Aziz and R. Savani. 2016. Hedonic Games. In *Handbook of Computational Social Choice*, F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia (Eds.). Cambridge University Press, Chapter 15.

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

[6] P. Berman, M. Karpinski, and A. Scott. 2003. *Approximation hardness of short symmetric instances of MAX-3SAT*. Technical Report TR03-049. Electronic Colloquium on Computational Complexity (ECCC).

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

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

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

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

[11] F. Brandt, M. Bullinger, and A. Wilczynski. 2021. Reaching Individually Stable Coalition Structures in Hedonic Games. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*. 5211–5218.

[12] F. Brandt and A. Wilczynski. 2019. On the Convergence of Swap Dynamics to Pareto-Optimal Matchings. In *Proceedings of the 15th International Conference on Web and Internet Economics (WINE) (Lecture Notes in Computer Science (LNCS))*. Springer-Verlag, 100–113.

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

[14] R. Carosi, G. Monaco, and L. Moscardelli. 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.

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

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

[17] J. Farrell and S. Scotchmer. 1988. Partnerships. *Quarterly Journal of Economics* 103 (1988), 279–297.

[18] M. Gairing and R. Savani. 2019. Computing Stable Outcomes in Symmetric Additively Separable Hedonic Games. *Mathematics of Operations Research* 44, 3 (2019), 1101–1121.

Draft – April 19, 2022

[19] M. Hoefer, D. Vaz, and L. Wagner. 2018. Dynamics in matching and coalition formation games with structural constraints. *Artificial Intelligence* 262 (2018), 222–247.

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

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

[22] W. Suksompong. 2015. Individual and Group Stability in Neutral Restrictions of Hedonic Games. *Mathematical Social Sciences* 78 (2015), 1–5.

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

APPENDIX: OMITTED PROOFS

In the appendix, we provide the proofs omitted in the main part of the paper.

A ANONYMOUS HEDONIC GAMES

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

We prove the two hardness results by providing separate reductions for each problem in the next two lemmas.

Lemma A.1. *$\exists$-IS-Sequence-AHG is* NP*-hard even for strict preferences.*

Proof. Let us perform a reduction from (3,B2)-SAT, a variant of the Satisfiability problem known to be NP-complete [6]. In an instance of (3,B2)-SAT, we are given a CNF propositional formula φ where every clause C_j, for $1 \leq j \leq m$, contains exactly three literals and every variable x_i, for $1 \leq i \leq p$, appears exactly twice as a positive literal and twice as a negative literal. From such an instance, we construct an instance of an anonymous game with initial partition as follows.

For each ℓ^{th} occurrence ($\ell \in \{1, 2\}$) of a positive literal x_i (resp., negative literal $\overline{x}_i$), we create a literal-agent y_i^ℓ (resp., $\overline{y}_i^\ell$). All literal-agents are singletons in the initial partition π_0. Let us consider four integers α, β^+, β^- and γ such that (1) $q \cdot \alpha + x \neq r \cdot \beta^+ + y \neq s \cdot \beta^- + z \neq t\gamma + w$ for every $r, s, t \in [n]$, $q \in [m]$, $x, y, z \in \{0, 1, 2\}$ and $w \in [7]$ and, w.l.o.g., $\alpha > \beta^+ > \beta^- > \gamma > 1$. For instance, we can set the following values: $\alpha = m^5$, $\beta^+ = m^4$, $\beta^- = m^3$ and $\gamma = m^2$ (condition (1) is satisfied since in a (3,B2)-SAT instance, it holds that $m \geq 4$ and $p = 3/4m$). For each clause C_j, we create $j\alpha$ dummy clause-agents who are all grouped within the same coalition K_j in the initial partition π_0. For each literal x_i (resp., $\overline{x}_i$), we create one variable-agent z_i (resp., $\overline{z}_i$) and $i\beta^+ - 1$ (resp., $i\beta^- - 1$) dummy variable agents who are all grouped within the same coalition Z_i (resp., $\overline{Z_i}$) in the initial partition π_0. Finally, for each variable x_i, we create $i\gamma$ dummy agents who are all grouped within the same coalition G_i^1 in the initial partition π_0, $i\gamma + 3$ dummy agents who are all grouped within the same coalition G_i^2 in the initial partition π_0 and $i\gamma + 5$ dummy agents who are all grouped within the same coalition G_i^3 in the initial partition π_0. These dummy agents are used as a gadget for a cycle. Although we have created many agents, the construction remains polynomial by considering reasonable values of α, β^+, β^- and γ, as previously described.

The preferences of the agents over sizes of coalitions are given in Table 2.

We claim that there exists a sequence of IS-deviations which leads to an IS partition iff formula φ is satisfiable.

Suppose first that there exists a truth assignment of the variables ϕ that satisfies all the clauses. Let us denote by ℓ_j a chosen literal-agent associated with an occurrence of a literal true in ϕ which belongs to clause C_j. Since all the clauses of φ are satisfied by ϕ, there exists such a literal-agent ℓ_j for each clause C_j. For every clause C_j, let literal-agent ℓ_j join coalition K_j. These IS-deviations make all the dummy clause-agents and the chosen literal-agents the most happy as possible, therefore

Draft – April 19, 2022

Table 2. Preferences of the agents in the reduced instance of Lemma A.1, for every $1 \leq i \leq p$, $1 \leq j \leq m$, $\ell \in \{1, 2\}$. Notation $cl(x_i^\ell)$ (resp., $cl(\overline{x}_i^\ell)$) stands for the index of the clause to which the ℓ^{th} occurrence of literal x_i (resp., $\overline{x}_i$) belongs, the framed value is the size of the initial coalition in partition π_0, and $[\dots]$ denotes an arbitrary order over the rest of the coalition sizes.

z_i :	$\lvert Z_i\rvert + 2 \succ \lvert G_i^1\rvert + 2 \succ \lvert G_i^2\rvert + 1 \succ \lvert G_i^3\rvert + 2 \succ \lvert G_i^3\rvert + 1 \succ \lvert G_i^1\rvert + 1 \succ \lvert Z_i\rvert + 1 \succ \boxed{\lvert Z_i\rvert} \succ [\dots]$
$\overline{z}_i$:	$\lvert \overline{Z_i}\rvert + 2 \succ \lvert G_i^3\rvert + 2 \succ \lvert G_i^2\rvert + 1 \succ \lvert G_i^1\rvert + 2 \succ \lvert G_i^1\rvert + 1 \succ \lvert G_i^3\rvert + 1 \succ \lvert \overline{Z_i}\rvert + 1 \succ \boxed{\lvert \overline{Z_i}\rvert} \succ [\dots]$
y_i^ℓ :	$\lvert K_{cl(x_i^\ell)}\rvert + 1 \succ \lvert Z_i\rvert + 2 \succ \lvert Z_i\rvert + 1 \succ \boxed{1} \succ [\dots]$
$\overline{y}_i^\ell$:	$\lvert K_{cl(\overline{x}_i^\ell)}\rvert + 1 \succ \lvert \overline{Z_i}\rvert + 2 \succ \lvert \overline{Z_i}\rvert + 1 \succ \boxed{1} \succ [\dots]$
K_j :	$\lvert K_j\rvert + 1 \succ \boxed{\lvert K_j\rvert} \succ [\dots]$
$Z_i \setminus \{z_i\}$:	$\lvert Z_i\rvert + 2 \succ \lvert Z_i\rvert + 1 \succ \boxed{\lvert Z_i\rvert} \succ \lvert Z_i\rvert - 1 \succ [\dots]$
$\overline{Z_i} \setminus \{\overline{z}_i\}$:	$\lvert \overline{Z_i}\rvert + 2 \succ \lvert \overline{Z_i}\rvert + 1 \succ \boxed{\lvert \overline{Z_i}\rvert} \succ \lvert \overline{Z_i}\rvert - 1 \succ [\dots]$
G_i^1 :	$\lvert G_i^1\rvert + 2 \succ \lvert G_i^1\rvert + 1 \succ \boxed{\lvert G_i^1\rvert} \succ [\dots]$
G_i^2 :	$\lvert G_i^2\rvert + 1 \succ \boxed{\lvert G_i^2\rvert} \succ [\dots]$
G_i^3 :	$\lvert G_i^3\rvert + 2 \succ \lvert G_i^3\rvert + 1 \succ \boxed{\lvert G_i^3\rvert} \succ [\dots]$

none of them will deviate afterwards. Then, let all remaining literal-agents y_i^ℓ (resp., $\overline{y}_i^\ell$) deviate by joining coalition Z_i (resp., $\overline{Z_i}$). Since ϕ is a truth assignment of the variables, for each variable x_i, there exists a coalition Z_i or $\overline{Z_i}$ that is joined by two literal-agents and thus reaches size the most preferred size $|Z_i| + 2$ or $|\overline{Z_i}| + 2$. For each variable, if coalition Z_i (resp., $\overline{Z_i}$) is not joined by two literal-agents, then it cannot be true for $\overline{Z_i}$ (resp., Z_i), and variable-agent z_i (resp., $\overline{z}_i$) then deviates for joining coalition G_i^2, and if one literal-agent previously joined coalition Z_i (resp., $\overline{Z_i}$) she deviates to be alone. No agent can then move in a IS-deviation because variable-agents in the gadget prefer sizes of coalitions which differ by at least two from the size of the current coalitions and no agent prefers a coalition with an intermediate size. This also holds for literal-agents and dummy agents. Therefore, the current partition is IS.

Suppose now that there exists no truth assignment of the variables that satisfies all the clauses. That means that it is not possible that one literal-agent joins each clause coalition while two literal-agents y_i^1 and y_i^2 join coalition Z_i or $\overline{y}_i^1$ and $\overline{y}_i^2$ join coalition $\overline{Z_i}$ for each variable x_i. Moreover, since each literal-agent prefers to join clause coalitions than variable coalitions, it means that in a maximal sequence of IS-deviations, all dummy clause-agents in each coalition K_j will be completely satisfied with a coalition size equal to $|K_j|+1$. However, in such a case, there exists a variable x_i such that at most one literal-agent joins coalition Z_i and at most one joins coalition $\overline{Z_i}$. It follows that both variable-agents z_i and $\overline{z_i}$ have an incentive to deviate to the gadget associated with variable x_i (their respective most preferred coalition sizes $|Z_i| + 2$ and $|\overline{Z_i}| + 2$ can never be reached). Within the gadget associated with variable x_i, variable-agents z_i and $\overline{z_i}$ are the only agents who can deviate and we necessarily reach the cycle illustrated in Figure 4.

It follows that no sequence of IS-deviations can reach an IS partition. □

Lemma A.2. *∀-IS-Sequence-AHG is* co-NP*-hard even for strict preferences.*

Proof. For this purpose, we prove the NP-hardness of the complement problem, which asks whether there exists a cycle in IS-deviations. Let us perform a reduction from (3,B2)-SAT [6]. In

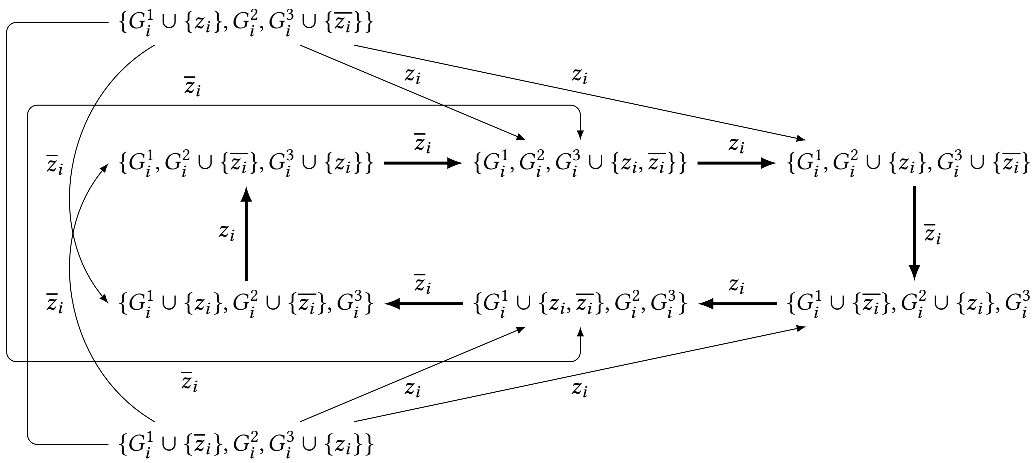


Fig. 4. Necessary cycle of IS-deviations within the gadget associated with variable x_i in the reduced instance of Lemma A.1.

an instance of (3,B2)-SAT, we are given a CNF propositional formula φ where every clause C_j, for $1 \leq j \leq m$, contains exactly three literals and every variable x_i, for $1 \leq i \leq p$, appears exactly twice as a positive literal and twice as a negative literal. From such an instance, we construct an instance of an anonymous game with initial partition as follows.

For each ℓ^{th} occurrence ($\ell \in \{1, 2\}$) of a positive literal x_i (resp., negative literal $\overline{x}_i$), we create a literal-agent y_i^ℓ (resp., $\overline{y}_i^\ell$). We create another agent t. All these agents are singletons in the initial partition π_0. Let us consider five integers α, β_1^+, β_1^-, β_2^+ and β_2^- such that (1) $q \cdot \alpha + x \neq r \cdot \beta_1^+ + y \neq s \cdot \beta_1^- + z \neq t \cdot \beta_2^+ + v \neq u \cdot \beta_2^- + w$ for every $r, s, t, u \in [n]$, $q \in [m]$ and $x, y, z, v, w \in \{0, 1, 2\}$ and, w.l.o.g., $\alpha > \beta_1^+ > \beta_1^- > \beta_2^+ > \beta_2^- > 1$. For instance, we can set the following values: $\alpha = m^5$, $\beta_1^+ = m^4$, $\beta_1^- = m^3$, $\beta_2^+ = m^2$, $\beta_2^- = m$ (condition (1) is satisfied since in a (3,B2)-SAT instance, it holds that $m \geq 4$ and $p = 3/4m$). For each clause C_j, we then create $j \cdot \alpha$ dummy clause agents grouped within the same coalition K_j in the initial partition π_0. We also create $(m + 1) \cdot \alpha$ dummy agents grouped within the same coalition K_{m+1} in initial partition π_0. Finally, for each literal x_i (resp., $\overline{x}_i$) and each $\ell \in \{1, 2\}$, we create $i \cdot \beta_\ell^+$ (resp., $i \cdot \beta_\ell^-$) dummy variable agents grouped within the same coalition Y_i^ℓ (resp., $\overline{Y_i^\ell}$) in the initial partition π_0. Although we have created many agents, the construction remains polynomial by considering reasonable values of α, β_1^+, β_1^-, β_2^+ and β_2^-, as previously described.

The preferences of the agents over sizes of coalitions are given in Table 3.

We claim that there exists a cycle of IS-deviations iff formula φ is satisfiable[4].

Suppose first that formula φ is satisfiable by a truth assignment of the variables denoted by ϕ. Let us denote by ℓ_j a chosen literal-agent associated with an occurrence of a literal true in ϕ which belongs to clause C_j. Since all the clauses of φ are satisfied by ϕ, there exists such a literal-agent ℓ_j for each clause C_j. Now let us denote by z_i^1 and z_i^2 the literal-agents associated with the two occurrences of the literal of variable x_i which is false in ϕ. In the same vein, let us denote by Z_i^1 and Z_i^2 the coalitions of dummy variable agents associated with z_i^1 and z_i^2, respectively. Since ϕ is a truth assignment of the variables, z_i^1, z_i^2, Z_i^1 and Z_i^2 all correspond to the same literal (either x_i

[4]Note that the singleton partition is nevertheless always IS.

Draft – April 19, 2022

Table 3. Preferences of the agents in the reduced instance of Lemma A.2, for every $1 \leq i \leq p$, $1 \leq i' < p$, $1 \leq j \leq m+1$, $\ell \in \{1,2\}$. Notation $cl(x_i^\ell)$ (resp., $cl(\overline{x}_i^\ell)$) stands for the index of the clause to which the ℓ^{th} occurrence of literal x_i (resp., $\overline{x}_i$) belongs, the framed value is the size of the initial coalition in partition π_0, and $[\dots]$ denotes an arbitrary order over the rest of the coalition sizes.

y_i^1 :	$\lvert K_{cl(x_i^1)}\rvert + 2 \succ \lvert K_{cl(x_i^1)+1}\rvert + 2 \succ \lvert K_{cl(x_i^1)+1}\rvert + 1 \succ \lvert K_{cl(x_i^1)}\rvert + 1 \succ \lvert Y_i^1\rvert + 2 \succ \lvert Y_i^2\rvert + 2 \succ \lvert Y_i^2\rvert + 1 \succ \lvert Y_i^1\rvert + 1 \succ \boxed{1} \succ [\dots]$
$y_{i'}^2$:	$\lvert K_{cl(x_{i'}^2)}\rvert + 2 \succ \lvert K_{cl(x_{i'}^2)+1}\rvert + 2 \succ \lvert K_{cl(x_{i'}^2)+1}\rvert + 1 \succ \lvert K_{cl(x_{i'}^2)}\rvert + 1 \succ \lvert Y_{i'}^2\rvert + 2 \succ \lvert Y_{i'+1}^1\rvert + 2 \succ \lvert Y_{i'+1}^1\rvert + 1 \succ \lvert \overline{Y_{i'+1}^1}\rvert + 2 \succ \lvert \overline{Y_{i'+1}^1}\rvert + 1 \succ \lvert Y_{i'}^2\rvert + 1 \succ \boxed{1} \succ [\dots]$
y_p^2 :	$\lvert K_{cl(x_p^2)}\rvert + 2 \succ \lvert K_{cl(x_p^2)+1}\rvert + 2 \succ \lvert K_{cl(x_p^2)+1}\rvert + 1 \succ \lvert K_{cl(x_p^2)}\rvert + 1 \succ \lvert Y_p^2\rvert + 2 \succ \lvert K_1\rvert + 2 \succ \lvert K_1\rvert + 1 \succ \lvert Y_p^2\rvert + 1 \succ \boxed{1} \succ [\dots]$
$\overline{y}_i^1$:	$\lvert K_{cl(\overline{x}_i^1)}\rvert + 2 \succ \lvert K_{cl(\overline{x}_i^1)+1}\rvert + 2 \succ \lvert K_{cl(\overline{x}_i^1)+1}\rvert + 1 \succ \lvert K_{cl(\overline{x}_i^1)}\rvert + 1 \succ \lvert \overline{Y_i^1}\rvert + 2 \succ \lvert \overline{Y_i^2}\rvert + 2 \succ \lvert \overline{Y_i^2}\rvert + 1 \succ \lvert \overline{Y_i^1}\rvert + 1 \succ \boxed{1} \succ [\dots]$
$\overline{y}_{i'}^2$:	$\lvert K_{cl(\overline{x}_{i'}^2)}\rvert + 2 \succ \lvert K_{cl(\overline{x}_{i'}^2)+1}\rvert + 2 \succ \lvert K_{cl(\overline{x}_{i'}^2)+1}\rvert + 1 \succ \lvert K_{cl(\overline{x}_{i'}^2)}\rvert + 1 \succ \lvert \overline{Y_{i'}^2}\rvert + 2 \succ \lvert Y_{i'+1}^1\rvert + 2 \succ \lvert Y_{i'+1}^1\rvert + 1 \succ \lvert \overline{Y_{i'+1}^1}\rvert + 2 \succ \lvert \overline{Y_{i'+1}^1}\rvert + 1 \succ \lvert \overline{Y_{i'}^2}\rvert + 1 \succ \boxed{1} \succ [\dots]$
$\overline{y}_p^2$:	$\lvert K_{cl(\overline{x}_p^2)}\rvert + 2 \succ \lvert K_{cl(\overline{x}_p^2)+1}\rvert + 2 \succ \lvert K_{cl(\overline{x}_p^2)+1}\rvert + 1 \succ \lvert K_{cl(\overline{x}_p^2)}\rvert + 1 \succ \lvert \overline{Y_p^2}\rvert + 2 \succ \lvert K_1\rvert + 2 \succ \lvert K_1\rvert + 1 \succ \lvert \overline{Y_p^2}\rvert + 1 \succ \boxed{1} \succ [\dots]$
t :	$\lvert K_{m+1}\rvert + 2 \succ \lvert Y_1^1\rvert + 2 \succ \lvert Y_1^1\rvert + 1 \succ \lvert \overline{Y_1^1}\rvert + 2 \succ \lvert \overline{Y_1^1}\rvert + 1 \succ \lvert K_{m+1}\rvert + 1 \succ \boxed{1} \succ [\dots]$
K_j :	$\lvert K_j\rvert + 2 \succ \lvert K_j\rvert + 1 \succ \boxed{\lvert K_j\rvert} \succ 1 \succ [\dots]$
Y_i^ℓ :	$\lvert Y_i^\ell\rvert + 2 \succ \lvert Y_i^\ell\rvert + 1 \succ \boxed{\lvert Y_i^\ell\rvert} \succ 1 \succ [\dots]$
$\overline{Y_i^\ell}$:	$\lvert \overline{Y_i^\ell}\rvert + 2 \succ \lvert \overline{Y_i^\ell}\rvert + 1 \succ \boxed{\lvert \overline{Y_i^\ell}\rvert} \succ 1 \succ [\dots]$

or $\overline{x}_i$) and it holds that $\bigcup_{1\leq j\leq m} \ell_j \cap \bigcup_{1\leq i\leq p}\{z_i^1, z_i^2\} = \emptyset$. We will construct a cycle in IS-deviations involving, as deviating agents, the literal-agents ℓ_j, for every $1 \leq j \leq m$, the literal-agents z_i^1 and z_i^2, for every $1 \leq i \leq p$, and agent t. Since m is even in a (3,B2)-SAT (recall that $m = 4/3p$), there is an odd number of deviating agents in total.

First of all, let agent z_p^2 and then agent ℓ_1 join coalition K_1. For each $1 < i \leq p$, let agent z_{i-1}^2 and then agent z_i^1 join coalition Z_i^1. Let agent t and then agent z_1^1 join coalition Z_1^1. For each even j such that $3 < j \leq m$, let agent ℓ_{j-1} and then agent ℓ_j join coalition K_j. Finally, let agent ℓ_2 join coalition K_3. The reached partition is $\pi := \{K_1 \cup \{\ell_1, z_p^2\}, K_2, K_3 \cup \{\ell_2\}, K_4 \cup \{\ell_3, \ell_4\}, K_5, K_6 \cup \{\ell_5, \ell_6\}, K_7, \dots, K_m \cup \{\ell_{m-1}, \ell_m\}, K_{m+1}, Z_1^1 \cup \{t, z_1^1\}, Z_1^2, Z_2^1 \cup \{z_1^2, z_2^1\}, Z_2^2, \dots, Z_p^1 \cup \{z_{p-1}^2, z_p^1\}, Z_p^2, \overline{Z_1^1}, \overline{Z_1^2}, \dots, \overline{Z_p^1}, \overline{Z_p^2}\}$, where coalition $\overline{Z_i^\ell}$ refers to $\overline{Y_i^\ell}$ if $Z_i^\ell = Y_i^\ell$ and to Y_i^ℓ if $Z_i^\ell = \overline{Y_i^\ell}$. Partition π is the first step of the cycle.

From partition π, let literal-agent ℓ_3 deviate from current coalition $K_4 \cup \{\ell_3, \ell_4\}$ to join coalition $K_3 \cup \{\ell_2\}$. This deviation makes literal-agent ℓ_4 worse off, who then deviates to join coalition K_5. Then, the same deviations occur for literal-agents ℓ_5 and ℓ_6, and so on. More generally, for every odd j such that $3 \leq j \leq m$ by increasing order of indices, literal-agent ℓ_j leaves coalition $K_{j+1} \cup \{\ell_j, \ell_{j+1}\}$ to join coalition $K_j \cup \{\ell_{j-1}\}$ and then literal-agent ℓ_{j+1}, who is worse off by this deviation, deviates to join coalition K_{j+2}. After that, agent t deviates from coalition $Z_1^1 \cup \{t, z_1^1\}$ to join coalition $K_{m+1} \cup \{\ell_m\}$, which makes literal-agent z_1^1 worse off. Then, for each $1 \leq i \leq p$ by increasing order of indices, let literal-agent z_i^1 deviate from coalition $Z_i^1 \cup \{z_i^1\}$ to join coalition Z_i^2 and then literal-agent z_i^2 deviate from coalition $Z_{i+1}^1 \cup \{z_i^2, z_{i+1}^1\}$ (or $K_1 \cup \{z_i^2, \ell_1\}$ if $i = p$) to join coalition $Z_i^2 \cup \{z_i^1\}$, which makes literal-agent z_{i+1}^1 (or ℓ_1 if $i = p$) worse off. We thus reach partition $\pi_1 := \{K_1 \cup \{\ell_1\}, K_2, K_3 \cup \{\ell_2, \ell_3\}, K_4, K_5 \cup \{\ell_4, \ell_5\}, K_6, K_7 \cup \{\ell_6, \ell_7\}, \dots, K_{m-1} \cup \{\ell_{m-2}, \ell_{m-1}\}, K_m, K_{m+1} \cup \{\ell_m, t\}, Z_1^1, Z_1^2 \cup \{z_1^1, z_1^2\}, Z_2^1, Z_2^2 \cup \{z_2^1, z_2^2\}, \dots, Z_p^1, Z_p^2 \cup \{z_p^1, z_p^2\}, \overline{Z_1^1}, \overline{Z_1^2}, \dots, \overline{Z_p^1}, \overline{Z_p^2}\}$.

Then, from partition π_1, for every odd j such that $1 \leq j < m$ by increasing order of indices, let literal-agent ℓ_j deviate from coalition $K_j \cup \{\ell_j\}$ to join coalition K_{j+1} and then literal-agent ℓ_{j+1} deviate from $K_{j+2} \cup \{\ell_{j+1}, \ell_{j+2}\}$ (or $K_{m+1} \cup \{\ell_m, t\}$ if $j = m-1$) to join coalition $K_{j+1} \cup \{\ell_j\}$, which makes literal-agent ℓ_{j+2} (or agent t if $j = m-1$) worse off. Let agent t then deviate from

Draft – April 19, 2022

coalition $K_{m+1} \cup \{t\}$ to join coalition Z_1^1 and literal-agent z_1^1 deviate from coalition $Z_1^2 \cup \{z_1^1, z_1^2\}$ to join coalition $Z_1^1 \cup \{t\}$, which makes literal-agent z_1^2 worse off. Then, for each $1 \leq i < p$ by increasing order of indices, let literal-agent z_i^2 deviate from coalition $Z_i^2 \cup \{z_i^2\}$ to join coalition Z_{i+1}^1, and literal-agent z_{i+1}^1 deviate from coalition $Z_{i+1}^2 \cup \{z_{i+1}^1, z_{i+1}^2\}$ to join coalition $Z_{i+1}^1 \cup \{z_i^2\}$, which makes literal-agent z_{i+1}^2 worse off. And then, let literal-agent z_p^2 deviate from coalition $Z_p^2 \cup \{z_p^2\}$ to join coalition K_1, and literal-agent ℓ_1 deviate from coalition $K_2 \cup \{\ell_1, \ell_2\}$ to join coalition $K_1 \cup \{z_p^2\}$, which makes literal-agent ℓ_2 worse off. Finally, let literal-agent ℓ_2 deviate from coalition $K_2 \cup \{\ell_2\}$ to join coalition K_3, and we have finally reached partition π, leading to a cycle.

Suppose now that there exists a cycle of IS-deviations. Observe first that no dummy agent can deviate. Indeed, the only coalition sizes that are preferred by a dummy agent to the size of her initial coalition are the size of the current coalition plus one and the size of the current coalition plus two. These sizes cannot be reached by joining other coalitions by condition (1), and the fact that the other coalitions do not want to integrate more than two additional agents in their coalition. Therefore, the only possible deviations are when the dummy agents let at most two agents join their coalition. It follows that no agent can belong to a coalition whose size is ranked after size 1 in her preferences, i.e., we do not have to care about the preferences within [. . .] in the preference ranking of the agents. Indeed, a literal-agent or agent t (for the sake of simplicity we also talk about agent t when referring to literal-agents since the behavior is similar), who starts from a coalition of size one, can join some coalitions of dummy agents and sometimes accepts one additional agent in such coalitions. So the worst thing which can happen to these deviating literal-agents is that the other literal-agent in the coalition leaves the coalition but in such a case, they are still in a coalition whose size is ranked before one in their preferences. Observe that literal-agents can only join coalitions of dummy agents with at most one other literal-agent in this coalition and that no literal-agent can join another literal-agent outside a coalition of dummy agents, because size two is not ranked before size one in the preferences of the agents (recall that a coalition of dummy agents cannot be of size smaller than 2). Moreover, since the literal-agents can never be in a coalition of size less preferred than one, once a literal-agent deviates from her initial coalition where she is alone, she has no incentive to come back to the coalition where she is alone. Hence, the deviations in the cycle must be performed by literal-agents who change of coalition of dummy agents.

Each deviating literal-agent i in the cycle must be left at some step otherwise, there would be no reason for her to come back to a previous coalition. Since she can be left only by one other literal-agent, it follows that the current coalition of agent i was of size $|K| + 2$ for a given coalition K of dummy agents and becomes of size $|K| + 1$. To be able to come back to a previous coalition, agent i must prefer $|K| + 2$ over $|K| + 1$ and there must be intermediate sizes in the preference ranking of agent i between $|K| + 2$ and $|K| + 1$. Therefore, if agent i is an agent y_i^ℓ (resp., $\overline{y}_i^\ell$), then agent i must be left by another literal-agent from the coalition of dummy agents (a) $K_{cl(x_i^\ell)}$ (resp., $K_{cl(\overline{x}_i^\ell)}$), i.e., the coalition associated with the clause to which the occurrence of the literal of the literal-agent belongs, or (b) Y_i^ℓ (resp., $\overline{Y_i^\ell}$), i.e., the coalition associated with the occurrence of the literal of the literal-agent.

(a) If this is the coalition of dummy agents $K_{cl(x_i^\ell)}$ (resp., $K_{cl(\overline{x}_i^\ell)}$), then the other literal-agent who leaves the coalition must not be associated with an occurrence of a literal belonging to this clause otherwise, it would be her most preferred size and she would not have incentive to leave the coalition. Therefore, according to the preferences of the literal-agents, the only possibility is that this other literal-agent who leaves the coalition $K_j := K_{cl(x_i^\ell)}$ (resp., $K_j := K_{cl(\overline{x}_i^\ell)}$) is associated with an occurrence of a literal belonging to C_{j-1} if $j > 1$ or is agent y_p^2 or $\overline{y}_p^2$ if $j = 1$. Due to the preferences of the literal-agents, if a literal-agent leaves such a coalition, it is necessarily for joining

Draft – April 19, 2022

the coalition of dummy agents K_{j-1} (which has an additional literal-agent) if $j > 1$ or Y_p^2 or $\overline{Y}_p^2$ (with an additional literal-agent) if $j = 1$. In the latter case ($j = 1$), this is the only possibility even if the associated deviating agent y_p^2 (resp., $\overline{y}_p^2$) prefers several coalition sizes over $|K_1| + 2$, because the other choices would prevent her to come back to size $|K_1| + 2$: the worst thing which can occur after some steps if she chooses other coalitions is that she is in a coalition of size $|K_{cl(x_p^2)}| + 1$ or $|K_{cl(x_p^2)+1}| + 1$ (resp., $|K_{cl(\overline{x}_p^2)}| + 1$ or $|K_{cl(\overline{x}_p^2)+1}| + 1$) which are both preferred to $|K_1| + 2$, contradicting the cycle.
$(b1)$ Otherwise, if this is the coalition of dummy agents Y_i^1 or $\overline{Y_i^1}$, then the only possibility is that the other literal-agent who leaves the coalition is literal-agent y_{i-1}^2 or $\overline{y}_{i-1}^2$ if $i > 1$, or agent t if $i = 1$. If $i > 1$, as mentioned earlier, literal-agent y_{i-1}^2 (resp., $\overline{y}_{i-1}^2$) cannot deviate to coalitions of dummy clause agents, otherwise she would never come back to the current coalition size. If we talk about coalition Y_i^1, then the only possibility is that literal-agent y_{i-1}^2 (resp., $\overline{y}_{i-1}^2$) deviates for joining coalition Y_{i-1}^2 (resp., $\overline{Y_{i-1}^2}$) (and an additional literal-agent). Otherwise, i.e., if we talk about $\overline{Y_i^1}$, then literal-agent y_{i-1}^2 (resp., $\overline{y}_{i-1}^2$) deviates to join Y_{i-1}^2 (resp., $\overline{Y_{i-1}^2}$) (and an additional literal-agent) but she could also deviate for joining coalition of dummy variable agents Y_i^1. But, in such a case, she would be joined or would join literal-agent y_i^1 who has no reason to deviate (to join dummy variable coalitions), therefore literal-agent y_{i-1}^2 (resp., $\overline{y}_{i-1}^2$) would still need to deviate to join coalition Y_{i-1}^2 (resp., $\overline{Y_{i-1}^2}$) (and an additional literal-agent). If $i = 1$, the same arguments can be applied and then agent t deviates to join coalition of clause agents K_{m+1}.
$(b2)$ If this is the coalition of dummy agents Y_i^2 (resp., $\overline{Y_i^2}$), then the only possibility is that the other literal-agent who leaves the coalition is literal-agent y_i^1 (resp., $\overline{y}_i^1$). Since, literal-agent y_i^1 (resp., $\overline{y}_i^1$) cannot deviate to join a coalition of dummy clause agents, she will necessarily join the coalition of dummy variable agents Y_i^1 (resp., $\overline{Y_i^1}$) (with an additional literal-agent).

To summarize, if there is a cycle, only the following can occur:

- agent t in coalition K_{m+1} can only be left by a literal-agent corresponding to an occurrence of a literal belonging to clause C_m who deviates to join coalition of dummy clause agents K_m;
- a literal-agent y_i^ℓ (resp., $\overline{y}_i^\ell$) in a coalition of dummy clause agents K_j (for $1 < j \leq m$), corresponding to the clause C_j to which the ℓ^{th} occurrence of x_i (resp., $\overline{x}_i$) belongs, can only be left by a literal-agent corresponding to an occurrence of a literal belonging to clause C_{j-1} who deviates to join coalition of dummy clause agents K_{j-1};
- a literal-agent y_i^ℓ (resp., $\overline{y}_i^\ell$) in coalition of dummy clause agents K_1, corresponding to the clause C_1 to which the ℓ^{th} occurrence of x_i (resp., $\overline{x}_i$) belongs, can only be left by literal-agent y_p^2 (or $\overline{y}_p^2$) who joins coalition of dummy variable agents Y_p^2 (or $\overline{Y_p^2}$);
- a literal-agent y_i^2 (resp., $\overline{y}_i^2$), for $1 \leq i \leq p$, in a coalition of dummy variable agents Y_i^2 (resp., $\overline{Y_i^2}$) can only be left by literal-agent y_i^1 (resp., $\overline{y}_i^1$) who joins coalition of dummy variable agents Y_i^1 (resp., $\overline{Y}_i^1$);
- a literal-agent y_i^1 (resp., $\overline{y}_i^1$), for $1 < i \leq p$, in a coalition of dummy variable agents Y_i^1 (resp., $\overline{Y_i^1}$) can only be left by a literal-agent y_{i-1}^2 or $\overline{y}_{i-1}^2$ who joins coalition of dummy variable agents Y_{i-1}^2 or $\overline{Y}_{i-1}^2$;
- literal-agent y_1^1 (resp., $\overline{y}_1^1$) in coalition of dummy variable agents Y_1^1 (resp., $\overline{Y_1^1}$) can only be left by agent t who joins coalition of dummy clause agents K_{m+1};

Draft – April 19, 2022

It follows that, for the cycle to occur, we need, as deviating agents, for each clause, a literal-agent corresponding to an occurrence of a literal belonging to this clause who alternates between coalitions of dummy clause agents and, for each variable, two literal-agents corresponding to the same literal, who alternate between coalitions of dummy variable agents. Note that a literal-agent belonging to the cycle can only alternate between coalitions of dummy clause agents or between coalitions of dummy variable agents, but not both. Hence, by setting to true the literals associated with the literal-agents alternating between coalitions of dummy clause agents in the cycle, we get a valid truth assignment of the variables which satisfies all the clauses. ◻

B HEDONIC DIVERSITY GAMES

In this section, we provide the missing proofs for hedonic diversity games. First, we consider the remaining two examples of cycling under strong assumptions.

Theorem 4.2. *The dynamics of IS-deviations may cycle in HDGs even if any three of the following restrictions apply:*

(1) preferences are naturally single-peaked,
(2) preferences are strict,
(3) the starting partition is the singleton partition, or
(4) all deviations satisfy solitary homogeneity.

Proof. ($\neg$1) Our next example satisfies all properties except single-peakedness of preferences. Let us consider an HDG with 12 agents: 3 red agents and 9 blue agents. There are two deviating agents in the cycle: blue agents b_1 and b_2. In the cycle, there are three fixed coalitions C_1, C_2 and C_3 such that:

- C_1 contains 1 red agent and 2 blue agents;
- C_2 contains 1 red agent and 1 blue agent;
- C_3 contains 1 red agent and 4 blue agents.

The relevant part of the preferences of the agents is given below.

$$
\begin{array}{ll}
b_1: & \frac{1}{5} \succ \frac{1}{3} \succ \frac{1}{7} \succ \frac{1}{6} \succ \frac{1}{4} \succ 0 \\
b_2: & \frac{1}{7} \succ \frac{1}{3} \succ \frac{1}{5} \succ \frac{1}{4} \succ \frac{1}{6} \succ 0 \\
C_1: & \frac{1}{5} \succ \frac{1}{4} \succ \frac{1}{3} \succ \frac{1}{2} \succ [1 \text{ if red, } 0 \text{ otherwise}] \\
C_2: & \frac{1}{3} \succ \frac{1}{2} \succ [1 \text{ if red, } 0 \text{ otherwise}] \\
C_3: & \frac{1}{7} \succ \frac{1}{6} \succ \frac{1}{5} \succ \frac{1}{4} \succ \frac{1}{3} \succ \frac{1}{2} \succ [1 \text{ if red, } 0 \text{ otherwise}]
\end{array}
$$

Consider the sequence of IS-deviations in Figure 5 that describe a cycle in the dynamics. The two deviating agents of the cycle b_1 and b_2 are marked in bold and the specific deviating agent between two states is indicated next to the arrows.

To show that this cycle can be reached from the singleton partition, it suffices to observe that the two deviating agents b_1 and b_2 prefer to join the fixed coalitions than being alone and that each fixed coalition can be formed from the singleton partition: the red agent of each future fixed coalition joins first a blue agent and then all the other blue agents of the future fixed coalition successively join. Note that all deviations result in non-homogeneous target coalitions, and therefore satisfy solitary homogeneity.

($\neg$2) Our final example satisfies all properties except strictness of preferences. Let us consider an HDG with 10 agents: 4 red agents and 6 blue agents. There are two deviating agents: red agent r and blue agent b, and three fixed coalitions C_1, C_2 and C_3 such that:

- C_1 contains 2 red agents;
- C_2 contains 1 red agent and 3 blue agents;

Draft – April 19, 2022

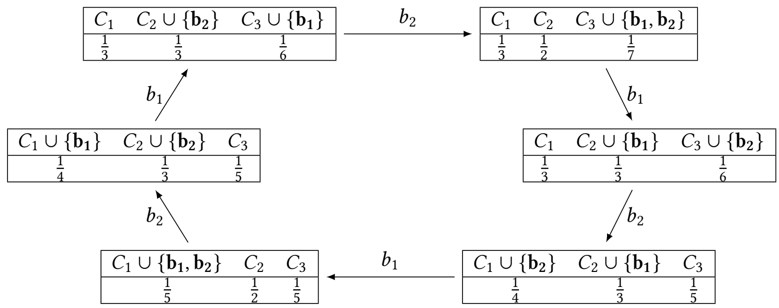


Fig. 5. Possibility of cycling of IS-dynamics starting from the singleton partition in HDG under strict preferences. All deviations satisfy solitary homogeneity.

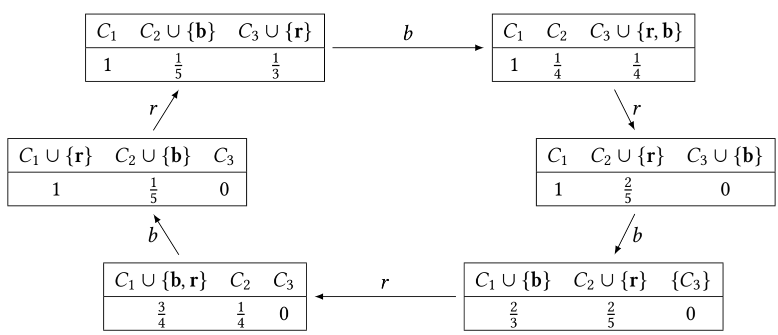


Fig. 6. Possibility of cycling of IS-dynamics starting from the singleton partition in HDG under strict preferences. All deviations satisfy solitary homogeneity.

- C_3 contains 2 blue agents.

The relevant part of the preferences of the agents is given below.

$$\begin{array}{rl}
r: & \frac{3}{4} \succ \frac{2}{5} \succ \frac{1}{4} \sim \frac{1}{3} \succ 1 \\
b: & \frac{1}{4} \succ \frac{1}{5} \succ \frac{1}{2} \sim \frac{2}{3} \sim \frac{3}{4} \succ 0 \\
C_1: & \frac{3}{4} \succ \frac{2}{3} \succ \frac{1}{2} \sim \frac{1}{3} \succ 1 \\
C_2: & \frac{2}{5} \succ \frac{1}{3} \sim \frac{1}{4} \sim \frac{1}{5} \succ \frac{1}{2} \succ [1 \text{ if red, } 0 \text{ otherwise}] \\
C_3: & \frac{1}{4} \succ \frac{1}{3} \succ \frac{1}{2} \succ 0
\end{array}$$

Consider the sequence of IS-deviations in Figure 6 that describe a cycle in the dynamics. The two deviating agents of the cycle r and b are marked in bold and the specific deviating agent between two states is indicated next to the arrows.

To show that this cycle can be reached from the singleton partition, it suffices to observe that partition $\{C_1 \cup \{b\}, C_2 \cup \{r\}, C_3\}$ belonging to the cycle can be reached from the singleton partition. Indeed, agent b can join a red agent from the future fixed coalition C_1 while the other red agent of the future fixed coalition C_1 can join a blue agent from the future fixed coalition C_3. The second blue agent of the future fixed coalition C_3 then joins them and afterwards, the red agent leaves them to join b and the other red agent of C_1. For forming coalition C_2, the red agent joins one of the blue agents, and then the two remaining blue

Draft – April 19, 2022

agents join them. Agent r can then join coalition C_2. Note that all deviations result in non-homogeneous target coalitions, and therefore satisfy solitary homogeneity.

□

Next, we provide a proof for the lemma used for the part of the proof of the previous theorem given in the body of the paper.

LEMMA 4.1. *Given a set R_a of red (respectively, set B_a of blue) agents whose preferences satisfy $\frac{2}{3} \succ 1 \succ \frac{1}{2}$ (respectively, $\frac{1}{3} \succ 0 \succ \frac{1}{2}$), it is possible to create the homogeneous coalition R_a (respectively, B_a) by means of an individual dynamics starting from singleton coalitions.*

PROOF. In the following proof, we consider various sets of auxiliary agents. We assume that we take *new* agents in every step of the constructions.

We show the statement for homogeneous coalitions of blue agents. The statement for red agents is completely symmetric, by reversing the respective roles. We use a few types of auxiliary agents with extreme preferences that have their peaks at the largest or smallest ratios, except for liking homogeneous coalitions the worst. Specifically, we consider four sets R_x, B_x, R_y, and B_y of agents with the following preferences. Note that the sets R_y and B_y are only needed for the statement about red agents, but we state them for completeness.

$$R_x, B_x: \quad f \succ g \text{ if and only if } 0 < f < g \text{ or } f > g = 0$$

$$R_y, B_y: \quad f \succ g \text{ if and only if } 1 > f > g \text{ or } f < g = 1$$

Now, let B_a be a set of blue agents such that every agent in B_a has preferences satisfying $\frac{1}{3} \succ 0 \succ \frac{1}{2}$. Suppose that $|B_a| = k$ and $B_a = \{b_{a,i} : i \in [k]\}$. Let $r_{x,i} \in R_x$ for $i \in [2]$ and $b_{x,i} \in B_x$ for $i \in [k+2]$. We create a *trash coalition* (used to get rid of agents not needed anymore) by creating $C_a = \{r_{x,2}, b_{x,k}, b_{x,k+1}\}$. To create C_a, we let the blue agents of this coalition join the red agent $r_{x,2}$.

Now, we perform for any $i \in [k-1]$ the following steps (in increasing order of indices): $b_{x,i}$ joins $r_{x,1}$, then $b_{a,i}$ joins this coalition. Then, $b_{x,i}$ leaves this coalition to join $C_a \cup \{b_{x,j} : j \in [i-1]\}$, and finally $b_{a,i}$ joins $b_{a,k} \cup \{b_{a,j} : j \in [i-1]\}$. Note that these are all IS-deviations. In particular, $b_{a,i}$'s deviation to join $b_{a,4}$'s coalition is feasible, because she leaves coalition $\{r_{x,1}, b_{a,i}\}$ which has a ratio of $1/2$, which is strictly worse for her than being in a homogeneous coalition. After this procedure, we have obtained the coalition B_a. □

THEOREM 4.5. *∃-IS-SEQUENCE-HDG is NP-hard and ∀-IS-SEQUENCE-HDG is co-NP-hard, even for strict preferences.*

We prove the two hardness results by providing separate reductions for each problem in the next two lemmas.

LEMMA B.1. *∃-IS-SEQUENCE-HDG is NP-hard even for strict preferences.*

PROOF. Let us perform a reduction from (3,B2)-SAT [6]. In an instance of (3,B2)-SAT, we are given a CNF propositional formula φ where every clause C_j, for $1 \le j \le m$, contains exactly three literals and every variable x_i, for $1 \le i \le p$, appears exactly twice as a positive literal and twice as a negative literal. From such an instance, we construct an instance of a hedonic diversity game with initial partition as follows. The proof works in the same way as the proof of Lemma A.1 except that we have to ensure appropriate ratios of red agents in each coalition.

For each ℓ^{th} occurrence ($\ell \in \{1,2\}$) of a positive literal x_i (resp., negative literal $\overline{x}_i$), we create a red literal-agent y_i^ℓ (resp., a blue literal-agent $\overline{y}_i^\ell$). All literal-agents are singletons in the initial partition π_0. Let us consider three integers α, β and γ such that (1) $\alpha > 2m-1$, $\beta > \max\{3p-2; 3p\alpha + 3p - 2\}$, $\gamma > \max\{12p-2; 6p\beta + 12p - 1\}$. For instance, we can set the following values: $\alpha = m^2$, $\beta = m^4$ and $\gamma = m^7$ (one can verify that condition (1) is satisfied, especially because in a (3,B2)-SAT instance,

Draft – April 19, 2022

Table 4. Preferences of the agents in the reduced instance of Lemma B.1, for every $1 \leq i \leq p$, $1 \leq j \leq m$, $\ell \in \{1, 2\}$. Notation $cl(x_i^\ell)$ (resp., $cl(\overline{x}_i^\ell)$) stands for the index of the clause to which the ℓ^{th} occurrence of literal x_i (resp., $\overline{x}_i$) belongs, the framed value corresponds to the ratio of the initial coalition in partition π_0, and $[\dots]$ denotes an arbitrary order over the rest of the coalition ratios.

$$
\begin{array}{rl}
z_i: & \frac{3i}{\beta+2} \succ \frac{6(p-i)+2}{i\gamma+2} \succ \frac{6(p-i)+4}{i\gamma+1} \succ \frac{6(p-i)+6}{i\gamma+2} \succ \frac{6(p-i)+6}{i\gamma+1} \succ \frac{6(p-i)+2}{i\gamma+1} \succ \frac{3i-1}{\beta+1} \succ \boxed{\frac{3i-2}{\beta}} \succ [\dots] \\
\overline{z}_i: & \frac{3i-2}{\beta+2} \succ \frac{6(p-i)+6}{i\gamma+2} \succ \frac{6(p-i)+3}{i\gamma+1} \succ \frac{6(p-i)+2}{i\gamma+2} \succ \frac{6(p-i)+1}{i\gamma+1} \succ \frac{6(p-i)+5}{i\gamma+1} \succ \frac{3i-2}{\beta+1} \succ \boxed{\frac{3i-2}{\beta}} \succ [\dots] \\
y_i^\ell: & \frac{2cl(x_i^\ell)}{\alpha+1} \succ \frac{3i}{\beta+2} \succ \frac{3i-1}{\beta+1} \succ \boxed{1} \succ [\dots] \\
\overline{y}_i^\ell: & \frac{2cl(\overline{x}_i^\ell)-1}{\alpha+1} \succ \frac{3i-2}{\beta+2} \succ \frac{3i-2}{\beta+1} \succ \boxed{0} \succ [\dots] \\
\hline
K_j: & \frac{2j}{\alpha+1} \succ \frac{2j-1}{\alpha+1} \succ \boxed{\frac{2j-1}{\alpha}} \succ [\dots] \\
Z_i \setminus \{z_i\}: & \frac{3i}{\beta+2} \succ \frac{3i-1}{\beta+1} \succ \boxed{\frac{3i-2}{\beta}} \succ \frac{3i-3}{\beta-1} \succ [\dots] \\
\overline{Z_i} \setminus \{\overline{z}_i\}: & \frac{3i-2}{\beta+2} \succ \frac{3i-2}{\beta+1} \succ \boxed{\frac{3i-2}{\beta}} \succ \frac{3i-2}{\beta-1} \succ [\dots] \\
G_i^1: & \frac{6(p-i)+2}{i\gamma+2} \succ \frac{6(p-i)+2}{i\gamma+1} \succ \frac{6(p-i)+1}{i\gamma+1} \succ \boxed{\frac{6(p-i)+1}{i\gamma}} \succ [\dots] \\
G_i^2: & \frac{6(p-i)+4}{i\gamma+1} \succ \frac{6(p-i)+3}{i\gamma+1} \succ \boxed{\frac{6(p-i)+3}{i\gamma}} \succ [\dots] \\
G_i^3: & \frac{6(p-i)+6}{i\gamma+2} \succ \frac{6(p-i)+6}{i\gamma+1} \succ \frac{6(p-i)+5}{i\gamma+1} \succ \boxed{\frac{6(p-i)+5}{i\gamma}} \succ [\dots]
\end{array}
$$

it holds that $m \geq 4$ and $p = 3/4m$). For each clause C_j, we then create α dummy clause-agents with among them $2j - 1$ red agents. They are all grouped within the same coalition K_j in the initial partition π_0. For each literal x_i (resp., $\overline{x}_i$), we create a red variable-agent z_i (resp., a blue variable-agent $\overline{z}_i$) and $\beta - 1$ (resp., $\beta - 1$) dummy variable-agents with among them $3i - 2$ (resp., $3i - 2$) red agents (z_i included). They are all grouped within the same coalition Z_i (resp., $\overline{Z_i}$) in the initial partition π_0. Finally, for each variable x_i, we create three coalitions in partition π_0 of dummy agents G_i^1, G_i^2 and G_i^3 of size $i\gamma$ with among them, $6(p-i)+1$, $6(p-i)+3$ or $6(p-i)+5$ red agents, for each coalition respectively. These dummy agents are used as a gadget for a cycle. Although we have created many agents, the construction remains polynomial by considering reasonable values of α, β and γ, as previously described.

The preferences of the agents over ratios of red agents are given in Table 4.

We claim that there exists a sequence of IS-deviations which leads to an IS partition iff formula φ is satisfiable.

Suppose first that there exists a truth assignment of the variables ϕ that satisfies all the clauses. Let us denote by ℓ_j a chosen literal-agent associated with an occurrence of a literal true in ϕ which belongs to clause C_j. Since all the clauses of φ are satisfied by ϕ, there exists such a literal-agent ℓ_j for each clause C_j. For every clause C_j, let literal-agent ℓ_j join coalition K_j. These IS-deviations make the chosen literal-agents reach their most preferred ratio so none of them will deviate afterwards.

Draft – April 19, 2022

For the clause-agents, they all reach either their first or second most preferred ratio but have no possibility to improve their satisfaction in the latter case so none of them will deviate afterwards neither. Then, let all remaining literal-agents y_i^ℓ (resp., $\overline{y}_i^\ell$) deviate by joining coalition Z_i (resp., $\overline{Z_i}$). Since ϕ is a truth assignment of the variables, for each variable x_i, there exists a coalition Z_i or $\overline{Z_i}$ that is joined by two literal-agents and thus reaches the most preferred ratio $\frac{3i}{\beta+2}$ or $\frac{3i-2}{\beta+2}$. For each variable, if coalition Z_i (resp., $\overline{Z_i}$) is not joined by two literal-agents, then it cannot be true for $\overline{Z_i}$ (resp., Z_i), and variable-agent z_i (resp., $\overline{z}_i$) then deviates for joining coalition G_i^2, and if one literal-agent previously joined coalition Z_i (resp., $\overline{Z_i}$) she deviates to be alone. No agent can then move in a IS-deviation because variable-agents z_i (resp., $\overline{z}_i$) in the gadget prefer ratios which differ by one blue agent (resp., red agent) from the ratio of the current coalitions. This also holds for literal-agents and dummy agents. Therefore, the current partition is IS.

Suppose now that there exists no truth assignment of the variables that satisfies all the clauses. That means that it is not possible that one literal-agent joins each clause coalition while two literal-agents y_i^1 and y_i^2 join coalition Z_i or $\overline{y}_i^1$ and $\overline{y}_i^2$ join coalition $\overline{Z_i}$ for each variable x_i. By construction of the preferences, the only agents who want to join a coalition K_j are literal-agents associated with a literal belonging to clause C_j and the only agents who want to join a coalition Z_i (resp., $\overline{Z_i}$) are literal-agents y_i^1 and y_i^2 (resp., $\overline{y}_i^1$ and $\overline{y}_i^2$). Moreover, since each literal-agent prefers to join clause coalitions than variable coalitions, it means that in a maximal sequence of IS-deviations, all clause-agents in K_j will reach one of the two most preferred ratios, $\frac{2j}{\alpha+1}$ in case a red literal-agent joined or $\frac{2j-1}{\alpha+1}$ in case a blue literal-agent joined. In both cases, they have no incentive to deviate afterwards. However, in such a case, there exists a variable x_i such that at most one literal-agent joins coalition Z_i and $\overline{Z_i}$. It follows that both variable-agents z_i and $\overline{z_i}$ have an incentive to deviate to the gadget associated with variable x_i (their respective most preferred ratios $\frac{3i}{\beta+2}$ and $\frac{3i-2}{\beta+2}$ can never be reached). Within the gadget associated with variable x_i, variable-agents z_i and $\overline{z_i}$ are the only agents who can deviate and we necessarily reach a cycle, which is the same as described in Figure 4 for the proof of Lemma A.1.

Finally, we must verify that all the fractions described in the preferences with different variables are indeed different. First of all, for gadget coalitions, since $\gamma > 12p-2$, it holds that $\frac{6(p-i)+6}{i\gamma+1} > \frac{6(p-i)+6}{i\gamma+2} > \frac{6(p-i)+5}{i\gamma} > \frac{6(p-i)+5}{i\gamma+1} > \frac{6(p-i)+4}{i\gamma+1} > \frac{6(p-i)+3}{i\gamma} > \frac{6(p-i)+3}{i\gamma+1} > \frac{6(p-i)+2}{i\gamma+1} > \frac{6(p-i)+2}{i\gamma+2} > \frac{6(p-i)+1}{i\gamma} > \frac{6(p-i)+1}{i\gamma+1}$ for every $i \in [p]$. Moreover, it holds that $\frac{6(p-i)+6}{i\gamma+1} > \frac{6(p-(i+1)+1}{(i+1)\gamma+1}$ for every $i \in [p-1]$ so all the values associated with ratios preferred to the initial ones are different for all gadget coalitions. For variable coalitions, since $\beta > 3p-2$, it holds that $\frac{3i}{\beta+2} > \frac{3i-1}{\beta+1} > \frac{3i-2}{\beta} > \frac{3i-2}{\beta+1} > \frac{3i-2}{\beta+2}$ for every $i \in [p]$. Moreover, it holds that $\frac{3i}{\beta+2} < \frac{3(i+1)-2}{\beta+2}$ for every $i \in [p-1]$ so all the values associated with ratios preferred to the initial ones are different for all variable coalitions. For clause coalitions, since $\alpha > 2m-1$, it holds that $\frac{2j-1}{\alpha+1} < \frac{2j-1}{\alpha} < \frac{2j}{\alpha+1}$ for every $j \in [m]$. Moreover, it holds that $\frac{2j}{\alpha+1} < \frac{2(j+1)-1}{\alpha+1}$ for every $j \in [m-1]$ so all the values associated with ratios preferred to the initial ones are different for all clause coalitions. It remains to check that the ratios associated with clause, variable or gadget coalitions do not interfere with each other. Since $\gamma > 6p\beta + 12p - 1$, it holds that the highest reachable ratio associated with a gadget coalition is smaller than the smallest reachable ratio associated with a variable coalition, i.e., $\frac{6p}{\gamma+1} < \frac{1}{\beta+2}$. Since $\beta > 3p\alpha + 3p - 2$, it holds that the highest reachable ratio associated with a variable coalition is smaller than the smallest reachable ratio associated with a clause coalition, i.e., $\frac{3p}{\beta+2} < \frac{1}{\alpha+1}$. Therefore, all the reachable ratios are indeed different for clause, variable and gadget coalitions. It follows that the previously described deviations are indeed the only possible ones and hence no sequence of IS-deviations can reach an IS partition. $\square$

Draft – April 19, 2022

LEMMA B.2. *$\forall$-IS-SEQUENCE-HDG is* co-NP*-hard even for strict preferences.*

PROOF. For this purpose, we prove the NP-hardness of the complement problem, which asks whether there exists a cycle in IS-deviations. Let us perform a reduction from (3,B2)-SAT [6]. In an instance of (3,B2)-SAT, we are given a CNF propositional formula φ where every clause C_j, for $1 \leq j \leq m$, contains exactly three literals and every variable x_i, for $1 \leq i \leq p$, appears exactly twice as a positive literal and twice as a negative literal. From such an instance, we construct an instance of a hedonic diversity game with initial partition as follows. The proof works in the same way as the proof of Lemma A.2 except that we have to ensure appropriate ratios of red agents in each coalition.

For each ℓ^{th} occurrence ($\ell \in \{1,2\}$) of a positive literal x_i (resp., negative literal $\overline{x}_i$), we create a red literal-agent y_i^ℓ (resp., a blue literal-agent $\overline{y}_i^\ell$). We create another red agent t. All these agents are singletons in the initial partition π_0. Let us consider four integers α, β_1^+, β_1^- and β_2 such that (1) $\alpha > 6m+2$, $\beta_1^+ > \max\{4p-2; (2p+1)\alpha + 4p\}$, $\beta_1^- > \max\{4p-2; 2p\beta_1^+ + 2p - 2\}$ and $\beta_2 > \max\{3p-2; 3p\beta_1^- + 4p\}$. For instance, we can set the following values: $\alpha = m^3$, $\beta_1^+ = m^5$, $\beta_1^- = m^7$ and $\beta_2 = m^9$ (one can verify that condition (1) is satisfied, especially because in a (3,B2)-SAT instance, it holds that $m \geq 4$ and $p = 3/4m$). For each clause C_j, we then create α dummy clause-agents with among them $3j-2$ red agents. They are all grouped within the same coalition K_j in the initial partition π_0. We also create α dummy agents with among them $3m+1$ red agents, they are all grouped within the same coalition K_{m+1} in initial partition π_0. For each first occurrence of literal x_i (resp. $\overline{x}_i$), we create β_1^+ (resp., β_1^-) dummy variable agents with among them $2i-1$ red agents, they are all grouped within the same coalition Y_i^1 (resp., $\overline{Y_i^1}$) in the initial partition π_0. Finally, for each second occurrence of literal x_i (resp. $\overline{x}_i$), we create β_2 dummy variable agents with among them $3i-2$ red agents, they are all grouped within the same coalition Y_i^2 (resp., $\overline{Y_i^2}$) in the initial partition π_0. Although we have created many agents, the construction remains polynomial by considering reasonable values of α, β_1^+, β_1^- and β_2, as previously described.

The preferences of the agents over coalition ratios are given in Table 5.

We claim that there exists a cycle of IS-deviations iff formula φ is satisfiable. We omit the formal proof of equivalence which follows exactly the same arguments as the proof of Lemma A.2 with even the same name of agents and fixed coalitions. When given a truth assignment of the variables which satisfies formula φ, it is easy to see that the cycle described in the first part of the proof of Lemma A.2 can also occur in this instance, proving the if part. For the only if part, the same arguments as the ones given in the second part of the proof of Lemma A.2 also hold, except that we need to adapt to the context of evaluations of coalitions based on red agent ratios. The only point that must be additionally checked is that all the fractions described in the preferences with different variables are indeed different.

First of all, for clause coalitions, since $\alpha > 6m+2$, it holds that $\frac{3j}{\alpha+2} > \frac{3j-1}{\alpha+1} > \frac{3j-1}{\alpha+2} > \frac{3j-2}{\alpha} > \frac{3j-2}{\alpha+1} > \frac{3j-2}{\alpha+2}$ for every $j \in [m+1]$. Moreover, it holds that $\frac{3j}{\alpha+2} < \frac{3(j+1)-2}{\alpha+2}$ for every $j \in [m]$ so all the values associated with ratios preferred to the initial ones are different for all clause coalitions. For variable coalitions associated with the first positive occurrence of a variable, since $\beta_1^+ > 4p-2$, it holds that $\frac{2i+1}{\beta_1^++2} > \frac{2i}{\beta_1^++1} > \frac{2i}{\beta_1^++2} > \frac{2i-1}{\beta_1^+} > \frac{2i-1}{\beta_1^++1}$ for every $i \in [p]$. Moreover, it holds that $\frac{2i+1}{\beta_1^++2} < \frac{2(i+1)-1}{\beta_1^++1}$ for every $i \in [p-1]$ so all the values associated with ratios preferred to the initial ones are different for all variable coalitions associated with the first positive occurrence of a variable. For variable coalitions associated with the first negative occurrence of a variable, since $\beta_1^- > 4p-2$, it holds that $\frac{2i}{\beta_1^-+2} > \frac{2i}{\beta_1^-+2} > \frac{2i-1}{\beta_1^-} > \frac{2i-1}{\beta_1^-+1} > \frac{2i-1}{\beta_1^-+2}$ for every $i \in [p]$. Moreover, it holds that $\frac{2i}{\beta_1^-+2} < \frac{2(i+1)-1}{\beta_1^-+2}$ for every $i \in [p-1]$ so all the values associated with ratios preferred to the

Draft – April 19, 2022

Table 5. Preferences of the agents in the reduced instance of Lemma B.2, for every $1 \le i \le p$, $1 \le i' < p$, $1 \le j \le m+1$, $\ell \in \{1,2\}$. Notation $cl(x_i^\ell)$ (resp., $cl(\overline{x}_i^\ell)$) stands for the index of the clause to which the ℓ^{th} occurrence of literal x_i (resp., $\overline{x}_i$) belongs, the framed value corresponds to the ratio of the initial coalition in partition π_0, and $[\dots]$ denotes an arbitrary order over the rest of the coalition ratios.

$$y_i^1: \frac{3cl(x_i^1)}{\alpha+2} \succ \frac{3cl(x_i^1)-1}{\alpha+2} \succ \frac{3(cl(x_i^1)+1)}{\alpha+2} \succ \frac{3(cl(x_i^1)+1)-1}{\alpha+2} \succ \frac{3(cl(x_i^1)+1)-1}{\alpha+1} \succ \frac{3cl(x_i^1)-1}{\alpha+1} \succ \frac{2i+1}{\beta_1^++2} \succ \frac{2i}{\beta_1^++2} \succ \frac{3i}{\beta_2+2} \succ \frac{3i-1}{\beta_2+1} \succ \frac{2i}{\beta_1^++1} \succ \boxed{1} \succ [\dots]$$

$$y_{i'}^2: \frac{3cl(x_{i'}^1)}{\alpha+2} \succ \frac{3cl(x_{i'}^1)-1}{\alpha+2} \succ \frac{3(cl(x_{i'}^1)+1)}{\alpha+2} \succ \frac{3(cl(x_{i'}^1)+1)-1}{\alpha+2} \succ \frac{3(cl(x_{i'}^1)+1)-1}{\alpha+1} \succ \frac{3cl(x_{i'}^1)-1}{\alpha+1} \succ \frac{3i'}{\beta_2+2} \succ \frac{2(i'+1)+1}{\beta_1^++2} \succ \frac{2(i'+1)}{\beta_1^++2} \succ \frac{2(i'+1)}{\beta_1^++1} \succ \frac{2(i'+1)+1}{\beta_1^-+2} \succ \frac{2(i'+1)}{\beta_1^-+2} \succ \frac{2(i'+1)}{\beta_1^-+1} \succ \frac{3i'-1}{\beta_2+1} \succ \boxed{1} \succ [\dots]$$

$$y_p^2: \frac{3cl(x_p^1)}{\alpha+2} \succ \frac{3cl(x_p^1)-1}{\alpha+2} \succ \frac{3(cl(x_p^1)+1)}{\alpha+2} \succ \frac{3(cl(x_p^1)+1)-1}{\alpha+2} \succ \frac{3(cl(x_p^1)+1)-1}{\alpha+1} \succ \frac{3cl(x_p^1)-1}{\alpha+1} \succ \frac{3p}{\beta_2+2} \succ \frac{3}{\alpha+2} \succ \frac{2}{\alpha+2} \succ \frac{2}{\alpha+1} \succ \frac{3p-1}{\beta_2+1} \succ \boxed{1} \succ [\dots]$$

$$\overline{y}_i^1: \frac{3cl(\overline{x}_i^1)-1}{\alpha+2} \succ \frac{3cl(\overline{x}_i^1)-2}{\alpha+2} \succ \frac{3(cl(\overline{x}_i^1)+1)-1}{\alpha+2} \succ \frac{3(cl(\overline{x}_i^1)+1)-2}{\alpha+2} \succ \frac{3(cl(\overline{x}_i^1)+1)-2}{\alpha+1} \succ \frac{3cl(\overline{x}_i^1)-2}{\alpha+1} \succ \frac{2i}{\beta_1^-+2} \succ \frac{2i-1}{\beta_1^-+2} \succ \frac{3i-2}{\beta_2+2} \succ \frac{3i-2}{\beta_2+1} \succ \frac{2i-1}{\beta_1^-+1} \succ \boxed{0} \succ [\dots]$$

$$\overline{y}_{i'}^2: \frac{3cl(\overline{x}_{i'}^1)-1}{\alpha+2} \succ \frac{3cl(\overline{x}_{i'}^1)-2}{\alpha+2} \succ \frac{3(cl(\overline{x}_{i'}^1)+1)-1}{\alpha+2} \succ \frac{3(cl(\overline{x}_{i'}^1)+1)-2}{\alpha+2} \succ \frac{3(cl(\overline{x}_{i'}^1)+1)-2}{\alpha+1} \succ \frac{3cl(\overline{x}_{i'}^1)-2}{\alpha+1} \succ \frac{3i'-2}{\beta_2+2} \succ \frac{2(i'+1)}{\beta_1^++2} \succ \frac{2(i'+1)-1}{\beta_1^++2} \succ \frac{2(i'+1)-1}{\beta_1^++1} \succ \frac{2(i'+1)}{\beta_1^-+2} \succ \frac{2(i'+1)-1}{\beta_1^-+2} \succ \frac{2(i'+1)-1}{\beta_1^-+1} \succ \frac{3i'-2}{\beta_2+1} \succ \boxed{0} \succ [\dots]$$

$$\overline{y}_p^2: \frac{3cl(\overline{x}_p^1)-1}{\alpha+2} \succ \frac{3cl(\overline{x}_p^1)-2}{\alpha+2} \succ \frac{3(cl(\overline{x}_p^1)+1)-1}{\alpha+2} \succ \frac{3(cl(\overline{x}_p^1)+1)-2}{\alpha+2} \succ \frac{3(cl(\overline{x}_p^1)+1)-2}{\alpha+1} \succ \frac{3cl(\overline{x}_p^1)-2}{\alpha+1} \succ \frac{3p-2}{\beta_2+2} \succ \frac{2}{\alpha+2} \succ \frac{1}{\alpha+2} \succ \frac{1}{\alpha+1} \succ \frac{3p-2}{\beta_2+1} \succ \boxed{0} \succ [\dots]$$

$$t: \frac{3m+3}{\alpha+2} \succ \frac{3m+2}{\alpha+2} \succ \frac{3}{\beta_1^++2} \succ \frac{2}{\beta_1^-+2} \succ \frac{2}{\beta_1^++1} \succ \frac{2}{\beta_1^-+1} \succ \frac{3m+2}{\alpha+1} \succ \boxed{1} \succ [\dots]$$

$$K_j: \frac{3j}{\alpha+2} \succ \frac{3j-1}{\alpha+2} \succ \frac{3j-2}{\alpha+2} \succ \frac{3j-1}{\alpha+1} \succ \frac{3j-2}{\alpha+1} \succ \boxed{\frac{3j-2}{\alpha}} \succ [\dots]$$

$$Y_i^1: \frac{2i+1}{\beta_1^++2} \succ \frac{2i}{\beta_1^++2} \succ \frac{2i}{\beta_1^++1} \succ \frac{2i-1}{\beta_1^++1} \succ \boxed{\frac{2i-1}{\beta_1^+}} \succ [\dots]$$

$$\overline{Y_i^1}: \frac{2i}{\beta_1^-+2} \succ \frac{2i-1}{\beta_1^-+2} \succ \frac{2i}{\beta_1^-+1} \succ \frac{2i-1}{\beta_1^-+1} \succ \boxed{\frac{2i-1}{\beta_1^-}} \succ [\dots]$$

$$Y_i^2: \frac{3i}{\beta_2+2} \succ \frac{3i-1}{\beta_2+1} \succ \boxed{\frac{3i-2}{\beta_2}} \succ [\dots]$$

$$\overline{Y_i^2}: \frac{3i-2}{\beta_2+2} \succ \frac{3i-2}{\beta_2+1} \succ \boxed{\frac{3i-2}{\beta_2}} \succ [\dots]$$

initial ones are different for all variable coalitions associated with the first negative occurrence of a variable. For variable coalitions associated with the second occurrence of a literal, since $\beta_2 > 3p-2$, it holds that $\frac{3i}{\beta_2+2} > \frac{3i-1}{\beta_2+1} > \frac{3i-2}{\beta_2} > \frac{3i-2}{\beta_2+1} > \frac{3i-2}{\beta_2+2}$ for every $i \in [p]$. Moreover, it holds that $\frac{3i}{\beta_2+2} < \frac{3(i+1)-2}{\beta_2+2}$ for every $i \in [p-1]$ so all the values associated with ratios preferred to the initial ones are different for all variable coalitions associated with the second occurrence of a literal.

It remains to check that the ratios associated with clause or variable coalitions do not interfere with each other. Since $\beta_2 > 3p\beta_1^- + 4p$, it holds that the highest reachable ratio associated with a variable coalition related to the second occurrence of a literal is smaller than the smallest reachable ratio associated with a variable coalition related to the first negative occurrence of a variable, i.e., $\frac{3p}{\beta_2+2} < \frac{1}{\beta_1^-+2}$. Since $\beta_1^- > 2p\beta_1^+ + 2p - 2$, it holds that the highest reachable ratio associated with a variable coalition related to the first negative occurrence of a variable is smaller than the smallest reachable ratio associated with a variable coalition related to the first positive occurrence of a variable, i.e., $\frac{2p}{\beta_1^-+2} < \frac{1}{\beta_1^++1}$. Since $\beta_1^+ > (2p+1)\alpha + 4p$, it holds that the highest reachable ratio associated with a variable coalition related to the first positive occurrence of a variable is smaller than the smallest reachable ratio associated with a clause coalition, i.e., $\frac{2p+1}{\beta_1^++2} < \frac{1}{\alpha+2}$. Therefore, all the reachable ratios are indeed different for all clause and variable coalitions. It follows that the deviations described in the second part of the proof of Lemma A.2 are the only possible ones. Hence the described cycle is actually the only possible one. $\square$

C FRACTIONAL HEDONIC GAMES

The hardness reductions in this section are from the NP-complete problem exact 3-Cover [?]. An instance of exact 3-cover consists of a tuple (R, S), where R is a ground set together with a set S

Draft – April 19, 2022

of 3-element subsets of R. A 'yes'-instance is an instance so that there exists a subset $S' \subseteq S$ that partitions R.

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.*

We provide separate reductions for the two hardness results in the next lemmas.

Lemma C.1. *$\exists$-IS-Sequence-FHG is* NP*-hard even in symmetric FHGs with non-negative weights where the initial partition is the singleton partition.*

Proof. We provide a reduction from exact 3-cover.

Let (R, S) be an instance of exact 3-cover. We may assume that every $r \in R$ occurs in at least one set of S. Let $m_r := |\{s \in S : r \in s\}| - 1 \geq 0$, for $r \in R$. We define the symmetric FHG on agent set N, where the underlying graph consists of a 4-clique for every set in s, and m_r copies of a non-negative version of the example from Theorem 5.1. Formally, $N = \bigcup_{s \in S}(\{t_s\} \cup \{s^i : i \in s\}) \cup \bigcup_{r \in R} \bigcup_{v=1}^{m_r} \{a_w^{r,v}, b_w^{r,v}, c_w^{r,v} : w = 1, \dots, 5\}$, and non-negative, symmetric weights are given by

- For all $r \in R$, $v \in \{1, \dots, m_r\}$, and $w \in \{1, \dots, 5\}$,
  - $v(a_w^{r,v}, b_w^{r,v}) = v(b_w^{r,v}, c_w^{r,v}) = v(a_w^{r,v}, c_w^{r,v}) = 228$,
  - $v(a_w^{r,v}, a_{w+1}^{r,v}) = 436$, $v(a_w^{r,v}, b_{w+1}^{r,v}) = 228$, $v(a_w^{r,v}, c_{w+1}^{r,v}) = 248$,
  - $v(b_w^{r,v}, a_{w+1}^{r,v}) = 223$, $v(b_w^{r,v}, b_{w+1}^{r,v}) = 171$, $v(b_w^{r,v}, c_{w+1}^{r,v}) = 236$, and
  - $v(c_w^{r,v}, a_{w+1}^{r,v}) = 223$, $v(c_w^{r,v}, b_{w+1}^{r,v}) = 171$, $v(c_w^{r,v}, c_{w+1}^{r,v}) = 188$.
- $v(t_s, s^i) = 304$, $s \in S, i \in s$,
- $v(s^j, s^i) = 304$, $s \in S, i, j \in s$,
- $v(s^i, a_1^{i,v}) = 304$, $s \in S, i, j \in s, v \in \{1, \dots, m_r\}$, and
- $v(x, y) = 0$ for all agents $x, y \in N$ such that the weight is not defined, yet.

In the above definition, all indices are to be read modulo 5 (where the modulo function is assumed to map to $\{1, \dots, 5\}$). For $s \in S$, define $N^s = \{t_s\} \cup \{s^i : i \in s\}$.

Assume first that (R, S) is a 'yes'-instance and let $S' \subseteq S$ be a partition of R. For $r \in R$, let $\sigma_r : \{s \in S \setminus S' : r \in s\} \to \{1, \dots, m_r\}$ be a bijection. Note that the domain and image of σ_r have the same cardinality for every $r \in R$, because S' is a partition of R. Consider the partition $\pi = \bigcup_{r \in R} \bigcup_{v=1}^{m_r} \{\{a_2^{r,v}, b_2^{r,v}, c_2^{r,v}, a_3^{r,v}, b_3^{r,v}, c_3^{r,v}\}, \{a_4^{r,v}, b_4^{r,v}, c_4^{r,v}, a_5^{r,v}, b_5^{r,v}, c_5^{r,v}\}, \{b_1^{r,v}, c_1^{r,v}\}\} \cup \bigcup_{s \in S'} \{N^s\} \cup \bigcup_{s \in S \setminus S'} \{\{t_s\}\} \cup \{\{s^i, a_1^{i,\sigma_i(s)}\} : i \in s\}$. It is quickly checked that π is IS. Moreover, π can be reached by deviations starting from the singleton partition, by forming the coalitions one by one. In particular, coalitions of the type $\{a_2^{r,v}, b_2^{r,v}, c_2^{r,v}, a_3^{r,v}, b_3^{r,v}, c_3^{r,v}\}$ can be formed by having $a_3^{r,v}$ join $b_3^{r,v}$, forming a coalition that is subsequently joined by $c_3^{r,v}$, $a_2^{r,v}$, $b_2^{r,v}$, and finally $c_2^{r,v}$. Hence, it is possible to reach an IS partition with IS-deviations, starting with the singleton partition.

Now, assume that it is possible to reach an IS partition π by starting the dynamics from the singleton partition. Define by $G = (N, E)$ the graph with edge set $E = \{\{d, e\} : v(d, e) > 0\}$, a combinatorial representation of the unweighted version of the FHG under consideration. Note that all coalitions of π are cliques in G, because all agents that get part of a coalition of size at least 2 have positive utility and would block any further agent that does not award them positive utility. Now, consider a set of agents $D := \{a_w^{r,v}, b_w^{r,v}, c_w^{r,v} : w = 1, \dots, 5\}$ for some $r \in R$, $v \in \{1, \dots, m_r$. Assume for contradiction that for all agents $d \in D$, $\pi(d) \subseteq D$. This yields an IS partition of the game considered in Theorem 5.1, because the agents would also form an IS partition in this game. This is due to the fact that no agents with mutual negative utility would form a coalition, and a deviation with negative weights would still be a deviation if these weights are set to 0. This is a contradiction. Hence, some agent in D forms a coalition with an agent outside D. By the fact that all coalitions in

Draft – April 19, 2022

π are cliques in G, the only such agent can be $a_1^{r,v}$. By the same fact, $\pi(a_1^{r,v}) \cap D = \{a_1^{r,v}\}$ and there exists a unique $s \in S$ with $r \in s$ such that $\pi(a_1^{r,v}) = \{a_1^{r,v}, s^r\}$.

Next, let $s \in S$. We claim that $\{t^s\} \in \pi$ or $N^s \in \pi$. Otherwise, consider $r \in s$ with $s^r \notin \pi(t^s)$. Again by the clique property, $v_{s^r}(\pi) \leq \frac{304}{2}$ and $\pi(t^s) \subseteq N^s$. Hence, $v_{s^r}(\pi(t^s) \cup \{s^r\}) \geq \frac{608}{3}$, and every agent in $\pi(t^s)$ would welcome s^r. This contradicts the individual stability of π.

Consider the set $T = \{s \in S \colon \{t^s\} \in \pi\}$. Then, for every $s \in T$ and $r \in s$, there exists $v \in \{1, \dots, m_r\}$ with $\pi(s^r) = \{a_1^{r,v}, s^r\}$. Otherwise, $\pi(s^r) \subseteq N^s$ and t^s can perform a deviation by joining $\pi(s^r)$. Hence, the sets in T cover every element in $r \in R$ exactly m_r times (in order to form all the required coalitions of the type $\{a_1^{r,v}, s^r\}$). Since S covers every $r \in R$ exactly $m_r + 1$ times, the set $S' = S \setminus T$ forms a partition of R. Hence, (R, S) is a 'yes'-instance. $\square$

Lemma C.2. *$\forall$-IS-Sequence-FHG is* co-NP*-hard even in symmetric FHGs with non-negative weights.*

Proof. For this purpose, we prove the NP-hardness of the complement problem, which asks whether there exists a cycle in IS-deviations. We provide a reduction from exact 3-cover.

Let (R, S) be an instance of exact 3-cover. Let $l = |S| - |R|/3$. Choose α with a polynomial-size representation in the input size satisfying $\frac{l}{l+1}\alpha < 152 < \frac{l+1}{l+2}\alpha$. For the reduction to work, any number satisfying these boundaries suffices, and for a polynomial-size representation, one can for example use the midpoint of the boundaries.

Define the symmetric FHG on agent set N where $N = R \cup \{r^s \colon s \in S, r \in s\} \cup \{s_1, s_2 \colon s \in S\} \cup \{a_w, b_w, c_w \colon w = 1, \dots, 5\}$. We define $C = \{a_w, b_w, c_w \colon w = 1, \dots, 5\}$. The utilities are given as follows.

- For all $w \in \{1, \dots, 5\}$, reading indices modulo 5 (where the modulo function is assumed to map to $\{1, \dots, 5\}$),
  - $v(a_w, b_w) = v(b_w, c_w) = v(a_w, c_w) = 228$,
  - $v(a_w, a_{w+1}) = 436$, $v(a_w, b_{w+1}) = 228$, $v(a_w, c_{w+1}) = 248$,
  - $v(b_w, a_{w+1}) = 223$, $v(b_w, b_{w+1}) = 171$, $v(b_w, c_{w+1}) = 236$, and
  - $v(c_w, a_{w+1}) = 223$, $v(c_w, b_{w+1}) = 171$, $v(c_w, c_{w+1}) = 188$.
- For all $s \in S$,
  - $v(a_1, s_2) = v(s_1, s_2) = \alpha$,
  - $v(s_1, r^s) = \alpha$, $v(r^s, r) = 2\alpha$, $r \in S$, and
- $v(x, y) = 0$ for all agents $x, y \in N$ such that the weight is not defined, yet.

The reduction is illustrated in Figure 7. Finally, define $\pi = \{\{r\} \colon r \in R\} \cup \{\{s_1, i^s, j^s, k^s\} \colon \{i, j, k\} = s \in S\} \cup \{\{a_1\} \cup \{s_2 \colon s \in S\}\} \cup \{\{b_1, c_1\}, \{a_2, b_2, c_2, a_3, b_3, c_3\}, \{a_4, b_4, c_4, a_5, b_5, c_5\}\}$.

We claim that (R, S) is a 'yes'-instance if and only if the IS dynamics starting with π can cycle.

First assume that (R, S) is a 'yes'-instance and let $S' \subseteq S$ be a partition of R by the sets in S. We consider three stages of deviations. In the first stage, the agents in a coalition with some s_1 for $s \in S'$ join the agents of type r_v. This will leave all agents in $\{s_1 \colon s \in S'\}$ in singleton coalitions. In the second stage, agents s_2 for $s \in S'$ join their copies s_1. This leaves the agent a_1 with a utility of $\frac{l}{l+1}\alpha < 152 = v_{a_1}(\{a_1, b_1, c_1\})$. Therefore, we can have a_1 join $\{b_1, c_1\}$. From now on, we consider the subgame induced by the agents in C. We start to let a_5, b_5, and c_5 join $\{a_1, b_1, c_1\}$. Then, we reach essentially the same partition (differing only in an index shift), so we can repeat these deviations.

Conversely, assume that there exists an infinite sequence of deviations starting from π. Agents of the type r^s can perform at most one deviation joining the agent r if she is still in a singleton coalition. After this deviation, they land in a coalition that cannot be altered anymore. Therefore, agents of the type r for $r \in R$ will never deviate, because they cannot receive positive utility, unless joining an agent of the type r^s, which will never leave her coalition with s_1 unless joining r. Agents of the type s_1 will never perform a deviation, because every agent that leaves her coalition can

Draft – April 19, 2022

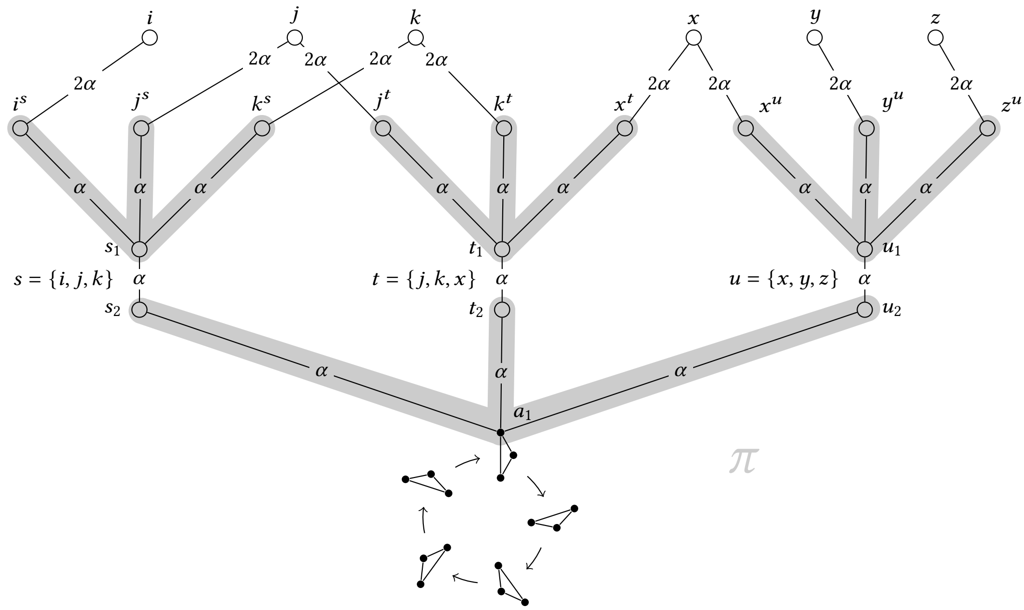


Fig. 7. Schematic of the symmetric FHG of the hardness construction. The figure is based on the instance $(\{i, j, k, x, y, z\}, \{\{i, j, k\}, \{j, k, x\}, \{x, y, z\})$. The non-singleton coalitions above a_1 of the initial partition π are depicted in gray. The only possibility for a_1 to deviate is if two of s_2, t_2, or u_2 perform a deviation, which in turn can only happen if the coalition partners of their respective partners s_1, t_1, or u_1 have been deviating before.

never be joined again, and the agent s_2 can only perform a deviation by joining s_1. In turn, agents of the type s_2 can only deviate if their copy s_1 is forced into a singleton coalition. At this point, they can deviate exactly once, forming a coalition that can never be changed again.

Agents in $C\setminus\{a_1\}$ can only perform a deviation after a_1 has performed a deviation. Thus, the only possibility for an infinite length of deviations is if a_1 performs a deviation. Since a_1 cannot join the coalition of agents of the type s_2 again, once they left her coalition, the only possible deviation is by joining the coalition $\{b_1, c_1\}$, obtaining a utility of 152. The utility of a_1 for any subset $C \subseteq \pi(a_1)$ that can arise as her coalition before she deviated for the first time is $v_{a_1}(C) = \frac{h}{1+h}\alpha$ for $h = C \cap \{s_2 \colon s \in S\}$. It follows that a_1 can only deviate once all except l agents of the type s_2 have left her coalition. Now let π' be the partition right before the first deviation of a_1 and define $S' = \{s \in S \colon s_2 \in \pi'(s_1)\}$. Then, S' consists of exactly $|R|/3$ elements. The only way that all except l agents of type s_2 have left $\pi(a_1)$ is if S' covers precisely the elements of R. Hence, S' forms a partition of R. Consequently, (R, S) is a 'yes'-instance. □

Theorem 5.6. *∃-IS-Sequence-FHG is* NP*-hard and ∀-IS-Sequence-FHG is* co-NP*-hard, even in asymmetric FHGs.*

We prove the two hardness results by providing separate reductions for each problem in the next two lemmas.

Lemma C.3. *∃-IS-Sequence-FHG is* NP*-hard even in asymmetric FHGs.*

Proof. We provide a reduction from exact 3-cover.

Draft – April 19, 2022

Let (R, S) be an instance of exact 3-cover. We may assume that every $r \in R$ occurs in at least one set of S. Let $m_r := |\{s \in S : r \in s\}| - 1 \geq 0$, and $l = |S| - |R|/3$. Define the simple, asymmetric FHG based on the directed graph $G = (V, A)$, where $V = \bigcup_{r \in R}\{r_1, \dots, r_{m_r}\} \cup S \cup \bigcup_{s \in S}\{r^s : r \in s\} \cup \bigcup_{v=1}^{l}\{a_1^v, a_2^v, a_3^v\}$ and $A = \bigcup_{s \in S}(\{(s, r^s), (r^s, r_1), \dots, (r^s, r_{m_r}) : r \in S\} \cup \{(s, a_1^1), \dots, (s, a_1^l)\}) \cup \bigcup_{v=1}^{l}\{(a_1^v, a_2^v), (a_2^v, a_3^v), (a_3^v, a_1^v)\}$. Finally, define the partition $\pi = \bigcup_{a \in V \setminus (S \cup \{r^s : s \in S, r \in s\})}\{\{a\}\} \cup \{\{s, i_1^s, j_1^s, k_1^s\} : \{i, j, k\} = s \in S\}$. The reduction is illustrated in Figure 8. It depicts the asymmetric, directed graph corresponding to the instance $(\{i, j, k, x, y, z\}, \{\{i, j, k\}, \{j, k, x\}, \{x, y, z\})$ together with the initial partition π.

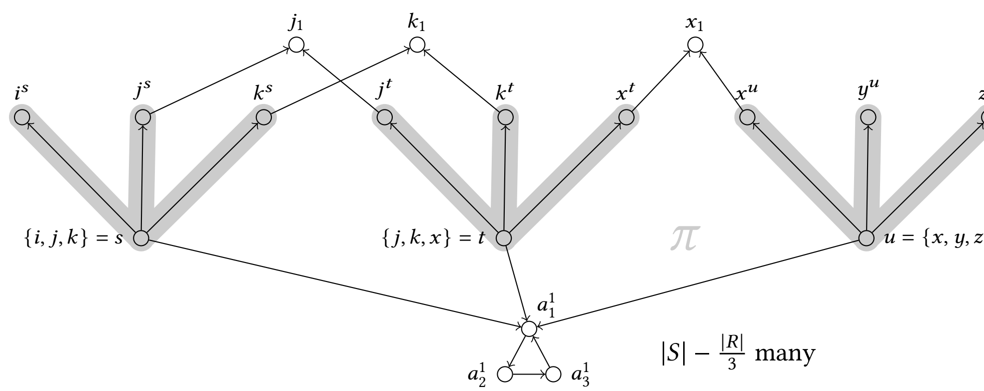


Fig. 8. Schematic of the simple, asymmetric FHG of the hardness construction. The figure is based on the instance $(\{i, j, k, x, y, z\}, \{\{i, j, k\}, \{j, k, x\}, \{x, y, z\})$. The non-singleton coalitions of the initial partition π are depicted in gray.

We claim that (R, S) is a 'yes'-instance if and only if the IS dynamics starting with π can converge.

Assume first that (R, S) is a 'yes'-instance and let $S' \subseteq S$ be a partition of R by sets in S. Consider the following deviations. First, the agents in the set $\bigcup_{s \in S \setminus S'}\{r^s : r \in s\}$ join one by one the agents in $\bigcup_{r \in R}\{r_1, \dots, r_{m_r}\}$ to end up in coalitions of size 2. Since S' covers every element of R exactly once, this step can be performed. Next, the agents $\{s \in S \setminus S'\}$ join the agents $\{a_1^1, \dots, a_1^l\}$ in an arbitrary bijective way. Finally, agents a_2^v join agents a_3^v. It is quickly checked that the resulting partition is IS.

Conversely, assume that there exists a converging sequence of deviations starting with the partition π_0 and terminating in partition π^*. Then, one agent of every set $\{a_1^v, a_2^v, a_3^v\}$ must form a coalition with an agent outside of this set. The only possibility for this is if a_1^v is joined by an agent corresponding to a set in S. Every such agent can only perform a deviation if all the other agents in her initial coalition have deviated before. Define $S' = \{s \in S : \pi_0(s) = \pi^*(s)\}$. It can only happen that $3|S| - |R|$ agents corresponding to the sets in $S \setminus S'$ deviate if S' forms a partition of R. Hence, (R, S) is a 'yes'-instance. □

Lemma C.4. *$\forall$-IS-Sequence-FHG is* co-NP*-hard even in asymmetric FHGs.*

Proof. For this purpose, we prove the NP-hardness of the complement problem, which asks whether there exists a cycle in IS-deviations. We provide a reduction from exact 3-cover.

Let (R, S) be an instance of exact 3-cover. We may assume that every $r \in R$ occurs in at least one set of S. Let $m_r := |\{s \in S : r \in s\}| - 1 \geq 0$, and $l = |R|/3$. Define the simple, asymmetric FHG based on the graph $G = (V, A)$, where $V = \{r_1, \dots, r_{m_r} : r \in R\} \cup \{r^s : s \in S, r \in s\} \cup \{s_1, s_2 : s \in S\} \cup \{b_1, b_2, b_3\} \cup \{a_1, \dots, a_l\}$, and $A = \bigcup_{s \in S}(\{(r^s, r_1), \dots, (r^s, r_{m_s}), (s_1, r^s) : r \in s\} \cup$

$\{(s_1,s_2),(b_1,s_2)\}\}\cup\{(a_v,b_1)\colon v=1,\dots,l\}\cup\{(b_2,b_3),(b_3,b_1)\}$. The reduction is illustrated in Figure 9. Finally, define $\pi=\{\{r_1\},\dots,\{r_{m_r}\}\colon r\in R\}\cup\{\{s_1,i^s,j^s,k^s\}\colon \{i,j,k\}=s\in S\}\cup\{\{b_1\}\cup\{s_2\colon s\in S\}\cup\{a_1,\dots,a_l\}\}\cup\{\{b_2\},\{b_3\}\}$.

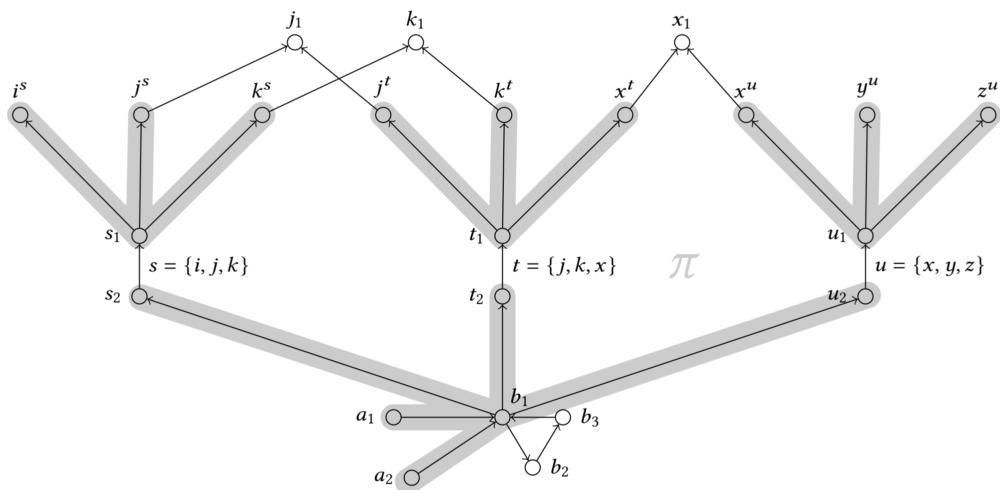


Fig. 9. Schematic of the simple, asymmetric FHG of the hardness construction. The figure is based on the instance $(\{i,j,k,x,y,z\},\{\{i,j,k\},\{j,k,x\},\{x,y,z\}\})$. The non-singleton coalitions of the initial partition π are depicted in gray. The only possibility for b_1 to deviate is if one of s_2, t_2, or u_2 performs a deviation, which in turn can only happen if the coalition partners of her respective partner s_1, t_1, or u_1 have been deviating before.

We claim that (R,S) is a 'yes'-instance if and only if the IS dynamics starting with π can cycle.

First assume that (R,S) is a 'yes'-instance and let $S'\subseteq S$ be a partition of R by the sets in S. We consider three stages of deviations. In the first stage, the agents in a coalition with some s_1 for $s\notin S'$ join the agents of type r_v. This will leave all agents in $\{s_1\colon s\notin S'\}$ in singleton coalitions. In the second stage, agents s_2 for $s\notin S'$ join their copies s_1. This leaves the agent b_1 with a utility of $l/(2l+1)<\frac{1}{2}$. Hence, we start cycling in the final stage by having b_1 join b_2, b_2 join b_3, b_3 join b_1, and repeating these deviations.

Now, assume that there exists an infinite sequence of deviations starting from π. Agents of the type r_v for $v=1,\dots,m_r$ will never deviate, because they cannot receive positive utility. Agents of the type r^s for $s\in S, r\in s$ can only deviate once to join an agent of the former type. Then, no agent can join their coalition, because the only agents r^s would allow cannot deviate. In addition, r^s can never improve her utility again. Hence, this coalition will stay the same for the remainder of the heuristics. Agents of the type s_1 will never deviate, because they are initially in their best coalition, and every agent that leaves can never be joined again. Next, agents of the type s_2 can only deviate if their copy s_1 is forced into a singleton coalition. At this point, they can deviate exactly once, forming a coalition that can never be changed again.

Agents a_v can never deviate unless b leaves their coalition. Agents b_2 and b_3 can only be involved in a deviation at most once until b_1 forms a coalition of her own or performs a deviation. Since b_1 can never form a coalition of her own, the only possibility for an infinite length of deviations is if b_1 performs a deviation. Since b_1 cannot join the coalition of agents of the type s_2 again, once they left her coalition, the only possible deviation is by joining the agent b_2 obtaining a utility of $\frac{1}{2}$. The utility of b_1 for any subset $C\subseteq\pi(b_1)$ that can arise before she deviated for the first time is

Draft – April 19, 2022

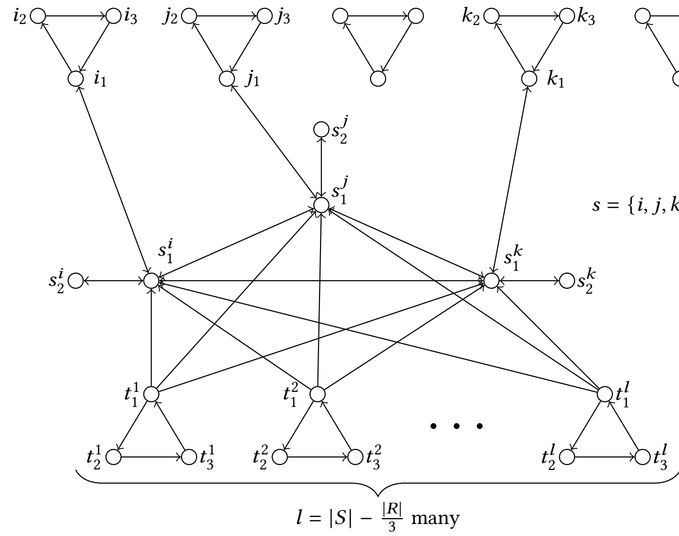


Fig. 10. Schematic of the simple FHG of the hardness construction. Bidirected edges indicate a mutual utility of 1.

$v_{b_1}(C) = \frac{h}{l+1+h}$ for $h = C \cap \{s_2 \colon s \in S\}$. It follows that b_1 can only deviate once all except l agents of the type s_2 have left her coalition. Now let π' be the partition right before the first deviation of b_1 and define $S' = \{s \in S \colon v_{s_1}(\pi') > 0\}$. Then, S' consists of exactly $|R|/3$ elements. The only way that all except l agents of type s_2 have left $\pi(b_1)$ is if S' covers precisely the elements of R. Hence, S' forms a partition of R. Consequently, (R, S) is a 'yes'-instance. □

Theorem 5.7. *∃-IS-Sequence-FHG is* NP*-hard even in simple FHGs when starting from the singleton partition.*

Proof. The reduction is from exact 3-cover.

Let an instance (R, S) of exact 3-cover be given and set $l = |S| - \frac{|R|}{3}$. We construct the simple FHG induced by the following directed graph $G = (V, A)$. Let $V = \{r_1, r_2, r_3 \colon r \in R\} \cup \{s_v^i \colon v = 1, 2, i \in s \text{ for } s \in S\} \cup \{t_v^w \colon v = 1, 2, 3, w = 1, \dots, l\}$ and edges given by $A = \{(r_1, r_2), (r_2, r_3), (r_3, r_1) \colon r \in R\} \cup \{(r_1, s_1^r), (s_1^r, r_1) \colon r \in s \text{ for } s \in S\} \cup \{(s_1^i, s_2^i), (s_2^i, s_1^i) \colon i \in s \text{ for } s \in S\} \cup \{(t_1^w, t_2^w), (t_2^w, t_3^w), (t_3^w, t_1^w) \colon w = 1, \dots, l\} \cup \{(t_1^w, s_1^i) \colon w = 1, \dots, l, i \in s \text{ for } s \in S\}$. The construction is illustrated in Figure 10. We define $T^w = \{t_1^w, t_2^w, t_3^w\}$ for $w = 1, \dots, l$.

Assume first that there exists a 3-cover of R through sets in S and let $S' \subseteq S$ be a set of 3-element sets partitioning R. Let $\sigma \colon \{1, \dots, l\} \to S \setminus S'$ be a bijection and let $\tau \colon R \to S'$ be the function defined by $\tau(r) = s$ for the unique $s \in S'$ with $r \in s$, i.e., the function that maps an element of R to its partition class. We define the partition of agents $\pi = \{\{r_2, r_3\}, \{r_1, \tau(r)_1^r\} \colon r \in R\} \cup \{\{t_1^w\} \cup \{\sigma(w)_1^i \colon i \in \sigma(w)\} \colon w = 1, \dots, l\} \cup \{\{s_2^i\} \colon s \in S\} \cup \{\{t_2^w, t_3^w\} \colon w = 1, \dots, l\}$.

Note that π is IS. Agents of the type i_2 or t_2^w are in their best coalitions. Agents of the type i_3, t_3^w, or s_2^i could only obtain positive utility by joining a coalition of which at least one agent would get worse if they joined. Agents of the type i_1 or t_1^w cannot join another coalition that gives them positive utility because this would be blocked by an agent in that coalition. In particular, i_1 cannot join a coalition $\{t_1^i, t_1^j, t_1^k, \tau^{-1}(t)\}$ for $t \in S \setminus S'$ with $i \in t$, because $\tau^{-1}(t)$ blocks this. Similarly, t_1^w cannot join a coalition $\{s_1^i, i_1\}$ for $i \in R$ or a coalition $\{t_1^x\} \cup \{\sigma(x)_1^i \colon i \in \sigma(x)\}$ for $x \neq w$, because

Draft – April 19, 2022

this is blocked by i_1 and t_1^x, respectively. Finally, agents of the type s_1^i obtain utility $1/2$ and cannot join s_2^i. Any other deviation to a coalition that gives them positive utility is blocked. Hence, π is an IS partition of agents.

Note that π can be obtained by IS-deviations from the singleton partition by forming each of the coalitions in π. In particular, coalitions of the type $\{t_1^w\} \cup \{\sigma(w)_1^i : i \in \sigma(w)\}$ are formed by letting t_1^w join $\sigma(w)_1^i$ for an arbitrary $i \in \sigma(w)$ and then the two $\sigma(w)_1^j$ for $j \in \sigma(w)\setminus\{i\}$ join one after another. This shows that we find a converging sequence if (R, S) is a 'yes'-instance.

Conversely, assume that there exists an IS partition π of the agents that can be reached by IS-deviations starting from the singleton partition. We denote the sequence of partitions by $\pi^0, \dots, \pi^l$ for some integer l, where $\pi^0 = \{\{v\} : v \in V\}$ is the singleton partition, $\pi^l = \pi$, and partition π^{p+1} can be reached from partition π^p by an IS-deviation of agent z^p for $0 \le p \le l - 1$.

We start with a technical invariant of the IS heuristics that turns out to be very useful in determining the structure of the coalitions that agents of the type r_1 and t_1^w eventually will be part of.

To formulate the claim, denote $S_1 = \{s_1^i : i \in s \text{ for } s \in S\}$ and denote $\mathcal{N} = \{r_1 : r \in R\} \cup \{t_1^w : w = 1, \dots, l\} \cup \{s_2^i : i \in s \text{ for } s \in S\}$. The set $\mathcal{N}$ contains precisely the agents that have a directed edge to or from an agent in S_1, i.e. the outgoing and incoming neighbors of agents in S_1. We simultaneously pose the following claims for $0 \le p \le l$:

- $\pi^p(r_3) \subseteq V^r$ for $r \in R$,
- $\pi^p(r_2) \subseteq V^r$ or $\pi^p(r_2) \subseteq \{r_1, r_2\} \cup \{s_1^r : s \in S, r \in s\}$ for $r \in R$,
- $\pi^p(t_v^w) \subseteq T^w$ for $v = 1, 2$, $w = 1, \dots, l$,
- $V^r, T^w \notin \pi^l$ for $r \in R$ and $w = 1, \dots, l$,
- $\pi^p(s_2^i) \subseteq \{s_1^i, s_2^i\}$ for $s \in S, i \in s$,
- $\pi^p(a) \cap \mathcal{N} = \{a\}$, for $a \in \mathcal{N}$, and
- $\pi^p(a) \cap S_1 \neq \emptyset$ implies $v_a(\pi^p) > 0$, for $a \in \mathcal{N}$.

The claim is initially true for the singleton partition π^0. Assume that it holds after iteration p for $0 \le p \le l - 1$. Consider the agent z^p that performs the IS-deviation to reach π^{p+1}. If $z^p \notin S_1 \cup \mathcal{N}$, the claim holds for $p + 1$, because these agents can only join the coalition with agents in their 3-cycle and if they want to join the coalition of an agent in $\mathcal{N}$, this agent will block it if she already forms a coalition with an agent in S^1. If $z^p \in \mathcal{N}$, she will only deviate if she receives positive utility afterwards. The claim is true by induction if this positive utility comes from an agent outside S_1. Otherwise, she joins the coalition of $x \in S_1$. Then, $\pi^l(x) \cap \mathcal{N} = \emptyset$, because every agent $y \in \pi^l(y) \cap \mathcal{N}$ would block the inclusion of agent z^p (by the final claim). In addition, since y^s is the deviating agent, she will receive positive utility after this deviation. Hence, all claims hold. Finally, if $z^p \in S_1$, she joins an agent in $\mathcal{N}$ (otherwise she would not receive positive utility in π^{p+1}). If she joins an agent of type s_2^i, the claim follows, because by induction $\{s_2^i\} \in \pi^p$. If she joins an agent of type i_1 where $i \in s$, then $i_3 \notin \pi^l(i^1)$ (this agent would block the change). Hence, the claim for the agent i_2 follows by induction. In addition, the claim for the agent i_1 follows because no agent from $\mathcal{N}$ joins and she receives positive utility through s_1^i afterwards. Other IS-deviations for the agents in S_1 are not possible. Together, the claims are established. In particular, they all hold for the IS partition π.

We apply the claims to show that for every $w \in \{1, \dots, l\}$, there exists a $s \in S$ and $i \in s$ with $s_1^i \in \pi(t_1^w)$. Otherwise, $\pi(t_v^w) \subseteq T^w$ for $v = 1, 2, 3$ and $T^w \notin \pi$. Hence, π is not IS.

Now, fix $w \in \{1, \dots, l\}$ and let $s \in S$ and $i \in s$ with $s_1^i \in \pi(t_1^w)$. We claim that $\pi(t_1^w) = \{t_1^w\} \cup \{s_1^j : j \in s\}$. By the claims, $\pi(t_1^w) \subseteq \{t_1^w\} \cup S_1$. Under this condition, $v_{s_1^u}(\pi) \le \frac{1}{2}$ and $v_{s_1^u}(\pi) = \frac{1}{2}$ only if $\pi(t_1^w) = \{t_1^w\} \cup \{s_1^j : j \in s\}$. Note that $\{s_2^i\} \in \pi$. Hence, $v_{s_1^u}(\pi) \ge \frac{1}{2}$ since otherwise π is not IS. Hence the claim follows.

Draft – April 19, 2022

Table 6. Coalitions approved by the agents in the reduced instance of Lemma D.1, for every $1 \leq j \leq m$, $1 \leq i \leq p$ and $\ell \in \{1, 2\}$. Notation lit_j^r stands for the literal-agent associated with the r^{th} literal of clause C_j ($1 \leq r \leq 3$) and $cl(x_i^\ell)$ (resp., $cl(\overline{x}_i^\ell)$) denotes the index of the clause to which literal x_i^ℓ (resp., $\overline{x}_i^\ell$) belongs. All the coalitions that are not mentioned are disapproved by the agents.

Agents	Approved coalitions
k_j	$\{k_j, lit_j^1\}$, $\{k_j, lit_j^2\}$, $\{k_j, lit_j^3\}$, $\{k_j, k_j^2\}$
v_i	$\{v_i, y_i^1, y_i^2\}$, $\{v_i, \overline{y}_i^1, \overline{y}_i^2\}$, $\{v_i, v_i^2\}$
y_i^ℓ	$\{y_i^\ell, y_i^{3-\ell}\}$, $\{y_i^\ell, y_i^{3-\ell}, v_i\}$, $\{y_i^\ell, k_{cl(x_i^\ell)}\}$
$\overline{y}_i^\ell$	$\{\overline{y}_i^\ell, \overline{y}_i^{3-\ell}\}$, $\{\overline{y}_i^\ell, \overline{y}_i^{3-\ell}, v_i\}$, $\{\overline{y}_i^\ell, k_{cl(\overline{x}_i^\ell)}\}$
k_j^2	$\{k_j^2, k_j^3\}$
k_j^3	$\{k_j, k_j^3\}$
v_i^2	$\{v_i^2, v_i^3\}$
v_i^3	$\{v_i, v_i^3\}$

Define $S' = S \setminus \{s \in S \colon t_1^w \in \pi(s_1^i) \text{ for } i \in s\}$. The coalitions of type $\{t_1^w\} \cup \{s_1^j \colon j \in s\}$ imply that $|S'| = |S| - (|S| - |R|/3) = |R|/3$.

By the above claims, for every $r \in R$, there exists $s \in S$ with $r \in s$ and $s_1^r \in \pi(r_1)$. In particular, $s \in S'$. Hence, $\bigcup_{s \in S'} s = R$ and since $|S'| = |R|/3$ and $|s| = 3$ for all $s \in S'$, the sets in S' must be disjoint. Hence, (R, S) is a 'yes'-instance. □

D DICHOTOMOUS HEDONIC GAMES

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.*

We prove the two hardness results by providing separate reductions for each problem in the next two lemmas.

Lemma D.1. *$\exists$-IS-Sequence-DHG is* NP*-hard even when starting from the singleton partition.*

Proof. Let us perform a reduction from (3,B2)-SAT [6]. In an instance of (3,B2)-SAT, we are given a CNF propositional formula φ where every clause C_j, for $1 \leq j \leq m$, contains exactly three literals and every variable x_i, for $1 \leq i \leq p$, appears exactly twice as a positive literal and twice as a negative literal. From such an instance, we construct an instance of a dichotomous hedonic game with initial partition as follows.

For each clause C_j, for $1 \leq j \leq m$, we create a clause-agent k_j and agents k_j^2 and k_j^3. For each variable x_i, for $1 \leq i \leq p$, we create a variable-agent v_i and agents v_i^2 and v_i^3. The agents k_j^2 and k_j^3 (resp., v_i^2 and v_i^3) are used to form a gadget involving clause-agent k_j (resp., variable-agent v_i) to reproduce the counterexample provided in the proof of Proposition 6.1. For each ℓ^{th} occurrence ($\ell \in \{1, 2\}$) of a positive literal x_i (resp., negative literal $\overline{x}_i$), we create a literal-agent y_i^ℓ (resp., $\overline{y}_i^\ell$). The initial partition π_0 is the singleton partition, i.e., every agent is initially alone. The dichotomous preferences of the agents are described in Table 6.

We claim that there exists a sequence of IS-deviations ending in an IS partition iff formula φ is satisfiable.

Suppose first that there exists a truth assignment of the variables ϕ such that formula φ is satisfiable. Let us denote by ℓ_j a chosen literal-agent associated with an occurrence of a literal true

Draft – April 19, 2022

in ϕ which belongs to clause C_j. Since all the clauses of φ are satisfied by ϕ, there exists such a literal-agent ℓ_j for each clause C_j. Now let us denote by z_i^1 and z_i^2 the literal-agents associated with the two occurrences of the literal of variable x_i which is false in ϕ. Since ϕ is a truth assignment of the variables that satisfies all the clauses of formula φ, it holds that $\bigcup_{1\le j\le m}\{\ell_j\}\cap\bigcup_{1\le i\le n}\{z_i^1, z_i^2\} = \emptyset$. Let us consider the following sequence of IS-deviations starting from the singleton partition where every agent has utility 0:

- For every $1 \le j \le m$, literal-agent ℓ_j joins clause-agent C_j, which makes both agents happier since they now belong to an approved coalition;
- For every $1 \le i \le n$, literal-agent z_i^1 joins literal-agent z_i^2, which makes both agents happier since they now belong to an approved coalition (they correspond to two occurrences of the same literal), and then variable-agent v_i joins them, which makes v_i happier without deteriorating the satisfaction of agents z_i^1 and z_i^2;
- For every two agents y_i^1 and y_i^2 (resp., $\overline{y}_i^1$ and $\overline{y}_i^2$) who were not involved in the previous deviations (i.e., literal x_i (resp., $\overline{x_i}$) is true in ϕ but the two occurrences of this literal have not been used for satisfiability of formula φ), literal-agent y_i^1 joins literal-agent y_i^2, which makes both agents happier since they now belong to an approved coalition;
- For every $1 \le j \le m$, agent k_j^2 joins agent k_j^3, which makes agent k_j^2 happier and does not deteriorate the satisfaction of agent k_j^3 who still belongs to a disapproved coalition;
- For every $1 \le i \le n$, agent v_i^2 joins agent v_i^3, which makes agent v_i^2 happier and does not deteriorate the satisfaction of agent v_i^3 who still belongs to a disapproved coalition.

We claim that the resulting partition is IS. Observe that the only dissatisfied agents (who are the only ones who would have an incentive to still perform an IS-deviation) are the literal-agents who remained alone, agents k_j^3 for every $1 \le j \le m$ and agents v_i^3 for every $1 \le i \le n$. The only better coalition for agent k_j^3 is the one she would form with only clause-agent k_j. However, there is no clause-agent k_j still alone since all the clauses are satisfied by truth assignment ϕ. The only better coalition for agent v_i^3 is the one she would form with only variable-agent v_i. However, there is no variable-agent v_i still alone since ϕ is a truth assignment of all variables. For remaining literal-agents, they must correspond to a true literal in ϕ for which the literal-agent associated with the other occurrence of the literal already forms a pair with a clause-agent. Therefore, they cannot join this other literal-agent. Moreover, they cannot join their associated clause-agent because she is not alone anymore. Hence, there is no IS-deviation from this partition, which is then IS.

Suppose now that there does not exist a truth assignment of the variables that satisfies all the clauses of formula φ. Observe that if a clause-agent k_j does not form a coalition with one of the literal-agents associated with the literals of her clause, then there will be a cycle among the agents k_j, k_j^2 and k_j^3, as described in the proof of Proposition 6.1. Additionally, if a variable-agent v_i does not form a coalition with either y_i^1 and y_i^2 or $\overline{y}_i^1$ and $\overline{y}_i^2$, then there will be a cycle among the agents v_i, v_i^2 and v_i^3, as described in the proof of Proposition 6.1. Therefore, since there is no possibility to find a truth assignment of the variables which satisfies all the clauses, we necessarily get a cycle in a sequence of IS-deviations starting from the singleton partition. □

Lemma D.2. *∀-IS-Sequence-DHG is* co-NP*-hard.*

Proof. For this purpose, we prove the NP-hardness of the complement problem, which asks whether there exists a cycle in IS-deviations. Let us perform a reduction from the Satisfiability problem which asks the satisfiability of a CNF propositional formula φ given by a set of clauses $C_1, \dots, C_m$ over variables $x_1, \dots, x_p$. We construct an instance of a dichotomous hedonic game with initial partition as follows.

Draft – April 19, 2022

For each clause C_j, for $1 \leq j \leq m$, we create two clause-agents k_j and k'_j. Let us denote by p_i^+ (resp., p_i^-) the number of positive (resp., negative) literals of variable x_i in formula φ. For each k^{th} occurrence of literal x_i (resp., $\overline{x_i}$) of variable x_i, we create a literal-agent y_i^k (resp., $\overline{y}_i^k$). The initial partition is given by $\pi^0 := \{\{x_i^1, \dots, x_i^{p_i^+}, \overline{x}_i^1, \dots, \overline{x}_i^{p_i^-}\}_{1 \leq i \leq p}, \{k_j, k'_j\}_{1 \leq j \leq m}\}$. The dichotomous preferences of the agents over the coalitions to which they belong are summarized below.

- Each literal-agent y_i^k (resp., $\overline{y}_i^k$), for $1 \leq i \leq p$ and $1 \leq k \leq p_i^+$ (resp., $1 \leq k \leq p_i^-$), gives utility 1 to the coalitions where agent k'_j belongs, where k'_j refers to the clause C_j to which the k^{th} occurrence of literal x_i (resp., $\overline{x}_i$) belongs, and to all coalitions only composed of literal-agents associated with variable x_i where some literal-agents associated with $\overline{x_i}$ (resp., x_i) are missing. All the other coalitions are valued 0.
- Each clause-agent k_j, for $1 \leq j \leq m$, only gives utility 1 to the coalitions which contain clause-agent k'_{j+1} and one literal-agent associated with a literal belonging to clause C_{j+1} (where $m + 1$ refers to 1). All the other coalitions are valued 0.
- Each clause-agent k'_j, for $1 \leq j \leq m$, only gives utility 1 to the coalitions which contain agent k_j. All the other coalitions are valued 0.

We claim that there exists a cycle of IS-deviations iff formula φ is satisfiable.

Suppose first that formula φ is satisfiable by a truth assignment of the variables denoted by ϕ. For each clause C_j, for $1 \leq j \leq m$, we choose a literal-agent y_i^k (resp., $\overline{y}_i^k$) such that the k^{th} occurrence of literal x_i (resp., $\overline{x}_i$) belongs to clause C_j and literal x_i (resp., $\overline{x}_i$) is true in ϕ. By satisfiability of formula φ, there always exists such a literal-agent. Then, literal-agent y_i^k (resp., $\overline{y}_i^k$) deviates from her coalition of literal-agents associated with variable x_i to coalition $\{k_j, k'_j\}$. This deviation is beneficial for the literal-agent because she values her new coalition with utility 1 since k'_j belongs to it and her old coalition with utility 0 since no literal-agent associated with her opposite literal has left the coalition of literal-agents associated with variable x_i (we have chosen only literal-agents associated with literals true in ϕ). Moreover, this deviation does not decrease the utility of the agents of the joined coalition: agent k'_j still values the coalition with utility 1 since agent k_j belongs to it and agent k_j still values the coalition with utility 0. Therefore, this deviation is an IS-deviation. After all these deviations, we reach a partition π which contains the coalitions $\{k_j, k'_j, \ell_j\}$ for every $1 \leq j \leq m$, where ℓ_j denotes a literal-agent associated with a literal true in ϕ which belongs to clause C_j. Then, for each $1 \leq j \leq m$, by increasing order of the indices, clause-agent k_j deviates to coalition $\{k_{j+1}, k'_{j+1}, \ell_{j+1}\}$ (where $m + 1$ refers to 1). This deviation is beneficial for clause-agent k_j since she deviates to a coalition containing k'_{j+1} and a literal-agent associated with a literal belonging to clause C_{j+1}. Moreover, this deviation does not hurt the joined coalition: literal-agent ℓ_{j+1} still values the coalition with utility 1 since agent k'_{j+1} belongs to it, clause-agent k'_{j+1} still values the coalition with utility 1 since agent k_{j+1} belongs to it and clause-agent k_{j+1} still values the coalition with utility 0. Therefore, this deviation is an IS-deviation. However, when clause-agent k_j has left her old coalition, this old coalition becomes either $\{k'_j, \ell_j\}$ if $j = 1$ or $\{k'_j, \ell_j, k_{j-1}\}$ otherwise. Therefore, this deviation hurts clause-agent k'_j from the old coalition. After all these deviations, we reach a partition which contains the coalitions $\{k_j, k'_{j+1}, \ell_{j+1}\}$ for every $1 \leq j \leq m$ (where $m + 1$ refers to 1). At this point, for each $1 \leq j \leq m$, by increasing order of the indices, clause-agent k'_j deviates to coalition $\{k_j, k'_{j+1}, \ell_{j+1}\}$ (where $m + 1$ refers to 1), in order to recover utility 1 by belonging to the same coalition as clause-agent k_j. This deviation does not hurt the joined coalition because literal-agent ℓ_{j+1} still values the coalition with utility 1 since agent k'_{j+1} belongs to it, clause-agent k'_{j+1} still values the coalition with utility 0 since clause-agent k_{j+1} has previously left the coalition (in the previous “round” of deviations) and clause-agent k_j still values the coalition

Draft – April 19, 2022

with utility 1 since agents k'_{j+1} and ℓ_{j+1} belong to it. Therefore, this deviation is an IS-deviation. However, when clause-agent k'_j has left her old coalition, this old coalition becomes either $\{k_{j-1}, \ell_j\}$ if $j = 1$ or $\{k_{j-1}, k'_{j-1}, \ell_j\}$ otherwise. Therefore, this deviation hurts clause-agent ℓ_j from the old coalition. After all these deviations, we reach a partition which contains the coalitions $\{k_j, k'_j, \ell_{j+1}\}$ for every $1 \leq j \leq m$ (where $m + 1$ refers to 1). At this point, for each $1 \leq j \leq m$, by increasing order of the indices, literal-agent ℓ_j deviates to coalition $\{k_j, k'_j, \ell_{j+1}\}$, in order to recover utility 1 by belonging to the same coalition as clause-agent k'_j. This deviation does not hurt the joined coalition because literal-agent ℓ_{j+1} still values the coalition with utility 0 since agent k'_{j+1} does not belong to it, clause-agent k_j still values the coalition with utility 0 since agent k'_{j+1} does not belong to it, and clause-agent k'_j still values the coalition with utility 1 since agent k_j belongs to it. Therefore, this deviation is an IS-deviation. After all these deviations, we reach again partition π and thus there is a cycle in the sequence of IS-deviations.

Suppose now that there exists a cycle of IS-deviations. From π^0, no clause-agent has incentive to deviate: each clause-agent k'_j already values her current coalition with utility 1 since agent k_j belongs to it, and each clause-agent k_j values her current coalition with utility 0 but there is no coalition containing both agent k'_{j+1} and a literal-agent associated with a literal belonging to clause C_{j+1}. Therefore, some literal-agents must deviate and leave their initial coalition, that they value with utility 0 since no literal-agent has left it yet. Observe that once a literal-agent associated with variable x_i has left her initial coalition, no literal-agent associated with the opposite literal can leave the coalition because she values it with utility 1. If a literal-agent deviates, this is for joining coalition $\{k_j, k'_j\}$ where clause C_j refers to the clause where her associated literal occurrence appears. After such a deviation which is an IS-deviation because it does not decrease the utility of the members of the joined coalition, the only agents with incentive to deviate are clause-agents k_j if a literal-agent ℓ_{j+1} has joined coalition $\{k_{j+1}, k'_{j+1}\}$. Suppose that there exists a clause coalition $\{k_j, k'_j\}$ such that no literal-agent has joined it. Then, consider a clause index j such that no literal-agent has joined coalition $\{k_{j+1}, k'_{j+1}\}$ and a clause index j' such that for all clause coalitions $\{k_r, k'_r\}$, with $j' \leq r \leq j$, a literal-agent ℓ_r has joined the coalition but this is not the case for coalition $\{k_{j'-1}, k'_{j'-1}\}$ ($m + 1$ refers to 1, and 0 to m). Then, by progressive IS-deviations, all agents belonging to clause coalitions from j' to j will deviate for joining coalition $\{k_j, k'_j\}$. Indeed, clause-agent $k_{j'}$ will deviate to coalition $\{k_{j'+1}, k'_{j'+1}, \ell_{j'+1}\}$, and then clause-agent $k'_{j'}$ will follow her in this coalition, and then literal-agent $\ell_{j'}$ will also follow $k'_{j'}$ in this coalition. But agent $k_{j'+1}$ has incentive to do the same for coalition $\{k_{j'+2}, k'_{j'+2}, \ell_{j'+2}\}$, which leads agents $k'_{j'+2}$ and $\ell_{j'+2}$ to follow her, as well as agents $k_{j'}$, $k'_{j'}$ and $\ell_{j'}$. This process of IS-deviations then continues in the same way until all these agents group in coalition $\{k_j, k'_j, \ell_j\}$. However, since clause-agent k_j can never leave this coalition (there is no coalition containing both agent k'_{j+1} and a literal-agent associated with a literal belonging to clause C_{j+1}), no other agent will leave this coalition neither. We will therefore reach a stable state, a contradiction. It follows that each clause coalition $\{k_j, k'_j\}$, for $1 \leq j \leq m$, must be joined by a literal-agent ℓ_j associated with a literal belonging to clause C_j. Therefore, by setting to true the literals associated with literal-agents who have joined clause coalitions (we have previously said that no two literal-agents associated with opposite literals can both leave their initial coalition), we get a truth assignment of the variables which satisfies all the clauses of formula φ. $\square$