Single-Agent Dynamics in Additively Separable Hedonic Games

Felix Brandt, Martin Bullinger, and Leo Tappe

Institut für Informatik, Technische Universität München
brandtf@in.tum.de, bullinge@in.tum.de, leo.tappe@tum.de

Abstract

The formation of stable coalitions is a central concern in multiagent systems. A considerable stream of research defines stability via the absence of beneficial deviations by single agents. Such deviations require an agent to improve her utility by joining another coalition while possibly imposing further restrictions on the consent of the agents in the welcoming as well as the abandoned coalition. While most of the literature focuses on unanimous consent, we also study consent decided by majority vote, and introduce two new stability notions that can be seen as local variants of popularity. We investigate these notions in additively separable hedonic games by pinpointing boundaries to computational complexity depending on the type of consent and restrictions on the utility functions. The latter restrictions shed new light on well-studied classes of games based on the appreciation of friends or the aversion to enemies. Many of our positive results follow from the *Deviation Lemma*, a general combinatorial observation, which can be leveraged to prove the convergence of simple and natural single-agent dynamics under fairly general conditions.

Introduction

Coalition formation is a central concern in multi-agent systems and considers the question of grouping a set of agents, e.g., humans or machines, into coalitions such as teams, clubs, or societies. A prominent framework for studying coalition formation is that of hedonic games, where agents' utilities are solely based on the coalition they are part of, and which thus disregards inter-coalitional relationships (Drèze and Greenberg 1980). Hedonic games have been successfully used to model various scenarios evolving from operations research or the mathematical social sciences such as research team formation (Alcalde and Revilla 2004), task allocation (Saad et al. 2011), or community detection (Newman 2004; Aziz et al. 2019). Identifying desirable coalition structures is often based on the prospect of coalitions to stay together. To this end, various notions of stability have been introduced and studied. A coalition structure (henceforth partition) is stable when no individual or group of agents benefits by joining another coalition or by forming a new coalition.

In this paper, we focus on deviations by single agents. The simplest example is a Nash deviation where some agent unilaterally decides to leave her current coalition in order to join another coalition. While such a deviation clearly captures the incentive of single agents to perform deviations, it completely ignores the other agents' opinions about the deviation. To overcome this shortcoming, various restrictions of Nash deviations have been proposed. This has motivated stability notions such as individual stability or contractual Nash stability, which consider the unanimous consent of some or all of the coalitions directly affected by the deviation. While unanimous consent is in fact used in the formation process of international bodies like the EU or the NATO, it might be impractical and even undesirable in small- or medium-scale coalition formation scenarios. As a compromise, we also study intermediate notions of stability based on majority votes among the involved coalitions. This setting has received little attention so far (Gairing and Savani 2019), and we will also define some new majority-based stability notions.

The study of hedonic games was initiated by Drèze and Greenberg (1980), and later popularized by Banerjee, Konishi, and Sönmez (2001), Cechlárová and Romero-Medina (2001), and Bogomolnaia and Jackson (2002). Since then, a large body of research has been devoted to defining suitable game representations and solution concepts. An overview of many important aspects is provided in the survey by Aziz and Savani (2016). One prominent, natural, and arguably simple type of hedonic games is given by additively separable hedonic games (Bogomolnaia and Jackson 2002). In these games, agents entertain cardinal utilities for other agents and the utility for a coalition is defined by taking the sum of the individual utility values. This game representation allows, for instance, the modeling of settings where agents have friends and enemies, and their goal is to simultaneously maximize the number of friends and minimize the number of enemies, while one of these two goals can have higher priority than the other one (Dimitrov et al. 2006). Our work provides exact boundaries for the computational tractability of stability concepts based on single-agent deviations in additively separable hedonic games, showing a clear cut between Nash stability and stability notions under consent. We give simple and precise conditions for restricted classes of utility functions that pinpoint the boundaries of computational tractability. This includes well-studied classes of games where agents only distinguish between friends and enemies.

A more recent line of research on stability notions focuses on the dynamical aspects leading to the formation of stable outcomes (e.g., Bilò et al. 2018; Hoefer, Vaz, and Wagner 2018; Carosi, Monaco, and Moscardelli 2019; Brandt, Bullinger, and Wilczynski 2021). This yields an important distributed perspective on the coalition formation process. The value of some positive computational results in the context of hedonic games is diminished by the fact that they implicitly assume that a central authority has the means to collect all individual preferences, compute a stable partition, and enforce this partition on the agents. In contrast, simple dynamics based on single-agent deviations provide a much more plausible explanation for the formation of stable partitions. A versatile tool to prove the convergence of dynamics are potential functions, which guide the dynamics towards stable states (e.g., Bogomolnaia and Jackson 2002; Suksompong 2015). We extend the applicability of this approach by considering *non-monotonic* potential functions, i.e., potential functions that might decrease in some rounds of the dynamic process. This is possible because the total number of rounds can be bounded by observing the potential function from a global perspective using a new general combinatorial insight that we call the Deviation Lemma. The Deviation Lemma is not restricted to additively separable utilities or the specific type of single-agent deviations. For instance, the combinatorial relationship of the lemma also arises naturally in the analysis of deviation dynamics in classes of games beyond the scope of this paper, such as anonymous hedonic games (Bogomolnaia and Jackson 2002). In fact, the lemma holds for every sequence of partitions such that each partition evolves from its predecessor by having one element move to another partition class. It establishes a relationship between the development of the sizes of coalitions involved in deviations to information solely based on the starting partition and the terminal partition of the sequence.

For the special case of symmetric utility functions, additively separable hedonic games are well understood: the standard notion of utilitarian social welfare represents an increasing potential function for the dynamics induced by Nash stability (Bogomolnaia and Jackson 2002), but finding stable states (even under unanimous consent of the welcoming coalition) leads to PLS-complete problems (Gairing and Savani 2019). As we will see, this implies worst-case exponential running time of the dynamics. By contrast, our results hold for restricted sets of non-symmetric utility functions and our computational boundaries lie between polynomial-time computability and NP-completeness. In fact, whenever we identify a potential function guaranteeing the existence of stable outcomes, we are also able to prove that, from any starting partition, the corresponding simple dynamics of single-agent deviations converges to a stable partition in a polynomial number of rounds.

Preliminaries and Model

In this section we introduce hedonic games and our stability concepts. We use the notation $[k] = \{1, \ldots, k\}$ for any positive integer k.

Hedonic Games

Throughout the paper, we consider settings with a set $N = [n]$ of n agents. The goal of coalition formation is to find a partition of the agents into different disjoint coalitions according to their preferences. Hence, we search a partition $\pi : N \to 2^N$ such that $i \in \pi(i)$ for every agent $i \in N$ and either $\pi(i) = \pi(j)$ or $\pi(i) \cap \pi(j) = \emptyset$ holds for every pair of agents i and j, where $\pi(i)$ denotes the coalition to which agent i belongs. We refer to the partition π given by $\pi(i) = \{i\}$ for every agent $i \in N$ as the *singleton partition*, and to $\pi = \{N\}$ as the *grand coalition*.

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 $(N, \succsim)$, where N is an agent set and $\succsim = (\succsim_i)_{i\in N}$ is a tuple of weak orders $\succsim_i$ over $\mathcal{N}_i$ which represent the preferences of the respective agent i. Hence, agents express preferences only over the coalitions which they are part of without considering externalities. The generality of the definition of hedonic games gives rise to many interesting subclasses of games that have been proposed in the literature. Many of these classes rely on cardinal utility functions $v_i \colon N \to \mathbb{R}$ for every agent i, which are aggregated in various ways (Aziz et al. 2019; Bogomolnaia and Jackson 2002; Olsen 2012). One particularly natural and prominent such model considers aggregation by taking the sum of individual utilities. Formally, following Bogomolnaia and Jackson (2002), an *additively separable hedonic game* (ASHG) (N, v) consists of an agent set N and a tuple $v = (v_i)_{i\in N}$ of utility functions $v_i \colon N \to \mathbb{R}$ such that $\pi(i) \succsim_i \pi'(i)$ iff $\sum_{j\in\pi(i)} v_i(j) \geq \sum_{j\in\pi'(i)} v_i(j)$. Clearly, ASHGs are a subclass of hedonic games, and we can assume without loss of generality that $v_i(i) = 0$ (or set the utility of an agent for herself to an arbitrary constant). ASHGs have a natural representation by a complete directed graph $G = (N, E)$ with weight $v_i(j)$ on arc (i, j). An ASHG is called *symmetric* if $v_i(j) = v_j(i)$ for every pair of agents i and j, and it can then be represented by a complete undirected graph with weight $v_i(j)$ on edge $\{i, j\}$. There are various classes of ASHGs with certain restrictions for the utility functions that allow a natural interpretation in terms of friends and enemies. An agent j is called *friend* (respectively, *enemy*) of agent i if $v_i(j) > 0$ (respectively, $v_i(j) < 0$). An ASHG is called *friends-and-enemies game* (FEG) if $v_i(j) \in \{-1, 1\}$ for every pair of agents i and j. Further, following Dimitrov et al. (2006), an ASHG is called an *appreciation of friends game* (AFG) (respectively, an *aversion to enemies game* (AEG)) if $v_i(j) \in \{-1, n\}$ (respectively, $v_i(j) \in \{-n, 1\}$). In all of these games, agents pursue the objective to maximize their number of friends while minimizing their number of enemies. In the case of an FEG, these two goals have equal priority, while there is a strict priority for one of the goals in AFGs and AEGs.

Stability Based on Single-Agent Deviations

We want to study stability under single agents' incentives to perform deviations. A *single-agent deviation* performed by agent i transforms a partition π into a partition π' where $\pi(i) \neq \pi'(i)$ and, for all agents $j \neq i$,

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

We write $\pi \xrightarrow{i} \pi'$ to denote a single-agent deviation performed by agent i transforming partition π to partition π'.

We consider the case of myopic rational agents who only engage in a deviation if it immediately makes them better off. Formally, a *Nash deviation* is a single-agent deviation performed by agent i making agent i better off, i.e., $\pi'(i) \succ_i \pi(i)$. Any partition in which no Nash deviation is possible is called *Nash stable* (NS).

This concept of stability is very strong and comes with the drawback that only the preferences of the deviating agent are considered. Therefore, various refinements have been proposed which additionally require the consent of the abandoned and the welcoming coalition. For a compact representation, we introduce them via the notion of favor sets.

Let $C \subseteq N$ be a coalition and $i \in N$ be an agent. The *favor-in set* of C with respect to i is the set of agents in C (excluding i) that strictly favor having i inside of C than outside, i.e., $F_{\text{in}}(C, i) = \{j \in C \setminus \{i\} \colon C \cup \{i\} \succ_j C \setminus \{i\}\}$. Similarly, the *favor-out set* of C with respect to i is the set of agents in C (excluding i) that strictly favor having i outside of C than inside, i.e., $F_{\text{out}}(C, i) = \{j \in C \setminus \{i\} \colon C \setminus \{i\} \succ_j C \cup \{i\}\}$.

Following Bogomolnaia and Jackson (2002) and Dimitrov and Sung (2007), an *individual deviation* (respectively, *contractual deviation*) is a Nash deviation $\pi \xrightarrow{i} \pi'$ such that $F_{\text{out}}(\pi'(i), i) = \emptyset$ (respectively, $F_{\text{in}}(\pi(i), i) = \emptyset$). A single-agent deviation that is both an individual and a contractual deviation is called *contractual individual deviation*. All of these deviation concepts give rise to a respective stability concept. A partition is called *individually stable* (IS), *contractually Nash stable* (CNS), or *contractually individually stable* (CIS) if it allows for no individual, contractual, or contractual individual deviations, respectively.

While these stability concepts include agents affected by the deviation, they require unanimous consent, which might be unnecessarily strong in some settings. Based on this observation, we define several hybrid stability concepts where the possibility of a deviation by some agent is decided via majority votes of the involved agents.

A Nash deviation $\pi \xrightarrow{i} \pi'$ is called *majority-in deviation* (respectively, *majority-out deviation*) if $|F_{\text{in}}(\pi'(i), i)| \geq |F_{\text{out}}(\pi'(i), i)|$ (respectively, $|F_{\text{out}}(\pi(i), i)| \geq |F_{\text{in}}(\pi(i), i)|$). A single-agent deviation that is both a majority-in deviation and a majority-out deviation is called *separate-majorities deviation*. As before, a partition is called *majority-in stable* (MIS), *majority-out stable* (MOS), or *separate-majorities stable* (SMS) if it allows for no majority-in, majority-out, or separate-majorities deviations, respectively. The concepts MIS and MOS are a special case of voting-based stability notions by Gairing and Savani (2019) for a threshold of $1/2$.

Finally, it is possible to relax SMS by performing one joint vote instead of two separate votes. A Nash deviation $\pi \xrightarrow{i} \pi'$ is called a *joint-majority deviation* if $|F_{\text{out}}(\pi(i), i)| + |F_{\text{in}}(\pi'(i), i)| \geq |F_{\text{in}}(\pi(i), i)| + |F_{\text{out}}(\pi'(i), i)|$. A partition is then called *joint-majority stable* (JMS) if it allows for no joint-majority deviations. JMS is particularly interesting as it is a natural local version of popularity (Pop), an axiom recently studied in the context of hedonic games (Gärdenfors 1975; Cseh 2017; Brandt and Bullinger 2020).[1]

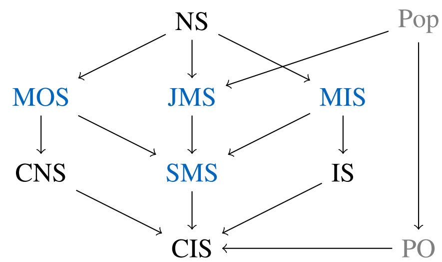

Figure 1: Logical relationships between stability notions and other solutions concepts. An arrow from concept α to concept β indicates that if a partition satisfies α, it also satisfies β. Majority-based stability notions are highlighted in blue, other single-agent based stability notions in black.

Also note that while CIS is a refinement of Pareto optimality (PO), there is no logical relationship between other (majority-based) stability concepts and PO. In particular, we denote the stability concepts based on single-agent deviations by $\mathcal{S}$, i.e., $\mathcal{S} = \{\text{NS}, \text{IS}, \text{CNS}, \text{CIS}, \text{MIS}, \text{MOS}, \text{SMS}, \text{JMS}\}$. A taxonomy of our related solution concepts is provided in Figure 1. For a more concise notation, we refer to deviations with respect to stability concept $\alpha \in \mathcal{S}$ as α*-deviations*, e.g., IS-deviations for $\alpha = \text{IS}$.

All these stability concepts naturally induce dynamics where we choose some starting partition and obtain a successor partition by having some agent perform a deviation from the current partition. More precisely, given a stability concept $\alpha \in \mathcal{S}$, an execution of α*-dynamics* is an infinite or finite sequence $(\pi_j)_{j \geq 0}$ of partitions and a corresponding sequence $(i_j)_{j \geq 1}$ of (deviating) agents such that $\pi_{j-1} \xrightarrow{i_j} \pi_j$ is an α-deviation for every j. The partition π_0 is then called the *starting partition*. Given a hedonic game G, and a stability concept $\alpha \in \mathcal{S}$, we say that the dynamics *converges for starting partition* π_0 if every execution of the α-dynamics on G with starting partition π_0 is finite. Additionally, the α-dynamics *converges* on G if it converges for every starting partition.

Proving convergence of dynamics is a very natural way to prove the existence of stable states and underlines the robustness of the stability concept. It complements a static solution concept with a decentralized process to reach a solution.

Results

In this section, we present our results.

Computational Boundaries for Nash Stability

First, we consider the notion of Nash stability. In the absence of negative utility values, the partition consisting solely of the grand coalition is Nash stable. Conversely, in the absence of

[1] Informally speaking, a partition is popular if there is no other partition preferred by a majority of the agents. JMS partitions can only be challenged by partitions evolving through Nash deviations.

positive utility values, the singleton partition is Nash stable. It is therefore necessary for an ASHG to have both positive and negative utility values in order to admit a non-trivial Nash stable partition (see also Gairing and Savani 2019).

Sung and Dimitrov (2010) showed that deciding whether an ASHG has an NS partition is NP-hard by a reduction from EXACT3COVER. This reduction produces an ASHG with four distinct positive utility values and one negative utility value. We improve upon this result by showing that a reduction is possible with only one positive and one negative utility value. Moreover, it is possible for *any* choice of these two utility values, as long as the absolute value of the negative utility value is at least as large as the positive utility value. We state the theorem in a general way allowing the positive and negative utility value to be dependent on the number of agents of the particular instance. In this way, we simultaneously cover several important cases. For instance, the hardness holds for fixed constant positive and negative utility values as in FEGs, or for AFGs and AEGs. Note that for all of our stability notions, a stable partition is a polynomial-time verifiable certificate: one can simply check whether any agent can perform a deviation, and if no one can, the partition is stable. Therefore, we omit the proof of membership in NP in all of our reductions. The proof of the next and some subsequent results are omitted due to space restrictions.

Theorem 1. *Let $f^+ : \mathbb{N} \to \mathbb{Q}_{>0}$ and $f^- : \mathbb{N} \to \mathbb{Q}_{<0}$ be two polynomial-time computable functions satisfying $|f^-(m)| \geq f^+(m)$ for all $m \in \mathbb{N}$. Then, the problem of deciding whether an ASHG with utility values restricted to $\{f^-(n), f^+(n)\}$ has an NS partition is* NP*-complete.*

Theorem 1 requires the negative utility value to be at least as large in absolute value as the positive utility value. While we leave open the computational complexity for completely arbitrary pairs of negative and positive values, we can show that the problem is also hard when the positive utility value is significantly larger than the absolute value of the negative utility value. The reduction is a variant of the reduction in Theorem 1.

Theorem 2. *Deciding whether an AFG has an NS partition is* NP*-complete.*

Deviation Lemma and Applications

By contrast, restricting the utility values to one positive and one negative value leads to positive results for other notions of stability. These results can be shown in a unified manner using a potential function argument that crucially hinges on the following general observation.

Lemma 1 (Deviation Lemma)**.** *Let $\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k$ be a sequence of k single-agent deviations. Then, the following identity holds:*

$$\sum_{j \in [k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)| = \frac{1}{2} \sum_{i \in N} |\pi_k(i)| - |\pi_0(i)|. \quad (1)$$

Proof. Let $\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k$ be a sequence of k single-agent deviations and fix some $j \in [k]$. Then, the following facts hold:

$$|\pi_j(i_j)| = \left(\sum_{i \in \pi_j(i_j) \setminus \{i_j\}} |\pi_j(i)| - |\pi_{j-1}(i)| \right) + 1,$$

$$|\pi_{j-1}(i_j)| = \left(\sum_{i \in \pi_{j-1}(i_j) \setminus \{i_j\}} |\pi_{j-1}(i)| - |\pi_j(i)| \right) + 1,$$

$$\pi_j(i) = \pi_{j-1}(i) \quad \forall i \in N \setminus (\pi_j(i_j) \cup \pi_{j-1}(i_j)).$$

Combining these facts allows us to express the difference of the deviator's coalition sizes as follows:

$$\begin{aligned}
|\pi_j(i_j)| - |\pi_{j-1}(i_j)| &= \left(\sum_{i \in \pi_j(i_j) \setminus \{i_j\}} |\pi_j(i)| - |\pi_{j-1}(i)| \right) \\
&\quad - \left(\sum_{i \in \pi_{j-1}(i_j) \setminus \{i_j\}} |\pi_{j-1}(i)| - |\pi_j(i)| \right) \\
&\quad + \sum_{i \in N \setminus (\pi_j(i_j) \cup \pi_{j-1}(i_j))} |\pi_j(i)| - |\pi_{j-1}(i)| \\
&= \sum_{i \in N \setminus \{i_j\}} |\pi_j(i)| - |\pi_{j-1}(i)|.
\end{aligned}$$

Adding $|\pi_j(i_j)| - |\pi_{j-1}(i_j)|$ to both sides yields

$$2 \left(|\pi_j(i_j)| - |\pi_{j-1}(i_j)| \right) = \sum_{i \in N} |\pi_j(i)| - |\pi_{j-1}(i)|.$$

Summing these terms for all $j \in [k]$, interchanging summation order, and telescoping gives

$$\begin{aligned}
\sum_{j \in [k]} 2 \left(|\pi_j(i_j)| - |\pi_{j-1}(i_j)| \right) &= \sum_{j \in [k]} \sum_{i \in N} |\pi_j(i)| - |\pi_{j-1}(i)| \\
2 \sum_{j \in [k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)| &= \sum_{i \in N} \sum_{j \in [k]} |\pi_j(i)| - |\pi_{j-1}(i)| \\
2 \sum_{j \in [k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)| &= \sum_{i \in N} |\pi_k(i)| - |\pi_0(i)|.
\end{aligned}$$

Dividing both sides by 2 completes the proof. □

The Deviation Lemma is especially useful as the right-hand side of Equation (1) does not depend on k, and we can therefore also find bounds for its left-hand side solely depending on the number of players n.

Lemma 2. *Consider a sequence of k successive single-agent deviations*

$$\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k.$$

Then, the following bounds hold:

$$-\frac{n(n-1)}{2} \leq \sum_{j \in [k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)| \leq \frac{n(n-1)}{2}.$$

Proof. Observe that for all $i \in N$ and all partitions π, we have

$$1 \leq |\pi(i)| \leq n.$$

Thus, we can find the bounds

$$-n(n-1) \leq \sum_{i \in N} |\pi_k(i)| - |\pi_0(i)| \leq n(n-1).$$

Applying Lemma 1 yields the desired result. □

We demonstrate the power of the Deviation Lemma by proving convergence of the dynamics for a variety of deviation types and classes of ASHGs.

Theorem 3. *The dynamics of IS-deviations always converges in ASHGs with at most one nonnegative utility value.*

Proof. Let (N, v) be an ASHG such that the v_i take on at most one nonnegative value. If there are no nonnegative valuations, all IS-deviations are singleton formations, so after at most n deviations, we reach a stable partition. Now, suppose that there is exactly one nonnegative utility value $x \geq 0$. If there are no negative valuations, then in case $x = 0$ we terminate immediately, and in case $x > 0$ the grand coalition will form after at most n^2 deviations. The latter holds because every deviation increases the number of pairs of agents which are part of the same coalition. Thus, we will now assume that in addition to the single nonnegative utility value x, there is at least one negative utility value, and we denote the largest absolute value of a negative utility value by y. Further, define $\Delta = \min\{v_i(C) - v_i(C') \colon i \in N,\, C, C' \in \mathcal{N}_i,\, v_i(C) > v_i(C')\}$. Intuitively, $\Delta > 0$ is the minimum improvement any agent is guaranteed to have when making a NS-deviation. Further, consider the potential function Φ defined by the social welfare of a partition as $\Phi(\pi) = \sum_{i \in N} v_i(\pi)$.

Let us investigate how this potential changes for a single IS-deviation $\pi \xrightarrow{i} \pi'$.

$$\begin{aligned}
\Phi(\pi') - \Phi(\pi) &= \underbrace{v_i(\pi') - v_i(\pi)}_{\text{deviator}} \\
&+ \underbrace{\sum_{j \in \pi'(i) \setminus \{i\}} v_j(\pi') - v_j(\pi)}_{\text{welcoming coalition}} + \underbrace{\sum_{j \in \pi(i) \setminus \{i\}} v_j(\pi') - v_j(\pi)}_{\text{abandoned coalition}} \\
&= v_i(\pi') - v_i(\pi) + \sum_{j \in \pi'(i) \setminus \{i\}} v_j(i) - \sum_{j \in \pi(i) \setminus \{i\}} v_j(i) \\
&= v_i(\pi') - v_i(\pi) + x\left(|\pi'(i)| - 1\right) - \sum_{j \in \pi(i) \setminus \{i\}} v_j(i) \\
&\geq \Delta + x\left(|\pi'(i)| - 1\right) - x\left(|\pi(i)| - 1\right) \\
&= \Delta + x\left(|\pi'(i)| - |\pi(i)|\right).
\end{aligned}$$

The third equality comes from the fact that i performs an IS-deviation, so all agents $j \in \pi'(i) \setminus \{i\}$ must accept i, which means they must have $v_j(i) = x$. Now, let π_0 be any initial partition and consider any sequence of k successive IS-deviations

$$\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \ldots \xrightarrow{i_k} \pi_k.$$

Telescoping and termwise application of the above inequality yields $\Phi(\pi_k) - \Phi(\pi_0) = \sum_{j \in [k]} \Phi(\pi_j) - \Phi(\pi_{j-1}) \geq \sum_{j \in [k]} \Delta + x\left(|\pi_j(i_j)| - |\pi_{j-1}(i_j)|\right) = k\Delta + x \sum_{j \in [k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)|$. We recognize the sum from the Deviation Lemma, which can be bounded from below using Lemma 2:

$$\Phi(\pi_k) - \Phi(\pi_0) \geq k\Delta - x\frac{n(n-1)}{2}. \tag{2}$$

As the right hand side is unbounded in k, the sequence must be finite. To be precise, we can bound the potentials of the initial and final partitions by

$$\Phi(\pi_0) \geq -n(n-1)y, \quad \Phi(\pi_k) \leq n(n-1)x.$$

Substituting in these bounds and rearranging for k gives

$$k \leq \frac{(2y + 3x)n(n-1)}{2\Delta}. \tag{3}$$

□

There are a few important insights gained by the previous proof. First, the bound obtained via the Deviation Lemma does not mean that the potential function Φ is increasing in every round. In fact, since utilities are not necessarily symmetric, the deviating agent might move from a rather large coalition to a smaller coalition only improving her utility by Δ whereas the utility of all agents in the abandoned coalition are decreased by x. In fact, the Deviation Lemma does not give us control of the potential function in a single round. Also, it does not control the utility changes caused by the deviator. We apply it to control the utility changes of agents involved in deviations except for the deviator to obtain Equation (2). Hence, we can bound their utility changes by a global constant solely depending on input data. The utility changes caused by the deviator will then eventually lead to the potential reaching a local maximum.

Second, we can easily obtain polynomial bounds on the running time of the dynamics. If x and y are polynomially bounded in n and all valuations are integer, polynomial running time is directly obtained from Equation (3). In particular, this is the case for FEGs, AFGs, and AEGs, so individually stable partitions can be found in polynomial time for these games. After showing two more applications of the Deviation Lemma for other types of deviations, we will capture this observation in Corollary 1.

Third, the previous theorem is tight in the sense that the dynamics can cycle if we have two nonnegative utility values. Indeed, in an instance with agent set $N = [3]$ and utility values $v_i(j) = 1, v_j(i) = 0$ for $(i, j) \in \{(1, 2), (2, 3), (3, 1)\}$, the dynamics can infinitely cycle among the partitions $\{\{1, 2\}, \{3\}\}$, $\{\{1\}, \{2, 3\}\}$, and $\{\{1, 3\}, \{2\}\}$. However, the partition consisting of the grand coalition is individually stable and can be reached through the dynamics.

Our next application of the Deviation Lemma considers contractual Nash stability, where we obtain a similar result if we allow at most one nonpositive value. The proof is completely analogous and is therefore omitted. Note that this result also breaks down if we simultaneously allow the utility values -1 and 0 by constructing a similar cycle as in the previous example.

Theorem 4. *The dynamics of CNS-deviations always converges in ASHGs with at most one nonpositive utility value.*

Theorems 3 and 4 use the Deviation Lemma to derive positive results for the single-sided unanimity-based stability notions IS and CNS. In a third application of the deviation lemma, we show that this technique is also applicable to majority-based stability notions, at least when we involve both the welcoming and the abandoned coalition in the vote. The key idea is a suitable arrangement of the terms occurring in the difference of the potential with respect to the agents affected by a deviation.

Theorem 5. *The dynamics of JMS-deviations always converges in ASHGs with at most two distinct utility values.*

Note that since every JMS-deviation is also an SMS-deviation, the previous result holds for SMS as well. As in the discussion after Theorem 3, we obtain a polynomial running time of the dynamics for appropriate restrictions of the cases. We collect important consequences in the following corollary. In particular, we extend results by Dimitrov et al. (2006) and Aziz and Brandl (2012) who proved the existence of IS partitions for AFGs and AEGs, respectively.[2]

Corollary 1. *The dynamics of IS-, CNS-, and JMS-deviations always converges in polynomial time in AFGs, AEGs, and FEGs.*

We would like to stress that convergence of the dynamics does not guarantee a polynomial running time in general. An example is the case of symmetric utility values in ASHGs. For NS this can be directly inferred from the PLS-reduction by (Gairing and Savani 2019), which satisfies *tightness*, a property of reductions defined by Schäffer and Yannakakis (1991).

Proposition 1. *The dynamics of NS-deviations in symmetric ASHGs may require exponentially many rounds before converging to an NS partition.*

Proof. It is easy to verify that the PLS-reduction from PartyAffiliation under the Flip neighborhood by Gairing and Savani (2019, Observation 2) is tight. Schäffer and Yannakakis (1991, Lemma 3.3) showed that tight reductions preserve the existence of exponentially long running times of the standard local search algorithm, i.e., the NS-dynamics in our case. Note that the standard local search algorithm of the source problem can have an exponential running time, because PartyAffiliation is a generalization of MaxCut whose standard local search algorithm can run in exponential time with respect to the flip neighborhood (Schäffer and Yannakakis 1991, Theorem 5.15).[3] □

While the previous proposition uses a nonconstructive argument avoiding to construct an explicit example with an exponential running time, it is possible to construct such an example even in the more restricted case of IS-dynamics. To this end, it is possible to modify an example for MaxCut provided by Monien and Tscheuschner (2010) by essentially reverting the sequence of flips for MaxCut to obtain an execution of the IS-dynamics. Thus, we generalize the previous proposition via a constructive proof.

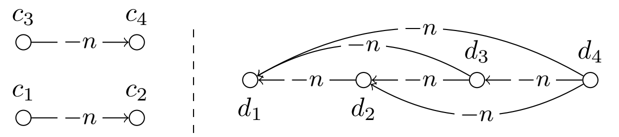

Figure 2: The *aversion to enemies* games without MIS partition (left) and MOS partition (right) from Proposition 3. Omitted edges have weight 1.

Proposition 2. *The dynamics of IS-deviations in symmetric ASHGs may require exponentially many rounds before converging to an IS partition.*

Stability under Majority Consent

In this section, we study stability under majority consent. First, the existential results of Theorem 3 and Theorem 4 are contrasted with the non-existence of stable partitions in AEGs under the majority-based relaxations of the respective stability concepts.

Proposition 3. *There exists an AEG which contains no MIS (respectively, MOS) partition.*

Proof. First, we provide an AEG with no MIS partition. Let $N = \{c_1, c_2, c_3, c_4\}$, i.e., there are $n = 4$ agents, and valuations defined as $v_{c_1}(c_2) = v_{c_3}(c_4) = -n$ and all other valuations set to 1. The AEG is illustrated in Figure 2 (left).

Assume for contradiction that there exists an MIS partition π. Then, $c_1 \notin \pi(c_2)$ and $c_3 \notin \pi(c_4)$. Also, $|\pi(c_1)| \leq 1$ (respectively, $|\pi(c_3)| \leq 1$), because otherwise, c_2 (respectively, c_4) would join via an MIS-deviation). But then $\pi(c_1) = \{c_1\}$ and $\pi(c_3) = \{c_3\}$, and c_1 could deviate to join $\pi(c_3)$, a contradiction.

Second, we provide an AEG without MOS partition. Let $N = \{d_1, d_2, d_3, d_4\}$, and define valuations for all $i, j \in [4]$ with $i < j$ as $v_{d_i}(d_j) = 1$ and $v_{d_j}(d_i) = -4$. An illustration is provided in Figure 2 (right).

Assume for contradiction that there exists an MOS partition π. Then, every coalition $C \in \pi$ must fulfill $|C| \leq 2$. Otherwise, the agent of C with the second smallest index would form a singleton via an MOS-deviation. In addition, there cannot be a singleton, because if some agent is in a singleton, there must be a second such agent, and then the one with the smaller index would join the other one. Hence, π consists of two pairs. But then d_1 would deviate to the pair not containing her, a contradiction. □

We can leverage the AEGs provided in the previous proposition as gadgets in reductions to show hardness of the associated decision problems. This can be interpreted as a more exact boundary (compared to Theorem 1) of the tractabilities encountered in Theorem 3 and Theorem 4 for the special case of AEGs.

Theorem 6. *It is* NP*-complete to decide if there exists an MIS (respectively, MOS) partition in AEGs.*

[2] These contributions actually show existence of partitions satisfying properties stronger than IS.

[3] We refer to the respective references for formal definitions of the involved combinatorial problems.

The utility restrictions in Theorem 6 are not as flexible as in the negative result for Nash stability in Theorem 1 or the positive results for unanimity-based dynamics in Theorem 3 and Theorem 4. In fact, the picture for majority-based notions is more diverse, because we obtain another positive result in the class of AFGs.

Theorem 7. *When starting from the grand coalition, the dynamics of MIS-deviations converges after at most n rounds in AFGs.*

Proof. The key insight is that there can only be deviations to form a new singleton coalition yielding no more than n deviations. Let $\pi_0 = \{N\}$ be the initial partition, and consider a sequence of k MIS-deviations

$$\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k.$$

We inductively define coalitions evolving from the grand coalition if removing the deviator as $G_0 = N$, and $G_j = G_{j-1} \setminus \{i_j\}$ for $j > 0$.

Now, we proceed to simultaneously prove the following claims by induction:

1. $\forall j \in [k] : \pi_{j-1}(i_j) = G_{j-1}$.
2. $\forall j \in [k] : \pi_j(i_j) = \{i_j\}$.
3. $\forall j \in [k] : \{i \in \pi_{j-1}(i_j) \colon v_{i_j}(i) = n\} = \emptyset$.

The base case $j = 1$ is immediate. For the induction step, let $2 \leq j \leq k$ and suppose the claims are true for all $1 \leq l < j$. We start with the first claim. By the induction hypothesis, $\pi_{j-1} = \{G_{j-1}\} \cup \{\{i_l\} \colon 1 \leq l < j\}$. This means that if $\pi_{j-1}(i_j) \neq G_{j-1}$, we must have $\pi_{j-1}(i_j) = \{i_j\}$, indicating $i_j = i_l$ for some $l < j$. Then, the welcoming coalition cannot be G_{j-1}, as i_j, by induction hypothesis, abandoned G_{l-1} due to not having any friends in G_{l-1}, and thus has, by $G_{j-1} \subseteq G_{l-1}$, no friends in G_{j-1}, either. The alternative is that i_j joins another singleton coalition $\{i_m\}$ to form a pair. However, if i_m abandoned G_m at some point $m < l$, then she dislikes i_j, and won't allow her to join. If i_m abandoned G_m at some point $m > l$, then i_j dislikes i_m, and has no incentive to join. Hence, $\pi_{j-1}(i_j) = G_{j-1}$. For the second claim, note that i_j cannot join another singleton $\{i_m\}$, because i_m abandoned G_{m-1} at some point $m < j$ and thus dislikes i_j. Hence, i_j must form a singleton $\pi_j(i_j) = \{i_j\}$, which she only wants to do if $\{i \in \pi_{j-1}(i_j) \colon v_{i_j}(i) = n\} = \emptyset$. This accomplishes the third claim, and completes the induction proof.

Finally, as there can be at most n singletons, the dynamics must terminate after at most n rounds. □

The computational boundaries in this section encountered so far only hold for one-sided stability notions where either the welcoming or the abandoned coalition takes a vote. On the other hand, Theorem 5 shows that these are opposed by tractabilities under two-sided majority consent.

For general utilities, existence of SMS (and therefore JMS) partitions is not guaranteed anymore, and we show that the tractabilities break down.

Theorem 8. *Deciding whether an ASHG contains an SMS (respectively, JMS) partition is* NP*-complete.*

Conclusion and Discussion

We studied stability based on single-agent deviations in additively separable hedonic games with a particular focus on games with restricted utility functions that can be naturally interpreted in terms of friends and enemies. We identified a computational boundary between Nash stability and stability with unanimous consent. The picture is less clear when deviations are governed by majority consent. While stable partitions always exist when considering both the abandoned and the welcoming coalition of the deviating agent, we obtain both positive and negative results if only one of these coalitions is considered. Table 1 summarizes our results and compares them with related results from the literature. Notably, we obtain all of our positive results through the convergence of simple and natural dynamics. This also extends previously known results about IS. Aziz and Brandl (2012) obtain a polynomial algorithm essentially by running a dynamics from the singleton partition, whereas Dimitrov et al. (2006) take a different, graph-theoretical approach considering strongly connected components. The construction of CIS partitions by Aziz, Brandt, and Seedig (2013) is done by iteratively identifying specific coalitions, and it is not known whether CIS-dynamics converge in polynomial time for natural starting partitions such as the singleton partition or grand coalition. An important tool in establishing our results concerning convergence of dynamics is the Deviation Lemma, a general combinatorial insight that allows us to study dynamics from a global perspective.

Table 1: Overview of our computational results. A red cell means existence of games without stable partition and usually comes with computational intractability. A green cell means that a stable partition can be constructed in polynomial time (Function-P), and in the case of our results even by executing a dynamics. A white cell means that it is unknown whether a stable partition always exists.
[a]: Aziz and Brandl (2012), [b]: Aziz, Brandt, and Seedig (2013), [c]: Dimitrov et al. (2006), [d]: Sung and Dimitrov (2010)

	General	$FEGs$	$AEGs$	$AFGs$
NS	NP-c[d]	NP-c (Th. 1)	NP-c (Th. 1)	NP-c (Th. 2)
IS	NP-c[d]	FP (Th. 3)	FP[a] (Th. 3)	FP[c] (Th. 3)
CNS	NP	FP (Th. 4)	FP (Th. 4)	FP (Th. 4)
CIS	FP[b]	FP[b]	FP[b]	FP[b]
MIS	NP-c (Th. 6)	?	NP-c (Th. 6)	FP (Th. 7)
MOS	NP-c (Th. 6)	?	NP-c (Th. 6)	?
JMS	NP-c (Th. 8)	FP (Th. 5)	FP (Th. 5)	FP (Th. 5)
SMS	NP-c (Th. 8)	FP (Th. 5)	FP (Th. 5)	FP (Th. 5)

Our work offers a wide range of interesting follow-up questions. First, Table 1 contains some problems left open in our analysis. Specifically, despite the existence of partitions without CNS partitions, the complexity of the existence problem of CNS partitions remains open for general utilities. Also, the voting-based stability notions deserve further investigation, and might even lead to interesting discoveries in other classes of hedonic games. Lastly, an intriguing further direction is to study further applications of the Deviation Lemma, particularly in domains other than coalition formation.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft under grant BR 2312/12-1.

References

Alcalde, J.; and Revilla, P. 2004. Researching with whom? Stability and manipulation. *Journal of Mathematical Economics*, 40(8): 869–887.

Aziz, H.; and Brandl, F. 2012. Existence of Stability in Hedonic Coalition Formation Games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 763–770.

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

Aziz, H.; Brandt, F.; and Seedig, H. G. 2013. Computing Desirable Partitions in Additively Separable Hedonic Games. *Artificial Intelligence*, 195: 316–334.

Aziz, H.; and Savani, R. 2016. Hedonic Games. In Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D., eds., *Handbook of Computational Social Choice*, chapter 15. Cambridge University Press.

Banerjee, S.; Konishi, H.; and Sönmez, T. 2001. Core in a simple coalition formation game. *Social Choice and Welfare*, 18: 135–153.

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

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

Brandt, F.; and Bullinger, M. 2020. Finding and Recognizing Popular Coalition Structures. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 195–203.

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

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

Cechlárová, K.; and Romero-Medina, A. 2001. Stability in Coalition Formation games. *International Journal of Game Theory*, 29: 487–494.

Cseh, Á. 2017. Popular Matchings. In Endriss, U., ed., *Trends in Computational Social Choice*, chapter 6. AI Access.

Dimitrov, D.; Borm, P.; Hendrickx, R.; and Sung, S. C. 2006. Simple Priorities and Core Stability in Hedonic Games. *Social Choice and Welfare*, 26(2): 421–433.

Dimitrov, D.; and Sung, S. C. 2007. On top responsiveness and strict core stability. *Journal of Mathematical Economics*, 43(2): 130–134.

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

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

Gärdenfors, P. 1975. Match Making: Assignments based on bilateral preferences. *Behavioral Science*, 20(3): 166–173.

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

Karp, R. M. 1972. Reducibility among Combinatorial Problems. In Miller, R. E.; and Thatcher, J. W., eds., *Complexity of Computer Computations*, 85–103. Plenum Press.

Monien, B.; and Tscheuschner, T. 2010. On the power of nodes of degree four in the local max-cut problem. In *Proceedings of the 7th International Conference on Algorithms and Complexity (CIAC)*, number 6078 in Lecture Notes in Computer Science (LNCS), 264–275. Springer-Verlag.

Newman, M. E. J. 2004. Detecting community structure in networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, 38(2): 321–330.

Olsen, M. 2012. On defining and computing communities. In *Proceedings of the 18th Computing: Australasian Theory Symposium (CATS)*, volume 128 of *Conferences in Research and Practice in Information Technology (CRPIT)*, 97–102.

Saad, W.; Han, Z.; Basar, T.; Debbah, M.; and Hjorungnes, A. 2011. Hedonic Coalition Formation for Distributed Task Allocation among Wireless Agents. *IEEE Transactions on Mobile Computing*, 10(9): 1327–1344.

Schäffer, A. A.; and Yannakakis, M. 1991. Simple Local Search Problems that are Hard to Solve. *SIAM Journal on Computing*, 20(1): 56–87.

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

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

Missing Proofs

In the appendix, we provide the proofs missing in the body of the paper.

Theorem 1. *Let $f^+ : \mathbb{N} \to \mathbb{Q}_{>0}$ and $f^- : \mathbb{N} \to \mathbb{Q}_{<0}$ be two polynomial-time computable functions satisfying $|f^-(m)| \geq f^+(m)$ for all $m \in \mathbb{N}$. Then, the problem of deciding whether an ASHG with utility values restricted to $\{f^-(n), f^+(n)\}$ has an NS partition is* NP*-complete.*

Proof. Let f^+, f^- be two functions as defined above and consider the class of ASHGs with utility values restricted to $\{f^-(n), f^+(n)\}$. We provide a reduction from the NP-complete problem EXACT3COVER (E3C) (Karp 1972). An instance of EXACT3COVER consists of a tuple (R, S), where R is a ground set together with a set S of 3-element subsets of R. A 'yes'-instance is an instance such that there exists a subset $S' \subseteq S$ that partitions R. Given an instance (R, S) of E3C, for every $r \in R$, we define $S_r = \{s \in S : r \in s\}$, i.e., S_r comprises the elements of S containing r, and $n_r = |S_r|$.

Now, let (R, S) be an instance of E3C. We produce an ASHG (N, v) satisfying $v_i(j) \in \{f^-(n), f^+(n)\}$ for all $i, j \in N$ such that (R, S) has an exact cover if and only if (N, v) has an NS partition π. Define the agent set as $N = \{c\} \cup \bigcup_{s \in S} A^s \cup \bigcup_{r \in R}\{b_i^r : i \in [n_r - 1]\}$, where $A^s = \left\{a_{r_1}^s, a_{r_2}^s, a_{r_3}^s, a^s\right\}$ for $s = \{r_1, r_2, r_3\} \in S$. Hence, the agent set consists of copies of the elements in R corresponding to the frequency they occur in the sets of S minus 1, copies for the elements in sets of S together with one specific agent for each such set, and an auxiliary agent c. Now, define the following valuations v:

- For each $s \in S, a \neq a' \in A^s : v_a(a') = f^+(n)$.
- For each $r \in R, s \in S_r, i \in [n_r - 1] : v_{a_r^s}(b_i^r) = v_{b_i^r}(a_r^s) = v_{b_i^r}(c) = f^+(n)$.
- All other valuations are $f^-(n)$.

This reduction can be performed in polynomial time, as there are at most $4|S| + |R||S| + 1$ agents, and f^+, f^- can be computed in polynomial time. We claim that (R, S) admits an exact cover $S' \subseteq S$ if and only if (N, v) has an NS partition π.

$\Longrightarrow$: Suppose (R, S) has an exact cover $S' \subseteq S$. We construct an NS partition π.

- We have coalitions corresponding to the cover, i.e., for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \dots, s_{n_r - 1}$ and define for each $i \in [n_r - 1]$ the coalition $\{a_r^{s_i}, b_i^r\}$.
- All agents a^s with $A^s \notin \pi$ are in a singleton: $\pi(a^s) = \{a^s\}$.
- Agent c is also in a singleton: $\pi(c) = \{c\}$.

To see that this partition is NS, we perform a case analysis for the various types of agents in order to show that no agent has an incentive to deviate.

- An agent a with $\pi(a) = A^s$ has $v_a(\pi) = 3f^+(n)$, whereas every other coalition contains at most one agent she likes. So she has no incentive to deviate.
- An agent a_r^s with $\pi(a_r^s) \neq A^s$ is in a pair with an agent b_i^r, and so are the other two agents $a_{r'}^s$ from A^s. Thus, $v_{a_r^s}(\pi) = f^+(n)$, whereas every other coalition contains at most one agent she likes. So she has no incentive to deviate.
- An agent a^s with $\pi(a^s) \neq A^s$ is alone, but all other agents $a_r^s \in A^s$ are in a pair with an agent b_i^r that she dislikes, and as $f^+(n) + f^-(n) \leq 0$, she has no incentive to deviate.
- An agent b_i^r is in a pair with an agent a_r^s, so she has $v_{b_i^r}(\pi) = f^+(n)$. The best alternative would be joining c, which does not yield an improvement in utility, so she has no incentive to deviate.
- Finally, c has $v_c(\pi) = 0$, which is her most desired outcome, as she dislikes all other agents.

Together, we conclude that π is NS.

$\Longleftarrow$: Suppose (N, v) contains an NS partition π. We show that then there must be an exact cover $S' \subseteq S$ of R. We begin with some observations:

1. Agent c must be in a singleton coalition, otherwise she would deviate to a singleton coalition.
2. Agents b_i^r must have utility $v_{b_i^r}(\pi) \geq f^+(n)$, otherwise they would join $\{c\}$.
3. Coalitions of agents a^s satisfy $\pi(a^s) \cap A^{s'} = \emptyset$ for $s' \neq s$. Suppose for contradiction that there is an agent $a \in \pi(a^s) \cap A^{s'}$. Consider the sets $A = \{i \in \pi(a^s) : v_a(i) = f^+(n)\}$ and $A' = \{i \in \pi(a^s) : v_{a^s}(i) = f^+(n)\}$. Then, we have $A \cap A' = \emptyset$. If $|A| \leq |A'|$, then a has an incentive to deviate to a singleton as she dislikes all agents from A' as well as a^s. Similarly, if $|A'| \leq |A|$, then a^s has an incentive to form a singleton coalition as she dislikes all agents from A as well as a.
4. Using observation 3, we must have $\pi(a^s) \neq \pi(b_i^r)$, as otherwise $v_{b_i^r}(\pi) \leq 0$, contradicting observation 2. Hence, we have $\pi(a^s) \subseteq A^s$ for all $s \in S$.
5. Now, consider an agent b_i^r. Define the sets $A = \{a_r^s : s \in S_r\}$ and $B = \left\{b_j^r : j \in [n_r - 1]\right\}$. By observation 2, we must have $|A \cap \pi(b_i^r)| \geq |\pi(b_i^r) \setminus A|$. We show that we must have $|A \cap \pi(b_i^r)| = |\pi(b_i^r) \setminus A|$. Suppose for contradiction that $|A \cap \pi(b_i^r)| > |\pi(b_i^r) \setminus A|$. Then, each agent $a_r^s \in A \cap \pi(b_i^r)$ has $v_{a_r^s}(\pi) \leq 0$ and would, by observation 4, rather deviate to $\pi(a^s)$. Moreover, we show that we must have $\pi(b_i^r) \setminus A \subseteq B$. Suppose for contradiction that this is not true. Then there are two cases. In the first case, there is an agent $b_j^{r'} \in \pi(b_i^r) \setminus A$ with $r \neq r'$. This agent dislikes all agents in A, and so would rather deviate to join $\{c\}$. In the second case, there is an agent $a_{r'}^s \in \pi(b_i^r) \setminus A$ with $r \neq r'$. This agent dislikes all but one agent from A as well as b_i^r, so would rather deviate to join $\pi(a^s)$.

Observation 5 shows that coalitions of agents b_i^r are of the form $A \uplus B$, where $A \subseteq \{a_r^s : s \in S_r\}, B \subseteq \left\{b_j^r : j \in [n_r - 1]\right\}$ and $|A| = |B|$. This leaves for each $r \in R$ exactly one agent a_r^s that is not in such a coalition.

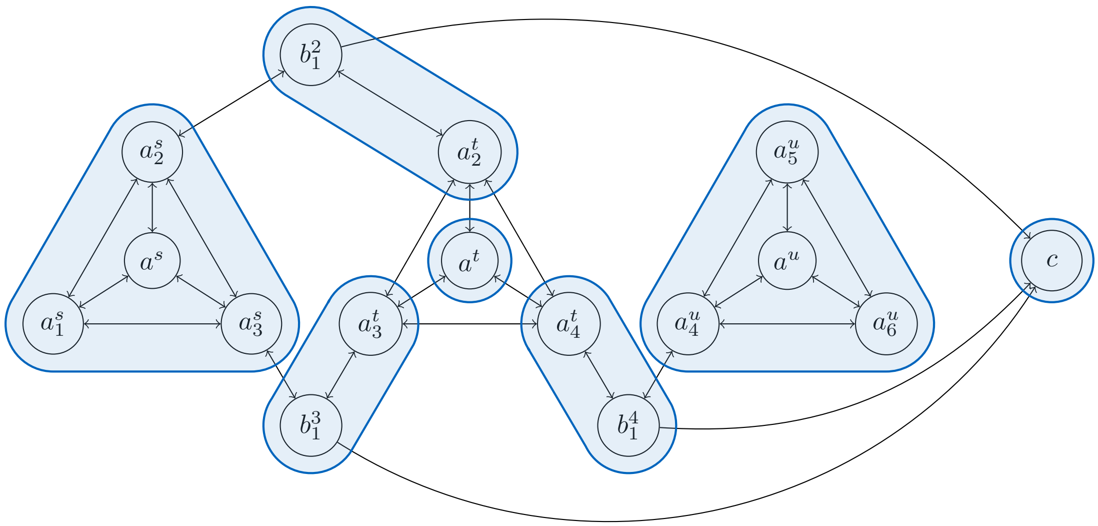

Figure 3: The reduction from the proof of Theorem 1 for the Yes-instance of E3C $(\{1,\dots,6\},\{s,t,u\})$ with $s=\{1,2,3\}, t=\{2,3,4\}$ and $u=\{4,5,6\}$. Drawn edges have weight $f^+(n)$, omitted edges have weight $f^-(n)$. The partition corresponding to the exact cover $\{s,u\}$ is highlighted.

For these agents we have $\pi(a_r^s) = A^s$, yielding a cover $S' = \{s \in S \colon A^s \in \pi\}$. $\square$

The proof of the next theorem is similar to the proof of Theorem 1. The essential difference is that we represent now every element in the ground set of an E3C-instance by a pair of agents.

Theorem 2. *Deciding whether an AFG has an NS partition is* NP*-complete.*

Proof. We provide another reduction from E3C. Let (R,S) be an instance of E3C. We produce an AFG (N,v) such that (R,S) has an exact cover if and only if (N,v) has a NS partition π. Define the agent set $N = \{d\} \cup \bigcup_{s\in S} A^s \cup \bigcup_{r\in R}(\{c_1^r, c_2^r\} \cup \{b_i^r \colon i \in [n_r-1]\})$, where $A^s = \{a_r^s \colon r \in s\}$ for $s \in S$.

Also, define the following valuations v:

- For each $s \in S, a \neq a' \in A^s : v_a(a') = n$.
- For each $r \in R, s \in S_r, i \in [n_r - 1] : v_{a_r^s}(b_i^r) = v_{b_i^r}(a_r^s) = v_{b_i^r}(d) = n$.
- For each $r \in R, s \in S_r : v_{c_1^r}(a_r^s) = v_{c_1^r}(c_2^r) = v_{c_2^r}(c_1^r) = v_{c_2^r}(d) = n$.
- All other valuations are -1.

This reduction can be performed in polynomial time, as there are only polynomially many agents. We now claim that (R,S) has an exact cover $S' \subseteq S$ if and only if (N,v) has a NS partition π.

$\implies$: Suppose (R,S) has an exact cover $S' \subseteq S$. We construct a NS partition π.

- We have coalitions corresponding to the cover, i.e., for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \dots, s_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalition $\{a_r^{s_i}, b_i^r\}$.
- For each $r \in R$ with $n_r > 1$, the agents c_1^r and c_2^r are in a pair $\{c_1^r, c_2^r\}$.
- Agent d is in a singleton $\{d\}$.

In this partition, each agent is together with some number of friends and no enemies. Every alternative coalition has at most as many friends as the current coalition, so no agent has incentive to deviate.

$\impliedby$: Suppose (N,v) has a NS partition π. We show that then there must be an exact cover $S' \subseteq S$ of R. We begin with some observations:

1. Agent d must be in a singleton coalition, because her value for any other agent is negative.
2. An agent c_2^r must be in a pair with c_1^r, otherwise she would join $\{d\}$.
3. An agent b_i^r must be in a coalition with at least one agent a_r^s, otherwise she would join $\{d\}$.
4. Agents a_r^s and $a_r^{s'}$ with $s \neq s'$ must be in distinct coalitions, otherwise c_1^r would join them.
5. Combining observations 3 and 4, we get that each agent b_i^r must be in a pair with exactly one agent a_r^s.

We now know that for each $r \in R$, exactly $n_r - 1$ of the agents a_r^s must be in pairs with agents b_i^r. This leaves exactly one agent a_r^s not in a pair. For these agents we have $\pi(a_r^s) = A^s$, yielding a cover $S' = \{s \in S \colon A^s \in \pi\}$. □

Theorem 5. *The dynamics of JMS-deviations always converges in ASHGs with at most two distinct utility values.*

Proof. Let (N, v) be an ASHG such that the v_i take on at most two distinct values, and consider once again the potential

$$\Phi(\pi) = \sum_{i \in N} v_i(\pi).$$

If the v_i take on only one value or both values are nonnegative (resp., nonpositive), convergence is clear, as Φ increases with every JMS-deviation. So suppose that the v_i are restricted to $\{-y, x\}$ with $y > 0$ and $x > 0$. As in the proof of Theorem 3, set $\Delta = \min\{v_i(C) - v_i(C') \colon i \in N,\, C, C' \in \mathcal{N}_i,\, v_i(C) > v_i(C')\}$.

Let us now investigate a single JMS-deviation $\pi \xrightarrow{i} \pi'$. To reduce notational clutter, set $F_{\text{in}} = F_{\text{in}}(\pi(i), i), F_{\text{out}} = F_{\text{out}}(\pi(i), i), F'_{\text{in}} = F_{\text{in}}(\pi'(i), i)$, and $F'_{\text{out}} = F_{\text{out}}(\pi'(i), i)$. Note that, by definition of a JMS-deviation, we have $|F'_{\text{in}}| + |F_{\text{out}}| \geq |F'_{\text{out}}| + |F_{\text{in}}|$, from which we can conclude

$$|F'_{\text{in}}| - |F_{\text{in}}| \geq \frac{|F'_{\text{in}}| - |F_{\text{in}}| + |F'_{\text{out}}| - |F_{\text{out}}|}{2} \geq |F'_{\text{out}}| - |F_{\text{out}}|.$$

Further, note that due to restriction of the utility values to $\{-y, x\}$, we have

$$\forall j \in F_{\text{in}} \cup F'_{\text{in}} : v_j(i) = x, \forall j \in F_{\text{out}} \cup F'_{\text{out}} : v_j(i) = -y$$

and

$$|F_{\text{in}}| + |F_{\text{out}}| = |\pi(i)| - 1, \quad |F'_{\text{in}}| + |F'_{\text{out}}| = |\pi'(i)| - 1.$$

Combining with our inequality from above, we obtain

$$|F'_{\text{in}}| - |F_{\text{in}}| \geq \frac{|\pi'(i)| - |\pi(i)|}{2} \geq |F'_{\text{out}}| - |F_{\text{out}}|.$$

The change in Φ through the JMS-deviation can then be bounded as

$$\begin{aligned}
&\Phi(\pi') - \Phi(\pi) = \underbrace{v_i(\pi') - v_i(\pi)}_{\text{deviator}} \\
&+ \underbrace{\sum_{j \in \pi'(i) \setminus \{i\}} v_j(\pi') - v_j(\pi)}_{\text{welcoming coalition}} + \underbrace{\sum_{j \in \pi(i) \setminus \{i\}} v_j(\pi') - v_j(\pi)}_{\text{abandoned coalition}} \\
&= v_i(\pi') - v_i(\pi) + \sum_{j \in \pi'(i) \setminus \{i\}} v_j(i) - \sum_{j \in \pi(i) \setminus \{i\}} v_j(i) \\
&= v_i(\pi') - v_i(\pi) + x|F'_{\text{in}}| - y|F'_{\text{out}}| - x|F_{\text{in}}| + y|F_{\text{out}}| \\
&= v_i(\pi') - v_i(\pi) + x\left(|F'_{\text{in}}| - |F_{\text{in}}|\right) - y\left(|F'_{\text{out}}| - |F_{\text{out}}|\right) \\
&\geq \Delta + x\frac{|\pi'(i)| - |\pi(i)|}{2} - y\frac{|\pi'(i)| - |\pi(i)|}{2}.
\end{aligned}$$

Now, let π_0 be any initial partition and consider any sequence of k successive JMS-deviations

$$\pi_0 \xrightarrow{i_1} \pi_1 \xrightarrow{i_2} \dots \xrightarrow{i_k} \pi_k.$$

Telescoping and termwise application of the above inequality gives

$$\begin{aligned}
&\Phi(\pi_k) - \Phi(\pi_0) = \sum_{j \in [k]} \Phi(\pi_j) - \Phi(\pi_{j-1}) \\
&\geq \sum_{j \in [k]} \Delta + x\frac{|\pi_j(i_j)| - |\pi_{j-1}(i_j)|}{2} - y\frac{|\pi_j(i_j)| - |\pi_{j-1}(i_j)|}{2} \\
&= k\Delta + \frac{x - y}{2} \sum_{j \in [k]} |\pi_j(i_j)| - |\pi_{j-1}(i_j)|.
\end{aligned}$$

The sum from Lemma 1 appears for prefactors of different sign, and can be bounded using Lemma 2:

$$\begin{aligned}
\Phi(\pi_k) - \Phi(\pi_0) &\geq k\Delta - \frac{x + y}{2}\frac{n(n-1)}{2} \\
&= k\Delta - \frac{(x + y)n(n-1)}{4}.
\end{aligned}$$

As the right hand side is unbounded in k, the sequence must be finite. To be precise, we can bound the potentials of the initial and final partitions by

$$\Phi(\pi_0) \geq -n(n-1)y, \quad \Phi(\pi_k) \leq n(n-1)x.$$

Substituting in these bounds and rearranging for k gives

$$k \leq \frac{(5x + 5y)n(n-1)}{4\Delta}.$$

□

Proposition 2. *The dynamics of IS-deviations in symmetric ASHGs may require exponentially many rounds before converging to an IS partition.*

Proof. Let the agent set be $N = \{u_1, u_2, v_0\} \cup_{i=1}^{n} N_i$ with $N_i = \{v_i, u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4}\}$, and consider the symmetric ASHG on this set of agents with utility values induced by the graph presented in Figure 4, where the weights of the building component G_i are depicted in Figure 5. More precisely, the weight function is given by $f_n^i(k) = k + 5(2^{n-i+1} - 1)$.[4] All weights on missing edges are 0.

The underlying combinatorial structure consists of a short path G_0 together with n copies of the same graph with exponentially growing weights. Graph G_{i-1} and G_i are connected by an edge $\{v_{i-1}, u_{i,1}\}$.

Consider the partition of N indicated by the blue and green vertices and defined by $\pi = \{\{u_1, v_0\} \cup \bigcup_{i=1}^{n}\{v_i, u_{i,2}, u_{i,4}\}, \{u_2\} \cup \bigcup_{i=1}^{n}\{u_{i,1}, u_{i,3}\}\}$. We claim that there is an execution of the IS-dynamics starting with π where agent v_i performs 2^{i+1} deviations for $i \in \{0, 1, \dots, n\}$.

We will recursively construct such a sequence of deviations. In the i-th step of the recursion, agent v_i will already perform 2^{i+1} deviations, and no agent in N_j will performs a deviation for $j > i$. Then, we will insert appropriate subsequences propagating through the graph. These insertions

[4] Note that there is a typo in the weight function by Monien and Tscheuschner (2010). Probably they meant a similar weight function as the one used here.

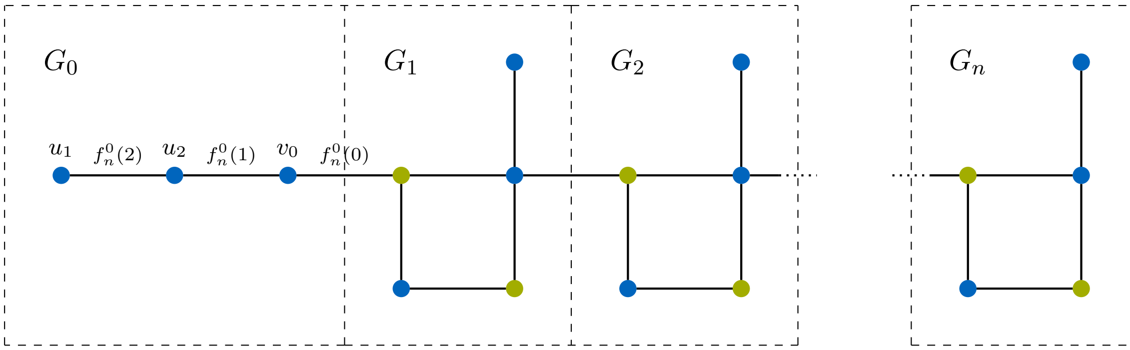

Figure 4: Exponential length IS-dynamics inspired by Monien and Tscheuschner (2010).

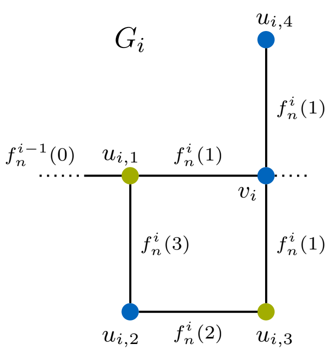

Figure 5: Weights of the building component G_i

change the coalition $u_{i+1,1}$ was part of when v_i performs an IS-deviation. However, this is not a problem, because the IS-deviations of v_i are valid independently of the coalition that $u_{i+1,1}$ is part of. For $i = 0$, consider the sequence of deviations performed by (v_0, u_2, v_0), where v_0 performs $2 = 2^{0+1}$ deviations.

Now, let $k \geq 1$ and assume that the sequence is constructed for $k-1$. We extend the sequence of deviations by inserting suitable subsequences. right before v_{k-1} performs her m-th deviation, then we insert

$$\begin{cases} (v_k, u_{k,3}, v_k, u_{k,2}, u_{k,3}, v_k, u_{k,1}) & \text{if } m \text{ odd} \\ (u_{k,2}, v_k, u_{k,1}) & \text{if } m \text{ even} \end{cases}$$

By the choice of the utility values and the initial partition, this sequence consists of NS-deviations. Since all edge utility values are nonnegative, the sequence consists indeed of IS-deviations. The most interesting deviations to check are the ones performed by agents $u_{i,1}$. Whenever they perform a deviation, they leave the coalition of $u_{i,2}$ and v_i to join the coalition of v_{i-1}. Indeed, this yields an improvement in utility, because $f_n^{i-1}(0) = 5(2^{n-i+2}-1) > 4+10(2^{n-i+1}-1) = f_n^i(3) + f_n^i(1)$. Note that after every even m, the subpartition of vertices in G_k is the same as in the initial partition π. Moreover, the agent v_k performs 2^{k+1} deviations.

In particular, for $k = n$, we have found an ASHG with a number of agents linear in n and (exponential) utility values which also require polynomial space. However, the constructed execution of the IS-dynamics takes exponentially many rounds.[5] □

Theorem 6. *It is* NP*-complete to decide if there exists an MIS (respectively, MOS) partition in AEGs.*

We split the proof into two separate reductions provided in Lemma 3 and Lemma 4. We start with the proof for MIS-stability.

Lemma 3. *It is* NP*-complete to decide if there exists an MIS partition in AEGs.*

Proof. By reduction from E3C. Let (R, S) be an instance of E3C. We produce an AEG (N, v) such that (R, S) admits an exact cover if and only if (N, v) contains an MIS partition. Define $N = \bigcup_{s\in S} A^s \cup \bigcup_{r\in R} \bigcup_{i=1}^{n_r-1} B_i^r$, where $A^s = \{a_{r_1}^s, a_{r_2}^s, a_{r_3}^s, a^s\}$ for $s = \{r_1, r_2, r_3\} \in S$, and $B_i^r = \{b_{i,j}^r : j \in [4]\}$ for $r \in R, i \in [n_r - 1]$. Define valuations v as:

- For each $s \in S, a \neq a' \in A^s$: $v_a(a') = 1$.
- For each $r \in R, s \in S_r, i \in [n_r - 1]$: $v_{a_r^s}(b_{i,1}^r) = v_{b_{i,1}^r}(a_r^s) = 1$.
- Each B_i^r has internal valuations as in the first example of Proposition 3, i.e., if v' denotes the valuations of this example, then $v_{b_{i,j}^r}(b_{i,k}^r) = v'_{c_j}(c_k)$, where the negative valuations are adapted to the specific number of agents in the instance.
- All other valuations are $-n$.

We proceed to prove correctness of the reduction.

$\Longrightarrow$: Suppose (R, S) has an exact cover $S' \subseteq S$. We construct an MIS partition π as follows.

- We have coalitions corresponding to the cover, i.e. for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \ldots, s_{n_r-1}$ and define for each $i \in [n_r - 1]$ the coalitions $\{a^{s_i}\}$, $\{a_r^{s_i}, b_{i,1}^r\}$, $\{b_{i,2}^r, b_{i,4}^r\}$, and $\{b_{i,3}^r\}$.

[5] Note that it is necessary in this example that the edge weights grow exponentially. If they were polynomially bounded, then the IS-dynamics would run in polynomial time, because every deviation increases the social welfare.

No agent has an incentive to deviate, making the partition NS and thus MIS.

$\Longleftarrow$: Suppose (N, v) has an MIS partition π. We construct an exact cover $S' \subseteq S$. We begin with some observations:

1. No agent is in a coalition with someone she dislikes, otherwise she would deviate to a singleton coalition. In particular, this means $\pi(a^s) \subseteq A^s$ and $\pi(b^r_{i,j}) \subseteq B^r_i$ for $j \in \{2,3,4\}$.
2. Each agent of type $b^r_{i,1}$ must be in a coalition with exactly one agent a^s_r. If $\pi(b^r_{i,1}) \subseteq B^r_i$, we would contradict the fact that the subgame induced by B^r_i has no stable partition (see Proposition 3). As $b^r_{i,1}$ cannot form a coalition with someone she dislikes, at least one agent c of the type a^s_r must be in her coalition. Finally, no other agent giving positive utility to $b^r_{i,1}$ can be in a common coalition with c.

Now, we know that for each $r \in R$, exactly $n_r - 1$ of the agents a^s_r must be in pairs with $b^r_{i,1}$. This leaves exactly one agent a^s_r not in a pair. We claim that for these agents we have $\pi(a^s_r) = A^s$. Indeed, it is clear that we then must have $\pi(a^s_r) \subseteq A^s$. If $\pi(a^s_r) = \{a^s_r\}$, she would deviate to join $\pi(a^s)$. Then, $|\pi(a^s_r)| \geq 2$, and members from $A^s \setminus \pi(a^s_r)$ would have an incentive to join $\pi(a^s_r)$. It follows that $A^s \setminus \pi(a^s_r) = \emptyset$, and therefore $\pi(a^s_r) = A^s$. Hence, we obtain a cover $S' = \{s \in S \colon A^s \in \pi\}$. $\square$

Note that it can be shown that a partition in the reduced instances in the reduction of the previous lemma is NS if and only if it is MIS. Hence, the lemma provides yet another proof to the respective statement about Nash stability first shown by Sung and Dimitrov (2010) (and already revisited in Theorem 1). We proceed with the complementary proof for MOS-stability.

Lemma 4. *It is* NP*-complete to decide if there exists an MOS partition in AEGs.*

Proof. Again, we reduce from E3C. Let (R, S) be an instance of E3C. We produce an AEG (N, v) with agent set $N = \bigcup_{s \in S} A^s \cup \bigcup_{r \in R} \bigcup_{i=1}^{n_r - 1} B^r_i$, where $A^s = \left\{a^s_{r_1}, a^s_{r_2}, a^s_{r_3}, a^s\right\}$ for $s = \{r_1, r_2, r_3\} \in S$, and $B^r_i = \left\{b^r_{i,j} \colon j \in [4]\right\}$ for $r \in R, i \in [n_r - 1]$. Define the following valuations v:

- For each $s \in S, a \neq a' \in A^s$: $v_a(a') = 1$.
- For each $r \in R, s \in S_r, i \in [n_r - 1]$: $v_{a^s_r}(b^r_{i,1}) = 1$.
- Each B^r_i has internal valuations as in the second example constructed in the proof of Proposition 3, i.e., if v' are the valuations from this example, then $v_{b^r_{i,j}}(b^r_{i,k}) = v'_{d_j}(d_k)$, where the negative valuations are adapted to the specific number of agents in the instance.
- All other valuations are $-n$.

We claim that (R, S) has an exact cover if and only if (N, v) has an MOS partition.

$\Longrightarrow$: Suppose (R, S) has an exact cover $S' \subseteq S$. We construct an MOS partition π.

- We have coalitions corresponding to the cover, i.e. for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \dots, s_{n_r - 1}$ and define for each $i \in [n_r - 1]$ the coalitions $\{a^{s_i}\}$, $\{a^{s_i}_r, b^r_{i,1}\}$, $\{b^r_{i,2}, b^r_{i,3}\}$, and $\{b^r_{i,4}\}$.

The only agents that have an incentive to deviate are agents of types $b^r_{i,1}$ and $b^r_{i,3}$. However, there is some $s \in S$ such that $\pi(b^r_{i,1}) = \{b^r_{i,1}, a^s_r\}$, and a^s_r ensures that b^r_i cannot leave. Similarly, $\pi(b^r_{i,3}) = \{b^r_{i,3}, b^r_{i,2}\}$, and $b^r_{i,2}$ ensures that $b^r_{i,3}$ cannot leave. Note that agents a^s for $s \notin S'$ cannot deviate, because all their friends form a coalition with an enemy. Hence, π is MOS.

$\Longleftarrow$: Suppose now that (N, v) has an MOS partition π. We construct an exact cover $S' \subseteq S$. First, we make some observations:

1. Agents $b^r_{i,2}$ must have $\pi(b^r_{i,2}) \subseteq B^r_i$. If there was an agent $a \in \pi(b^r_{i,2}) \setminus B^r_i$, then, as $v_{b^r_{i,2}}(a) = -n$, $b^r_{i,2}$ would rather be in a singleton, and could form one, as $\left|F_{\text{out}}(\pi(b^r_{i,2}), b^r_{i,2})\right| \geq |\{a\}| = 1 = \left|\{b^r_{i,1}\}\right| \geq \left|F_{\text{in}}(\pi(b^r_{i,2}), b^r_{i,2})\right|$.
2. Using observation 1, we can conclude that agents $b^r_{i,3}$ must also have $\pi(b^r_{i,3}) \subseteq B^r_i$.
3. Using observations 1 and 2, we can conclude that agents $b^r_{i,4}$ must also have $\pi(b^r_{i,4}) \subseteq B^r_i$.
4. Agents $a \in A^s$ and $a' \in A^{s'}$ with $s \neq s'$ satisfy $\pi(a) \neq \pi(a')$. For contradiction, suppose this is not the case, i.e., there are $a \in A^s$ and $a' \in A^{s'}$ with $s \neq s'$ such that $\pi(a) \neq \pi(a') =: C$. Clearly, both prefer to be in a singleton coalition. Further, we can assume without loss of generality that $|A^s \cap C| \leq \left|A^{s'} \cap C\right|$ (otherwise, we can just swap them). Then, as $|F_{\text{out}}(C, a)| \geq \left|A^{s'} \cap C\right| \geq |A^s \cap C| > |F_{\text{in}}(C, a)|$, agent a could deviate to form a singleton coalition, a contradiction.
5. Agents $b^r_{i,1}$ must be in a coalition with no other agents from B^r_i and at least one other agent from $N \setminus B^r_i$. This follows from observations 1, 2, and 3 in conjunction with the fact that the subgame induced by B^r_i is identical to the example from the second part of Proposition 3 which has no MOS partition. Due to the valuations for agent $b_{i,1}$, some agent a^s_r must be in her coalition, and due to observation 4, there can be at most one such agent in her coalition. If there were further agents from A^s in her coalition, $b^r_{i,1}$ could deviate to a singleton coalition. Thus, the only possibility is that $b^r_{i,1}$ is in a pair with exactly one agent a^s_r.

We now know that for each $r \in R$, exactly $n_r - 1$ of the agents a^s_r must be in pairs with $b^r_{i,1}$. This leaves exactly one agent a^s_r not in a pair. For these agents we have $\pi(a^s_r) \subseteq A^s$. Also, $\pi(a^s) \subseteq A^s$, as any agent outside would like to leave

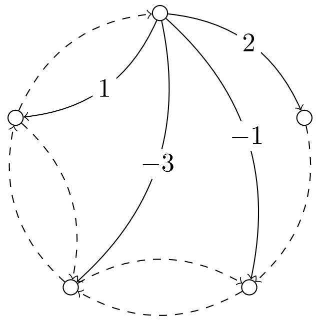

Figure 6: The ASHG without SMS partition from Proposition 4. Outgoing edges with weights have been drawn explicitly only for one agent, they are the same for each agent (up to rotation).

and there is at most 1 vote for her to stay. Consequently, $|\pi(a_r^s)| \geq 2$, and members from $A^s \setminus \pi(a_r^s)$ would have an incentive to join $\pi(a_r^s)$. Hence, $\pi(a_r^s) = A^s$, and we obtain a cover $S' = \{s \in S \colon A^s \in \pi\}$. □

Before we proceed with the proof of Theorem 8, we want to provide an example of a game where no SMS partition exists.

Proposition 4. *There exists an ASHG without SMS partition.*

Proof. Let $N = [5]$ and consider the utilities according to Table 2 below.

Table 2: Valuations for an ASHG without SMS partition.

v	1	2	3	4	5
1	0	2	−1	−3	1
2	1	0	2	−1	−3
3	−3	1	0	2	−1
4	−1	−3	1	0	2
5	2	−1	−3	1	0

See Figure 6 for a graphical representation of this example. We show that no partition can be SMS by an exhaustive case analysis. Let $+_{[5]}$ denote addition modulo 5, mapping to the representative in $[5]$. Assume for contradiction that π is SMS, and $C \in \pi$ is a coalition of largest cardinality.

- Suppose $|C| = 5$. Then $\pi = \{N\}$, and all agents can form a singleton via an SMS-deviation.
- Suppose $|C| = 4$. Then we can write it as $\{i, i +_{[5]} 1, i +_{[5]} 2, i +_{[5]} 3\}$ for some $i \in N$, and agent i can form a singleton via an SMS-deviation.
- Suppose $|C| = 3$. Then it is either of the form $\{i, i +_{[5]} 1, i +_{[5]} 2\}$ or of the form $\{i, i +_{[5]} 1, i +_{[5]} 3\}$ for some $i \in N$. In the first case, agent $i +_{[5]} 2$ can form a singleton coalition, in the second case, agent $i +_{[5]} 3$ can form a singleton coalition.
- Suppose $|C| = 2$. Then π also has to contain a singleton $\{i\}$. If $\pi(i +_{[5]} 1) \in \{\{i +_{[5]} 1\}, \{i +_{[5]} 1, i +_{[5]} 2\}\}$, then i can join $i +_{[5]} 1$ via an SMS-deviation. If $\pi(i +_{[5]} 1) \in \{\{i +_{[5]} 1, i +_{[5]} 3\}, \{i +_{[5]} 1, i +_{[5]} 4\}\}$, then $i +_{[5]} 1$ can join i via an SMS-deviation.
- Suppose $|C| = 1$. Then any agent i can join $i +_{[5]} 1$ via an SMS-deviation.

□

Theorem 8. *Deciding whether an ASHG contains an SMS (respectively, JMS) partition is* NP*-complete.*

Proof. We provide a polynomial-time reduction from E3C that simultaneously works for JMS and SMS. Let (R, S) be an instance of E3C. We produce an ASHG (N, v) such that for all $\alpha \in \{\text{JMS, SMS}\}$, (R, S) has an exact cover if and only if (N, v) has a partition that is α. Define the agent set $N = \bigcup_{s \in S} A^s \cup \bigcup_{r \in R} \bigcup_{i=1}^{n_r} B_i^r$, where $A^s = \{a_r^s \colon r \in s\}$ for $s \in S$ and $B_i^r = \{b_{i,j}^r \colon j \in [5]\}$ for $r \in R, i \in [n_r - 1]$.

Also, define utilities v as follows:

- For each $s \in S, a \neq a' \in A^s : v_a(a') = 2$.
- For each $r \in R, s \in S_r, i \in [n_r - 1] : v_{a_r^s}(b_{i,1}^r) = 1, v_{b_{i,1}^r}(a_r^s) = 0$.
- Each B_i^r has internal utilities as in the example constructed in Proposition 4, i.e., if v' are the utilities in the example, then $v_{b_{i,j}^r}(b_{i,k}^r) = v'_j(k)$.
- All other valuations are $-M$, where $M = |S| + 5$ (can be thought of as $-\infty$).

The reduction is visualized in Figure 7. Note that the it can be performed in polynomial time, as there are at most $3|S| + 5|R||S|$ agents. We proceed with the proof of the correctness of the reduction and show that if (R, S) has an exact cover, then (N, v) also has a JMS and SMS partition, and conversely if (N, v) has a partition π that is either JMS or SMS, then there is an exact cover in the instance (R, S).

$\Longrightarrow$: Suppose (R, S) has an exact cover $S' \subseteq S$. We construct a stable partition π.

- We have coalitions corresponding to the cover, i.e., for each $s \in S : A^s \in \pi \iff s \in S'$.
- This leaves for each $r \in R$ exactly $n_r - 1$ sets $s \in S_r$ such that $A^s \notin \pi$. Arbitrarily number these sets $s_1, \ldots, s_{n_r - 1}$ and define for each $i \in [n_r - 1]$ the coalitions $\{a_r^{s_i}, b_{i,1}^r\}, \{b_{i,2}^r, b_{i,3}^r\}, \{b_{i,4}^r, b_{i,5}^r\}$.

We claim that this partition is JMS and SMS. To see this, note that the only agents that have incentive to deviate are agents of type $b_{i,1}^r$ who would prefer to join $\{b_{i,2}^r, b_{i,3}^r\}$. Fix any such agent $b_{i,1}^r$. The agent a_r^s she is paired with would vote against her leaving, so the partition is MOS and thus SMS. To see that it is also JMS, note that even though $b_{i,2}^r$ would vote in favor of the deviation, $b_{i,3}^r$ is against it, which together with the against-vote of a_r^s ensures that there is a strict joint majority against the deviation.

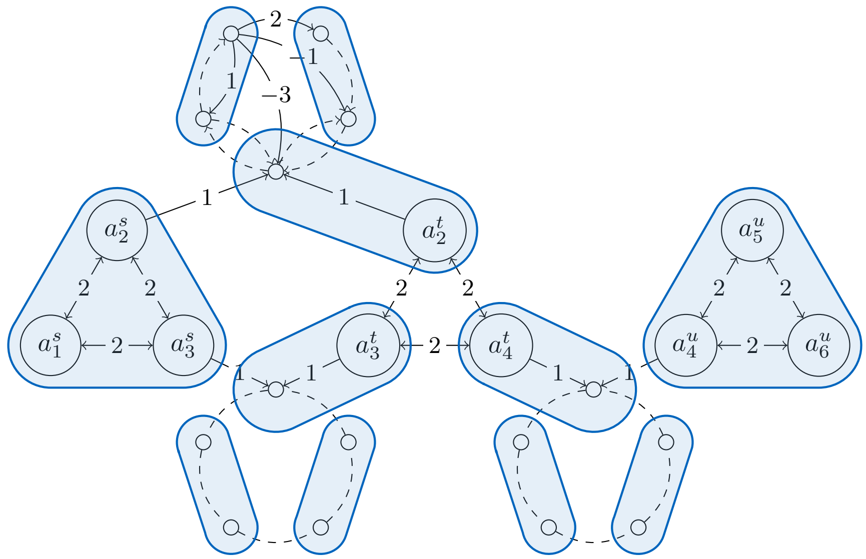

Figure 7: Schematic of the reduction from the proof of Theorem 8 for the Yes-instance of E3C $(\{1,\dots,6\},\{s,t,u\})$ with $s=\{1,2,3\}, t=\{2,3,4\}$ and $u=\{4,5,6\}$. Some edges have been omitted for clarity. The indicated partition is both SMS and JMS.

$\Longleftarrow$: Suppose there is a partition π that is JMS or SMS. We show that then there must be an exact cover $S' \subseteq S$ of R. We begin with some observations:

1. Agents $b^r_{i,j}$ with $j \in \{2,\dots,5\}$ must have $\pi(b^r_{i,j}) \subseteq B^r_i$. For contradiction, suppose this is not so. Consider first the case that there is exactly one outside agent $a \in \pi(b^r_{i,j}) \setminus B^r_i$. Then, as $v_a(b^r_{i,j}) = -M$, a has incentive to form a singleton coalition, and this is a valid SMS-deviation (and therefore JMS-deviaton). The other case is that there are at least two agents $a \neq a' \in \pi(b^r_{i,j}) \setminus B^r_i$. Then, as $v_{b^r_{i,j}}(a) = -M$ and $\left|F_{\text{out}}(\pi(b^r_{i,j}), b^r_{i,j})\right| \geq |\{a,a'\}| = 2 = \left|\left\{b^r_{i,j+_{[5]}1}, b^r_{i,j+_{[5]}4}\right\}\right| \geq \left|F_{\text{in}}(\pi(b^r_{i,j}), b^r_{i,j})\right|$, $b^r_{i,j}$ can form a singleton coalition.
2. Agents a^s_r and $a^{s'}_{r'}$ with $s \neq s'$ have $\pi(a^s_r) \neq \pi(a^{s'}_{r'})$. For contradiction, suppose the contrary, i.e., suppose that there are a^s_r and $a^{s'}_{r'}$ with $s \neq s'$, but $\pi(a^s_r) = \pi(a^{s'}_{r'}) =: C$. As $v_{a^s_r}(a^{s'}_{r'}) = v_{a^{s'}_{r'}}(a^s_r) = -M$, both would rather be in a singleton coalition. Further, we can assume without loss of generality that $|A^s \cap C| \leq \left|A^{s'} \cap C\right|$ (otherwise, we can just swap them). Then, as $|F_{\text{out}}(C, a^s_r)| \geq \left|A^{s'} \cap C\right| \geq |A^s \cap C| > |F_{\text{in}}(C, a^s_r)|$, a^s_r can deviate to form a singleton coalition.
3. Agents $b^r_{i,1}$ must be in a pair with exactly one agent a^s_r. Fix such an agent $b^r_{i,1}$. First, due to observation 1, she cannot be alone, and no other agents from B^r_i can be in her coalition, as the example constructed in Proposition 4 has no SMS partition. Consequently, she must form a coalition with at least one agent outside of B^r_i, and no agents from B^r_i. Next, due to observation 2, she can be together with at most one agent of type a^s_r. If there was another member from A^s (other than a^s_r), $b^r_{i,1}$ could deviate to a singleton coalition.

We now know that for each $r \in R$, exactly $n_r - 1$ of the agents a^s_r must be in pairs with agents $b^r_{i,1}$. This leaves exactly one agent a^s_r not in a pair. We claim that for these agents we have $\pi(a^s_r) = A^s$, yielding a cover $S' = \{s \in S \colon A^s \in \pi\}$. Suppose that a^s_r is such an agent not in a pair. Then, $\pi(a^s_r) \subseteq A^s$. If the other two agents from A^s form a pair, then a^s_r has an incentive to join them. Otherwise, the other two agents would have an incentive to join a^s_r. In any case, the only stable situation is $\pi(a^s_r) = A^s$. $\square$